SOPHOS

Security made simple.

Intercept X Deep Learning

Intercept X combines deep learning with best-in-class anti-exploit technology, CyptoGuard anti-ransomware, root cause analysis, and more to form the industry's most comprehensive endpoint protection. This unique combination of features allows Intercept X to stop the widest range of endpoint threats.

Highlights

- The number one performing malware detection engine
- Prevents both known and never-seen-before malware
- Blocks malware before it executes
- Does not rely on signatures
- Protects even when the host is offline
- Detects malware in approximately 20 milliseconds
- Trained on hundreds of millions of samples
- Proven on VirusTotal since August of 2016
- Classifies files as malicious, potentially unwanted apps (PUA), or benign
- Works out of the box with no additional training needed
- Extremely small footprint (under 20MB)
- Focused on Windows portable executables

Much of today's security is reactive and far too slow. As the volume and complexity of endpoint attacks has continued to grow, legacy approaches have struggled to keep pace. For example, SophosLabs analyzes over 400,000 new malware samples every day. To make meeting this challenge even more difficult, SophosLabs found that 75% percent of malware is unique to a single organization.

Deep learning, an advanced form of machine learning, is helping to change the way we approach endpoint security, and Intercept X is leading the charge. By integrating deep learning, Intercept X is changing endpoint security from a reactive to a predictive approach to protect against unknown threats.

Deep Learning vs. Other Types of Machine Learning

> *"Intercept X uses a deep learning neural network that works like the human brain… This results in a high accuracy rate for both existing and zero-day malware, and a lower false positive rate."*
>
> ESG Lab Report, December 2017

While many products claim to use machine learning, not all machine learning is created equally. At Sophos, we use deep learning to detect malware. Also referred to as 'deep learning neural networks' or 'neural networks', deep learning was inspired by the way the human brain works. It is the same type of machine learning often used for facial recognition, natural language processing, self-driving cars, and other advanced fields of computer science and research.

Deep learning has consistently outperformed other machine learning models, including random forest, k-means clustering, or Bayesian networks, but requires vast amounts of data and computational power to build an effective model. At Sophos, this has been made simple thanks to the malware collection and analysis efforts of SophosLabs over the past 30 years and the telemetry we receive from our 100+ million endpoints every single day.

Deep learning has several inherent benefits compared to other types of machine learning commonly used in endpoint security:

Smarter: Deep learning models process data through multiple analysis layers, just like neurons in the human brain, each layer making the model considerably more powerful. It analyzes complex relationships between different input features. This allows it to automatically uncover the best combination and manipulation of inputs that would otherwise be impossible for humans to determine. This means that the Sophos deep learning malware detection model will be able to detect malware that would go unnoticed by other machine learning engines.

More Scalable: Deep Learning elegantly scales to hundreds of millions of training samples. This is important considering that SophosLabs analyzes 2.8 million new malware samples every week. Because it can continue to ingest massive amounts of training data our model can 'memorize' the entire observable threat landscape as part of its training process. Since it can process significantly more input, deep learning can more accurately predict threats today while continuing to stay up-to-date over time.

Lighter: Traditional machine learning approaches result in huge model sizes, which can sometimes take many gigabytes on disk. However, Sophos' deep learning approach results in highly compressed models. The Sophos deep learning model is incredibly small, less than 20MB on the endpoint, with almost zero impact on performance.

Sophos Deep Learning Capabilities

Sophos provides deep learning expertise with industry's highest-performing malware detection engine:

Experienced: Unlike the competition, we have been cybersecurity machine learning experts for a long time, and have had our malware detection deep learning models in production environments for years. The Sophos malware detection model was created by our team of data scientists with DARPA driven technology. In 2010, the US Defense Advanced Research Projects Agency (DARPA), created their Cyber Genome Program to uncover the 'DNA' of malware and other cyber threats. This was the origin of what is now the algorithm embedded in Intercept X.

Proven: We have been open and transparent with our models. In addition to presenting details of our methodology at industry conferences such as Black Hat, we also have not shied away from allowing our model to be tested by independent third parties. The model has been proven on VirusTotal since August of 2016, and has received high scores from third-party testers such as NSS Labs. In all cases, it has proven to be extremely effective while having low false positives.

> *"One of the best performance scores we have ever seen in our tests"*
>
> Maik Morgenstern, CTO, AV-TEST

Performance: Sophos' Deep Learning technology is incredibly fast. In less than 20 milliseconds the model is able to extract millions of features from a file, conduct a deep analysis, and determine if a file is benign or malicious. This entire process happens before the file executes.

SophosLabs: One of the most important aspects to any model is the data that used for training. Our team of data scientists are part of the SophosLabs group, granting them access to hundreds of millions of samples. This allows them to create the best possible predictions in our models. The integration between the two groups also leads to better data labeling (and therefore better modeling). The bi-directional sharing of threat intelligence and real-world feedback between the team of data scientists and threat researchers continuously improves the accuracy of our models.

> *"Intercept X stopped every complex, advanced attack we threw at it"*
>
> ESG Lab Report, December 2017