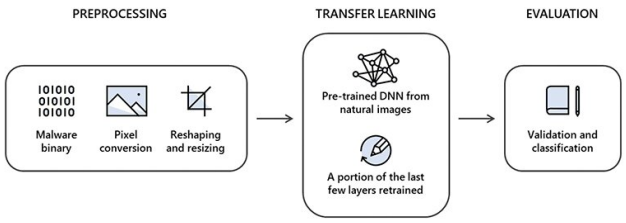
Microsoft-Intel project converts malware into images to cut threats

12 May 2020, by Peter Grad



Credit: Microsoft

Two industry giants are working to get a clearer picture of how to combat malware—literally.

Members of Microsoft's Threat Protection Intelligence Team have joined representatives of Intel Labs to create images out of malware samples that can be used to detect malicious code .

Using an approach called static malware-as-image network analysis (STAMINA), researchers feed malware samples into a program that converts the data into grayscale images. They then analyze the samples for structural patterns that can be used to distinguish between benign and malicious code, and then rank the malicious suspects into degree of threat.

The study relied on earlier work by Intel on deep transfer learning for static malware classification. Deep learning is a component of artificial intelligence relying on machine learning, smart computer networks that learn on their own.

Static analysis permits malware detection without having to execute code or monitor runtime behavior.

Drawing on Microsoft's massive dataset of malware code collected through its Defender security system, the researchers say they achieved "high accuracy" in detecting malware and "low false positives."

With static analysis, most threats are detected before they are triggered, according to the Microsoft report posted on its security blog about STAMINA on May 8.

"While static analysis is typically associated with traditional detection methods," the report says, "it remains to be an important building block for AI-driven detection of malware. It is especially useful for pre-execution detection engines: static analysis disassembles code without having to run applications or monitor runtime behavior."

The study consisted of three steps: image conversion, transfer learning and evaluation. In a process that included pixel conversion and resizing, malware code drawn from 2.2 million infected files was converted into two-dimensional images. The next step used transfer learning to apply knowledge obtained about detected malware in one task to similarly structured unidentified code. The last step was evaluation.

The report states the STAMINA program achieved an accuracy of more than 99 percent identifying and categorizing malware samples, with a false positives rate of 2.6 percent.

In a white paper distributed by Intel, researchers explain: "As malware variants continue to grow, traditional signature-matching techniques cannot keep up. We looked to applying deep-learning techniques to avoid costly feature engineering and used machine-learning techniques to learn and build classification systems that can effectively identify malware program binaries."

For now, the program works best with smaller file sizes.

"For bigger size applications, STAMINA becomes less effective due to limitations in converting billions of pixels into JPEG images and then resizing them," the report says.

Microsoft Defender began as an anti-spyware program first offered with Windows XP and has subsequently expanded into a full anti-virus and anti-malware system as part of the Windows Security package included with Windows 10. In a 2018 study, leading spyware research lab AV-TEST found Defender achieved a 100 percent detection rate of malicious URL samples, and three false positives.

More information:
www.microsoft.com/security/blo … ware-classification/

www.intel.com/content/www/us/e … tion-whitepaper.html

APA citation: Microsoft-Intel project converts malware into images to cut threats (2020, May 12) retrieved 27 September 2020 from https://techxplore.com/news/2020-05-microsoft-intel-malware-images-threats.html