PATHFINDER REPORT

451 Research® | Advisory

Machine Learning Raises Security Teams to the Next Level

MAY 2017

COMMISSIONED BY

McAfee™
Together is power.

About this paper

A Pathfinder paper navigates decision-makers through the issues surrounding a specific technology or business case, explores the business value of adoption, and recommends the range of considerations and concrete next steps in the decision-making process.

About 451 Research

451 Research is a preeminent information technology research and advisory company. With a core focus on technology innovation and market disruption, we provide essential insight for leaders of the digital economy. More than 100 analysts and consultants deliver that insight via syndicated research, advisory services and live events to over 1,000 client organizations in North America, Europe and around the world. Founded in 2000 and headquartered in New York, 451 Research is a division of The 451 Group.

NEW YORK
1411 Broadway
New York NY 10018
+1 212 505 3030

SAN FRANCISCO
140 Geary Street
San Francisco, CA 94108
+1 415 989 1555

LONDON
Paxton House
(Ground floor)
30, Artillery Lane
London, E1 7LS, UK
P +44 (0) 207 426 1050

BOSTON
75-101 Federal Street
5th Floor
Boston, MA 02110
Phone: +1 617.598.7200
Fax: +1 617.357.7495

While machine learning can detect patterns hidden in the data at rapid speeds, the less obvious value of machine learning is providing enough automation to allow humans the time and focus to initiate creative responses when responses are less obvious.

INTRODUCTION

Machine learning is all around us, enriching our online lives every day. We see it with our own eyes when search engines accurately predict what we're looking for after we type only a few letters. We feel it protecting our bank accounts evaluating credit card transactions for signs of fraud. We notice it in selections of articles and ads in online newspapers. We no longer think twice about these conveniences; in fact, it's hard to imagine online life without machine learning.

In relation to cybersecurity, machine learning has been changing the game as a means of managing the massive amounts of data within corporate environments. However, machine learning lacks the innately human ability to creatively solve problems and intellectually analyze events. It has been said time and again that people are a company's greatest asset.

Machine learning makes security teams better, and vice versa. Human-machine teams deliver the best of both worlds:

- **Machine learning means security teams are better informed so can, therefore, make better decisions.** Security executives realize that the intelligence and creativity of their security operations experts are critical business resources. Machine learning is a technology that allows chief security officers (CSOs) to get the most out of human and security product assets.
- **Adversaries are human, continuously introducing new techniques**. Creative new tactics and strategies dealt by adversaries force security teams to employ machine learning to automate the discovery of new attack methods – creative problem solving and the unique intellect of the security team strengthens the response.
- **Machine learning becomes more accurate as more data is available to feed its algorithms**. Enhancements in handling big data using high-performance and massive-capacity storage architectures have enabled the growth of artificial intelligence.
- **IT teams need help analyzing faults**. In those rare instances when endpoint security cannot prevent damage from an attack, machine learning accumulates relevant data elements into one place, placing it at the fingertips of security analysts when needed.
- **Human-machine teaming makes for sustainable endpoint security.** As new threats are introduced, security teams alone cannot sustain the volume, and machines alone cannot issue creative responses. Human-machine teams make endpoint security more effective without draining performance or inhibiting the user experience.

This Pathfinder Report summarizes the key technical and use-case attributes of machine learning before making recommendations on how to evaluate the capability. This report is sponsored by McAfee.

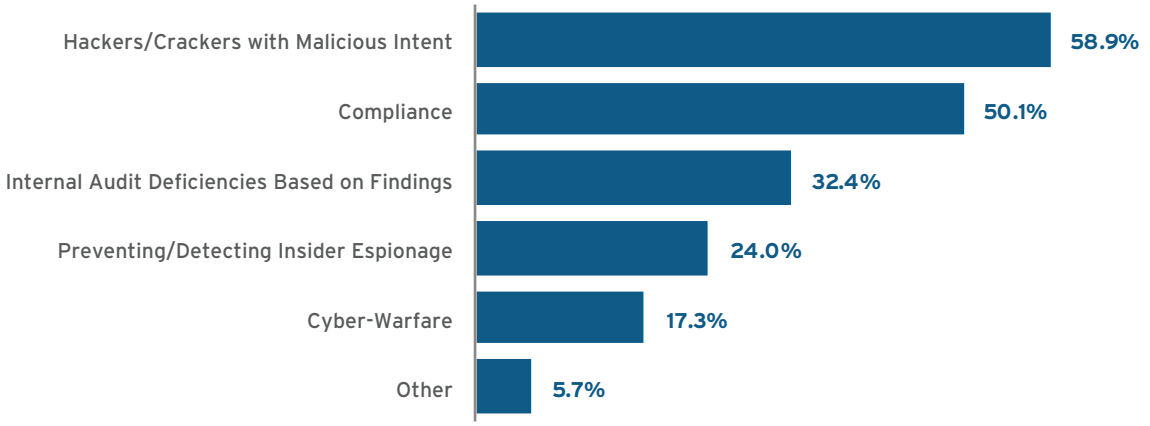
Machine learning allows endpoint security to continually evolve to stop new attack tactics

We see attackers focusing on vulnerable endpoints as the preferred point of entry for malware. Popular tactics include phishing, downloaded code that executes in the browser, infected email attachments, user-installed rogue executable programs, and open stolen account credentials. One of the challenges for IT operations is that endpoints are not constrained in the datacenter where they can be surrounded by layers of security defenses under the vigilance of security teams. Rather, they are constantly on the move, in and out of the network.

Market data confirms that the most important job of enterprise security teams is to prevent cyberattacks from penetrating the infrastructure. It makes sense, because all of the other security products are there to support that one central mission of protecting the business. According to respondents in the recent 451 Research Voice of the Enterprise: Information Security survey, hackers remain at the top of information security concerns, outpacing compliance, internal audit deficiencies, espionage and cyber warfare.

Figure 1: Top information security concerns

What were your top general information security concerns during the last 90 days?



Source: 451 Research, Voice of the Enterprise: Information Security, Budgets and Outlook, April 2017

Endpoint security is in a constant state of stepwise refinement, embracing new prevention techniques to thwart new attacker tactics. Machine learning is a natural extension to other malware-prevention methods.

Endpoint security has been protecting our devices for decades in a constant back-and-forth conflict with hackers and attackers. Malware developers create a new threat; endpoint security deploys an antidote; malware developers refine their attacks – and around we go.

Typical protection methods include:

- **Search for exploit code using signatures and patterns**. This is the classic blacklist approach – explicitly identifying attacks that can then be blocked – that started the antivirus market. Today's endpoint security products tend to also download active code snippets to identify classes of attacks and remove them from the device.
- **Whitelists specify programs that can execute on the endpoint.** Anything not on the approved list is considered a threat and is not allowed to run. This is effective in environments where the endpoint configuration is relatively stable and leads to a quick decision.
- **IP reputation filters are consulted before allowing connections to or from sites in the cloud.** This blacklist approach forbids communications to sites with questionable histories or geographies.
- **Host intrusion prevention systems adds a time dimension and brief layer of machine learning to endpoint security**. For instance, while it is reasonable for a user to open an email attachment, click on a link and save an encrypted file, it is not reasonable for this to occur within 300msec.
- **Personal firewalls enforce security policies for network zones**. This method controls access to servers and protocols.

The dark web is driven by intelligent bad actors who are often financially motivated to create new threats with new attack techniques. Security becomes personal when considering the people behind the attacks, making the human-machine team the best sustaining defense. CSOs empower security operations to blend the best elements of art and science, where security team employees provide creative responses and leverage machine learning to provide high-performance scientific responses. While machine learning can detect patterns hidden in the data at rapid speeds, the less obvious value of machine learning is providing enough automation to allow humans the time and focus to initiate creative responses when responses are less obvious. By using a filter for optimization across the best advantages of human and machine elements, it's easier to evaluate the relationship between them.

Machine learning adds critical capability to security strategies

Computers are dumb. Computers cannot think for themselves; they can only blindly follow the instructions that developers give them. However, machines are excellent at repetitive tasks, such as performing calculations across massive amounts of data. That is the secret of machine learning – its ability to draw statistical inferences and construct models from crunching big data. The results of machine-learning calculations can drive endpoint security protection from before malware can execute through cleanup and remediation phases.

The key differentiator in incorporating machine learning into endpoint security is the amount of relevant data consumed by the algorithms. The data requirements are far beyond the ability of any endpoint to process. Most likely, processing must be performed in cloud-based datacenters with results condensed for endpoint usage. Anything less either doesn't produce enough data for machine learning to be accurate or adversely affects user performance. Diverse sources from large endpoint security vendors will process millions of malware samples per day, and we believe they are best positioned to bring the benefits of machine learning to endpoint security.

Machine learning starts in the datacenters of endpoint security vendors. The millions of malware samples are analyzed to find common traits, new tactics, and updated understanding of how malware classes access the network, memory and file systems. It is all about using machine-learning algorithms to search and compare every detail about each malware sample to uncover information allowing malware to be more readily and accurately identified.

The process of security researchers analyzing malware to develop signatures is still important, but only as a capability to address the large volume of *known* malware because it cannot be expected to evolve quickly enough to meet the rapid pace of malware being introduced to the wild. Machine learning becomes the fastest way to identify new attacks and to push that information out to endpoint security platforms.

We find machine learning manifests itself in multiple ways in helping save security teams' time and energy:

- **User experience is optimized.** Machine-learning algorithms feed information to the endpoint about file attributes that indicate the presence of malware. These attributes may be related to type, size and source, as well as header anomalies and detected sequences of operating system calls. A quick scan before execution allows security to perform its preliminary triage without souring the user experience.
- **Suspicious behavior flagged automatically**. Once the program is running, machine learning on the endpoint monitors behavior for signs of an attack. This runtime detection is keyed by information on attack tactics again uncovered by machine-learning analysis of malware samples in the datacenter. While pre-execution checks file attributes to make a malware decision, runtime execution requires knowledge of specific actions attackers are likely to use. For example, ransomware can render your files useless in less than a minute. Machine-learning analysis of ransomware attacks may uncover timing and access patterns of file shares that would indicate an attack is underway – allowing endpoint security to stop the threat before all files are encrypted.
- **Highly valuable investigation and response data available automatically**. Helping security teams respond to an incident, machine learning can identify suspicious connects and create alerts based on equations. In this case, security analysts need precise information on the threat such as files touched, registry changes, server connections, etc. Because machine learning looks across multiple dimensions, much of the data that incident response teams require is already available, but has traditionally required extensive manual correlation. Ideally, highly valuable investigation and response data would be available through the already-present endpoint management console. The presence of machine-learning technology results in significant time savings – by a factor of 10 is not uncommon – that can help security teams keep the business running.

The enterprise is seldom directly exposed to the working benefits of machine-learning computations; realizing the benefits is good enough. If machine learning is driven by massive amounts of data in cloud-based datacenters, ease of deployment for security operations cannot be beat. There is no extra effort involved in software deployment to protect endpoints when they reside on the network and when they are individually exposed on the web. Results should be transparently communicated to endpoint enforcement points.

Elevate security teams with machine learning

Corporate executives are quick to say that their people make the company special, and rightfully so. Successful security executives support this mantra with strategies to energize their security operations teams to be more effective and efficient through the use of technology. People matter the most, but combining human intelligence with machine-learning technology creates strong security teams.

Thousands of hours per year can be spent devolving into one day after another chasing alerts of skullduggery. If a security analyst requires 15 minutes to investigate and clear a security alert, then that person can only process about 30 alerts per day. This formula dooms security teams into unsustainable reactionary patterns, and it fails to allow security personnel to develop problem-solving skills. Attackers use automated practices to discover what works and then relaunch tactics for maximum effect. The best way for security teams to get ahead in this game is to allocate time for people to use their intelligence and creativity to enhance security practices, and to leverage efficiencies gained from machine-learning technology to make that time.

The visibility into tactics throughout the entire attack chain that machine learning affords is critical to enhancing the relationship between security teams and technology. Machine learning enables security teams to devise new defenses quickly to adapt to attackers' automated processes and make it more difficult for them to be effective. Remember, machine learning places the time sequence of activity observed between security products. With machine-learning assistance, security teams have greater insight into who the attacker is, the methods being used, where the attacks are coming from and how they are spreading, as well as which security measures are working and which are being defeated.

Most importantly, the presentation of machine-learning results enables people in security teams to do what they do best – create intelligent, innovative and effective solutions to new threats before significant damage is done to the business. If people are the company's greatest assets, then machine learning helps make them even greater.

Conclusion

Machine learning is a critical component of an enterprise endpoint security strategy. Given the volume and evolution of attacks hammering away at endpoints, security must be able to adapt without human intervention, and must provide the visibility and focus to enable humans to make more informed decisions. Machine learning has come of age with big data driving accuracy up and false positives down. The proof of successful human and technology teaming will be seen in the ability to rapidly dismiss alerts and accelerate solutions to thwart new threats. Your users deserve the best that cybersecurity has to offer, and today the best endpoint security products leverage machine learning.