The Problem

The cybersecurity industry is now over 30 years old. And just like people, with each passing decade, we realize that what worked for us in our 20s, simply won't work for us now or going forward. In fact, carrying forward the mindset and behaviors of those first 20 years exposes us to countless problems in health and long-term solvency. We learn that to survive in the world, we must adapt and evolve to a higher form of existence. The antiquated and archaic practices of our past limit our visibility into the future in detecting, and thereby avoiding, maliciousness. Consequently, they have given rise to a freight train-sized hole of opportunity for the cybercriminals and nation states that wish to exploit our blindspots in the cyberworld.

Blacklisting (and Signatures) Can Be Compromised

Blacklisting technologies rely almost 100% on signature-based techniques for detecting bad files and have been at the heart of our industry since the beginning when we had only rare outbreaks like Michelangelo, Stoned, and the Morris Worm. The grossly unfortunate fact is that blacklisting remains the predominant form of detection, and thereby prevention, in the market today. Signature-based approaches to security served us well when the number of bad objects (files, network traffic, and vulnerabilities) was small and the techniques to alter those files to bypass detection were non-existent.

Today, however, countless techniques exist to avoid these once stalwart protection technologies, including packers, mutation engines, obfuscators, encryption, and virtualization bypass techniques. Within milliseconds, a once easily-detected malicious file can be altered to be completely invisible to even today's best traditional detection technologies while remaining functionally identical to its original maliciousness. This allows the bad guys to easily bypass security infrastructure that once detected them with ease.

The sheer number of files submitted to security vendors today for analysis (over 100,000 daily) is so overwhelming that most vendors simply cannot handle the volume. Their methods and manpower become easily overwhelmed. The scale of the problem outpaces the industry's capacity for maintenance. As a result, most blacklisting products have rampant miss rates.

Whitelisting Can Be Compromised

Whitelisting technologies were developed in response to what the blacklisting world falls victim to - low detection rates. In other words, blacklisting alone detects only 5-10% of malicious files out there. The reason whitelisting was so promising for so long was that it effectively did the opposite of blacklisting. Rather than stopping every known bad file, which is difficult and can involve a large list, whitelisting only allows files to run which are known to be good. Only allowing good files to run is presumably an easier task to manage because it involves maintaining a much shorter list. This technique has been applied to security through identification of permissible URLs and files that are known, or perceived, to be clean and safe. But these solutions have some fundamental problems as well.

The first challenge with solutions that rely heavily on whitelisting is that one must simply trust what the vendor or your operations staff has designated as good. We have seen this model fail time and again with security and software vendors who have their development environments compromised and their private signing certificates stolen (e.g. Adobe, Bit9, and Opera Software). When these attacks occur, it allows the thief to sign their own malicious files as if they came from the trusted vendor. And because whitelisting solutions rely so heavily on this trust model, it allows the bad guys to easily bypass the technology.

Trust Can Be Compromised

As a consequence to the identified gaps of blacklisting and whitelisting, numerous technologies have crept up to fill in the gaps of signature technology including host intrusion protection systems (HIPS), heuristics, behavioral, and both hardware and software sandboxing. But all of these techniques have two core weaknesses - foundational signature elements and reliance on trust.

Technologies such as HIPS, heuristics and behavioral engines remain at their core, signature-based. They rely on knowing what is bad and creating a signature for each threat. Even sandboxing technologies which claim to not use signatures auto-detonate captured files and binaries and still rely on signatures to enable alerting and blocking the next time it sees the threat.

For these technologies to know if something is good or bad, they must map them to a list of known good or bad

behaviors which can take minutes, hours, or days using manual verification. Even then, the attack has already happened and the detonation may not discern the maliciousness of the malware.

Security Must Evolve

Bad guys have the advantage with more resources and time to outwit the various detection schemes of security vendors. Additionally, many security models, like signatures, require the engagement of a human. Human involvement is fallible and limited in scale to the speed and sophistication of advanced threats.

We as an industry must evolve from this outlived model to a new and ever-evolving technique; one that abandons signatures and blind trust; one that relies on a mathematical, algorithmic, and scientific approach to better effectiveness and measurable accuracy.

In short, we must evolve to trust math and science.

Introducing CylanceINFINITY ENGINE

CylanceINFINITY ENGINE is a fundamental and epic shift from traditional security methods of detecting good and bad. It is a highly intelligent, machine learning, data analysis platform.

As battle-tested security industry veterans, we know previous approaches can never cope with the volume and variety of advanced threats. In response, Cylance® designed CylanceINFINITY ENGINE to make intelligent decisions without relying on signatures. It does this by taking a predictive and actuarial approach to data on a network to determine good from bad.

This model exists in many other industries. Insurance companies use actuarial science to determine the likelihood of a risk event for the insured person at a surprisingly high rate of accuracy. This concept relies on advanced models of likely outcomes based on a variety of factors. For a standard insurance policy, they may consider twenty to thirty facts to determine the most likely outcome and charge appropriately. CylanceINFINITY ENGINE uses tens of thousands of measured facts harnessed across millions of objects to make its decisions, in near real time.

CylanceINFINITY ENGINE, at its heart, is a massively scalable data processing system capable of generating highly efficient mathematical models for any number of problems.

Cylance applies these models to big data to solve very difficult security problems with highly accurate results at exceptionally rapid rates. This is done by applying data science and machine learning on a massive scale. Coupled with world-class subject matter experts, Cylance cybersecurity is able to leap ahead of threats.

While CylanceINFINITY ENGINE is problem agnostic, correctly designing solutions to difficult problems takes time, knowledge, and effort. The Cylance team has focused all of their efforts on detecting advanced threats, in near real time, correctly, without signatures.

This problem is one that has long plagued the entire Internet. The existing solutions involving humans, trust models, or signatures have proven vastly incapable of solving this problem, resulting in massive infections, data loss, and a hostile environment for business, consumers, and the Internet at large.

What Is Machine Learning?

Machine learning is a formal branch of artificial intelligence and computational learning theory that focuses on building computer systems that can learn from data and make decisions about subsequent data. In 1950, Alan Turing first proposed the question, "Can computers think?" However, rather than teaching a computer to think in a general sense, the science of machine learning is about creating a system to computationally do what humans do, as thinking entities, in specific contexts.

Machine learning and big data analytics go hand-in-hand. Machine learning focuses on prediction based on properties learned from earlier data. This is how CylanceINFINITY ENGINE identifies malicious versus legitimate files. Data mining focuses on the discovery of previously unknown properties of data, so those properties can be used in future machine-learning decisions. This means CylanceINFINITY ENGINE learns on a continual basis, even as attacker methodologies change over time!

How It Works

CylanceINFINITY ENGINE collects data, trains, and learns from the data, then calculates likely outcomes based on what it sees. It's constantly getting smarter from environmental feedback and a constant stream of new data from all around the world.

To achieve its magic, CylanceINFINITY ENGINE first collects vast amounts of data from every conceivable source. Next, it extracts features that we have defined to be uniquely atomic characteristics of the file depending on its type (.exe, .dll, .com, .pdf, .java, .doc, .xls, .ppt, etc.). Then, it constantly adjusts to the real-time threatscape and trains the machine learning system to make better decisions. Finally, for each query to CylanceINFINITY ENGINE, we classify the data as good or bad.

CylanceINFINITY ENGINE – The Rubber Meets the Road

CylanceINFINITY ENGINE can be used to supercharge decision making at endpoints and woven tightly into existing security systems via a variety of integration options. It is cloud-enabled, but not cloud dependent, to support advanced detection on a global scale in limited form factor environments, or can operate autonomously while still achieving a stunning rate of protection. The breadth of deployment options helps to solve several fundamental problem points on a modern network.

CylanceV™

CylanceV is a REST SSL application programing interface designed to integrate with CylanceINFINITY ENGINE's intelligent cybersecurity decision making. Through the API and specially developed utilities, IT departments executing incident response and forensics can take the tedium out of tracking down malware and determining what is truly bad.

CylanceV enables a starting point for forensic analysis and timely remediation through an automated and highly efficient approach.

Tying other security tools like SIEM, log analysis, host and network monitoring, HIPS/NIDS, and investigation tools including antivirus, anti-malware, and forensics, into CylanceV provides contextual intelligence for more accurate and effective malware identification.

The CylanceV API allows utilities to be developed in most popular frameworks (.NET, Python, etc.) and invoked through HTTPS using tools such as CURL or WGET in order to make the data segmentation easier and more efficient.

Integrating third-party functionality, like Python scripts, Splunk, or C# to CylanceINFINITY ENGINE quickly determines what is safe and what is a threat, making smart security smarter. Together, they reduce the total number of prospective compromised machines to something manageable.

CylancePROTECT

CylancePROTECT is our host-based security solution built on CylanceINFINITY ENGINE technology. It leverages algorithmic science to greatly increase the speed and accuracy of host protection without reliance on signatures, heuristics or behavior modeling. It offers a real-time protection layer on the endpoint that can make decisions about the nature of malware independent of connecting to CylanceINFINITY ENGINE, and has a stunningly low impact on endpoint performance. CylancePROTECT offers a powerful front line of defense, whether your assets are behind your corporate firewall or in a coffee shop. Its extensive management capabilities easily blend CylanceINFINITY ENGINE's pervasive protection into your existing security workflow.

Summary

With CylanceINFINITY ENGINE, enterprises can definitively determine which files are good or bad in milliseconds with extraordinarily high detection accuracy and extremely low false positive rates. Because the system is self-collecting, self-training, and self-learning, it always stays ahead of the changes and unknowns. With such a strong mathematical approach, Cylance is changing the game of security… infinitely.

CYLANCE™