International Journal of Science Engineering and Advance Technology, IJSEAT, Vol. 13, Issue 1 | ISSN 2321-6905 May- 2025

Integrating Artificial Intelligence into Human-In-The-Loop Cybersecurity: Techniques and Advances

T.Poonkodi
Assistant Professor, Department of Computer Science,
Hindusthan College of Science and Commerce,
Ingur, Perundurai.
ponrohit.0707@gmail.com

Abstract - The process of defending computer networks against cyberattacks or unintentional unauthorized access has become known as cybersecurity. For businesses and security experts who are constantly searching for improved ways to combat these threats, this constantly changing threat landscape presents difficulties. By integrating cybersecurity with artificial intelligence, security professionals can better protect sensitive networks and data from online threats. For identifying malware attacks, network intrusions, phishing and spam emails, and data violations, among other threats, AI-based approaches can offer strong and efficient cyber defence solutions that can notify security problems as quickly as they happen. The objective of this research paper was to evaluate the impact of AI-based technologies on organizational cyber security and determine their effectiveness compared to traditional cyber security approaches. The proposed Enhanced Human-in-the-Loop Cybersecurity Techniques effectively prevent and detect uncertain cyber events.
.

Keywords*: cyber attacks, malware attacks, network intrusions, phishing emails, Enhanced Human-in-the-Loop*

I. Introduction

In order to prevent attacks, damage, or unauthorised access to networks, devices, software, and data, a collection of technologies, procedures, and practices is referred to as cybersecurity [1]. The rapid development of interconnected gadgets, systems, and networks is making cybersecurity more complicated.

Cybersecurity is the term for the tools and practices used to guard against harm or uncontrolled access to networks and data. It is essential because a lot of data is gathered, processed, and stored by businesses, governments, and military institutions. There are several types of cybersecurity, including information systems, infrastructure, intelligence, judicial, military, law enforcement, and commerce. Cybersecurity is a dynamic, multidisciplinary field that encompasses computer science, criminology, and information systems. The security goals have been integrity, nonrepudiation, secrecy, availability, and authentication.

Businesses, banks, corporations, and government networks are all at risk from cyberattacks. They range from organised activities (terrorists) to individual citizen crimes (hacking). Malware, phishing, denial-of-service attacks, social engineering, and man-in-the-middle assaults are examples of cyberthreats. The goal of cybersecurity is to lower the likelihood of cyberattacks. The management should take a proactive approach to managing cyber risks. Firewalls and other cybersecurity technology are generally accessible [3].

II. Related Works

Bhardwaj et al., [1] proposed a secure framework based on attack trees that uses the most important exploitable flaws to carry out the attacks. This study used the suggested framework in the form of a two-phase procedure to mimic real-time exploitation of vulnerabilities on CPS robotic systems. During real-time cyberattacks, this verifies the improved privacy of the data output of the incorporated sensor and real-world nodes with the controller system health monitor and intelligent monitor. This study used telnet pivoting and cross-site scripting to mimic conventional cyberattacks on cyber-physical controller systems. Using a tree-based attack method, the authors collected known and unknown vulnerabilities. The typical time it takes for cybercriminals with varying levels of expertise to attempt to breach CPS systems and devices.

M. Zajko et al., [2] proposed the techniques uses machine learning, where experts in IT create algorithms using data gathered over time. Through the method in which the algorithm is constructed, it can recognise and differentiate between legitimate and fraudulent access. Machine learning technology improves an organization's security by making threats and anomalies more predictable. Humans cannot

match the speed and accuracy of threat identification. Consequently, machine learning and artificial intelligence technology can prevent cyberattacks that could cost an organisation millions of dollars.

R. Winkels et al., [3] proposed the technique to find and stop any improper network behavior. The system has been designed allows it to recognize threats without the need for human assistance. Because of its independence, there is less chance of compromise, which could lead to illegal access to the company's databases. Despite being largely autonomous, the AI does permit the addition of supervised instructions when needed. This is crucial for categorizing the hazards facing the company. For example, the system might be told to classify ransomware and malware attacks according to specific features.

H. Bidgoli et al., [4] proposed the changes in technology that have affected cyber security. A collection of tools and techniques that are augmented by artificial intelligence (AI) as a subgroup is one of the major game changers in the field of cyber security. The term artificial intelligence (AI) is no longer just a catchphrase; it is currently being used extensively in many different fields. AI has rapidly advanced a wide range of fields, including robotics, healthcare, and customer service. It is also playing a significant role in the ongoing fight against cybercrime.

H. Zhuge et al., [5] discover that the many advantages and applications of AI in a variety of fields, including cybersecurity. AI can help keep hackers awake, automate threat detection, and react more quickly than typical software and manual techniques in light of the fast evolving nature of cyberattacks and growing number of technology in today's world.

Wang Z et al., [6] proposed developed a mobile communication network privacy authorization system based on Support Vector Regression (SVR) and suggested the kernel function with both the local and global functions; nevertheless, the simulation utilized less data.

III. Proposed Methodology

Although there are a number of drawbacks that will make it difficult to make a knowledgeable choice, artificial intelligence (AI) is crucial for preventing and identifying high-risk network behaviors.

3.1 Enhanced human in the loop based cyber security techniques (EHLCST)

Improved Human-machine intelligence and human wisdom could be combined in the loop, which is a crucial approach to achieving the complementing benefits of both. AI has a strong recognition impact for particular scenarios and can handle enormous amounts of data rapidly. Humans are more adaptable than robots and can make decisions more quickly when the network is changing, but machines are also necessary to play a supporting role.

3.2 Structure design of cyber security based on Enhanced human in the Loop

AI technology has many benefits for cyber security applications, it also has drawbacks. Based on this fact, this research presents an Enhanced Human in the Loop based Cyber Security Techniques (EHLCST). The Device Detection Unit (DDU) and Physical Intervention Unit (PIU) are the two primary subunits that make up EHLCST. To stop and identify cyber uncertain events, the two subunits communicate with one another.

3.2.1 Device Detection Unit (DDU)

DDU has the "leading role" in EHLCST. DDU will preprocess the data, which may involve data normalization, data cleansing, and other procedures, when high-risk events occur. Extracting characteristics from data is crucial when the data is regular. The primary information about the type of occurrence is contained in the data, along with other information that is not really related. In order to finish jobs faster, it is vital to decrease dimensions and choose features as the volume of data increases. Only when the selected technique satisfies the requirements, the recognition accuracy will be higher. The Superiority Level Unit (SLU) will evaluate the running result following the recognition result, and its value will establish whether the final result is based on DDU.

Two ways of identification are employed in DDU. The current use of a single recognition method does not ensure total dependability of the output since neural networks can deploy backdoors (Gu et al. 2017). As a result, the knowledge base serves as the evaluation basis in DDU design, and two recognition techniques are employed to produce the evaluation outcomes concurrently. To improve the variety of evaluation methods, the evaluating approaches employed for these two ways should be as dissimilar as possible. The difficulty of elusion can be raised

and the accuracy of the outcome further enhanced by employing two recognition techniques. SLU will deal with both of the results that were generated.

3.2.2 Physical Intervention Unit (PIU)

PIU has the "auxiliary role" in EHLCST. PIU should be granted processing power when DDU yields an unsatisfactory result. Following information given to the safety specialists, the event will be managed based on their experience. The ultimate decision about the event's safety will be made directly by the safety experts; DDU will no longer become involved. Due to the unpredictability of network event types, DDU must be expanded using PIU processing results in order to further increase its processing capabilities. Experts must additionally calibrate the data after presenting the final conclusion. A new kind of event is added to the knowledge base by means of feature extraction.

3.2.3 Superiority Level Unit (SLU)

SLU is added to the design in order to link MDM and PIU and achieve the collaboration of the two modules. Determining if PIU has to be called in order to finish the event processing is SLU's primary duty. SLU integrates the results and provides the superiority level when the DDU recognition methods yield two processing outputs. When the level of superiority is great, DDU will provide the final result immediately, which may save labor, shorten the time needed for identification, and satisfy the needs of processing cyber occurrences. In order to reduce inaccuracies, the feedback will be analyzed by experts after it reaches PIU when the confidence level is low.

There are two cases of low Superiority level. First, DDU provides two entirely different evaluation results, which will not establish whether the event is safe or not; and second, the results of the two methods are identical but fall short of the theoretical threshold under the indexes like accuracy, which may indicate that the results' credibility is low and that they will not be accepted as the final evaluation result.

IV. Comparison of EHLCST Model

Both models are basically assessed from 10 aspects in Table 1 to more clearly represent the differences between the general model and the EHLCST model in this paper, including:

Scalability: DDU employs two complementary recognition techniques. Users can increase the amount of recognition techniques or select particular technical approaches based on particular conditions. Simultaneously, the knowledge base can be enlarged and benefit from increased while handling novel security challenges.

Maintainability: EHLCST introduces two additional modules (PIU and SLU) in comparison to the general model, and these modules must be able to work together. Compared to the generic model, maintenance is more difficult in the event of failure.

Closeness: With the ultimate goal of high human-machine integration, EHLCST integrated the closed loop concept into the design. Following processing, the outcome must be further examined to see if it can be used as the end result.

Integration: DDU reflects the four modules into which the general model might split the handling of cyber security challenges. Furthermore, there is a high degree of integration among the three main EHLCST modules.

Participatory: The safety experts can take part in the decision-making process because the core concept of EHLCST is human-in-the-loop.

Interpretability: SLU is a crucial link between safety experts and AI, which can partially address the issue of AI's poor interpretability and intervene in outcomes with low confidence.

Reliability: the safety experts process the experimental results, which are provided by two recognition techniques. When compared to the general model, the model's dependability is noticeably higher.

Complexity: only DDU covers the processing of the general model, indicating that EHLCST has a greater level of complexity.

Characteristics	General Model	EHLCST
Scalability	Mid	High
Maintainability	Easy	High
Closeness	Low	High
Integration	Mid	High
Participatory	Low	High
Interpretability	Low	High
Reliability	Low	High

Complexity	Mid	High
Security	Mid	High
Deployability	Easy	Complex

Table 1. Comparisons of EHLCST Model

Security: While all models contain security features, the security is enhanced by using EHLCST, which will further lessen interference from data forgeries and other causes.

Deployability: EHLCST is more challenging to deploy because of its bigger system size. To improve security, it is advised to implement recognition techniques independently in the cloud and on-premises because AI can be inserted through the backdoor. In conclusion, EHLCST clearly outperforms the present general model in other areas, despite its poor deployment and maintainability performance. In order to address the drawbacks of the existing application of AI in cyber security, EHLCST can be deployed as a new study concept.

Implementation Factor in EHLCST	Metric	Value
Formalized Method of Implementation	Improving operational capability	73% Faster
Federated Learning Architecture	Reduction of bandwidth requirements	81%
Well-designed Interfaces	Analyst feedback frequency increase	60%
Specialized Training AI	Threat identification effectiveness	67% higher
Human-Machine Collaboration	Accuracy novel patterns with attack	43% higher

Table 2. Key Performance Indicators for EHLCST Model

V. Conclusion

Evaluating the current situation and the related difficulties is crucial for the researcher and practitioner communities because AI has a lot of potential for use in cyber security applications. As a result, this paper examined studies that concentrated on the use of AI in four areas of cyber security: network state awareness, user access authentication, monitoring of harmful behavior, and anomalous traffic identification. The paper asserted the importance of human-in-the-loop, offered a conceptual model, and described its use, in addition to outlining research prospects and challenges. Therefore, working with an organization to adopt and assess the suggested conceptual model is a potential follow-up task.

REFERENCES

[1] Bhardwaj, M.D. Alshehri, K. Kaushik, H.J. Alyamani, M. Kumar, Secure framework against cyber-attacks on cyberphysical robotic systems, J. Electron. Imaging 31 (6) (2022), 061802-061802.

[2] M. N. O. Sadiku, S. Alam, S. M. Musa, and C. M. Akujuobi," A primer on cybersecurity," International Journal of Advances in Scientific Research and Engineering, vol. 3, no. 8, Sept. 2017, pp. 71-74.

[3] M. N. O. Sadiku, T. J. Ashaolu, and S. M. Musa," Artificial Intelligence in medicine: A primer," International Journal of Trend in Research and Development, vol. 6, no. 1, Jan.-Feb. 2019, pp. 270-272.

[4] Y. Mintz and R. Brodie, "Introduction to artificial intelligence in medicine," Minimally Invasive Therapy & Allied Technologies, vol. 28, no. 2, 2019, pp. 73-81.

[5] M. N. O. Sadiku, Y. Zhou, and S. M. Musa, "Natural language processing in healthcare," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 8, no. 5, May 2018, pp. 39-42.

[6] M. Taddeo, T. McCutcheon, and L. Floridi , "Trusting artificial intelligence in cybersecurity is a double-edged sword," Nature Machine Intelligence, vol. 1, November 2019, pp. 557–560.

[7] M. Zajko, "Canada's cyber security and the changing threat landscape", Critical Studies on Security, vol. 3, no. 2, pp. 147-161, 2015.

[8] R. Winkels, Eleventh International Conference on Artificial Intelligence and Law: proceedings: June 4-8, 2007, Stanford Law School, Stanford,

California. Place of publication not identified: ACM, 2007.
[9] H. Bidgoli, Handbook of information security. Hoboken, NJ: John Wiley, 2006.
[10] H. Zhuge, "Semantic linking through spaces for cyber-physical-socio intelligence: A methodology", Artificial Intelligence, vol. 175, no. 5- 6, pp. 988-1019, 2011.
[11] Wang Z, Fang B (2019) Application of combined kernel function artifcial intelligence algorithm in mobile communication network security authentication mechanism. J Supercomput 75(9):5946–5964.
[12] Hariyanto, Sudiro SA, Lukman S (2015) Minutiae matching algorithm using artifcial neural network for fngerprint recognition. In: 2015 3rd international conference on artifcial intelligence, modelling and simulation (AIMS), pp 37–41.