Solution Brief

Trend Micro

Tackling Unknown Threats Using Machine Learning With Next-Generation Intrusion Prevention

Combating dynamic threats

Modern threats are evanescent, incredibly short-lived and ever-changing. Signature-based detection, even using regular expressions, cannot identify a wide range of current threats. It is not the right toolset, as signatures would have to be developed, monitored, and updated for specific instances of a wide variety of threats. Even still, these signatures would often be outdated as soon as they are released because of the dynamic nature of modern threats. Statistical models effectively close these security gaps.

The ability to evaluate security filters that represent statistical models of malicious network data will greatly improve the security effectiveness of technologies like Next-Generation Intrusion Prevention Systems (NGIPS). IT security managers are faced with a series of challenges: increasingly sophisticated threats, riskier user behavior, and a lack of visibility across their security systems.

The following use cases address some of the security gaps that cannot be solved by traditional signature-based detection solutions: malicious HTML content including JavaScript, malicious files, and malicious Adobe objects including Flash and PDF. Statistical models can be used to identify all of these, in-line, in real time.

Exploit Kits and Other Malicious Obfuscated HTML

Live Stack testing from NSS Labs makes use of a number of exploit kits, which account for over 80% of Live Stack testing misses. These exploit kits deliver malicious obfuscated HTML, with code objects encoded in some manner within the HTML document to be decoded by later JavaScript or VBScript evaluation. Delivering protection against exploit kits, which are designed to evade detection by regular expressions, provides a substantial increase in security effectiveness to these prevalent and growing threats[1].

Malicious Obfuscated Javascript

JavaScript is increasingly used to deliver malicious content, including attacks that use JavaScript alone to accomplish malicious actions[2]. These scripts are obfuscated in order to evade detection by signatures or regular expressions.

Statistical models can identify malicious obfuscated JavaScript, and close this gap. This is similar to the detection of obfuscated HTML, but focuses on malicious obfuscated JavaScript. This can apply to importation of JavaScript files and not just to HTML documents <script> elements.

Malicious Flash Objects

One of the largest gaps in security effectiveness in the NGIPS is the lack of coverage in identifying malicious Flash objects. Malicious Flash objects and PDF files are widely used in attacks.[3] Many of the vulnerabilities disclosed to the Trend Micro™ Zero Day Initiative™ (ZDI) bug bounty program are within Adobe products. External research indicates that static analysis can detect malicious Flash files. Where static analysis is successful, statistical models can be applied.

Malicious PDF Files

Another large gap in security effectiveness in the NGIPS is the lack of coverage in identifying malicious PDF files. External research indicates that static analysis can detect malicious PDFs. Because malicious PDFs often incorporate malicious Flash objects or obfuscated JavaScript, completing 2.2 Malicious Obfuscated Javascript and 2.[3] Malicious Flash above is prerequisite to this effort.

Malicious Portable Executable (PE) Files

Polymorphic malware results in over one million new malware samples per day.[4] Using PE headers alone, statistical models can predict whether an executable is malicious with greater than 95% accuracy.[5] Internal research from the Trend Micro™ TippingPoint™ Digital Vaccine™ Labs (DVLabs) team has verified that creation of statistical models that identify malicious PE files is straightforward and effective.

Custom Packed Files

Over 75% of malware executables are packed[6]. Regular expression filters can block files packed with off-the-shelf packing utilities, but malware authors are increasingly using custom polymorphic packers to evade detection. Custom packed files can be detected by measuring the compressibility of the files (ibid). While compressibility will not be directly used because of the latency it would introduce, related complexity measures such as entropy may be used. This approach is simpler and more direct using a model based on PE imports because it relies on very few features, and possibly just one.

Using Machine Learning to Address Unknown Threats with NGIPS

Across many industries, machine learning techniques are being quickly adopted; however, Trend Micro is the first to leverage this capability to detect and eliminate some of the threats mentioned above in-line at wire speed through the TippingPoint Next-Generation Intrusion Prevention System (NGIPS) and Threat Protection System (TPS).[7] Our revolutionary approach powered by XGen™ security provides an additional measure of security on top of traditional signature-based approaches to intrusion prevention.

The following illustration details a very simplistic representation of machine learning capabilities using the Trend Micro TippingPoint solutions.

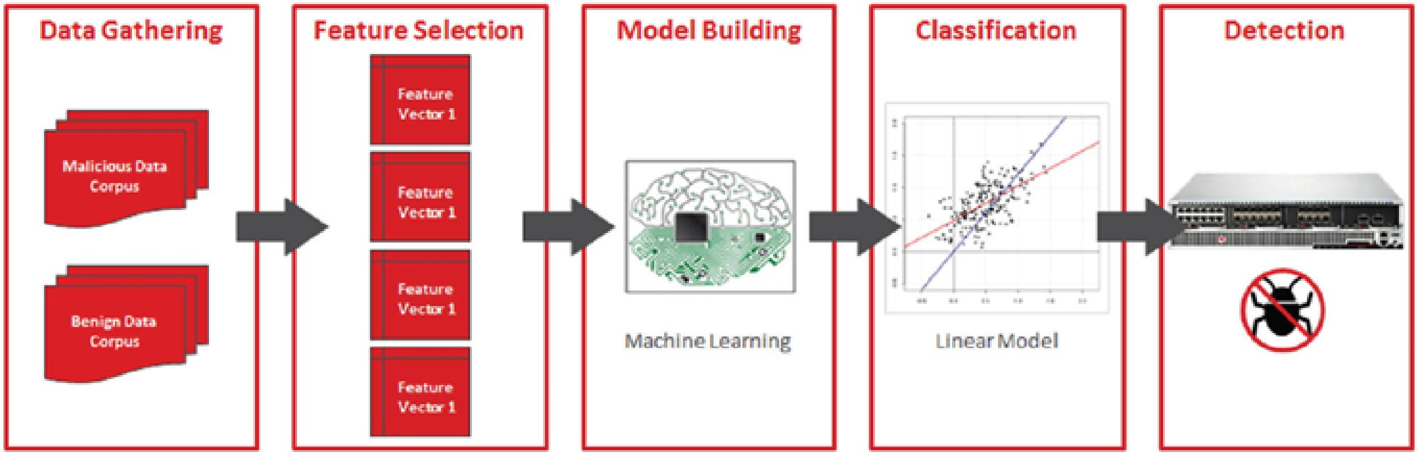

*Illustration: Basics of machine learning and application to the TippingPoint **NGIPS and TPS***

Digital Vaccine™ (DV) filter packages used by the TippingPoint NGIPS and TPS are a strong mechanism to detect network-based malicious activity, exploitation of vulnerabilities, and unwanted application use. However, as the TippingPoint solutions block these critical attacks more effectively, exploit kit authors adjusted their tactics to evade traditional signature-based techniques such as pattern-matching regular expressions. They now obfuscate content, including packing/compression, script obfuscation, encryption and much more. This makes classic detection mechanisms extremely difficult, often requiring multiple signatures and in many cases, only detecting a subset of the malicious content.

This is where machine learning and statistical data modeling become so effective. At a high level, machine learning works by training a machine by extracting "feature vectors" from a dataset of benign and malicious examples in order to compute a mathematical model. This model is evaluated against network traffic and, in the case of the TippingPoint solutions, can make a real-time decision about whether the content appears to be benign or malicious. If the content is determined to be malicious, the TippingPoint solutions block the content from entering the network. DV filters developed using the mathematical models operate without affecting network performance and without introducing a high amount of false positives.

Trend Micro TippingPoint also uses machine learning to detect Domain Generation Algorithms (DGAs) used in many malware families (for example, Conficker) to randomly generate domain names in order to contact their command and control (CnC) servers. TippingPoint Threat DV DGA filters include classifiers, developed using machine learning techniques across a significant DNS datasheet, that can detect families of DGAs using a combination of syntactical rules and logistic regression with over 95% accuracy. DGA filters are also in place to catch many types of malware whose domain names cannot be encompassed by a regular expression that would not generate a large number of false positives.

References:

1. https://www.trendmicro.com/cloud-content/us/pdfs/security- intelligence/white-papers/wp-evolution-of-exploit-kits.pdf
2. http://www.computerworld.com/article/3018972/security/ransom32- first-of-its-kind-javascript-based-ransomware-spotted-in-the-wild.html
3. http://www.computerworld.com/article/2521020/security0/rogue-pdfs- account-for-80--of-all-exploits--says-researcher.html
4. http://money.cnn.com/2015/04/14/technology/security/cyber-attack-hacks-security/
5. https://arxiv.org/ftp/arxiv/papers/1308/1308.2831.pdf
6. http://money.cnn.com/2015/04/14/technology/security/cyber-attack-hacks-security/
7. http://www.trendmicro.com/us/business/network-security/intrusion-prevention-system/