CYLANCE

CylanceINFINITYENGINE

CylanceINFINITYENGINE is an embeddable malware detection technology that uses Cylance's predictive models to classify files as good or bad by correlating them with the features found in millions of good and bad samples. This enables our models to detect even zero-day and previously unknown malware not in the original training set.

Traditional antivirus (AV) products, on the other hand, only evaluate a specific file against a finite list of signatures created through manual human analysis. Even if they use some automated techniques, these solutions are limited to signatures based on specific parts of files that were previously identified as known malware by a human.

Instead of using manually-created signatures, INFINITYENGINE computes a "confidence score" for every sample it processes. A Cylance confidence score of 0.80 means the model is 80% confident the file is good. A Cylance confidence score of -0.80 means the model is 80% confident the file is bad.

INFINITYENGINE also checks for various capabilities that are prevalent in malware and provides a threat indicator report to explain the classification. For example, the report will call out an executable's capabilities like logging of keystrokes, ability to inject code or terminate other processes, ability to tamper with windows firewall policy, etc. This provides a helpful data point to accelerate further analysis and response to detected threats.

Key Benefits

Deploy Game-Changing Malware Detection - INFINITYENGINE identifies threats in Windows portable executables, PDF's and Microsoft Office documents using predictive models which classify files as malware. Because it doesn't rely on traditional signatures or hash cloud lookups, it is the industry's leading solution to detect and classify zero-day and unknown malware.

Raise the Barrier on Attackers – Our predictive learning technology inspects 1000's of individual features in each file to classify the file as good or bad. This approach is superior to traditional signature-based detection because attackers and malware authors would require exponentially greater effort and resources to bypass detection.

Embed as a Component – INFINITYENGINE is a self-contained, small footprint service which can be integrated into SaaS products or appliances in less than a week. INFINITYENGINE runs locally and does not require any behavioral analysis or cloud lookups. In addition, cloud API's are available to offload classification or provide additional evidence.

Reduce Malware Analysis Time – INFINITYENGINE provides a score and threat indicators for detected malware. In addition, an evidence API reports on behavior observed when the malware is detonated inside a cloud sandbox.

Who embeds INFINITYENGINE?

Next-Gen Firewall/UTM/Email Gateways – detect on-application threats and conduct static analysis at network ingress points

SaaS Applications – scan and quarantine threats backed up by enterprise file synchronization and sharing products, as well as provide embedded scanning of cloud storage devices

Sandboxing/Forensics – pre-scan files prior to sandbox for prioritization or dynamic whitelisting and static indicators for malware

Portable Storage Devices – secure portable OS images like Windows to Go which are loaded on USB drives issued to employees or contractors

Vehicle Entertainment Systems - prevent accidental downloads of threats that might negatively impact vehicle entertainment systems

How It Works

Cylance is the only solution provider with this unique approach, allowing a lightweight, extremely fast, algorithmic analysis of what's safe and what's a threat. By examining tens of thousands of characteristics in each sample, Cylance makes instant decisions on what is good or bad, even the unknowns. Traditional security solutions rely on signatures and simple heuristics which are good at detecting malware they have analyzed before, but not new threats.

In addition, humans are incapable of leveraging vast amounts of data to make a determination on the maliciousness of a file due to the magnitude of the data involved, the tendency towards bias, and the number of computations required. Unfortunately, this is exactly the way that most security companies operate. They hire a large number of people to look through millions of files to determine which are good or bad. Humans have neither the brainpower nor the physical endurance to keep up with the overwhelming volume and sophistication of threats today. Advances have been made in behavioral analysis, vulnerability analysis, and identification of indicators of compromise, but these all suffer from the same fatal flaw. They are all based on a human perspective and analysis, which tends to over simplify the problem. Machines, however, do not suffer from the same challenges.

Machine learning focuses on prediction, based on properties learned from earlier data. This is how Cylance identifies malicious versus safe or legitimate files. Machine learning leverages a four-phase process of collection, extraction, learning and training, and classification:

Collection
Much like DNA analysis or an actuarial review, file analysis starts with the collection of a massive amount of data – in this case files of specific types (executables, dlls, .pdf, .doc, .xls, etc.). Hundreds of millions of files are collected via 'feeds' from industry sources, proprietary organizational repositories and live inputs from active computers being protected by the Cylance agent.

Extraction
Attribute extraction is the next phase in the machine learning process. This process is substantively different from the process of behavior identification or malware analysis currently conducted by threat researchers.

Rather than looking for things which people believe are suggestive of something that is malicious, Cylance leverages the computing capacity of machines and data mining techniques to identify the broadest possible set of characteristics of a file. These characteristics can be as basic as the PE file size or the compiler used and as complex as a review of the first logic leap in the binary. Cylance extracts the uniquely atomic characteristics of the file depending on its type (.exe, .dll, .com, .pdf, .java, .doc, .xls, .ppt, etc.).

By identifying the broadest possible set of attributes, Cylance removes the bias introduced by the manual classification of files. Use of thousands of attributes, also substantially increases the cost for an attacker to create a piece of malware that is not characterized by Cylance.

Learning and Training

Once the attributes are collected, they are normalized and converted to numerical values that can be used to build statistical models. Leveraging the millions of attributes of files identified in extraction, Cylance mathematicians then develop statistical models that accurately predict whether a file is valid or malicious.

It is important to remember that for each and every file, thousands of attributes are analyzed to differentiate between legitimate files and malware. This is how Cylance identifies malware - whether packed or not, known or unknown – and achieves an unprecedented and absolutely phenomenal level of accuracy. It divides a single file into an enormous number of characteristics, and analyzes each one against hundreds of millions of other files to make decisions about the normalcy of such characteristics.

Classification

INFINITYENGINE embeds the statistical models and enables customers to classify samples either locally or through a cloud lookup. This classification takes milliseconds and is very precise because of the breadth of the file characteristics analyzed. Additional API's are available to generate threat indicator and sandbox detonation evidence reports. These reports are generated post-classification to provide contextual information to validate the classification made by the statistical models. Threat indicators are common patterns that security analysts look for while analyzing new samples. We also have a sandbox in the cloud in which samples can be detonated to observe behavioral patterns.

Usage Guide

INFINITYENGINE uses the power of math to classify files as bad or good. SampleScoringTest analyzes a given file and as a result, provides a numerical score. The score is the confidence INFINITYENGINE has that a file is good or bad. Positive scores indicate good files. Negative scores indicate bad files.

INFINITYENGINE has two main components:

Service/InfinityService.sh – Executes the INFINITYSERVICE that provides the technology used to classify a sample. It will listen for connections from the client at the desired port (see examples below).

Client/SampleScoringTest.py – Leverages INFINITYENGINE technology to process samples and generate a report including score values. It connects to the INFINITYSERVICE at the desired port.

Requirements

Linux 64-bit (tested on Ubuntu 12.04.4-desktop-amd64)
Python 2.7.3 (older versions may still work)
Python modules: python-magic

How to Use

Run the INFINITYSERVICE at the desired port number (9002 by default). From the Service directory, run:

```
./InfinityService.sh
```

Use the Python script to classify samples. From the Client directory, run:

```
./SampleScoringTest.py --out ~/Desktop/results.csv ~/samples
```

Get usage instructions by executing:

```
./SampleScoringTest.py -h
```

Reviewer Guide

1. Install python 2.7.3 + (with python-magic) on a 64-bit linux system.

2. Get samples of malware from your internal repository or from private and public malware feeds. Feel free to contact Cylance for samples if you don’t have your own.

3. Review the results in the CSV file. A positive score means the file is good and a negative score means it is a threat.

4. Compare Cylance’s results against the results of other engines.

For support, please send an email to rpermeh@cylance.com.

About Cylance:

Cylance is the first company to apply artificial intelligence, algorithmic science and machine learning to cybersecurity and improve the way companies, governments and end-users proactively solve the world’s most difficult security problems. Using a breakthrough predictive analysis process, Cylance quickly and accurately identifies what is safe and what is a threat, not just what is in a blacklist or whitelist. By coupling sophisticated machine learning and artificial intelligence with a unique understanding of a hacker’s mentality, Cylance provides the technology and services to be truly predictive and preventive against advanced threats. For more information, visit cylance.com