Editorial

The Transformative Impact of Artificial Intelligence on Cybersecurity

Ioan-Cosmin MIHAI
The Romanian Association for Information Security Assurance

In the dynamic landscape of technology, Artificial Intelligence (AI) has emerged as a pivotal force, driving transformation across many sectors, especially cybersecurity. However, the accelerated evolution of AI technologies has inadvertently created a double-edged sword. While on the one hand, it has fortified our cyber defences, on the other, it has opened up unknown ways for exploitation by malicious actors.

AI can analyse and learn from vast amounts of data, revolutionising many sectors, including cybersecurity. AI-based systems can detect and respond to threats more quickly and accurately than human analysts. However, as AI technologies advance, they are also exploited by threat actors for illicit activities. This dual-use nature of AI has significant implications for cybersecurity and cybercrime.

Cybercriminals can use AI to analyse large amounts of data from various sources across the Internet, identify vulnerabilities of systems and users, and improve and intensify cyber-attacks. Traditional cyber-attacks involve an attacker's presence and interaction with specific computer systems to make certain decisions, often limiting his anonymity. AI obfuscates the connection and increases the distance between the victim and attacker, making it difficult to investigate and assign them.

Social engineering remains one of the main vectors of attack. Cybercriminals use AI to analyse vast amounts of data on the Internet and the dark net, build a social profile of potential victims based on their behaviour in cyberspace and create attractive emails and websites for targeted users. AI can gain users' confidence through sustained dialogue across various social platforms without the involvement of human attackers. This AI-driven social engineering represents a significant shift in the tactics of cybercriminals, making it more challenging to detect and prevent attacks.

AI can create malware applications that compromise information systems and data. Once AI makes these types of malware, they can propagate on a much broader scale and exploit any encountered vulnerabilities. AI can make decisions based on discovered vulnerabilities and the desired purpose, eliminating the need for command and control servers used by cybercriminals. This autonomous decision-making capability of AI-based malware represents a significant evolution in the threat landscape.

The release of GPT-4 was meant not only to improve the functionality of ChatGPT but also to make the model less likely to produce potentially harmful output. The

European Union Agency for Law Enforcement Cooperation (EUROPOL) identified various criminal use cases in GPT-3.5. However, a subsequent check of GPT-4 showed that all of them still worked. In some cases, the potentially harmful responses from GPT-4 were even more advanced.

ChatGPT excels at providing the user with ready-to-use information in response to various prompts. Suppose a potential criminal knows nothing about a particular crime area. In that case, ChatGPT can significantly speed up the research process by offering essential information that can be further explored in subsequent steps. As such, ChatGPT can be used to learn about many potential crime areas with no prior knowledge, ranging from how to break into a home to terrorism, cybercrime and child sexual abuse. The identified use cases that emerged from the workshops Europol carried out with its experts are incomplete. Instead, the aim is to give an idea of how diverse and potentially dangerous LLMs such as ChatGPT can be in the hands of malicious actors.

The beginning of AI tools like FraudGPT and WormGPT has further complicated the cybersecurity landscape. FraudGPT, an AI bot targeted for offensive purposes, is being sold on dark web marketplaces and Telegram channels. It can be used to craft spear phishing emails, create cracking tools, and perform carding. WormGPT, a black hat alternative to GPT models, is used by threat actors to launch business email compromise (BEC) attacks. It enables threat actors to automate convincing personalised fake emails, expanding the scope of BEC attacks and boosting the attack's success. These AI tools represent a significant advancement in the capabilities of cybercriminals, making it more difficult for organisations to defend against attacks.

While AI can potentially enhance cybersecurity, it also poses significant risks. The emergence of AI tools like FraudGPT and WormGPT highlights the urgent need for robust AI security measures. Implementing a defence-in-depth strategy with all the security telemetry available for fast analytics has become essential to finding these fast-moving threats before a phishing email can become ransomware or data exfiltration.

As AI continues to evolve, all actors involved in cybersecurity, including researchers, practitioners, and policymakers, must stay updated on the most recent findings and understand this technology's ramifications. Future research should develop more robust protection against AI-generated cyber-attacks and investigate how we can use the power of AI to strengthen cybersecurity efforts.

In terms of best practices, a multi-layered approach to cybersecurity is essential. We can use a variety of technical measures, including deploying antivirus software, firewalls, and intrusion detection systems. Additionally, it involves educating users to boost their awareness of potential cyber threats and integrating AI into cybersecurity solutions to improve their effectiveness. Cooperation and information sharing among organisations and between the public, private, and academic sectors can greatly enhance collective cybersecurity efforts. Despite the challenges, by staying informed and working together, we can harness the potential of AI to strengthen cybersecurity.