AI Cyber Defense (and Offense)

Ghostwritten for Michael MurrayAI Cyber Defense (and Offense)

Published on LinkedIn

I'm not imparting any great wisdom when I say that AI deployment is on a steep upward curve. In fact, AI adoption went from roughly one in five organizations in 2019 to almost two out of three in 2020. There's no reason to believe that number won't keep growing exponentially. While this is great news in many ways (especially to AI providers), opportunistic hackers are now using AI in their attacks.

Organizations are turning to AI cyber monitoring tools (like Darktrace, for example) to protect themselves from threats that overwhelm even the best analysts and existing systems. The number of new strategies and tactics seen from nation states, clandestine groups and individual bad actors have exploded in recent months as demand and access to global cloud-based operational systems increases.

Cyber analysts are finding it increasingly difficult to effectively monitor current levels of data volume, velocity, and variety across firewalls. The convergence of OT and IT networks is proving difficult for defense, government and critical infrastructure firms as previous commercial and non-AI based cybersecurity solutions are unlikely to deliver the requisite performance to detect new AI-based attack vectors.

To this end, it is becoming more difficult to identify critical threats without AI. This is a significant issue since many firms are assuming "AI will protect us" without realizing that AI Cybersecurity is also vulnerable to other AI which adversaries are deploying for attacks. Then the game becomes one of spending. It's an arms race to see who can spend the most for the most advanced AI Crypto solution. Who will learn the patterns of the other more quickly? Which will be the first to set dataset traps?

It's also important to note here that AI implementations are only as good as their setup and the people who train them. This means that an AI cyber monitoring tool, even a very good one, can still fail due to human error.

Examples of AI in Cybersecurity

Perhaps the best way to clarify the issue of AI in cybersecurity is through two examples.

In the first, AI learns patterns from data artifacts that are both obvious and hidden. In a game like checkers, there are a finite number of moves a human or machine can make on the checkerboard. Thus, playing checkers or chess against an average computer doesn't work out well for most of us.

This begs the question: are there finite numbers of moves to be made in a cyber attack or on a battlefield? Or are there an infinite quantity of moves to be made? I argue the latter, but if that's true then AI can only provide a portion of an organization's cybersecurity solution.

In the second example, let's look at IoT. Even with the advancements in AI, there are a staggering number of threat vectors (growing daily) sought and exploited through IoT devices. The reason for this is again tied to spending - the cost and complexity to attack is low, while the cost for defending is disproportionately high.

Further, these IoT access points are more difficult for AI to secure due to the fact they exist in the physical world where routines are less defined. These systems have a higher degree of freedom, autonomy and/or human interaction than pure digital systems.

How Hackers Foil AI Defenses

It is important to consider and even imagine how attackers might seek to destabilize AI models and learning systems. As new AI attacks emerge, data poisoning, clandestine learning/logic bombs and backdoors within training models and systems can and are being deployed to infect machine learning systems. The disastrous SolarWinds hack is one example of this technique. AI systems can be vulnerable to attacks from various training models, algorithm artifacts and the historic nature of previous attacks which can lead an attacker to create a simulation that takes the AI in a certain direction preferred by the attacker.

Machine-learning and deep-learning techniques will make sophisticated cyber-attacks easier and allow for faster, better targeted, and more destructive attacks. The impact of AI on cybersecurity will likely expand the threat landscape, introduce new threats, and alter the typical characteristics of threats. Besides, other than introducing new and powerful vectors to carry out attacks, AI systems will also become increasingly subject to manipulation themselves.

Compared to traditional hardware-software systems, AI-powered systems present specific features that can be attacked in non-traditional ways: in particular, the training data set may be compromised so that the resulting “learning” of the system is not as intended. Alternatively, external objects that will be sensed by the system can be tampered with so that the system fails to recognize them. It is therefore important to provide additional, ad hoc protection of AI systems, to ensure that they follow a secure development life cycle, from ideation to deployment and post-market surveillance, including runtime monitoring and auditing.

A “Good” AI Faces a Losing Fight

Using AI for defensive purposes faces several constraints, especially as governments including the European Union move to regulate high-risk applications and promote the

responsible use of AI. Meanwhile, on the attack side, the most pernicious uses are multiplying, the cost of developing applications is plummeting, and the “attack surface” is becoming denser every day, making any form of defense an uphill battle.

On the other hand, the IoT age will further densify the attack surface. AI is thus a ‘must’ to help companies and organizations manage this range of cybersecurity risk, technical challenges, and resource constraints. AI can improve systems’ robustness and resilience, but several conditions must be met with regards to edge node identity, visibility, and encryption.

As the maturity rate increases within AI systems, they will be incorporated into new products, services, and solutions to decrease the threats surrounding known attack scripts, programs, and routines. When used in conjunction with traditional methods, AI is a powerful tool for protecting against cybersecurity attacks. In the Internet Age, with hackers’ ability to commit theft or cause harm remotely, shielding assets and operations from those who intend harm has become more difficult than ever.

Employing AI judiciously as part of an overall cybersecurity regimen is a necessity, but caveat lector – it’s not as easy as setting your own AI against an attacker’s. Striking the right balance is essential, but far from simple. A sound cybersecurity combination of traditional and AI protection provides a holistic approach that reduces the overall exposure of the organization. The weakest link in your cybersecurity chain will always be human beings, so training and an understanding of the basic principles are necessary, but you also need to make the system as easy to follow as possible to prevent employees from doing an end run around your security.

It’s a delicate balancing act, and one that will only continue to grow in complexity, but defending against advanced AI cyber threats should be top of mind for organizations today.