Machine learning algorithm may be the key to timely, inexpensive cyber-defense

5 February 2021, by Matt Swayne

A machine learning algorithm may give organizations a powerful and cost-effective tool for defending against attacks on vulnerable computer networks and cyber-infrastructure, often called zero-day attacks, according to researchers.Image: Pixahive A machine learning algorithm may give organizations a powerful and cost-effective tool for defending against attacks on vulnerable computer networks and cyber-infrastructure, often called zero-day attacks, according to researchers. Credit: Pixahive

Attacks on vulnerable computer networks and cyber-infrastructure—often called zero-day attacks—can quickly overwhelm traditional defenses, resulting in billions of dollars of damage and requiring weeks of manual patching work to shore up the systems after the intrusion.

Now, a Penn State-led team of researchers used a machine learning approach, based on a technique known as reinforcement learning, to create an adaptive cyber defense against these attacks.

According to Minghui Zhu, associate professor of electrical engineering and computer science and Institute for Computational and Data Sciences co-hire, the team developed this adaptive machine learning-driven method to address current limitations in a method to detect and respond to cyber-attacks, called moving target defense, or MTD.

"These adaptive manual target-defense techniques can dynamically and proactively reconfigure deployed defenses that can increase uncertainty and complexity for attackers during vulnerability windows," said Zhu. "However, existing MTD techniques suffer from two limitations. First, manual selection can be very time consuming. Secondly, manually selected configurations might not be the most cost-effective method to handle this."

Typical responses to an attack can take up to 15 days, which can use up significant funds and resources for an organization, according to the researchers, who released their findings in the *ACM Transactions on Privacy and Security*.

Zhu said that zero-day attacks are among the most dangerous threats to computer systems and can cause serious and lasting damage. As an example, the WannaCry ransomware attack, which occurred in May 2017, targeted more than 200,000 Windows computers across 150 countries, and caused an estimated $4 billion to $8 billion worth of damage.

The team's approach relies on reinforcement learning, which, along with supervised and unsupervised learning, is one of the three main machine learning paradigms. According to the researchers, reinforcement learning is a way that a decision maker can learn to make the right choices by selecting actions that can maximize rewards by balancing exploitation—leveraging past experiences—and exploration—trying new actions, according to Peng Liu, the Raymond G. Tronzo, MD Professor of Cybersecurity in the College of Information Sciences and Technology.

"The decision maker learns optimal policies or actions through continuous interactions with an underlying environment, which is partially unknown," said Liu. "So, reinforcement learning is

particularly well-suited to defend against zero-day attacks when critical information—the targets of the attacks and the locations of the vulnerabilities—is not available."

The researchers tested their reinforcement learning algorithm in a 10-machine network. They added that although a 10-computer network may not seem very large, it is actually more than robust enough for the test. The setup also included web and mail servers, a Gateway server, SQL server, DNS server and admin server. A firewall was installed to prevent access to the internal hosts. The researchers also selected vulnerabilities that could produce multiple attack scenarios for the test.

The researchers added there is room for further improvement for their approach. For example, their algorithm relies on model-free reinforcement learning, which requires a large amount of data or a large number of iterations to learn a relatively good defense policy. In the future they would like to incorporate model-based approaches to accelerate the learning process.

More information: Zhisheng Hu et al. Adaptive Cyber Defense Against Multi-Stage Attacks Using Learning-Based POMDP, *ACM Transactions on Privacy and Security* (2020). DOI: 10.1145/3418897