The Rise of AI in Cyber Attacks and Defence

BY **TAN YUH WOEI,** General Manager – ASEAN, Google Cloud Security Sales

In 2024, artificial intelligence (AI) will be increasingly harnessed in both offensive and defensive capacities in cyber warfare. Sophisticated algorithms will be employed not only to orchestrate cyber attacks but also to bolster the resilience and efficacy of defensive measures against such threats.

From helping to summarise meetings to writing the software code to control machines on a factory floor, AI has taken great strides forward in the past year and captured the imagination like never before. Perhaps this is down to the convincing human-like responses that generative AI, such as intelligent chatbots, can produce when interacting with people.

At the same time, the development of AI has also been accelerated by the greater processing power available today via cloud computing. This has enabled AI to learn from larger and more varied sets of data, ensuring it is "smarter" through new ways of ingesting the data.

Clearly, businesses cannot ignore AI today. According to research firm IDC, its adoption is set to soar in Asia-Pacific, with 80 per cent of chief information officers harnessing it by 2028.

AI in cyber attacks

AI, like many advancing technologies, changes the equation for cyber security. Just as businesses are using AI to automate tasks and communicate better, cyber attackers are doing the same, but for malicious reasons.

AI is a tool that can be used by cyber criminals to create new threats for businesses, but it is also an innovation that can help shore up defences against online threats.

In 2024, cyber criminals are expected to use AI to make their attempts to break in more sophisticated, according to Google Cloud's latest cyber security forecast for the year.

In Singapore, the police reported that scams and cyber crimes increased by close to 50 per cent from a year earlier. Faced with such threats, it is imperative that businesses are aware of and take action to defend against emerging new threats that cyber criminals can unleash with the help of AI.

Phishing becomes more sophisticated

Phishing, which involves tricking a victim into giving up access, will be improved, professionalised and scaled up in the coming months with the help of AI. Generative AI and the large language models (LLMs) that it is trained on will be used by cyber criminals to create content that appears more legitimate to potential victims. This could be in the form of voice or video content.

Misspellings and grammar errors, which are often telltale signs of a fake message, say, from the authorities, will be eradicated with help from phishing content generated by AI. The same goes for cultural context – AI can help a hacker group that has no knowledge of a victim's societal norms craft a message that resonates deeply.

And LLMs can also help translate and clean up translations as well. Instead of hiring someone separately on the Dark Web to create messages for multiple locations across the globe, cyber criminals can now tailor them easily using free AI tools.

Even more insidious is the ability to weave in legitimate content. For example, a cyber-criminal could use AI to insert real passages previously used, for example, by a bank to communicate with customers. They can then generate a fake version that incorporates similar content and features. A phishing message will look, flow and read like the real thing when it is actually directing a victim to a fake website, for example.

AI for good and for bad

AI can be used in more bad (and ugly) ways than one. Fortunately, AI can also be used to do some good.

The box, "The Good, Bad and Ugly of AI in Cyber Warfare", describes these three different facets.

The Good, Bad and Ugly of AI in Cyber Warfare

The Good: Using AI to keep out malicious actors

Just as the bad guys can use AI to boost efficiency and scale up capabilities, businesses should harness AI to level the playing field. With generative AI and related technologies, organisations can improve the detection, response and attribution of cyber adversaries at scale.

Analysis of potential data breaches can be speeded up, as can reverse engineering of malicious software to unravel damage. Businesses can synthesise the large amounts of data at hand and contextualise it for threat intelligence detection and response.

AI can help cyber defenders better understand and prioritise the "flashing lights" or alerts on their cyber defence systems so they can target the most important and pertinent areas for remediation.

With generative AI augmenting human capabilities in inferring actions to take from large data sets, there will be new ways to overlay customer-specific data in a highly confidential manner. Businesses can glean insights that let them take action at speed and at scale.

AI can also help address the lack of talent in the cyber security field. By enabling human operators to make the big decisions based on actionable intelligence, AI can reduce the toil of manual analysis, such as analysing text logs to trace a cyber security incident.

The Bad: Scaling up attacks with AI's help

With generative AI, cyber criminals can scale up their malicious activities more easily. Phishing attempts can be executed by creating messages based on known names, job titles and even the health data of people from a target organisation.

There is no need for a specialised AI trained on a malicious LLM, because a regular, publicly available AI might think it is only generating content, say, to deliver regular marketing messages.

Besides phishing, cyber attackers can also use AI to create fake phone calls, for example, from a bank to trick victims into giving up personal information and possibly even login credentials to access their digital accounts.

Of concern as well is deepfake technology, which is becoming more commonly available. Hacker groups could use generative AI to create fake photos and videos to convince people of fake information.

What such deepfakes also accomplish when they become common is to make people uncertain of what they see and read. Eventually, public trust may be impacted, and it becomes more difficult for institutions, such as governments, banks and schools, to engage with their audiences to dispel scams and frauds.

The Ugly: AI as a service to mount attacks

The availability of more powerful deepfake tools that will become more persuasive over time means that even smaller, less sophisticated cyber criminals will be able to craft highly convincing scams that trick victims into losing personal data or more.

Many thrive in the underground economy that is behind many cyber attacks. In the dark corners of the Internet are hacker networks that help cyber criminals put together a campaign with the latest tools and exploits available.

Like in any production line, there is a clear separation of roles for each cyber attack. One contractor might be in charge of crafting phishing messages, while another prepares and tests a payload and yet another ensures that a ransom is paid to the right account.

Such arrangements have brought success to malicious actors and this as-a-service model is likely to escalate. Cyber criminals will also incorporate more sophisticated phishing and misinformation capabilities garnered through AI into their list of services.

A buyer could hire such cyber criminals to target a business, for example, with a stated aim of damaging its reputation or extracting a ransom by breaching its defences and stealing confidential data.

So, what does the future hold?

What to expect

Due to its geographic location and status as a financial and business hub for international business, Singapore will continue to be an attractive target for cyber criminals.

This is amid some important shifts in cyber security trends in Asia-Pacific, which will also face new types of tactics, techniques and procedures employed by cyber attackers. As more of the region raises its readiness through endpoint detection and response solutions, cyber criminals will also up their game, the *Google Cloud Cybersecurity Forecast* predicts.

As seen in other parts of the world, well-resourced threat actors will be constantly seeking ways to get past detection and defences. They can be expected to make use of zero-day vulnerabilities, so-called because they are not known to the public and thus difficult to defend against.

These vulnerabilities may present themselves in various parts of the digital infrastructure used by a business, including in security, networking and software apps. Businesses in Singapore and the region have to be prepared hacker groups can bypass their detection efforts and stand ready to remediate such threats to minimise damage.

In summary, cyber threats continue to be a spectre that follows the rapid digitalisation that tech-savvy countries such as Singapore have enjoyed in recent years.

Increasingly, cyber attackers will take advantage of advances in AI to deepen their capabilities and broaden the scale of their malicious efforts. This necessitates an urgent response by businesses to upgrade defences, improve readiness and draw on the strengths of AI to combat these new threats.