WiNG

Cyber Sherlocks And The Compelling Case For Hiring More Digital Watsons

SEPTEMBER 9, 2015 MARTIN GILES SECURITY

The following is a blog post that also appeared in edited form on the Huffington Post (link here (http://www.huffingtonpost.com/martin-giles/cyber-sherlocks-and-the-c_b_8084050.html)).

Sherlock Holmes and Doctor John Watson form one of literature's most memorable crime-fighting partnerships. When he dreamt up his characters, Sir Arthur Conan Doyle could not have foreseen the meteoric rise of computers, nor the fact that battling malicious hackers would become one of the world's most complex and pressing challenges. But if they are to foil the 21st century's Moriartys of malware, today's cyber defenders in companies will need to turn to a new generation of Watsons for assistance.

These new partners aren't fictional doctors; instead they are machine-learning platforms such as IBM's Watson (http://www.ibm.com/smarterplanet/us/en/ibmwatson/) , which hit the headlines in 2011 when it beat a set of human competitors to win the game show Jeopardy! A subset of artificial intelligence, machine learning involves creating powerful algorithms that spot patterns and relationships from historical data, and get better over time at making predictions about brand new data sets based on their experience. This approach lies at the heart of everything from Netflix's recommendation engine to the systems that banks and other financial firms use to spot fraud.

Several trends are likely to encourage wider use of the technology in cybersecurity too. The first of these is the soaring volume and velocity of attacks taking place. According to Symantec, more than 317 million pieces of malware (http://www.symantec.com/content/en/us/enterprise/images/mktg/Symantec/Email/12395/FY16_Symantec_WSTR_PT1_final.pdf) were created last year, a 26 percent increase over 2013. Brazen intrusions such as those at Sony Pictures, Anthem Healthcare and the U.S. government's Office of Personnel Management (http://www.nytimes.com/2015/07/10/us/office-of-personnel-management-hackers-got-data-of-millions.html?_r=0) are stark reminders of the fact that the opposition is becoming ever more sophisticated and persistent.

This escalation is happening at a time when security teams are facing a worrying shortage of skilled information-security professionals: the Information Systems Security Association recently put the size of this gap at one million jobs worldwide. Put starkly, there simply aren't enough cyber Sherlocks around. To make matters worse, some of the legacy tools they are using are becoming less and less effective. For instance, anti-virus systems that compare code with existing databases of dangerous "signatures" are of little use in spotting brand new malware.

Ready for training

Taken together, all these factors suggest that cyber threats are becoming too complex for humans to handle without more powerful tools to help them. Fortunately machine learning has developed to the point where it can be a formidable new assistant to defenders. In particular, the plummeting cost of data storage and computing power means that machine-learning models can now "train" cheaply on vast amounts of data, which greatly improves their ability to detect anomalies.

I have run into modern solutions that collect hundreds of millions of files via different feeds and then analyze an extremely broad set of their characteristics using a series of machine-learning techniques such as neural networks and gradient descent optimization. The goal is to distinguish "bad" characteristics of files from "good" ones, and the output is a model that analyzes all new files added to a customer's system and flags those considered malicious. This ability to classify things swiftly and at scale is a fundamental characteristic of many machine-learning models and explains why they have proven so popular in areas such as suppressing email spam.

Machine learning can also help with one of the toughest of all security challenges: spotting insiders who abuse their legitimate access to systems to steal intellectual property or other valuable and sensitive data. The most useful tools for defenders here are ones that can distinguish suspicious behavior of an employee or system from normal activity with very few false alarms or overlooked suspect events.

Services exist today that crunch large amounts of data such as information about the devices used by employees to access cloud services, the hours during which they are usually connected to the services and the actions they take on them. They then use this to construct a model of normal behavior over time and flag any deviations from the norm, such as a junior finance department

employee suddenly downloading reams of sensitive data from a cloud service late at night.

Learning opportunity

“Unsupervized learning” on unstructured data is quite often coupled with customer-driven, “supervised learning”. This can be in the form of specific security policies fed into a model, such as one that requires an alert to be issued when any user downloads a file that has the words “Highly Confidential” in it. Or it might involve security teams giving a thumbs-up or thumbs-down to alerts issued in order to train the model. Over time, this combination of supervized and unsupervized learning enables the service to figure out what really matters to an individual organization and minimizes false alarms.

As well as being used to spot risks, machine learning can also govern the way firms’ systems respond to them. Automating threat detection and response in this way may tempt some firms to downsize security teams. But a far more likely outcome is that machine learning will help in-house cyber sleuths deal with an overwhelming workload more effectively and free them up to focus on more strategic objectives.

The approach isn’t infallible. Machine learning models are only as good as the quality of the data sets on which they are based and the algorithms they employ. However, if the financial industry is any guide, the opportunity for security startups pioneering this approach is hugely exciting. Financial institutions were only able to rein in huge losses from credit-card fraud once they started using automated detection systems to augment human teams.

More automation is going to be a game-changer in cyber security too. Firms such as Cylance (http://www.cylance.com/) , Darktrace (https://www.darktrace.com/) and Palerra (http://palerra.com/) (a Wing portfolio company) are already pioneering the use of smart algorithms to combat cyber threats and others will surely follow. Even the great Sherlock Holmes himself would no doubt have welcomed machine learning’s assistance to crack a problem that is the polar opposite of elementary.

© 1990 — 2018 Wing, Menlo Park, CA