International Journal of Scientific Research & Engineering Trends
Volume 10, Issue 2, Mar-Apr-2024, ISSN (Online): 2395-566X

How AI and ML are Contributing to the Sophistication of Cyber Attacks

Kiran Sharma Panchangam Nivarthi
Algonquin Trail Chisago City, MN, USA

Abstract- Artificial Intelligence (AI) and Machine Learning (ML) are being extensively used to make cyber security more robust and adaptive to new forms of attack. At the same time, the two technologies have the potential to make several conventional cyber attack vectors significantly more potent and sophisticated. They are also introducing new attack vectors. With the help of AI and ML, malware can become more adaptive and harder to trace, and we already have an example in the form of IBM's Deep locker. They can also be used to identify better phishing targets and make phishing emails more believable by using freely available tools like Chat GPT. AI and ML can augment cyber attacks like DDoS that rely on botnets to adapt to a wide range of defensive measures and make coordinated botnet attacks easier to orchestrate. Man-in-the-middle attacks can become more potent with new AI and ML-augmented cryptanalysis and real-time spoofing. SQL injections can become more sophisticated by leveraging the right algorithm to generate queries that are more likely to bypass database security. The technologies can also be used for more comprehensive payload and traffic analysis of DNS servers for new generation of DNS tunneling attacks. AI and ML capabilities of adaptive behavior, more sophisticated automation, identifying patterns in data, and mimicking existing traffic/human patterns can lead to more sophisticated cyber attacks.

Index Terms- AI and ML, Cyber Attacks

I. INTRODUCTION

Artificial Intelligence (AI) and Machine Learning (ML) are often used interchangeably, and even though the specific differences between the two are a matter of some debate in both academia and the corporate world, it's generally accepted that the two are not the same [1]. Defining AI is relatively more complex than defining ML, and the lack of a standard definition is a problem [2]. The overarching definition that AI systems mimic human intelligence may not be relevant because AI outshines human intelligence in many areas.

A better approach would be to define AI from its features/elements that are common in most definitions - information processing, perception of the environment, decision-making, and achieving specific goals. ML, on the other hand, is the process of machine learning to perform certain operations without human guidance. It's primarily defined in the context of data, and ML algorithms are expected to identify patterns in data and glean insights without specific instructions. ML is generally considered a subfield of AI [3].

Both AI and ML have a wide range of applications in various industries, including cyber security. Machine learning algorithms and systems have become part of most standard and comprehensive cyber security arsenals and are used to safeguard against a wide range of cyber-attacks and threats [4]. A wide range of ML algorithms have been used for intrusion detection for years [5], and in this and several other cyber security dimensions, the line between ML and data science sometimes gets blurred [6].

The reason is that ML algorithms and cyber security features and systems built around them are able to process a significantly larger amount of data more efficiently than conventional tools relying on brute computing force. Machine learning techniques like clustering, self-organizing maps, and classification and regression trees (CARTs) can help with the identification and mitigation of a wide range of cyber-attacks [7]. Even though ML applications in cyber security and their overlap with the field are still in their infancy, they have already added several new dimensions to cyber security [8].

The overlap between AI and cyber security is more extensive, with a wider range of benefits and challenges, including several ethical challenges [9]. Like ML, AI has the potential to expand and augment the cyber security arsenal in a number of ways, including areas of solutions like asset management, governance, risk assessment, and risk management [10]. However, the primary benefit of using AI in cyber security is to prepare and augment it against the evolved, AI-powered cyber security threats and cyber-attacks [11].

© 2024 IJSRET

AI and ML have contributed significantly to the evolution of cyber attacks in the past few years, making them more sophisticated and more difficult to protect against. Just as AI and ML have introduced new dimensions to cyber security, they have also helped develop new classes/types of cyber-attacks unique to AI and ML or interrelated fields like data sciences. This includes data poisoning, which can corrupt an AI or ML model being trained on that data, and threats specific to generative AI. In this article, we will look into AI and ML's contribution to the evolution and sophistication of a range of cyber-attacks.

II. AI/ML AND A NEW ERA OF CYBER ATTACKS AND THREATS

Evolved/improved AI-powered cyber-attacks are a natural consequence of advances in AI and ML and easy access to powerful AI and ML models and systems. Not only has AI and ML strengthened classic cyber attacks and threats like malware, distributed denial of service (DDoS) attacks, man-in-the-middle (MitM) attacks, and phishing, but they have also led to the development of a new generation of cyber threats, including data misclassification and synthetic data generation [12].

Another set of AI-powered cyber-attacks included [13]:

- Reconnaissance attacks like intelligent target profiling and smart vulnerability detection.
- Access and penetration attacks like automated payload generation and smart fake review generation.
- Exploitation attacks (malware) like intelligent lateral movement and behavioral analysis.

One core advantage AI and ML-based attacks have on conventional cyber attacks is adaptability. They can adapt to the infrastructure they are attacking as well as its defenses, which sometimes allow them to penetrate the system without triggering any cyber security mechanism in place to stop such attacks [14]. Another dimension of threat is the access. Many AI and ML models have access to identifiable data like finances, healthcare, etc., and these models can be made to leak this data or share it with the wrong sources through a clever AI/ML-powered cyber-attack; it would be a significant cyber security breach [15].

AI and ML's impact on specific cyber attacks can provide us with more insights into how they have contributed to their evolution and sophistication.

1. Malware

Malware is one of the most common types of cyber-attacks. Malware, short for malicious software, is any piece of code or software that causes harm to a computing system or a network [16]. The two primary strengths of malware are its delivery, i.e., how it can transmit from one device/computer system to another and how much harm it can cause to a single device or a network. Malware is an all-encompassing term that also covers viruses, and its history can be traced back to the 1970s, when "Creeper" was introduced. However, malware has come a long way since then, especially now that AI and ML can power it.

One of the most well-known examples of how AI and ML can improve malware came from a legitimate source, i.e., IBM, in the form of Deep Locker. It's evasive malware, which means that it's designed to remain undetected by antivirus programs and other malware detection software. It hides in applications that have received a "pass" from cyber security elements (legitimate software) and only attacks when it detects the right conditions. This environmental awareness and delay tactics are just one example of how AI can make malware more sophisticated. However, the most terrifying part of Deep Locker is that it can also serve as a safe delivery method for a range of malware [17].

We can also look into different types of malware and how they are being augmented and improved by AI and ML.

Ransomware

Ransom AI is an example of AI-powered ransomware from academia. It is a framework to augment existing ransomware with AI, giving them the ability to adapt to protective measures and avoid detection as soon as possible [18]. There are several speculations on how AI can empower ransomware. It can remain undetected for longer and act at the most optimal time. It can also alter the ransom amount based on preset parameters, delivering a more significant impact. An AI-powered ransomware can also redefine and expand its scope, unlike conventional ransomware. It must be stated that the literature on the use of artificial intelligence and machine learning for developing ransomware is significantly lacking compared to the literature on using these technologies for detecting and classifying ransomware [19].

Trojans

Trojans are another type of malware that can be augmented and improved using AI and ML. There are also trojans that specifically target the neural networks underlying an AI model, like Po Trojan, that are triggered under very specific conditions and, once triggered, can cause the neural network to malfunction [20]. AI-based hardware trojans are another class of malware that is considered capable enough to bypass most state-of-the-art cyber security systems [21].

Virus

Computer viruses are a specific type of malware that replicates themselves into as many devices as possible. The most significant danger associated with an AI virus is its ability to adapt to the host device, identify its vulnerabilities, and determine the best course of action to replicate and activate.

© 2024 IJSRET

Spyware

Spyware is a specific class of malware that focuses on gathering data from the device they are on and passing it on to malicious entities. AI and ML-powered spyware can not just gather the relevant information but use it to attack the device or access more information from a device. An example is Vaspy, a spyware that learns a person's voice patterns and can mimic them for Voice Assistants (VA), and considering how much we rely upon Vas and voice activations, it can be devastating for an individual [22]. Another form of Spyware is keyloggers that learn the key patterns of a human user and pass this information on to hackers and other malicious actors. One example is Black Mamba, which uses AI models to differentiate between useful and irrelevant keystrokes and only passes on the most relevant information.

AI-powered malware can also be developed for specific use cases. One example is AI-powered GUI attacks that target some of the most common GUI elements people interact with on a daily basis on their electronic devices like web browsers and steal information like saved passwords in those browsers. The AI element is the detection of these elements and activity events, which may be different for each user and device [23]. One of the most significant threats AI-based/AI-powered malware poses is their ability to learn and study the target to identify the most opportune time/circumstances to attack, which can significantly broaden the scope of triggers that can activate malware [24]. Another threat that AI poses when it comes to weaponizing malware and making it more effective and sophisticated is that it can be trained to target vulnerabilities specific to a business, allowing the malware to slip through even in the presence of standard malware protection [25]. This vulnerability becomes even more significant if we take into account the inherent strength of AI models to learn from, understand, and adopt human behavior, which can be weaponized for malicious social engineering on a far larger scale than what human hackers can achieve [26].
In many industries, including defense, AI and ML have introduced various new attack vectors associated with malware. This includes corrupting AI and ML models and neural nets at the training stage (data poisoning), intelligent malware, vulnerability detection, etc. [27]. The use of smart and autonomous devices in the defense industry, including reconnaissance drones and self-targeting weapons, further expands the attack surface vulnerable to AI-powered malware attacks.

These are just some of the ways AI and ML are contributing to the sophistication of malware-based cyber-attacks that have been identified so far. At the pace at which AI tools are evolving, we may see new attack vectors and not just different breeds of current malware but unique malware products with no precedence in the existing cyber security paradigm. AI-based malware with loosely defined parameters and objectives can harm computer systems and networks in unprecedented ways. Another way AI and ML may contribute to the evolution of malware is their own presence. New breeds of malware may be developed for the sole purpose of corrupting or gaining control over AI and ML systems that are still in the learning phase.

2. Phishing

Phishing is a form of cyber attack where hackers and other malicious actors pose as legitimate companies and contacts to extract information from people. Conventional phishing attacks rely upon social engineering and leveraging the psychology of users, and while it's still a focal point, phishing has evolved. Earlier phishing relied upon the lack of relevant security measures and the unfamiliarity of a wide range of users with the digital systems they had to interact with. Now that people have become more digitally aware and safeguards against phishing are typically built into many digital and communication systems (like inboxes), phishing techniques have evolved and become more sophisticated to target people and steal useful information from them [28].

Artificial Intelligence and Machine Learning are two of the active "agents" contributing to this evolution of phishing. One of the most basic ways AI can make phishing more sophisticated is by learning from large sets of publicly available data and developing phishing attacks targeted at specific subsets or even individuals. AI has made spear phishing, i.e., phishing attacks targeting specific individuals, far easier because instead of learning about the targets themselves and sorting through the available information, the hackers can assign the task to an AI model. It will also become more efficient with every new target [29].

AI tools like ChatGPT that are freely available can be used to identify and replicate patterns in speech and communication and use this information to craft phishing emails/messages indistinguishable from a trusted human source [30]. AI models can also be used to launch sophisticated phishing attacks that learn from the behavior of spam filters and tailor each attempt to bypass the filter (without triggering it) until they find the right approach/configuration [31].

While it's still a developing area, with IoT and a wider range of smart devices induced in our everyday lives, bot-to-bot communication will grow significantly. It may lead to a new generation of phishing attacks where messages and other communication attempts are not deceiving humans but AI models (with autonomy and decision-making power). We can also expand this paradigm to include other areas of phishing, i.e., voice phishing and video phishing. Voice phishing is already a well-known form of fraud/cyber attack that relies upon fake voice messages or conversations [32]. It can be significantly enhanced with AI. Voice cloning is a rapidly evolving discipline in AI, and many tools can mimic not just voices but vocal expressions and ticks, making them virtually

© 2024 IJSRET

indistinguishable from the actual speaker [33] [34]. Combine that with a Natural Language Processing (NLP) algorithm like ChatGPT that can generate entire conversations, and the avenue for AI-powered voice phishing becomes quite significant. Real-time AI-powered video calls with clones of real humans may be years ahead, but they will further expand the possibilities for this particular avenue of AI-powered cyber-attacks.

3. Denial-of-Service (DOS) and Distributed Denial of Service (DDOS)

Any attack that aims to overwhelm a server and forces it to stop servicing its clients. It can take several different forms based on the environment it's attacking [35]. The definition has been expanded to include not just computer servers but any online service that a DoS attack may disrupt. A distributed denial-of-service or DDoS attack is one of the most common types of DoS attacks. A DDoS relies upon multiple computers/devices to disrupt an online service and prevents it from functioning as it was intended to function. DDoS attacks have been evolving since the early 90s and have become more sophisticated over time [36]. However, AI and ML have the potential to help DoS attacks evolve in unprecedented ways and scale.

Many DoS/DDoS attacks leverage a network of botnets to launch relatively large-scale attacks, and AI-powered botnets are one of the most significant threat dimensions to consider. One AI-powered botnet example was found on Twitter, which used a Large Language Model (LLM) Chat GPT to generate content that would be indistinguishable from humans [37]. While it's not uncommon for bots to be on social networking sites, the same strategy can be employed in smaller networks like work groups, universities, etc., and overwhelm the network capacity and bandwidth.

AI can also help make botnets more adaptive to their environment and resilient against cyber security measures. This can lead to adaptive DDoS attacks, with characteristic strengths like advanced reconnaissance of target networks, changing attack vectors, and using topologically adjacent attack infrastructure [38]. This is beyond the capabilities of conventional cyber security measures, and the recognition of this attack dimension has led to the development of adaptive DDoS attack detection using AI and ML [39].

Another way AI can empower DDoS attacks is with smart network monitoring and vulnerability capabilities, which, ironically, is also used in cyber security. The same AI algorithms that detect vulnerabilities in individual systems/devices and networks can also be weaponized to attack these systems or networks [40]. Suppose malicious actors can access the source code or start targeting open-source software applications. In that case, they can also identify vulnerabilities on the source code level using AI-powered vulnerability detection techniques [41]. There is also considerable research on simulating various AI-powered cyber-attacks, including DDoS attacks. When observed from the perspective of malicious actors/hackers, these simulation techniques, models, and tools can be used to identify new attack vectors that even the most advanced cyber security measures and tools are not equipped to handle [42].

Botnet management has always been an important factor behind the ingenuity and success rates of DDoS attacks, and well-managed botnets have proven resilient against dedicated cyber security efforts to shut them down [43]. Botnet topologies and botnet attack and defense algorithms have been studied for years, and by now, most of the conventional DDoS attack vectors have been classified and mapped out, and exceptions are likely to be rare [44] [45]. However, the right AI and ML models can lead to mostly automated and adaptive botnet management that may require far fewer human bot masters to control massive botnets. The AI models can even be trained to relay individualized instructions to bots in the botnets based on identified vulnerabilities and their individual level of penetration and control over the device or orchestrate botnet activity in complicated new patterns.

Another technological dimension that can exacerbate the situation and lead to larger, more sophisticated botnets and, consequently, DDoS attacks is the Internet of Things (IoT). Millions of unsupervised, autonomous devices with varying levels of cyber security measures and a plethora of new vulnerabilities will significantly enhance the attack surface, and AI will be crucial in mapping and exploiting these vulnerabilities for malicious actors [46]. Considering the scope of IoT and how it may revolutionize DDoS attacks, especially if they leverage the right AI techniques, the term Botnet of Things (BoT) has been coined [47].

AI models have become quite adept at mimicking human behavior, making them indistinguishable from human users, and botnets that leverage this strength can launch more sophisticated DDoS attacks without tripping any defense mechanisms.

In addition to this enhancement to existing DDoS attack dimensions, AI can lead to some new DDoS attack vectors. They can switch between network and application layers to form more sophisticated attack patterns, leverage different types of payloads, etc. The underlying factor common in all these layers of sophistication AI can add to the DDoS attack is its ability to adapt, learn, and react.

4. Man-in-the-Middle Attacks

Man-in-the-middle or MITM attacks have the potential to corrupt the integrity of the connection between two endpoints and exploit and divert the information traveling between those endpoints [48]. It's also studied widely in the context of encryption techniques and technologies. It's another area of cyber security where the Internet of Things (IoT) can be an

© 2024 IJSRET

exacerbating factor, as it's likely to increase device-to-device communication significantly, and with various levels of autonomy among these devices and unique cyber security vulnerabilities, there would be several new exploitation points for MITM attacks [49].

AI and ML can make these MITM attacks more sophisticated and damaging than they currently are in a number of different ways, starting with adaptive targeting. Complex AI algorithms can be deployed to identify, shortlist, or select high-value victims or more vulnerable victims for the MITM attacks. This can allow malicious actors/hackers to divert attention and resources for a selected number of victims, increasing their chances of success.

It wouldn't be a stretch to call encryption the first line of defense against MITM attacks, and it is already quite vulnerable to AI and ML. The situation may be further aggravated if quantum computing becomes more common and accessible. Encryption is evolving to deal with this threat, and technologies like block chain show promising results, but the threats are evolving just as rapidly [50]. There are already examples of ML algorithms being successfully employed for cryptanalysis, i.e., the process of decoding encrypted messages without breaking the cipher (discovering the key), and even though it's not understood how that algorithm and the underlying neural net achieved that, that doesn't diminish its ability to be employed in a sophisticated MITM attack [51]. A new dimension, i.e., quantum AI-based cryptanalysis, is also already being studied and may lead to a new generation of MITM attacks [52].

AI and ML's applications in behavior analysis can also be applied to make MITM attacks more sophisticated and potent. A comprehensive AI/ML model may also be capable of leveraging real-time spoofing to lead users to replicas of legitimate login windows and other online avenues where users are likely to share personal and sensitive information. Artificial Intelligence can also be used to learn from and corrupt/hijack the process of Deep Packet Inspection (DPI) via payload analysis, and if an MITM attack using this strategy targets a sensitive enough department/institution effectively, it can cause significant damage [53].

5. SQL Injection

An SQL injection attack is essentially the manipulation of databases underlying a software package or application by injecting them with malicious SQL commands [54]. Some web applications are more vulnerable to SQL injection attacks compared to others. SQL injection attacks can take several different forms, each with its own malicious intent and scope [55].

AI has been used for years to detect and prevent SQL injection attacks and to make cyber security systems more robust against this attack vector, and it has shown promising results for at least some classes of SQL injection attacks, like tautology [56]. These are the types of attacks that use queries that typically result in a true answer to bypass the restrictions imposed on the entry to the database. But AI can also be used for the attacks. One type of AI model - deep convolutional generative adversarial networks or GANs was used in conjunction with genetic algorithms, i.e., algorithms that mimic the biological evolutionary model of natural selection to combine and discard solutions to produce SQL injection attack samples to train the defensive/cyber security systems [57]. However, the same approach can be applied to identify SQL injections that are capable of bypassing the existing security systems without triggering an alert.

AI and ML models can also be used to make time-based blind SQL injection, which evaluates the duration of the delays in database response times to predict whether the response will be true or false, more sophisticated. The parallels that exist between how Large Language Model or LLM-based chat bots interact with backend databases and the conditions for executing a time-based blind SQL injection already offer training opportunities to make SQL injection attacks more sophisticated [58].

AI can also augment other specific types of SQL injection attacks like the UNION attack, which leverages the UNION operator, Boolean-based blind SQL injection, where attackers focus on manipulating a wide range of Boolean conditions (true/false statements) in the database, and out-of-band SQL injection attacks, that leverage other communication channels (like DNS) instead of database responses.

6. DNS Tunneling

DNS tunneling is a cyber attack type that exploits the limitations and weaknesses of the Domain Name System or DNS layer protocols to conduct malicious activities. The attackers leverage DNS queries, which are inherently simple information about a website and its associated IPs that a computer system/device can request from a DNS server and, normally, is not inspected or analyzed as thoroughly as usual internet traffic is. So, the people using DNS tunneling attacks leverage this status of DNS queries to transfer information and comments that would otherwise be blocked by firewalls and other cybersecurity measures [59].

Most successful DNS tunneling attacks have at least two characteristic features, i.e., payload analysis and traffic analysis [60]. An attacker must understand how a DNS server evaluates the payloads of the requests it gets to identify legitimate requests from malicious ones. They should also understand how the traffic of a DNS server typically behaves. Both of these areas can be outsourced to AI and ML models. AI and ML models trained in specific analysis techniques can go through data much faster than human attackers and recognize patterns

© 2024 IJSRET

that humans are unable to identify, leading to more sophisticated attacks.

From exploiting existing vulnerabilities and attack vectors like DNS over HTTPS abuse in unique ways to identifying new vectors, AI and ML models can augment DNS tunneling attacks in other, unique ways [61].

III. COMMON THEMES AND EXACERBATING FACTORS

The following themes can be found in most if not all, AI-powered cyber-attack enhancements and augmentations.

1. Automation
AI and ML models can help hackers and other malicious entities that design and deploy cyber-attacks become far more efficient by leveraging the power of automation. Compared to conventional (rule-based) automation techniques, AI and ML-based automation can be adaptive and self-guided. It's extensively studied in the context of industry 4.0/fourth industrial revolution [62], and malicious actors can benefit from AI-powered automation breakthroughs, freely available tools, and technologies as positive entities and actors in our society.

2. Data/Patterns (Vulnerability Detection)
AI and ML models are far more efficient than any human at analyzing data, whether it's traffic between two servers or personal data extracted from social media sites. They can identify patterns in data that hackers and other human malicious agents may not even think to look for, thus opening the doors for new attack vectors. This inherent AI/ML strength can be leveraged to detect vulnerabilities in systems and target the single most significant cyber defense vulnerability, i.e., humans. AI/ML can also be used to generate data that can be used in cyber-attacks, whether it's for overwhelming a system with requests and data overload or corrupting data that defensive AI and ML models are trained on.

3. Adaptive Behavior
The ability of AI and ML-based cyber attacks to adapt to the behavior and patterns of host devices and the cyber security measures in place makes them more difficult to detect and can significantly increase their chances of penetration and cause significant damage. These attacks can also alter the scope of their attack and may cause far more damage than intended.

4. Mimicking
AI and ML have already proven themselves to be quite adaptable at mimicking human speech and words (both in voice and text), and this can be used in a variety of cyber-attacks that rely on human behavior, like phishing emails that are impossible to differentiate from normal humans and voice messages from trusted sources that they didn't send. AI and ML models can also mimic network behavior and data patterns that servers and cyber security measures are used to bypass without triggering any alerts.

Quite a few factors have exacerbated and can exacerbate the weaponization of AI and ML for cyber-attacks.

5. Easy Access to AI Tools
Large Language Models (LLMs) like Chat GPT are freely available to the public, and it's just the beginning. It's already possible to find models similar to Chat GPT that are not similarly restricted and can be used in a number of malicious ways, but even if we don't entertain that possibility, the easily accessible and free-to-use AI tools can already be used by malicious actors in a number of ways, like generating human-like text and generating malicious code at a much faster rate and speed.

6. IOT
The Internet of Things (IoT) will be made up of millions and, in the future, billions of devices from hundreds of thousands of different manufacturers and vendors, each with its own tech stacks and unique set of vulnerabilities. These IoT devices will also communicate with each other and will have some level of autonomy, like smart fridges that are given permission to order certain groceries or thermostats that may alter temperature settings when certain conditions are met. Not all these devices may be adequately protected against cyber attacks, and they will significantly increase the overall attack surface.

6. Cloud-Based Computing Power
There is already a trend of crypto mining companies offering computing power to people who need to train their AI models, and going forward, when quantum computing becomes more commonplace, this may even include quantum computing power. This may significantly reduce the barrier of resources for malicious agents.

7. Defense Vectors
There is no doubt that AI can be weaponized in ways that we cannot even predict right now, let alone protect against. When this is financed and supported at a state level, with a state's resources, AI weaponization can grow to unprecedented scope. The most concerning part would be that for most state-funded cyber attackers, the goal may not just be exploitation for personal gain (like ransom ware for money and spyware for information) but destruction and damage, which may require far less sophistication.

8. AI/ML Models Attack Vectors
Finally, there are cyber attacks unique to AI and ML, like poisoning training data. This is currently the most understudied area of AI/ML-powered cyber-attacks because the technologies are still in their infancy, and their long-term impact on society

© 2024 IJSRET

and penetration is mostly unknown at the time. This is why it's difficult to predict how AI and ML-powered attacks on AI and ML systems will take shape in the coming future. Corrupting of diagnostic data may become the new version of ransom ware, and poisoning financial-oriented AI and ML models may lead to a new era of economic collapses.

IV. CONCLUSION

The academic literature on the topic indicates that AI and ML are more extensively used for new and improved cyber security measures and tools instead of developing more sophisticated cyber-attacks, but that's simply because most cutting-edge cyber threats are not analyzed and dissected until they are identified and handled. Millions of devices around the globe may already be under attack from AI and ML-powered cyber-attacks that are currently not even classified (unique attack vectors). Another factor to consider is that with AI and ML, a new generation of clever attacks and clever defenses may emerge. For now, most of the large organizations with significant computing resources at their disposal might be able to build or purchase top-of-the-line AI/ML-powered cyber defense. However, AI and ML models can be trained to identify vulnerabilities and weaknesses around these defenses at a fraction of the computing power. But it can also work the other way, i.e., AI/ML-powered cyber defenses that can withstand power conventionally and next-gen (AI/ML-augmented) attacks.

In conclusion, we can easily say that both AI and ML are contributing to and can contribute to making cyber-attacks more sophisticated, powerful, and damaging, and recognizing this reality and analyzing this problem not just from a defensive dimension but from an attack perspective is crucial.

REFERENCES

1. N. Kühl, M. Schemmer, M. Goutier and G. Satzger, "Artificial intelligence and machine learning," Electronic Markets, vol. 32, p. 2235–2244, 2022.
2. S. L. C. M. G. E. D. P. G. M.-P. F. a. D. B. Samoili, "AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence," Publications Office of the European Union, Luxembourg, 2020.
3. G. Rebala, A. Ravi and S. Churiwala, "Machine Learning Definition and Basics," in an Introduction to Machine Learning, 2019, p. 1–17.
4. A.Handa, A. Sharma and S. K. Shukla, "Machine learning in cybersecurity: A review," WIREs Data Mining and Knowledge Discovery, vol. 9, no. 4, 2019.
5. Y. Xin, L. Kong, Z. Liu, Y. Chen, Y. Li, H. Zhu, M. Gao, H. Hou, and C. Wang, "Machine Learning and Deep Learning Methods for Cybersecurity," IEEE Access, vol. 6, 2018.
6. I. H. Sarker, A. S. M. Kayes, S. Badsha, H. Alqahtani, P. Watters and A. Ng, "Cyber security data science: an overview from machine learning perspective," Journal of Big Data, vol. 7, 2020.
7. J. M. Torres, C. I. Comesaña and P. J. García Nieto, "Review: machine learning techniques applied to cyber security," International Journal of Machine Learning and Cybernetics, vol. 10, p. 2823–2836, 2019.
8. G. Apruzzese, P. Laskov, E. M. d. Oca, W. Mallouli, L. B. Rapa, A.V. Grammatopoulos and F. D. Franco, "The Role of Machine Learning in Cybersecurity," The Role of Machine Learning in Cybersecurity, vol. 4, no. 1, p. 1–38, 2023.
9. M. Taddeo, "Three Ethical Challenges of Applications of Artifcial Intelligence in Cybersecurity," Minds and Machines, vol. 29, p. 187–191, 2019.
10. R. Kaur, D. Gabrijelčič and T. Klobučar, "Artificial intelligence for cybersecurity: Literature review and future research directions," Information Fusion, vol. 97, 2023.
11. T. C. Truong, I. Zelinka, J. Plucar, M. Čandík and V. Šulc, "Artificial Intelligence and Cybersecurity: Past, Presence, and Future," Artificial Intelligence and Evolutionary Computations in Engineering Systems, p. 351–363, 2020.
12. M. M. Yamin, M. Ullah, H. Ullah, and B. Katt, "Weaponized AI for cyber attacks," Journal of Information Security and Applications, vol. 57, 2021.
13. B.Guembe, A. Azeta, S. Misra, V. C. Osamor, L. Fernandez-Sanz, and V. Pospelova, "The Emerging Threat of Ai-driven Cyber Attacks: A Review," Applied Artificial Intelligence, vol. 36, no. 1, 2021.
14. N. Kaloudi and J. Li, "The AI-Based Cyber Threat Landscape: ASurvey," ACM Computing Surveys, vol. 53, no. 1, p. 1–34, 2020.
15. Y. Miao, C. Chen, L. Pan, Q.-L. Han, J. Zhang and Y. Xiang, "Machine Learning–based Cyber Attacks Targeting on Controlled Information: A Survey," ACM Computing Surveys, vol. 54, no. 7, p. 1–34, 2020.
16. R. Tahir, "A Study on Malware and Malware Detection Techniques," I.J. Education and Management Engineering, pp. 20-30, 2018.
17. X. Yuan, J. Anderson, L. Yang, J. K. Joshua, and J. Land, "Survey of Recent Hacking Events," in 2019 SoutheastCon, Huntsville, AL, USA, 2019.
18. J. Assen, A. H. C. an, J. Luechinger, P. M. S. S´anchez, G. Bovet, G. M. P´erez and B. Stiller, "RansomAI: AI-powered Ransomware for Stealthy Encryption," arXiv preprint arXiv:2306.15559, 2023.
19. S. Razaulla, C. Fachkha, C. Markarian, A. Gawanmeh, W. Mansoor and B. C. M. Fung, "The Age of Ransomware: A Survey on the Evolution, Taxonomy, and Research Directions," IEEE Access, vol. 11, 2023.

© 2024 IJSRET

20. M. Zou, Y. Shi, C. Wang, F. Li, W. Song, and Y. Wang, "PoTrojan: Powerful neuron-level trojan designs in deep learning models," arXiv preprint arXiv:1802.03043, 2018.
21. Z. Pan and P. Mishra, "Design of AI Trojans for Evading Machine Learning-based Detection of Hardware Trojans," in 2022 Design, Automation & Test in Europe Conference & Exhibition (DATE) IEEE, 2022.
22. R. Zhang, X. Chen, S. Wen, X. Zheng, and Y. Ding, "Using AI to Attack VA: A Stealthy Spyware Against Voice Assistances in Smart Phones," IEEE Access, vol. 7, pp. 153542-153554, 2019.
23. N. Yu, Z. Tuttle, C. J. Thurnau and E. Mireku, "AI-Powered GUI Attack and Its Defensive Methods," in ACM Southeast Conference – ACMSE 2020 – Session 1, Tampa, FL, 2020.
24. K. Chung, X. Li, P. Tang, Z. Zhu, Z. T. Kalbarczyk, T. Kesavadas, and R. K. Iyer, "Machine Learning in the Hands of a Malicious Adversary: A Near Future If Not Reality," in Game Theory and Machine Learning for Cyber Security, John Wiley & Sons, Inc., 2021.
25. K. Renaud, M. Warkentin and G. Westerman, From ChatGPT to HackGPT: Meeting the Cybersecurity Threat of Generative AI, MIT Sloan Management Review, 2023.
26. A. D., V. K. K.A., S. C. S. b, and V. P., "Malware traffic classification using principal component analysis and artificial neural network for extreme surveillance," Computer Communications, pp. 50-57, 2019.
27. O. Illiashenko, V. Kharchenko, I. Babeshko, H. Fesenko and F. D. Giandomenico, "Security-Informed Safety Analysis of Autonomous Transport Systems Considering AI-Powered Cyberattacks and Protection," MDPI Entropy, vol. 25, no. 8, 2023.
28. Z. Alkhalil, C. Hewage, L. Nawaf and I. Khan, "Phishing Attacks: A Recent Comprehensive Study and a New Anatomy," Frontiers in Computer Science, vol. 3, 2021.
29. Microsoft, "How AI is changing phishing scams," 14 July 2023. [Online]. Available: https://www.microsoft.com/en-us/microsoft-365-life-hacks/privacy-and-safety/how-ai-changing-phishing-scams.
30. M. Gupta, C. Akiri, K. Aryal, E. Parker and L. Praharaj, "From ChatGPT to ThreatGPT: Impact of Generative AI in Cybersecurity and Privacy," IEEE Access, vol. 11, pp. 80218 - 80245, 2023.
31. F. Kamoun, F. Iqbal, M. A. Esseghir and T. Baker, "AI and machine learning: A mixed blessing for cybersecurity," in 2020 International Symposium on Networks, Computers and Communications (ISNCC), 2020.
32. K. Choi, J.-l. Lee, and Y.-t. Chun, "Voice phishing fraud and its modus operandi," Security Journal, vol. 30, p. 454–466, 2017.
33. W. Chen and X. Jiang, "Voice-Cloning Artificial-Intelligence Speakers Can Also Mimic Human-Specific Vocal Expression," 2023.
34. H. Malik and R. Changalvala, "Fighting AI with AI: Fake Speech Detection Using Deep Learning," in 2019 AES International Conference on Audio Forensics, 2019.
35. A. Cetinkaya, H. Ishii and T. Hayakawa, "An Overview on Denial-of-Service Attacks in Control Systems: Attack Models and Security Analyses," MDPI Entropy, vol. 21, no. 2, 2019.
36. J. Nazario, "DDoS attack evolution," Network Security, no. 7, pp. 7-10, 2008.
37. K.-C. Yang and F. Menczer, "Anatomy of an AI-powered malicious social botnet," arXiv:2307.16336, 2023.
38. R. Dobbins, "Adaptive DDoS Attacks and Learning How to Suppress Them," NETSCOUT, 2022.
39. R. Doriguzzi-Corin and D. Siracusa, "FLAD: Adaptive Federated Learning for DDoS attack detection," Computers & Security, vol. 137, 2024.
40. B.-X. Wang, J.-L. Chen and C.-L. Yu, "An AI-Powered Network Threat Detection System," IEEE Access, vol. 10, pp. 54029 - 54037, 2022.
41. S. Rajapaksha, J. Senanayake, H. Kalutarage and M. O. Al-Kadri, "AI-Powered Vulnerability Detection for Secure Source Code Development," Innovative Security Solutions for Information Technology and Communications, p. 275–288, 2022.
42. A. N. J. Alzarqawee and L. Fritsch, "Towards AI-powered Cybersecurity Attack Modeling with Simulation Tools: Review of Attack Simulators," Advances on P2P, Parallel, Grid, Cloud and Internet Computing Proceedings of the 17th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC-2022), 2023.
43. C. Y. Cho, J. Caballero, C. Grier, V. Paxson, and D. Song, "Insights from the Inside: A View of Botnet Management from Infiltration," LEET, p. 10, 2010.
44. G. Ollmann, Botnet communication topologies: Understanding the intricacies of botnet Command-and-Control, Damballa, 2009.
45. Y. Wang, J. Ma, L. Zhang, W. Ji, and X. H. Di Lu, "Dynamic game model of botnet DDoS attack and defense," Security and Communications Network, vol. 9, no. 16, 2016.
46. P. Wainwright and H. Kettani, "An Analysis of Botnet Models," in ICCDA '19: Proceedings of the 2019 3rd International Conference on Compute and Data Analysis, 2019.
47. V. A. Memos and K. E. Psannis, "AI-Powered Honeypots for Enhanced IoT Botnet Detection," in 2020 3rd World Symposium on Communication Engineering (WSCE), Thessaloniki, Greece, 2020.
48. M. Conti, N. Dragoni and V. Lesyk, "A Survey of Man In The Middle Attacks," IEEE Communications Surveys & Tutorials, vol. 18, no. 3, pp. 2027 - 2051, 2016.
49. Z. Cekerevac, Z. Dvorak, L. Prigoda and P. Cekerevac, "Internet Of Things And The Man-In-the-Middle Attacks – Security And Economic Risks," MEST Journal, pp. 15-25, 2017.

© 2024 IJSRET

50. A. H. Karbasi and S. Shahpasand, "A post-quantum end-to-end encryption over smart contract-based blockchain for defeating man-in-the-middle and interception attacks," Peer-to-Peer Networking and Applications, vol. 13, p. 1423–1441, 2020.
51. A. Benamira, D. Gerault, T. Peyrin and Q. Q. Tan, "A Deeper Look at Machine Learning-Based Cryptanalysis," EUROCRYPT 2021: Advances in Cryptology – EUROCRYPT 2021, p. 805–835, 2021.
52. H. Kim, S. Lim, A. Baksi, D. Kim, S. Yoon, K. Jang and H. Seo, "Quantum Artificial Intelligence on Cryptanalysis," Cryptology ePrint Archive, 2023.
53. A. Dijk, "Detection of Advanced Persistent Threats using Artificial Intelligence for Deep Packet Inspection," in 2021 IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, 2021.
54. L. Ma, D. Zhao, Y. Gao, and C. Zhao, "Research on SQL Injection Attack and Prevention Technology Based on Web," in 2019 International Conference on Computer Network, Electronic, and Automation (ICCNEA), Xi'an, China, 2019.
55. Y. Tiwari and M. Tiwari, "A Study of SQL of Injections Techniques and their Prevention Methods," International Journal of Computer Applications, vol. 114, no. 17, 2015.
56. J. Irungu, S. Graham, A. Girma, and T. Kacem, "Artificial Intelligence Techniques for SQL Injection Attack Detection," in ICIIT '23: Proceedings of the 2023 8th International Conference on Intelligent Information Technology, Da Nang; Vietnam, 2023.
57. D. Lu, J. Fei, L. Liu, and Z. Li, "A GAN-based Method for Generating SQL Injection Attack Samples," in 2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 2022.
58. G. Deng, Y. Liu, Y. Li, K. Wang, Y. Zhang, Z. Li, H. Wang, T. Zhang, and Y. Liu, "MasterKey: Automated Jailbreak Across Multiple Large Language Model Chatbots," arXiv:2307.08715 [cs.CR], 2023.
59. G. D'Angelo, A. Castiglione, and F. Palmieri, "DNS tunnels detection via DNS-images," Information Processing & Management, vol. 59, no. 3, 2022.
60. M. Sammour, B. Hussin, M. F. I. Othman, M. Doheir, B. AlShaikhdeeb and M. S. Talib, "DNS Tunneling: a Review on Features," International Journal of Engineering & Technology, vol. 7, pp. 1-5, 2018.
61. K. Hynek, D. Vekshin, J. Luxemburk, T. Cejka and A. Wasicek, "Summary of DNS Over HTTPS Abuse," IEEE Access, vol. 10, pp. 54668 - 54680, 2022.
62. D. Mathew, N. C. Brintha and J. T. W. Jappes, "Artificial Intelligence Powered Automation for Industry 4.0," New Horizons for Industry 4.0 in Modern Business, p. 1–28, 2023.

AUTHOR DETAIL

Kiran Sharma Panchangam Nivarthi. AUTHOR was born in Andhra Pradesh, India, in 1982. He received a Bachelor's in electronics and communications engineering (B. TECH) from Visveswaraya Institute of Technology in 2009 and a Master's in cybersecurity Law (MSL) from Francis King Carey College of Law, University of Baltimore, in 2023.

He has worked in the Cybersecurity and Data Privacy for over 16 years. He has been Senior Cybersecurity and Privacy Manager with Snap Finance LLC, Utah, since 2021. From 2008 to 2021, he worked in various cybersecurity and data privacy roles with FICO. He is the author of 4 articles related to privacy and security in the American Scientific Research Journal. His research interests include privacy in the age of artificial intelligence, governance, and AI security.

Mr. Nivarthi's awards and honors include the CSO50 Award in 2017, Royal Fellow (International Organization for Academic and Scientific Development, IOASD), SAS Eminent Fellow (Scholars Academic and Scientific Society, SEFM), Fellow of Information Privacy (FIP, IAPP) and Chair for International Association of Privacy Professionals (IAPP) Minneapolis/St. Paul Chapter. His memberships include the Information Systems Audit and Control Association (ISACA), the International Association of Privacy Professionals (IAPP), the Member (M) of IEEE in 2024, and the British Computer Society (BCS).

© 2024 IJSRET