Cybersecurity in the age of AI: Attack and Prevent

Pritam Pandit, *pre-final year student - BCA*, Tanisha Chakraborty, *pre-final year student - BCA*, Sumanta Bhattacharya, *Assistant Professor in Dept. of Computer Sc. & Tech. in iLEAD – Kolkata, India.*

***Abstract*—In the contemporary landscape of rapid technological advancement, artificial intelligence (AI) emerges as a pivotal concept within the realm of networking. Within the cyber domain, AI assumes significance for both offensive cyber operations and defensive cybersecurity measures. Leveraging AI algorithms, novel and potent methodologies for cyber incursions have been devised, alongside the development of robust security protocols. Subsequently, ongoing scholarly inquiry persists in this domain, marked by the emergence of diverse cyber threats including novel malware strains, viruses, phishing tactics, and sophisticated password intrusion techniques. Conversely, within the ambit of cybersecurity, AI-driven advancements have engendered innovative approaches for preemptive threat mitigation, detection, and response. This scholarly discourse reflects an enduring commitment to enhancing safeguards in light of the inherent implications for individual privacy and the preservation of societal well-being.**

***Keywords*—Artificial Intelligent, Cyberattack, Cybersecurity, DDOS, DeepPhish, DGA, DNN, IDF, LSTM, Malware**

I. Introduction

Artificial intelligence (AI) has seamlessly integrated into human daily life, pervading various domains from household chores to industrial operations. Notably, in the realm of industries, cybersecurity assumes paramount importance as it safeguards critical company data. The cybersecurity landscape, which originated in the 1980s, has witnessed persistent challenges, steadily escalating over time. AI's initial foray into cybersecurity occurred during the early 2000s, primarily through its utilization in Intrusion Detection Systems (IDS) to scrutinize network traffic and identify anomalies indicative of potential security breaches. However, with the rapid advancement of technology, AI's application has expanded beyond defense to encompass offensive cyber activities. Malicious actors now harness AI and machine learning algorithms to orchestrate cyber-attacks with heightened efficiency and speed. Consequently, while AI offers myriad benefits, it also poses inherent risks. Thus, as AI-driven cyber threats proliferate, imperative measures must be instituted to mitigate such risks and fortify cybersecurity defenses.

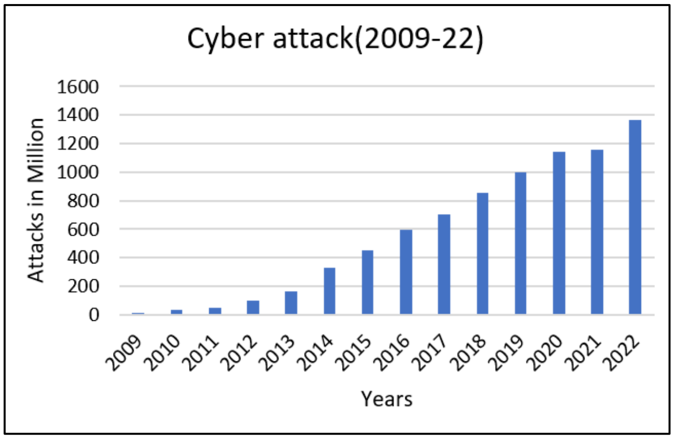

Fig.1. Raise of Cyberattack through years

II. AI in Cybersecurity and Cyberattack

A. *Reconnaissance Phase*

Attack method: In the Reconnaissance phase, the assailant initiates the pursuit of potential targets utilizing AI methodologies such as Long Short-Term Memory (LSTM), Neural Networks (NN), and Deep Neural Networks (DNN), coupled with end-to-end spear phishing tactics and intelligent profiling techniques [1]. Upon identifying a weakness within a system, the attacker employs end-to-end spear phishing methods to exploit it [1][7]. Subsequently, intelligent target profiling [7] strategies are utilized to obscure the payload within a video conferencing application, minimizing the risk of detection. The initial stage necessitates the organization and analysis of extensive datasets to identify system vulnerabilities [1]. Once the requisite data is acquired, DNN algorithms are leveraged to yield the most precise insights regarding these vulnerabilities [2][7].

F. Pritam Pandit, is a pre-final year student - Bachelor of Computer Application, Global College of Science and Technology (GCST) - Krishnanagar, Nadia, India (e-mail: panditpritam399@gmail.com).

S. Tanisha Chakraborty, is a pre-final year student - Bachelor of Computer Application, Global College of Science and Technology (GCST) - Krishnanagar, Nadia, India (e-mail: tanishachakraborty309@gmail.com).

T. Sumanta Bhattacharya is Asst. Professor - Computer Science & Technology in Institute of Leadership, Entrepreneurship and Development (iLEAD) – Kolkata, India (e-mail: sumanta.gcst@gmail.com).

Security Step: Due to its ability to employ algorithms and data analysis for simultaneous scanning of extensive datasets, AI holds potential for thwarting attacks during the reconnaissance phase [14]. By scrutinizing user behavior, verifying email content, and assessing whether a sender has previously accessed specific types of URLs within their peer groups [16], AI can effectively counter Spear Phishing attempts. Regular identification and remediation of vulnerabilities are essential to mitigate risks and minimize potential security gaps. Intelligent police agent gathers information from the network which gives early warning [15].

B. Access and Penetration Phase

Attack method: During this phase, which constitutes approximately 56% of attacks on average, significant emphasis is placed due to its critical nature. Various techniques are employed by attackers, including password guessing or cracking, captcha attacks, and the fabrication of fictitious reviews, all aimed at gaining unauthorized access to the target [1]. Phishing attempts are utilized alongside measures to evade detection by cyberattack detection systems [9], with the deployment of DeepPhish being notable in this regard. DeepPhish [13], employing the Long Short-Term Memory (LSTM) approach, serves to counteract attackers [1]. Noteworthy techniques within password guessing attacks include PassGAN [13] and offensive password authentication. PassGAN relies on the dissemination of real password leaks to infer passwords, while offensive password authentication predicts and appropriates genuine user passwords. Intelligent password brute force assaults utilize a Random Number Generator (RNN) in conjunction with past password sequences to systematically guess passwords character by character [1]. In efforts to enhance attack accuracy, expedite attack durations, and circumvent captcha mechanisms, intelligent captcha attacks [2][7] leverage Convolutional Neural Networks (CNN). Furthermore, the generation of counterfeit reviews employs Recurrent Neural Networks (RNN) to fabricate deceptive reviews [7].

Security Step: In safeguarding user accounts, AI plays a pivotal role by suggesting robust, distinctive passwords and securely storing login credentials [14]. This facilitates user authentication across various devices or locations, alongside enabling periodic password updates. Furthermore, the system's security is bolstered through the implementation of end-to-end encryption mechanisms [18]. Integrated Security Approach uses police agents; after detecting malicious work it sends decentralized instructions to prevent the attack [8].

C. Delivery Phase

Attack method: During this phase, attackers engage in endeavors to acquire internal system intelligence [8], often employing shortened URL facilitate malware dissemination upon interaction [12]. The attacker may manipulate system behavior and patterns to obfuscate its presence, utilizing shadow processes to enable malware mobility and alignment with typical system behavior, thereby evading detection by Intrusion Detection Systems (IDS) [12]. To elude detection by machine learning-based black box cyber threat detection systems, malicious actors leverage Long Short-Term Memory (LSTM) for crafting phishing URLs that remain undetected [7] and Generative Adversarial Networks (GAN) for generating adversarial malware with stealth characteristics [1]. Notably, DeepLocker [12] represents a methodology aimed at concealing its assault and selectively targeting specific entities for infection, employing Deep Neural Network (DNN) techniques for this purpose [1].

Security Step: Artificial intelligence (AI) offers the potential to expedite scanning processes and facilitate rapid tracking of system traffic, rendering it instrumental in the detection of novel and intricate malware. Through the analysis of extensive datasets using machine learning algorithms, AI can discern patterns and deviations indicative of malicious content. Furthermore, AI's capacity for real-time detection and classification of emerging threats is noteworthy, owing to its comprehensive training in identifying diverse forms of viruses and trojans [17]. ANN and DNN can help to know about the network traffic and also help to prevent these types of attack [8]

TABLE I

Name of the Phase	Pros	Cons
Reconnaissance	Using IDS can help to detect and block reconnaissance activities early	This could hinder legitimate network activities conducted by security teams
Access and Penetration	Multifactor authentication and strong password policies can prevent unauthorized access	Slow down workflow and may inconvenience legitimate users
Delivery	Web security can block malicious email and websites	Over reliance on automated filtering systems may result in false negatives
Exploitation	Regular updating and patching software vulnerabilities can prevent attackers from exploiting	It can lead delay in applying critical security updates
Installation	Endpoint protection solutions can detect and remove unauthorized software installation	Advance malware requires continuously update and improve their endpoint security measures
Command and Control	Implementing network segmentation and firewall rules can limit unauthorized communications	Attackers may use encrypted to cover communication channels
Action on Objectives	Quickly identify and mitigate the impact of cyber attacks	Lack of resources may delay detection

Fig.2. Pros & Cons of Cyber Security usign AI

D. *Exploitation Phase*

Attack method: The malicious actor secures authorized access to the target system, employing DeepHack to infiltrate web applications by leveraging artificial intelligence (AI) to scrutinize system behavior and identify vulnerabilities. DeepHack [13] integrates Neural Networks (NN) and reinforcement learning methodologies. Through the exploitation and compromise of the environmental control system, self-learning malware [7] manifests an unintended disruption of computer infrastructure, employing k-means clustering techniques. Machine-generated spear phishing aims to automate disinformation dissemination [1], employing Markov chains and Long Short-Term Memory (LSTM) algorithms.

Security Step: AI algorithms demonstrate the capacity to predict occurrences of data breaches and their probable locations, [14]. Through the analysis of patterns and behavior, AI can discern the most probable breach locations and proffer timely remedial measures [4]. Owing to its extensive training on numerous established patterns and solutions, AI has the capability to autonomously generate novel solutions. Additionally, the algorithmic capabilities of Artificial Intelligence (AI) enable the detection of automated disinformation production or synthetic data by comparing aberrant data with established data patterns; in such instances, the GAN method proves to be beneficial [16]. The attack can be detected or predicted in this phase by DNN [8]. Ai can detect small abnormalities in the system, like from where the user is logging in and the way of password typing [4].

E. *Installation Phase*

Attack method: After successfully evading detection by the Intrusion Detection System (IDS) [7], the malicious actor proceeds to implant malware within the target system, aiming to extract crucial system information and establish communication with the targeted entity. The AI-powered malware exhibits [2] enhanced mobility, facilitating the rapid infection of multiple devices within a condensed timeframe. By deploying malware, the hacker acquires remote access to the system [7], thus furthering their nefarious objectives.

Security Step: AI finds application in Endpoint security [16] by scrutinizing atypical behaviors indicative of potential threats to the system, thereby thwarting unauthorized access attempts by attackers seeking sensitive data. Through real-time analysis of behavior, AI promptly identifies anomalies and alerts security teams, enabling swift response [14] measures to avert potential damage.

F. *Command and Control Phase*

Attack method: During this phase, the malevolent actor endeavors to establish a communication conduit between the targeted system and itself, aiming to assume control over the system and interconnected systems within an internal network. The malicious actor may deploy remote-access Trojans to download and install [1], facilitating data exfiltration and commanding the target for botnet attacks, while also propagating malware to linked systems. In the data exfiltration phase, the Domain Generation Algorithm (DGA) is employed to generate numerous random domains [7], enabling malware to transmit and receive data. Although most of these domains may remain active, a single domain could establish communication with the malicious actor's Command and Control (C2) server [13]. Neural networks and deep learning techniques are harnessed within natural language processing (NLP) to enhance processing accuracy and efficiency [7]. However, adversarial manipulations of text data permit attackers to alter the meaning and structure of text, facilitating the production of artificial data [2]. This phase encompasses the implementation of a self-learning assault on the C2 channel through the utilization of the sophisticated self-learning attack methodology, referred to as DeepLocker, designed for self-evolving malware [13].

Security Step: AI demonstrates rapid system-wide scanning capabilities owing to its efficient algorithms.

Through comprehensive domain pattern learning, AI effectively identifies and blocks unauthorized malicious domains seeking access [5]. Furthermore, it inspects system files, removes any detected malware, and notifies the system accordingly. Notably, the AI methodology operates without necessitating access to the malware's source code; rather, it initially evaluates samples of Domain Generation Algorithm (DGA) and non-DGA domains to extract statistical characteristics for discerning DGA domains, [20].

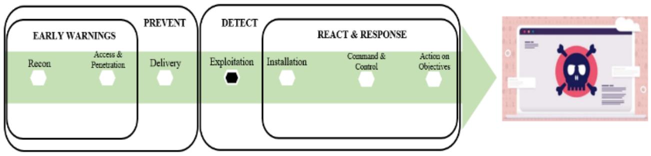

Fig.3. Stages of attack & prevention

G. *Action on Objectives Phase*

Attack method: During this phase, malware autonomously propagates, while vulnerabilities in the Command and Control (C2) server necessitate human intervention for mitigation. Conversely, Distributed Denial of Service (DDoS) attacks operate without human intervention [1]. In such attacks, the botnet directs traffic towards the system server's IP address, inundating it and impeding regular network traffic [3]. Should persistently requests from the botnet be denied, the malware promptly generates a new IP address [7] and transmits it to the server. These processes unfold seamlessly without human involvement.

Security Step: Artificial intelligence (AI) algorithms play a crucial role in the timely identification and targeted mitigation of denial-of-service (DDoS) attacks, [19]. The utilization of IP router telemetry enables the collection of extensive datasets. However, human intervention remains essential to instruct AI about botnet behaviors and recognize typical network traffic fluctuations, [19]. Expert System is used to solve reasoning of predefined knowledge as well as given problems which is checked take actions.

III. CONCLUSION

In conclusion to enhance system security and access control during the delivery phase, an interactive query should be displayed on the interface, set by an authorized user and updated daily. Any aberrant activity detected by the system should trigger alerts by the AI. Additionally, to fortify against exploitation, a multi-tiered authentication process should be implemented, with concurrent validation by the AI to ascertain authorization status before initiating antivirus scans. Daily software updates are imperative, alongside strict adherence to authorized software usage as dictated by the authorized user. AI-driven cyberattacks and cybersecurity inhabit a shared domain, employing AI or ML algorithms to fulfill their respective objectives. In cyberattacks, AI serves to enhance the stealth, accuracy, and complexity of attacks, rendering them more challenging to detect. Conversely, in cybersecurity, AI bolsters defense mechanisms, augmenting detection and response capabilities to counteract attacks effectively. Nonetheless, a notable challenge arises from the potential exploitation of AI by malicious actors, who may exploit vulnerabilities and biases inherent within AI systems to undermine cybersecurity defenses and execute more potent attacks. To address these challenges and mitigate the risks posed by AI-driven cyberattack threats, security organizations must adopt a multifaceted protective framework integrating AI-powered defense mechanisms with conventional security strategies, such as user training, network segmentation, and incident response planning. Additionally, collaborative efforts between industries and governmental entities can further enhance cybersecurity resilience. Safeguarding cybersecurity amidst the looming threat of cyberattacks necessitates continual innovation in both offensive and defensive strategies.

IV. ACKNOWLEDGMENT

We extend our sincere gratitude to all those who contributed to the completion of this research paper. First and foremost, we express our deepest appreciation to our mentor, Sumanta Bhattacharya for his invaluable guidance, encouragement, and support throughout the entire process. His expertise and constructive feedback were instrumental in shaping the direction and quality of this study.

Furthermore, we are grateful to the participants of this study for their time and cooperation, without which this research would not have been possible. Their contributions are deeply appreciated and have greatly enhanced the validity and relevance of our findings.

Finally, we acknowledge the support of our families and friends for their unwavering encouragement and understanding throughout this academic journey.

This work was made possible through the collective efforts and support of all those mentioned above, and we are sincerely thankful for their contributions.

V. REFERENCES

[1] Blessing Guembe, Ambrose Azeta, Sanjay Misra, Victor Chukwudi Osamor, Luis Fernandez-Sanz & Vera Pospelova "The Emerging Threat of Ai-driven Cyber Attacks: A Review, Applied Artificial Intelligence", 2022.

[2] Muhammad Mudassar Yamin1, Mohib Ullah1, Habib Ullah2, Basel Katt "Weaponized AI for Cyber Attacks", ResearchGate, Journal of Information Security and Application, 2021.

[3] Fabrizio Bertone, Francesco Lubrano, klodiana Goga "Artificial Intelligence Techniques to Prevent Cyber Attacks on Smart Grids" Vol 3, No 1 (2020): Special issue on cybersecurity of critical infrastructure.

[4] Ljubomir Lazic "BENEFIT FROM AI IN CYBERSECURITY" Conference Paper · October 2019.

[5] Feng Tao, Muhammad Shoaib Akhtar and Zhang Jiayuan "The future of Artificial Intelligence in Cybersecurity: A Comprehensive Survey", EAI Endorsed Transactions on

Creative Technologies | Volume 8 | Issue 28 | e3 July 2021.

[6] Rammanohar Das and Raghav Sandhane "Artificial Intelligence in Cyber Security" J. Phys.: Conf. Ser. 1964 042072, 2021.

[7] Erik Zouave, Marc Bruce, Kajsa Colde, Margarita Jaitner, Ioana Rodhe, Tommy Gustafsson "Artificially intelligent cyberattacks", FOI-R--4947--SE, March 2020.

[8] Nadine Wirkuttis and Hadas Klein "Artificial Intelligence in Cybersecurity", Cyber, Intelligence, and Security | Volume 1 | No. 1 | January 2017.

[9] Gregory Falco, Arun Viswanathan, Carlos Caldera, And Howard Shrobe "A Master Attack Methodology for an AI-Based Automated Attack Planner for Smart Cities", IEEE Access, VOLUME 6, 2018.

[10] Praveen Kumar Shukla, C. S. Raghuvanshi, Hari Om Sharan "Analysis of AI Based Approach to Prevent Cyber Attack on Web Applications in Contemporary Digital Era", ResearchGate, 2022.

[11] Triveni Krishnappa "A REVIEW ON ARTIFICIAL INTELLIGENCE TECHNIQUES IN PREVENTING CYBER THREATS", International Journal of Engineering Applied Sciences and Technology, Vol. 8, Issue 01, ISSN No. 2455-2143, Pages 185-189, 2023.

[12] Michal Markevych, Maurice Dawson "A REVIEW OF ENHANCING INTRUSION DETECTION SYSTEMSFOR CYBERSECURITY USING ARTIFICIAL INTELLIGENCE (AI)", International Conference Knowledge-Based Organization Vol. XXIX No 3 2023.

[13] Nektaria Kaloudi and Jingyue Li "The AI-Based Cyber Threat Landscape: A Survey", ResearchGate, Journal of Information Security and Application,2020.

[14] "Five ways AI can be used to prevent cyber-attacks"
https://aimagazine.com/ai-strategy/five-ways-ai-can-be-used-to-prevent-cyber-attacks

[15] "Defending Against AI-Based Cyber Attacks: A Comprehensive Guide"

https://scytale.ai/resources/defending-against-ai-based-cyber-attacks/

[16] "How Security Analysts Can Use AI in Cybersecurity"

https://www.freecodecamp.org/news/how-to-use-artificial-intelligence-in-cybersecurity/#how-ai-is-used-in-cybersecurity

[17] "AI in Cybersecurity: What You Need to Know"

https://www.analyticsvidhya.com/blog/2023/02/ai-in-cyber-security/

[18]"16 common types of cyberattacks and how to prevent them" https://www.techtarget.com/searchsecurity/tip/6-common-types-of-cyber-attacks-and-how-to-prevent-them

[19] "AI/ML for Better DDoS Detection"

https://www.darkreading.com/cyberattacks-data-breaches/how-ai-ml-can-thwart-ddos-attacks

[20] "USING ARTIFICIAL INTELLIGENCE/ MACHINE LEARNING TO DETECT DOMAIN GENERATION ALGORITHMS"

https://www.globaltechcouncil.org/artificial-intelligence/using-artificial-intelligence-machine-learning-to-detect-domain-generation-algorithms/

common-types-of-cyber-attacks-and-how-to-prevent-them

VI. Biographies

Pritam Pandit was born on 19-September-2004 in Karimpur, Nadia, India. He is currently pursuing his undergraduate degree in Bachelor of Computer Application from Global College of Science and Technology (GCST) - Krishnanagar, Nadia, India.

Tanisha Chakraborty was born on 26-December-2004 in Nabadwip, Nadia, India. She is currently pursuing her undergraduate degree in Bachelor of Computer Application from Global College of Science and Technology (GCST) - Krishnanagar, Nadia, India.

Sumanta Bhattacharya was born on 19-January-1975 in Serampore – Hooghly, India, is working as Assistant Professor in Department of Computer Science & Technology in Institute of Leadership, Entrepreneurship and Development (iLEAD) – Kolkata, under the Maulana Abul Kalam Azad University of Technology, West Bengal, India. Prior to this, he was associated with Global College of Science & Technology – Krishnagar as Assistant Professor and Head in Department of Computer Application for 1 year 6 months.

He has more than 19 years of Academic Experience, in including Deputation in Bhutan for 3 Years (under Indian Technical and Economic Cooperation Programme – Government of India) in Gyalpozhing College of Information Technology – Royal University of Bhutan. He has several Research Publications. He reviewed a Book on Software Engineering. He has attended several International Conferences, National Seminars/Workshops, Summer Schools, Refresher / Orientation course in India as well as abroad so far. He is a associated to many Professional Societies in various capacities as well as Lifetime Member, namely Computer Society of India (CSI), Kolkata, India, Forum of Scientists, Engineers and Technologists (FOSET), Kolkata, India International Association of Computer Science and Information Technology (IACSIT), Singapore, Indian Society for Technical Education (ISTE), New Delhi, India, Global Member of the Internet Society (ISoc) – USA; Chapter Member of the Internet Society (ISoc) – Kolkata Chapter, Kolkata, India, Life Member of the Indian Science Congress Association (ISCA), Kolkata, India, The Society of Digital Information and Wireless Communications (SDIWC), Computer Science Teachers Association (CSTA), New York – USA, International Association of Engineers (IAENG) - Hong Kong.