lastline®

WHITE PAPER

Using AI to Detect and Contain Cyberthreats

Combining Threat Detection and Network Anomaly Detection

Introduction

The term "artificial intelligence" (AI) has become ubiquitous in the description of modern applications, services, and user interfaces. The security domain is not an exception. Companies today are touting almost every new service or product as AI-based. The use of the term "AI" is so widespread, in fact, that it has become difficult to see beyond the hype and understand what "AI-based" really means.

This white paper provides a high-level description of what AI means, and explains some of the key terms surrounding AI and machine learning (ML) and their security applications, including anomaly detection. It also identifies some pitfalls in the use of AI in network security. Finally, it describes how we use AI at Lastline.

Artificial Intelligence and Machine Learning

As defined by Wikipedia, artificial intelligence is "intelligence demonstrated by machines, in contrast to the **natural intelligence** displayed by humans and other animals."

In general, there are two approaches to AI: emulation and simulation.

- Emulation tries to reproduce, in a synthetic way, the mechanisms that make humans intelligent (e.g., the human brain). This type of AI has produced very few practical applications.
- By contrast, simulation approaches try to reproduce the effects of intelligent behavior, regardless of how such behavior is achieved. This is the AI that we experience today.

Machines appear to be "intelligent" to the extent that they understand what we say or are able to solve complex problems requiring abstraction. However, the way in which they achieve this "intelligence" is different from the human thought process.

AI encompasses a number of different classes of techniques. However, it is ML that, in the past few years, has shown much promise. According to Wikipedia, ML is "a field of artificial intelligence that uses statistical techniques to give computer systems the ability to 'learn' (e.g., progressively improve performance on a specific task) from data without being explicitly programmed." A particular sub-field of ML is *deep learning*, which relies on neural networks to extract data representations and abstractions from unstructured data. As a result, deep learning techniques are more "autonomous," requiring less pre-processing of input data and tuning to achieve good performance.

ML techniques can operate in one of two modes: *supervised* or *unsupervised*. In supervised mode, analysts provide initial information about the data that is given as input to the learning process. For example, an ML algorithm is given labels that characterize the objects that are visible in a set of images. The ML algorithm then creates a classifier that is able to automatically recognize other instances of the data (that is, it is able

to recognize specific objects in new, previously unseen images.) In unsupervised mode, the data is simply characterized by a number of features (e.g., the colors of an object, or its shape), and then the algorithm is able to group, in clusters, objects that have similar features.

Both approaches have important applications to the field of security.

ML and Security

AI and specifically ML techniques have found widespread application in security. For example, supervised ML techniques are used to create malware detection systems. For this, an ML-based system could ingest large amounts of known-benign programs and known-malicious programs and generate a classifier that, given a program which has never been seen before, is able to determine whether the program is malicious.

Similarly, unsupervised ML is used to create clusters of similar data. These clusters can then be used to address similar events (e.g., the email messages used in a large-scale phishing campaign) or to identify anomalies. In particular, the use of unsupervised ML to identify anomalous behavior has been used (and touted) so widely that sometimes people associate anomaly detection with AI. However, anomaly detection is not, per se, part of AI.

Let's try to clarify this important point: Approaches to the identification of malicious behavior can be roughly categorized in two classes: Misuse Detection and Anomaly Detection. In the former case, a model of what is known to be malicious is created, and then instances of maliciousness are identified in the data. An example of a misuse detection approach is the use of signatures to detect network attacks. The advantage of these approaches is that they are generally very *precise* (that is, they don't make many mistakes), however, they can only detect malicious behavior that has been modeled, or seen before.

Anomaly detection uses a complementary approach. The idea is to create a model of what is normal and identify outliers that are outside the parameters of normality. The advantage of this approach is that it is possible to identify malicious behavior that has never been seen before. Even though this approach seems very promising, the approach is based on two important assumptions: *"what is anomalous is malicious and what is malicious will generate an anomaly."* Unfortunately, both assumptions do not hold at all times, causing both false positives and false negatives.

Given these two detection approaches (misuse detection vs. anomaly detection), it is now evident how supervised ML can be easily employed in misuse detection and unsupervised ML can be leveraged to perform anomaly detection, allowing for the automation of these techniques and supporting the scalable analysis of large datasets. For example, supervised ML could analyze a large dataset of known benign URLs and known phishing URLs to generate a classifier that is able to annotate URLs in emails with a probability value that the link will lead to a phishing site, so that a user is warned in advance about the possible attack. Unsupervised ML could be use to analyze the network activity history of hosts, to identify hosts that suddenly change their behavior performing actions never seen before, such as the upload of large amounts of information to an external host.

The Risks of ML and Security

While ML-based approaches seem a promising solution to many security problems, not everything is as simple as it seems. In fact, most ML techniques have been developed in fields such as computer vision, natural language processing, and genome analytics, and, in these fields, the analyzed information (images, text and DNA sequences) does not actively fight against the learning process.

However, this is exactly what happens in security. Threat actors continuously try to modify their malware samples so that they are (mis)classified as benign, while intruders try to cover their tracks by using traffic patterns similar to those regularly observed in the target network to avoid being identified as anomalous. Therefore, applying ML to security is not a traditional AI use case. To be effective, ML techniques need to be extended so that it is possible to perform *adversarial machine learning*. Failing to do so will make it easy for a motivated adversary to bypass ML-based products.

Adversarial ML requires the composition of multiple approaches (to avoid blind spots) and, in addition, the use of representative data in the ML process. If the right data in the right amount is not available, the ML process will fail miserably in an adversarial setting, in which an adversary is able to control the data being analyzed. In such cases, the attacker can show the ML system malicious input (such as programs or URLs) and pretend it is harmless, tricking the ML system in identifying similarly malicious inputs as harmless in the future (thereby bypassing detection).

Another problem of applying ML to security is that, in most cases, the malicious samples are much scarcer than the benign ones (there are less malware samples than benign samples, and less malicious network connections than benign ones). Since many ML techniques are based on statistical analysis, if the datasets are not balanced and representative of the domain being analyzed, it is possible to learn the wrong thing or fail to learn some important characteristics of the data, resulting in both false positives and false negatives.

False positives are particularly detrimental because of the base-rate fallacy. When a system has to identify few events (e.g., anomalies) in large amounts of data (e.g., terabytes of traffic observed on a large network), having a seemingly small false positive rate might result in an unmanageable number of false alerts. Consider, as an example, a network in which, every day, 100 million TCP connections are established. A false positive rate of 0.001%, which seems minuscule, would generate 1,000 alerts per day, which would be likely unmanageable for most organizations.

Size and Quality of Data

ML approaches are designed to extract high-level patterns, which support the creation of classifiers and clustering of data. However, the learning process is data-centric. Therefore, the resulting ML models are trustworthy only if the training data is "good."

Good data means two things. First, the data needs to be abundant. Without millions of examples, it is difficult, if not impossible, to create semantically rich abstractions of the data and identify characterizing patterns. Second, the data must be representative of the domain being analyzed. This means that all the aspects of a certain domain (e.g., network behaviors) need to be represented in a way that mirrors realistic deployments. In addition, the metadata associated to the various items need to be comprehensive. For example, learning from a program's external appearance (e.g., its at-rest binary representation – a PE file image in Windows.) without executing it can result in erroneous models. It is the behavior of a program that matters, after all, and not its external appearance (which can be easily changed using packing and encryption).

Employ Multiple Technologies for Higher Fidelity

Lastline uses AI techniques in many ways to protect its customers from data breaches and intrusions. In particular, Lastline trains its AI on both network behaviors and threat behaviors. This is different from network traffic analysis (NTA) solutions, which solely look at network (and user) behaviors. It is also different from advanced threat protection solutions, which focus on analyzing file and threat behaviors (often in a sandbox,

some also including AI-based threat classification). Lastline's unique approach of taking into account both network and threat behaviors for its AI-based analysis offers two critical benefits to its customers. First, the system can train on more and better data. This leads to better models and, in turn, to better detection results. Second, the composition of network and threat analysis creates more signals, which means more context and fewer false positives.

It Starts with the Data

ML algorithms get significantly better when they have more input data to work with. This is well expressed in a famous quote by Peter Norvig, Research Director at Google, who explained Google's superior results by saying that "we don't have better algorithms, we just have more data." By feeding our AI both network data as well as threat data, our solution can see with two eyes.

In addition to more data, Lastline has also better (more fine-grained) data that it can leverage as input to its algorithms. This is true both for the network data as well as for the threat data.

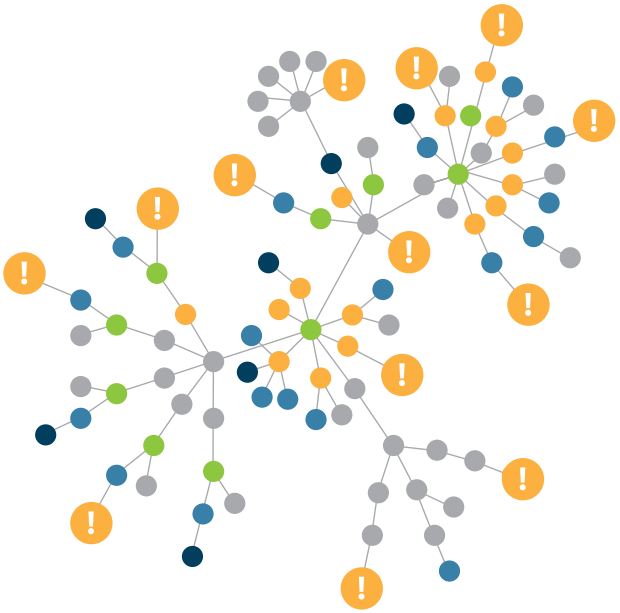
Lastline monitors both network traffic entering and exiting the network ("North-South" traffic) and traffic within the network ("East-West" traffic), as well as host activity. In addition, Lastline extracts events at several different abstraction levels, from raw packets to network flows, to detailed protocol decoding, such as HTTP and DNS requests, to host-based events. This is different from many network-based solutions that only extract some (high-level) metadata from network connections, such as IP addresses, ports and packet counts. By performing full packet inspection and application protocol decoding, Lastline obtains a very diverse and comprehensive network input dataset.

Lastline has created the most comprehensive threat analysis system, whose superiority has been demonstrated numerous times through independent testing and evaluations. By performing analysis based on Deep Content Inspection™ on millions of real-world programs, documents, web pages, emails and files, Lastline's sandbox is able to feed ML models with semantically rich data that is directly representative of the behavior of programs that were running.

Composing Network Anomaly Detection with Threat Classifiers

While more and better data is important, the full power of the Lastline AI solution gets unlocked through the composition of network and threat data analysis. For this analysis, Lastline uses a unique combination of threat detection techniques based on supervised ML and network anomaly detection techniques based on unsupervised ML.

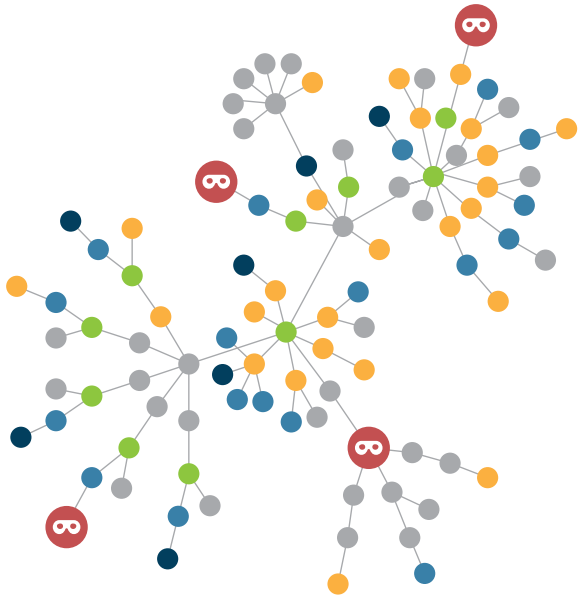
To understand how these two complementary aspects of AI are combined to achieve unsurpassed detection and response, consider the figures below.



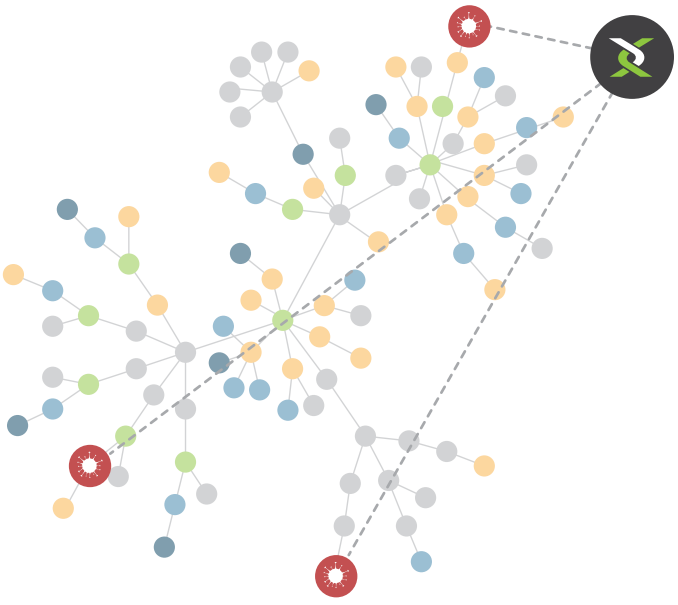
Network Anomalies

On the left, there is a network, with various anomalies associated with the hosts in the network. This is the view that other network traffic analytics (NTA) solutions provide: They tell you what events are not normal and unexpected, so you can answer questions like "what hosts uploaded a large file to an external host?" Or "What host just deployed a service on a non-standard port?" In a nutshell, it provides the answer to the question: "What is weird in my network?"

On the right, there is a picture of the same network, in which a number of hosts have been identified as being compromised using ML-based threat detection techniques. Being based on supervised ML and threat intelligence, these detections have little to no false positives. These hosts have been targeted with drive-by-download attacks, malware, or have been communicating with command-and-control servers controlled by cybercriminals. In a nutshell, this view provides the answer to the question: "What is bad in my network?"



Threat Detection



Combined Network Anomalies & Threat Detection

In the last picture, you can see Lastline's unique approach that is the result of the composition of the two views. By combining threat detections with anomalies, it is possible to understand what anomalies are actually associated with malicious behavior, reducing false positives. This allows the system to create "anomaly profiles" around compromised hosts, which can then be used to identify similarly anomalous hosts in the network. In a nutshell, Lastline answer the question: "What is the impact of this intrusion?"

The value of Lastline's approach is clear if one looks at each approach in isolation. On the one hand, by looking only at anomalies, one would incur the risk of being flooded by anomalous-yet-benign events, which are commonplace in most networks. At the same time, one would be blind to malicious events that do not generate any anomaly (many short-lived, low-traffic events fall in this category).

On the other hand, by looking only at threat detection, one would might miss important information around a compromised host that could provide vital information to determine the impact of a breach, and help in the hunt for other hosts in the network that have similar anomalous behavior.

It's only by combining these two views that it is possible to provide the best AI-based detection, by composing the advantages of both supervised and unsupervised ML techniques while mitigating their limitations.

Global Perspective

The combination of threat detection and network anomaly detection is achieved through a sophisticated correlation process, which collects related events and highlights relationships among them. Correlation identifies both relationships among events observed within a network (e.g., the steps carried out by an intruder in a multi-stage attack) and relationships with data in Lastline's Global Threat Intelligence Network. This is a large-scale threat intelligence knowledge base that is continuously updated with the behaviors of the millions of samples that Lastline observes on its customers' networks.

Highest Fidelity Insights into Malicious Behavior

The resulting process provides the most accurate AI-based solution on the market. Lastline combines multiple ML technologies and trains them on a massive data set of malicious behaviors. This has created a high-fidelity threat identification system that exhibits reduced false positives while providing rich contextualization to support incident triage and the ability to accelerate threat response.

Experience the Lastline Advantage

For more information please visit www.lastline.com

Lastline Corporate Headquarters
203 Redwood Shores Parkway, Suite 500
Redwood City, CA 94065

+1 877 671 3239
info@lastline.com
www.lastline.com