International Journal of Computer and Engineering Optimization (IJCEO)
Volume 02, Issue 01, January – February (2024)

RESEARCH ARTICLE

DEEP FIRE: HONEYPOT COMBINED DEEP LEARNING BASED FIREWALL TUNING FOR ENHANCING CYBERSECURITY

Vishnu Karthik Ravindran[1, *]

[1]Independent Researcher.

[*]*Corresponding e-mail: vkacharaya@gmail.com*

Abstract – Cyber security of internet-based cloud networks is a major concern for companies. Networks are always open to intrusions from both inside and outside of organizations. However, using traditional techniques to identify behavioral changes or malicious attacks is challenging. In this paper, a novel DEEP learning-based FIREwall tuning (Deep Fire) method has been developed to integrate deep learning and honeypot in order to mitigate and prevent the attacks of cloud systems. In order to identify ransomware activity and attack patterns, Apache spark engine is used which combines information from TrAck Replay Evaluate (TAPE) systems. The proposed technique uses Convolutional neural Network (CNN) for detecting the intrusion into two classes such as attack and normal. The proposed system has been evaluated using python simulator. The proposed technique has been evaluated using specific parameters such as detection accuracy and false alarm rate. The proposed system achieves higher detection accuracy of 22.6%,16.6% and 2.38% than the existing systems such as IDPS, virtual machine (ICI) and Honeypot based IDS technique respectively. By utilizing Apache Spark, the proposed DEEP-FIRE correlates network features, host attributes, and various events from other systems like TAPE and firewall to produce more accurate results.

Keywords – *Cybersecurity, Honeypots, Honeynet, Deep Learning, Firewall.*

1. INTRODUCTION

The Internet is the most influential technology that is being utilized in every aspect of day-to-day activities [1]. Various organizations use the Internet to accomplish their tasks. besides, many organizations expand their services and trust external service providers to attain their needs and requirements. The need of huge data analysis and computation power cannot be satisfied within a single company. Therefore, to satisfy and expand their service scopes, organizations rent and buy external services and processing applications. The Internet interconnects a number of computer networks and thus it simplified the interaction between different sectors across the world [2]. As a result, the need and dependency on the Internet are growing exponentially. The interconnection of a large number of networks brings emerging of various kinds of protocols and computer programs. These protocols and computer programs are the backbone of the operation of the Internet. But, as computer programs and protocols are getting more complex, they will have also implementation defects. Numbers of computer applications have vulnerabilities that can be exploited by attackers [3].

The harmful users use various attack techniques like sniffing, clear text traffic, password cracking, unencrypted, denial of service attacks, etc [4]. The security for the private resources through the internet had to tighten because most of the attacks are from an insider of the organization. However, using a firewall to secure their network won't produce the desired results every time [5]. Sometimes Firewall does not work with the attacks like denial of services which floods the request and the entire network is down. In these cases, Firewall will not act according to the situation.

To decrease the critical attacks, present in the organization, many security systems have been introduced. The security system used for scanning based on the rules (Firewall) can only prevent the attacks but the need for new attack detection is very important. To make such detection Honeypot technologies are widely used [6]. The honeypot is a computer resource that is used to monitor processes that require testing and compromise. Honeypot is characterized as "a data framework having the details of resources which are not authorized". The deployed Honeypot are designed in such a way that it can be accessed to the log and changes can be made in the program within Honeypot. The deployment of Honeypot can be made in and out of the network or even in the demilitarized zone depending on the application. Honeypot are additionally used to gain information and record details about the attackers. It is also used for tracing the stolen data and also obtains information about the attackers involved in malicious activities [7].

The attackers exploit the programs that are running in various systems to attain their malicious intents. In most of the cases, the vulnerabilities of programs are not discovered at the early stage. Firewalls have the ability to block and filter malicious traffics. The other network security solutions devised are IDSs. They are intended to detect previously

known assaults. nevertheless, the disadvantage of the IDSs is that they are unable to detect new kinds of attacks. Honeypots address the problems of IDS and firewalls [8]. They are network security entities that are able to detect new attacks and methods. They help network security experts and researchers to study almost all kinds of attacks, including the attackers' detail stepwise technique. In this paper, a novel DEEP learning-based FIREwall tuning (Deep Fire) method has been developed to integrate deep learning in order to mitigate and prevent the attacks of cloud systems. The key goals of the proposed technique have been given as follows.

- The aim of the proposed technique is to identify the intrusions and tune the firewall for securing the organization data.
- In order to identify ransomware activity and attack patterns, Apache spark engine is used which combines information from TrAck Replay Evaluate (TAPE) systems.
- The proposed technique uses Convolutional neural Network (CNN) for detecting the intrusion into two classes such as attack and normal.
- The proposed technique has been evaluated using specific parameters such as detection accuracy and false alarm rate.

The structure of this paper is as follows: Section II reviews related work. Section III outlines the proposed system model. Section IV presents the results, and Section V concludes the paper with recommendations for future research.

2. LITERATURE

In 2020, Masdari, M. and Khezri, H. [9] proposed a complete analysis of fuzzy misuse detection strategies meant to deal with various types of intrusions utilizing several ML and data mining techniques. The article then outlines the major contributions of fuzzy IDS approaches, as well as their benefits and drawbacks. In addition, comparisons of their performance evaluation factor, employed datasets, feature extraction methods, membership functions, and fuzzy logic controller (FLC) type are provided in each domain.

In 2021, S. Jin, et al. [10] created a signature-based light-weight intrusion detection system that can be quickly and directly deployed to car Electronic Control Units (ECUs). Experiment findings indicate that the suggested technique can successfully detect CAN traffic anomalies. A drop assault is detected 100% of the time, whereas a replay attack is detected 98.2% of the time [16]. However, only 66.2% of tempering attacks are detected. If the relationship between signals is explored, better results can be predicted.

In 2021, Rani, S. [11] proposed Intrusion Detection and Prevention System (IDPS) is a sophisticated cloud security solution that detects any anomalous network activities. This study examines the IDS, IPS, and hybrid methods to IDPS, as well as a comparison of traditional IDPS and cloud IDPS. The IDPS is designed to detect harmful assaults that include both known and unknown threats. Finally, the purpose of this work is to develop a method for detecting and inhibiting infiltration in the cloud environment, as well as to aid in the resolution of the problem of attacks in order to secure the cloud environment.

In 2021, Buzzio-Garcia, J. [12] recommends using Docker as a high-interaction honeypot to identify risks at both the network and host levels. In order to assure scalability, security, and dynamic functionality, it was created utilising open-source tools. It has been proven in a real-world test to be capable of capturing hazardous data for analysis on the network and host level using programmers like Virus Total.

In 2021, Feng, M et al. [13] propose a novel form of honeypot system based on deception-based technologies. Honeypots acquire idle IP addresses while keeping their identities hidden by using dynamic deception strategies. The experimental findings reveal that the suggested honeypot system successfully extends the monitoring range of conventional honeypots and has a great defensive impact versus unknown threats, thereby compensating for previous defense systems' inadequacies

In 2021, K.D. Singh [14] developed a corporate honeypot to protect cloud infrastructure virtual machines (VMs) (ICI). It has been proven in a real-world test to be capable of capturing hazardous data for analysis on the network and host level using programmes like Virus Total. The information gathered via snort is beneficial to the invader's actions. There have been two major discoveries. First, ICI must protect virtual machines from inside invaders. Second, the Cloud's infrastructure must be more secure. The findings indicate that more effective security warnings may be developed to improve security

In 2022, Sivamohan, S et al. [15] a honeynet-based IDS and Docker containers were utilised to build an effective active protection architecture. The development of honeynet technology is essential for threat detection and cloud security. According to the experiment's findings, it appears that this defence system can recognise and log the activities of the attacker, exposing fresh attack tactics and even zero-day vulnerabilities.

3. PROPOSED METHOD

To Intrusion detection is the process of preventing and detecting hostile attempts to breach the security of a network system. Anomaly intrusions are hard to detect. Network IDS (NIDS) as they are designed to control the entire network, they face a challenge of resource limitations. A NIDS must capture and handle massive amounts of data in real time. Sometimes, IDS can be a victim of attacks such as Denial of Services (DOS) attack, which may lead IDS to crash. Honeypots, which are popular network security mechanisms, facilitate monitoring and study newly emerging attacks, Honeypot systems may help to safeguard computer networks by informing users about security flaws in the network ahead of time. Therefore, in this work a novel DEEP learning-based FIREwall tuning (Deep Fire) method has been developed to integrate honeypot and deep learning in order to mitigate and prevent the attacks of cloud systems. In order to identify ransomware activity and attack patterns, Apache spark engine is used which combines information from TrAck

Replay Evaluate (TAPE) systems. The proposed technique uses Convolutional neural Network (CNN) for detecting the intrusion into two classes such as attack and normal. Figure 1 represents the block diagram of the DEEP-FIRE method.

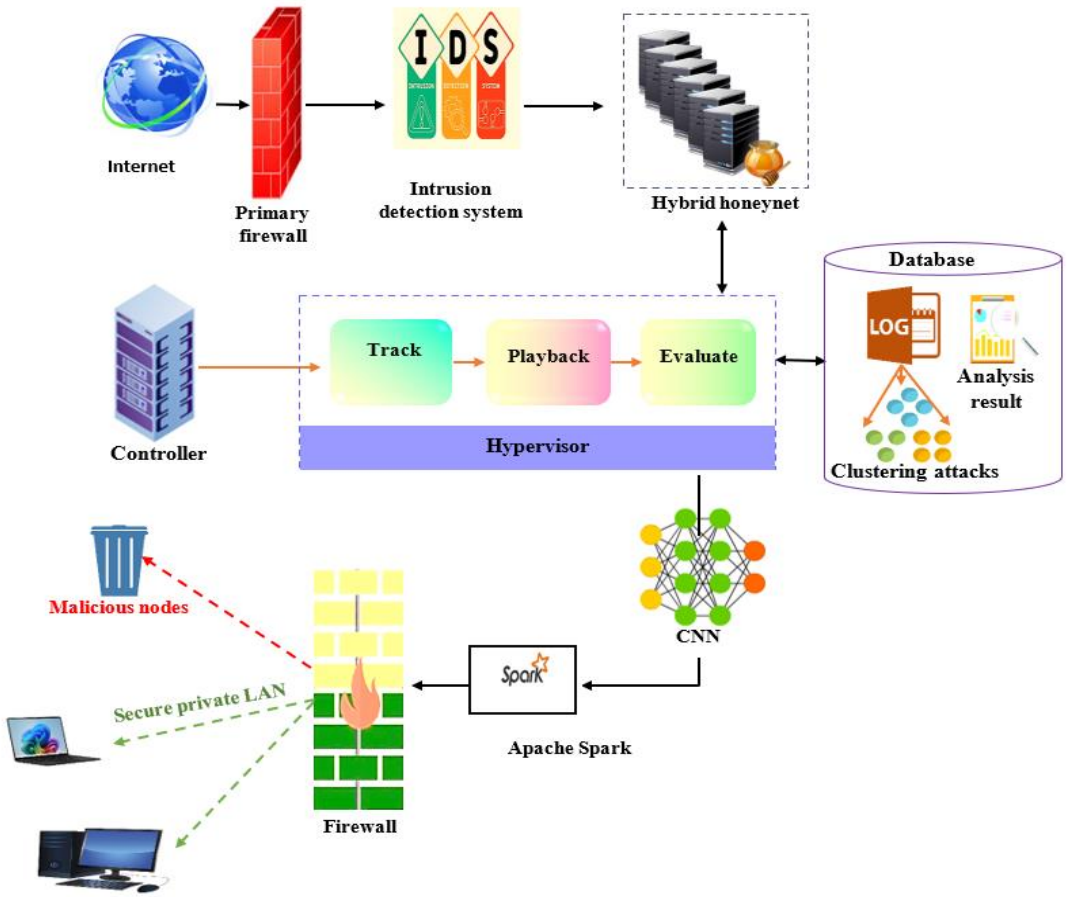


Figure 1. Architecture of proposed DEEP-FIRE method

3.1. Hybrid honeynet

Honeypots have mostly been employed to detect and analyze harmful behavior by risking their own resources in order to be attacked. In research, hybrid honeypot architectures made up of frontends and backends are frequently used, because to the benefits of high fidelity and scalability for thorough attacking data collection. It is often necessary for a hybrid honeypot system to have the ability to strictly regulate network traffic, for example, directing traffic from frontends to backends for comprehensive attack analysis. False alarms can occur while IDS are watching networks for possibly dangerous activity. Consequently, a hybrid honeynet system has been suggested in addition to IDS. To prevent honeypots from developing into autonomous entities, they are assembled into groups known as honeynets. Low- and high-interaction honeypots are combined in a hybrid honeynet.

3.2. TAPE framework

TAPE framework has been employed for tracking, playback and evaluate the attacks, and helps in identifying attacks. In honeypot operations, the task of monitoring non-deterministic occurrences falls to the Track Component. In order to replay the prior honeypot operation, the Playback Component creates a Honeypot Replier during the playback stage. It extracts non-deterministic events from log files and sends them to be played back via Hypervisor. The execution of the honeypot will be reconstructed based on the logged data. The rebuild data will be given to the Evaluator and the analyzer will compare the rebuild data with the real time data and the evaluated result will be stored in the database.

3.3. Apache Spark

The Apache Spark platform is used for streaming and storing events. Real-time data flows will be handled by a uniform, high-throughput, low-latency platform developed by the project. The process begins with the Spark Core API primarily in languages like Java, or Python which provides essential operations for distributed data handling across clusters. After establishing this core, data ingestion takes place, where data from logs, network packets, or security events is connected to Spark using sources like Apache Spark, HDFS, or cloud storage. This data is then loaded into Spark's Data Frame or RDD, enabling distributed storage and computation across nodes.

3.4. Convolutional neural network (CNN)

In proposed Deep learning-based firewall (Deep Fire) technique, an effective CNN is deployed to extract the specific data from the exploited data. The CNN models are well-known feature extractors with adaptable architectures that are utilized for a variety of duties and applications. A

basic CNN design consists of convolutional, activation, pooling, and fully connected layers arranged in a certain topology. From low-level features to more complicated features, the retrieved features can vary depending on the model's topological depth. When utilizing a fixed kernel, the CNN block executes a 1D convolution operation after receiving the pre-processed input dat., generating maps of activation.

Let's say that the CNN's input characteristic is the feature map of layer j, which is $Q_j(Q_0 = Y)$

$$Q_j = f(Q_{j-1} \otimes U_j + a_j) \quad (1)$$

where U_j is the convolution kernel weight vector of the j layer, accordingly; Convolution is represented by the operation symbol "⊗ "; The activation function is f (y), and a_j is the offset vector of the j layer. By defining various window values, the convolutional layer retrieves distinct feature information from the matrix of data Q_{j-1}, and it uses various convolution kernels to extract distinct features Q_j from the data. The same convolution kernel uses the "parameter sharing" approach in the convolution operation, Therefore, by using the same weight and offset, drastically reduces the neural network's overall number of parameters. The pooling layer typically follows the convolutional layer in sampling the feature map using various sampling techniques.

Assume that the pooling layer receives Q_j as input and outputs Q_{j-1}.

Thus, the pooling layer can be expressed as

$$Q_j = Subsampling(Q_j) \quad (2)$$

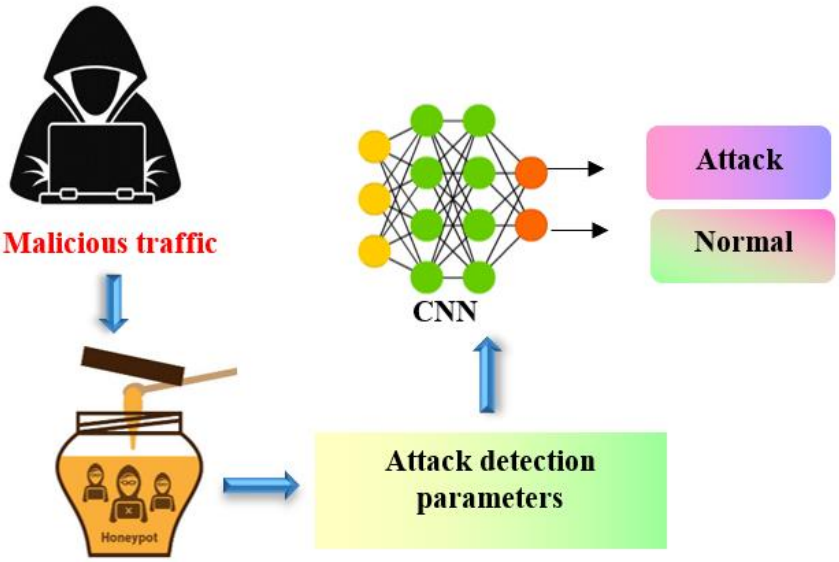


Figure 2. CNN in cyber-attack prediction

Typically, the sample criterion selects the mean or maximum value in the window region. The main function of the pooling layer is to decrease the dimension of redundant features in order to decrease their influence on the model. Figure 2 represents a CNN in cyber-attack prediction.

4. RESULT AND DISCUSSION

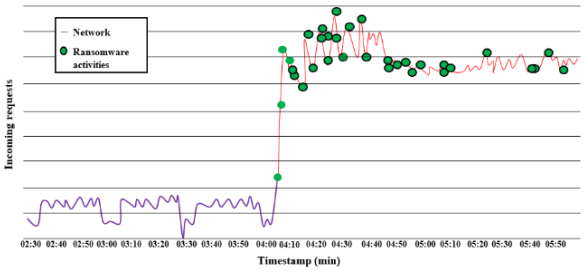
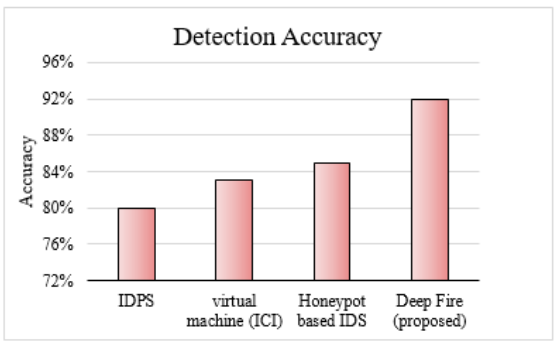
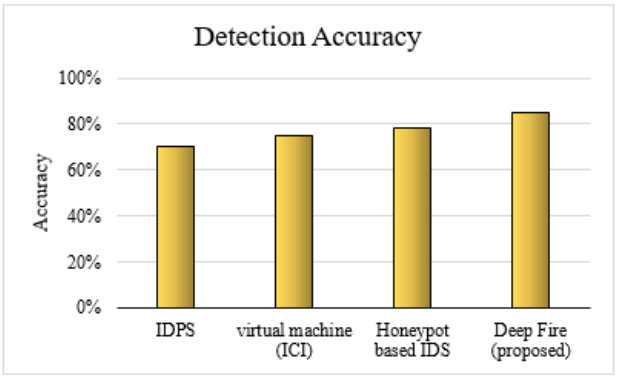


Figure 3. Ransome ware activities detected by deep learning

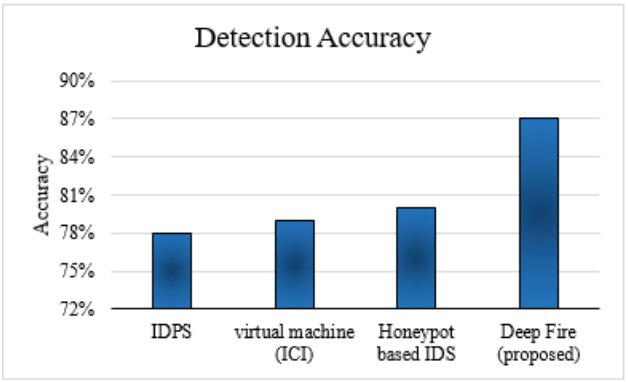
To The Ransome ware operations at a certain timestamp are shown in Figure3.Based on the findings, Ransome ware communicates with users using more than 200private proxies and 100C onion sites. However, all traffic coming from the aforementioned domains is filtered and blacklisted by Apache Spark. The firewall immediately terminates connectivity and warns the system when infected hosts try to connect to a forbidden domain.



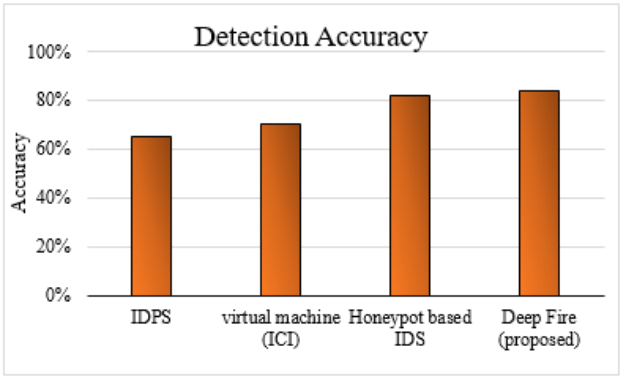
(i)



(ii)



(iii)



(iv)

Figure 4. Detection accuracy comparison

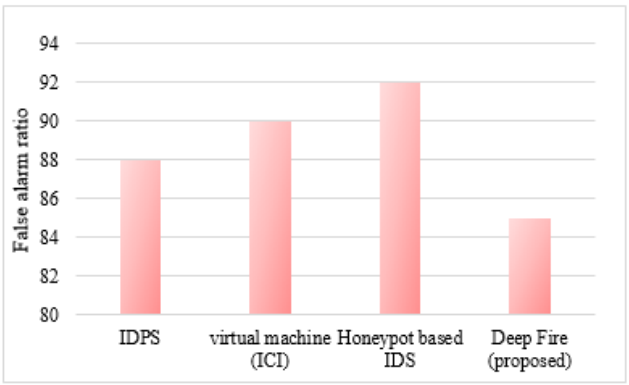


Figure 5. False alarm ratio

The detection accuracy of the proposed Deep Fire method is compared to that of current techniques like IDPS, Virtual machine (ICI) and Honeypot based IDS in Figure 4. The detection accuracy when there are 10 attackers is shown in Figure 4 (i), The detection accuracy when there are 20 attackers is shown in Figure 4 (ii), The detection accuracy when there are ten attackers is shown in Figure 4 (iii), The detection accuracy when there are twenty attackers is shown in Figure 4(iv).

Figure 5 represents a graphical representation of false alarm ratio comparison of the proposed Deep Fire technique with the existing technique. The proposed technique achieves 3.5%,11.7% and 8.2% has lower false alarm ratio than IDPS, virtual machine (ICI) and Honeypot based IDS technique respectively.

5. CONCLUSION

In this proposed work, a novel Deep Fire technique is implemented for intrusion detection in fire wall. The goal of the proposed Deep Fire is to protect and prevent the cloud systems from various attacks and to identify the ransomware activity. In order to generate more precise findings, the suggested Deep Fire correlates network data, host attributes, and different events from other systems like TAPE and firewall using Apache Spark. CNN implementation significantly improves network security. The experimental evaluation also shows that the proposed Deep Fire effectively restricts ransomware activity with little loss of data. The proposed system achieves higher accuracy of 22.6%,16.6% and 2.38% than the existing systems such as IDPS, virtual machine (ICI) and Honeypot based IDS technique respectively. In future, accuracy can be improved by implementing different deep learning algorithms.

CONFLICTS OF INTEREST

This paper has no conflict of interest for publishing.

FUNDING STATEMENT

Not applicable.

ACKNOWLEDGEMENTS

The author would like to express his heartfelt gratitude to the supervisor for his guidance and unwavering support during this research for his guidance and support.

REFERENCES

[1] F. Anjum, D. Subhadrabandhu, and S. Sarkar, "Signature based intrusion detection for wireless ad-hoc networks: A comparative study of various routing protocols," In *2003 IEEE 58th Vehicular Technology Conference. VTC 2003-Fall (IEEE Cat. No. 03CH37484)*, vol. 3, pp. 2152-2156, 2003, October. IEEE. [CrossRef] [Google Scholar] [Publisher Link]

[2] N. Weiler, "Honeypots for distributed denial-of-service attacks," In *Proceedings Eleventh IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises*, pp. 109-114, 2002, June. IEEE. [CrossRef] [Google Scholar] [Publisher Link]

[3] N.C. Paxton, "Development of a multi-layered botmaster based analysis framework (Doctoral dissertation," *The University of North Carolina at Charlotte*, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[4] M. Ring, S. Wunderlich, D. Scheuring, D. Landes, and A. Hotho, "A survey of network-based intrusion detection data sets," *Computers & Security*, vol. 86, pp.147-167, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[5] A. Mairh, D. Barik, K. Verma, and D. Jena, "Honeypot in network security: a survey," In *Proceedings of the 2011 international conference on communication, computing & security*, pp. 600-605, 2011, February. [CrossRef] [Google Scholar] [Publisher Link]

[6] M. Nawrocki, M. Wählisch, T.C. Schmidt, C. Keil, and J. Schönfelder, "A survey on honeypot software and data analysis," arXiv preprintarXiv:1608.06249, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[7] S. Guo, Y. Lin, N. Feng, C. Song, and H. Wan, "Attention based spatial-temporal graph convolutional networks for traffic flow forecasting," In *Proceedings of the AAAI conference on artificial intelligence*, vol. 33, no.01, pp. 922-929, 2019, July. [CrossRef] [Google Scholar] [Publisher Link]

[8] I. Kuwatly, M. Sraj, Z. Al Masri, and H. Artail, "A dynamic honeypot design for intrusion detection," In *the IEEE/ACS International Conference on Pervasive Services*, 2004. ICPS 2004. Proceedings, pp. 95-104, 2004, July. IEEE. [CrossRef] [Google Scholar] [Publisher Link]

[9] M. Masdari, and H. Khezri, "A survey and taxonomy of the fuzzy signature-based intrusion detection systems," *Applied Soft Computing*, vol. 92, pp. 106301, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] S. Jin, J. G. Chung, and Y. Xu, "Signature-Based Intrusion Detection System (IDS) for In-Vehicle CAN Bus Network," *2021 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1-5, 2021. doi: 10.1109/ISCAS51556.2021.9401087. [CrossRef] [Google Scholar] [Publisher Link]

[11] S. Rani, "A Perspective for Intrusion Detection & Prevention in Cloud Environment," *International Journal of Advanced Networking and Applications*, vol. 12, no. 6, pp. 4770-4775, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[12] J. Buzzio-Garcia, "Creation of a High-Interaction Honeypot System based-on Docker containers," In *2021 Fifth World Conference on Smart Trends in Systems Security and Sustainability (WorldS4)*, pp. 146-151, 2021, July. IEEE. [CrossRef] [Google Scholar] [Publisher Link]

[13] M. Feng, B. Xiao, B. Yu, J. Qian, X. Zhang, P. Chen, and B. Li, "A Novel Deception Defense-Based Honeypot System for Power Grid Network," In *International Conference on Smart Computing and Communication*, pp. 297-307, 2021, December. Springer, Cham. [CrossRef] [Google Scholar] [Publisher Link]

[14] K.D. Singh, "Securing of Cloud Infrastructure using Enterprise Honeypot," In *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, pp. 1388-1393, 2021, December. IEEE. [CrossRef] [Google Scholar] [Publisher Link]

[15] S. Sivamohan, S.S. Sridhar, and S. Krishnaveni, "Efficient Multi- platform Honeypot for Capturing Real-time Cyber Attacks," In *Intelligent Data Communication Technologies and Internet of Things*, pp. 291-308, 2022. Springer, Singapore. [CrossRef] [Google Scholar] [Publisher Link]

[16] S. Raja, S. Pran, N. Pandeeswari, P. Kiruthiga, D. Nithya, and G. Muthu Pandi, "Contemporary PCA and NBA based Hybrid Cloud Intrusion Detection System," *EAI Endorsed Transactions on Energy Web*. vol. 8, no. 36, 2021 Feb 19. [CrossRef] [Google Scholar] [Publisher Link]

AUTHORS

Vishnu Karthik Ravindran earned his Bachelor of Engineering in Computer Science and Engineering from Sri Ramakrishna Institute of Technology, India, in 2008. He later pursued a Master of Science in Computer Science at Syracuse University, New York, USA. Currently, he is a Software Engineer at a leading cloud-based company and actively contributes to open-source projects.

Arrived: 25.01.2024
Accepted: 20.02.2024