Secureworks®

WHITE PAPER

Leveraging AI to Modernize Vulnerability Management and Remediation

Introduction

In what seems like no time, artificial intelligence is directly influencing, or fully responsible for a number of products and services we are rapidly taking for granted, everything from self-driving vehicles, to medical imaging, to speech recognition, to the optimization of your playlist. AI technology, however, has been largely absent from one area: enterprise vulnerability management.

Defining AI

But before undertaking a discussion of AI's potential in vulnerability management, let's first take a moment to define "artificial intelligence," a phrase that today is used ubiquitously, if not always accurately.

Artificial Intelligence is an umbrella term that encompasses several areas of advanced computer science, everything from speech recognition to natural language processing, to robotics, to symbolic and deep learning. AI technologists are constantly striving to automate seemingly "intelligent" behavior, or put differently, programming computers to do historically human tasks.

One AI component used extensively in many applications is machine learning, algorithms that leverage historical data to make predictions or decisions. The more ample the historical data, the higher the probability the prediction will be useful or accurate. Thus, as more historical data is gathered, the machine learning engine's predictions improve, or in the vernacular of pop culture, the application "gets smarter." For example, a machine learning-based application that identifies the probability of lung cancer from an X-ray can make a prediction from a historical data set of 10 X-rays, but that prediction's accuracy will be negligible. As the historical data set expands from 10 to 10,000, the prediction becomes more reliable, and will improve again as the data set grows to 20,000, 30,000, and beyond.

Machine learning-based applications use historical data to make future predictions that improve as time progresses, without human intervention. As we'll discuss below, machine learning and other AI techniques can be powerful weapons in the battle to counter today's cyber bad actors, and there are many opportunities within the field of vulnerability management to leverage AI technology for better outcomes.

As mentioned previously, to date, the use of AI in vulnerability management has been largely inconsequential (there are some exceptions, for example, some pioneering work in scanning automation using expert systems), but that is unlikely to be the case for much longer. There are a number of elements of the vulnerability management process that can benefit greatly from the appropriate application of AI techniques, and we discuss several in this paper.

Technology in the vulnerability management community has, to this point, evolved little since the field's initial days, but given the abundance of rich, historic data, multi-dimensional risk elements, and a heretofore brute-force approach to remediation, the vulnerability management field appears ripe for AI exploitation.

There are many opportunities within the field of vulnerability management to leverage AI technology for better outcomes.

Developing a Meaningful Vulnerability Risk Score

The primary deficiency of the CVE risk score is its lack of context.

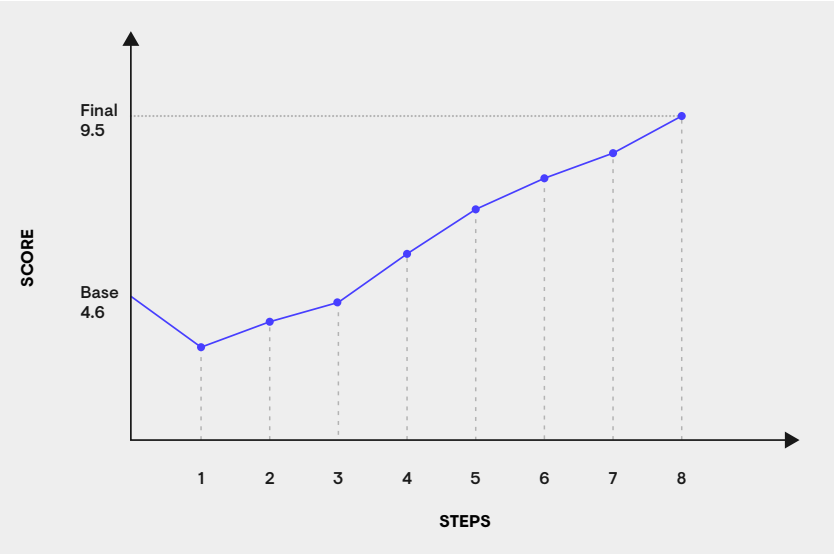
Even though context-based vulnerability risk scoring is a later-stage step in the overall vulnerability management process, it's the most important. All data gathered and used in the other vulnerability operations stages discussed in this paper contributes to the ultimate objective of modern vulnerability analysis: the context-based prioritization of each vulnerability to optimize remediation efforts and maximize risk reduction. Thus, we lead our discussion of AI in vulnerability management with intelligent vulnerability risk scoring.

Today, the phrase vulnerability "risk score" is largely synonymous with the risk score attached to each vulnerability in the CVE (Critical Vulnerabilities and Exposures) program, and although a useful starting point to assess the criticality of an individual vulnerability, it is a woefully inadequate measure of a vulnerability's risk to an enterprise in and of itself.

Olympic swimmer Michael Phelps was famous for eating 12,000 calories a day when training, including a daily breakfast of pancakes, French toast, fried egg sandwiches loaded with cheese, fried onions, and mayonnaise. Every doctor on the planet would – without context – assert that such a diet would be a recipe for heart disease and diabetes. But in context, for an Olympic swimmer in training, such a diet's risk to long-term health is negligible.

Similarly, a given vulnerability may earn a high CVE risk score, but on a given network, the affected asset may be completely isolated on a secure subnet, not connected to the Internet, on a device running no other services, and therefore presents very little risk to the business. Thus, to the individual organization, a high CVE risk vulnerability may – for that organization – be a much lower remediation priority than vulnerabilities scored lower by the CVE scale, but expose the organization more significantly given their context.

The CVE risk score is a useful starting point for a contextbased risk analysis, but not meaningful when applied without context.



Modern vulnerability management systems can leverage all the AI techniques discussed in this paper to collectively develop an in-depth understanding of the context of each asset. Once a sophisticated appreciation of the asset's context is acquired, it can be combined with in-depth knowledge of the specific vulnerability and the external threat environment to generate a "context-driven priority."

Establishing priorities and a plan to reduce risk while optimizing limited remediation resources is the ultimate goal of an intelligent vulnerability management program, and the only way to realistically accomplish that objective is with a context-sensitive assessment of risk.

Discerning Vulnerability Exploitation Trends

Brand marketing professionals are using AI-based sentiment analysis applications to evaluate countless posts on social media platforms that reference their products. Collecting this data and applying AI to it can provide insight into how a brand is perceived in the market, and how that changes – for better or worse – over time. Similarly, cyber security chat boards, media sites, and other on-line sources of cyber security conversations can be collected and analyzed to predict which vulnerabilities are the most likely to be exploited, which security experts are most concerned about, and how those sentiments change over time.

AI can interpret posts and blend the meanings of thousands to add context to any given vulnerability's practical risk, a risk that can change quickly as new exploits are created and distributed among the ever-growing community of bad actors.

Collecting and evaluating millions of such posts over time is impractical with human resources, but can be accomplished continuously with Neural Networks and Natural Language Processing (NLP) techniques, an AI technology that can discern meaning, positivity/negativity, and even more importantly, can extract precise technical information from text.

Performing Important Asset Detection

Certainly, finding all assets is the foundation of an effective vulnerability management program, especially those assets that may appear atypical in a given context. But, given the sheer number of assets in a typical network, using conventional detection mechanisms, it can be difficult to find network assets that are contextually out-of-the ordinary: for example, a server that hosts many websites or services, a workstation in a subnetwork full of servers, or a Linux server in a network of Windows machines with database services running. These kinds of assets should be considered particularly crucial, and as such deserve more attention from security teams.

To accomplish this, techniques from the field of Pattern Recognition, namely Novelty, Anomaly or Outlier Detection, can be employed to help uncover exceptional assets, or those that stray from the contextual "norm."

Manually comparing assets can be a useful technique when searching for unique assets. Yet, the number of possible characteristics of an individual asset can be overwhelming, making it difficult to compare assets with a single dimensional analysis; a multi-dimensional approach is therefore required.

One particular algorithm – Isolation Forest – can be particularly effective. In this process, several asset characteristics are compared using a multi-dimensional representation, and those that differ from contextual baselines are flagged (typically the top 10%). Using this filtering technique, the many "ordinary assets" are separated from the few "remarkable assets."

Assessing the Reliability of Detections

An element of vulnerability management that is often unappreciated by those outside the field is the challenge of vulnerability detection. AI can be employed in this part of the vulnerability management process to help reduce the number of false positives, essentially "detecting the mis-detections." Factors such as services running on the asset and the detection mechanism that flagged the vulnerability can be used to assess the probability that the identified vulnerability is, in fact, a legitimate one. And, as the experience of the AI system increases over time, its ability to accurately predict false positives versus legitimate vulnerabilities will improve.

The technique allows other observations to be included as evidence in the assessment, for example, how frequently does the detection mechanism being used generate false positives. Effectively, employing Bayesian networks allows a more intelligent analysis that balances imperfect scanning techniques with expert human knowledge.

Determining whether an asset is configured such that it has an exploitable vulnerability can be more of an art than a science, and the process is susceptible to a high frequency of false positives.

To improve the reliability of vulnerability detection, Bayesian networks can be used when there is uncertainty in an assessment, in this case, whether an identified vulnerability is legitimate.

Leveraging Industry Vulnerability Remediation Priority Data

All modern vulnerability management products today are either cloud-based or have a cloud-based component. Although there are myriad benefits to a cloud-based vulnerability management platform, one of the most valuable, yet typically under-appreciated, is the user data that can be anonymized and culled from the application. Every organization is often remediating vulnerabilities on multiple assets daily. Multiply several daily remediation activities across dozens, hundreds, or thousands of customers, and a cloud-based vulnerability management product has a rich data source on which to apply an AI engine. Using this ever-changing and growing data source can reinforce or contradict conventional vulnerability remediation prioritization. Which assets are enterprises patching the most frequently? Which vulnerabilities appear to be the most concerning to peer organizations? Which are lower priorities?

We all learned in high school that copying one classmate's answer on a test question is not only unethical, but a risky proposition given there's no assurance that you picked the right classmate to copy. However, if you could determine that 90% of the class chose a specific answer, you'd have significantly more confidence that the answer was the right one.

Using a machine-learning technique known as Gradient Boosted Tree Regression, user behaviors and preferences can be blended with their history of remediation to predict what is important (for example, clickthrough rate). Using this ever-expanding database of cloud-based users and their remediation activity, the contribution to the vulnerability risk score becomes a dynamic element that reflects the constantly changing nature of the threat landscape.

Applying AI to actual vulnerability remediation data across multiple organizations can yield insights based on the collective judgement of many hundreds or thousands of IT and security peers, and as discussed previously, the larger that peer group grows, the higher the probability the decisions are sound.

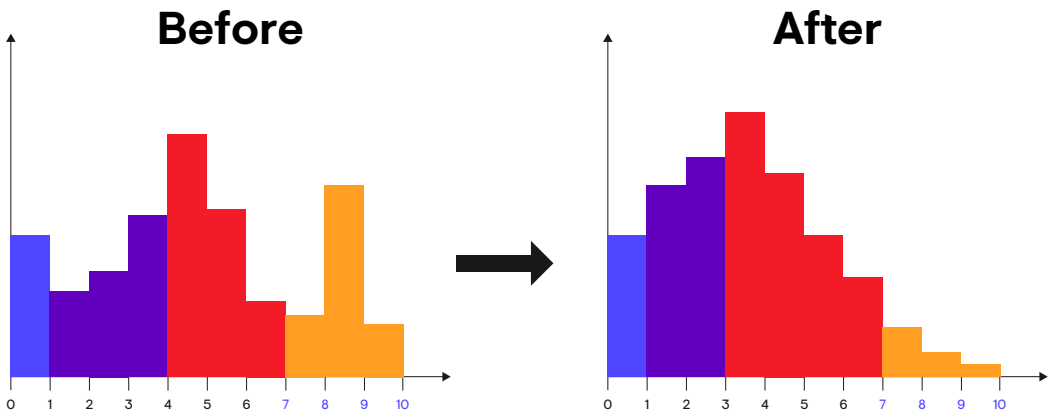
Developing Remediation Plan Recommendations

Once a context-driven priority list of vulnerabilities is established using the AI methodologies detailed here, optimizing remediation work is the final step in the vulnerability management process. Here, AI has a role to play as well.

Most medium to large enterprises can identify more vulnerabilities on their networks than could practically be remediated in any reasonable timeframe, so developing remediation plans that maximize risk reduction while minimizing remediation activity is essential to any modern vulnerability management program.

The objective of employing AI in vulnerability management is to give IT teams the risk insight necessary for them to focus only on legitimately critical vulnerabilities.

AI can be leveraged to address this challenge as well. Specifically, a Risk-Aware Recommender System – a hybrid between collaborative filtering and a content-based system – can be used to generate multiple remediation scenarios.



Similar to the algorithm used to make Amazon recommendations to consumers, a vulnerability management Recommender System would also take into account the risk reduction afforded by each remediation scenario using individual vulnerability risk scores generated using AI techniques as discussed at the beginning of this paper.

Conclusion

Advances in artificial intelligence afford IT and security teams a number of opportunities to reduce the human effort required to reduce the vulnerability risk of their networks. As the complexity of networks increases along with the number and sophistication of threat actors, AI technology can help alleviate the exploding burden on typical enterprise vulnerability management operations teams by enabling a combination of intelligent decision-making and automation, all of which is made possible by today's artificial intelligence technology.

Secureworks®

Secureworks® (NASDAQ: SCWX) a global cybersecurity leader, enables our customers and partners to outpace and outmaneuver adversaries with more precision, so they can rapidly adapt and respond to market forces to meet their business needs.

With a unique combination of cloud-native, SaaS security platform and intelligence-driven security solutions, informed by 20+ years of threat intelligence and research, no other security platform is grounded and informed with this much real-world experience.
www.secureworks.com

Corporate Headquarters

United States
1 Concourse Pkwy NE #500
Atlanta, GA 30328
+1 877 838 7947
www.secureworks.com

Europe & Middle East

France
8 avenue du Stade de France 93218
Saint Denis Cedex
+33 1 80 60 20 00

Germany
Main Airport Center,
Unterschweinstiege 10 60549
Frankfurt am Main Germany
069/9792-0

United Kingdom
One Creechurch Place,
1 Creechurch Ln
London EC3A 5AY
United Kingdom
+44(0)207 892 1000

1 Tanfield
Edinburgh EH3 5DA
United Kingdom
+44(0)131 260 3040

United Arab Emirates
Building 15, Dubai Internet City Dubai,
UAE PO Box 500111 00971 4 420
7000

Asia Pacific

Australia
Building 3, 14 Aquatic Drive Frenchs
Forest, Sydney NSW Australia 2086
1800 737 817

Japan
Solid Square East Tower 20F
580 Horikawa-cho, Saiwai-ku
Kawasaki, 212-8589
Japan
81-(44)556-4300
www.secureworks.jp