InfoWorld

November 10, 2015

GET TECHNOLOGY RIGHT®

SECURITY

Math to the rescue! Try this novel hacking defense

Can drilling into logs tell you whether you have been – or are being – hacked? Yes, if you use advanced math to look for weird variations like Anomaly Detective does

I CAME TO LOVE MATH LATER IN LIFE. IN JUNIOR high, I hated it so bad I had to take pre-algebra three times and my parents celebrated if I got a D. But I fell in love with it in my first year of college and I consider the A+ I got in my three-hour engineering calculus class to be one of the hardest-won achievements of my life.

My affection for math is one reason I'd like to introduce you to a company called Prelert, which uses math to detect anomalies and hackers. What's especially interesting for the math nerd in me is that Prelert's flagship product, Anomaly Detective, doesn't look for "malicious" items. You don't teach it what is or isn't malicious. Instead it uses math and statistics to look for occurrences and patterns that are anomalously above or below a normal baseline.

Anomaly Detective works by analyzing your logs in near-real time. Each field of information is a potential clue. The engine doesn't need to know what type of data it's working on or what each field means. It simply looks at the data and applies a mathematical analysis to identify anomalies. Anomaly Detective comes in two flavors: One works on top of Splunk, and another uses Elasticsearch, Hadoop, or one of several other big data platforms.

Anomalies are detected a few different ways. One might be the identification of a true statistical rarity — for example, the sudden appearance of FTP running between the DMZ and another network zone.

Individual "temporal deviations" models look for highs and lows compared to a baseline that Anomaly Detective algorithms have created. Anomaly Detective examines your data over time and applies machine learning to determine what is and isn't normal. It's smart enough to understand "work days" and other normal, reoccurring periods, where traffic and attributes go up or down in a repeating pattern.

Each data analysis scenario can have one or more analyzed dimensions. Dimension examples could be time, source, and destination IP addresses; source or destination network zones; application IDs; and so on. The higher the number of data analysis dimensions, the greater the chance an anomaly will be noticed. The more anomalous the dimensions detected, the stronger the confidence in the find will be.

Anomaly Detective analyzes the logs using custom-defined config files, which tell the engine what dimensions of data to look for, initiate lots of calculations, and produce anomaly scores. The industry calls what Anomaly Detective does "behavioral analytics." I call it cool math!

Prelert likes to tell a real-life DNS tunneling story. A subset of malware and hackers likes to tunnel their commands and/or victim's data (into and out of the victim's network) using the DNS protocol. They create fake DNS packets that will be passed along outside of the victim's network because network firewalls rarely block DNS traffic. But Anomaly Detective's engine noticed there were high levels of information content in the DNS query requests. Voilà! Anomaly and checkmate! The customer wasn't thinking about DNS tunneling. Anomaly Detective's engine wasn't thinking about DNS tunneling. It only cared that, suddenly, DNS Query Request packets were starting to look mathematically weird.

Think about how many Windows log events you record on a daily basis. I know customers who see terabytes logged every day. How can any single person analyze all of that? What events should you look for? Where might they be suspicious and when? How many bad logons in a single period is too many? Is that elevated logon allowed on that server, from that origination point?

That's the beauty of behavioral analytics. You don't have to figure out everything. Instead, simply tell Anomaly Detective to suck in your Windows event logs and let it do the rest.

How accurately can Anomaly Detective identify log data anomalies? I sent Prelert 15 different attack scenarios I see on a fairly regular basis, some of which can be incredibly hard for traditional systems to detect. Anomaly Detective detected 10 of the 15 — before it had been tuned.

I haven't tried Anomaly Detective in a large-scale environment or talked to any customers yet, but I'm infatuated with what Prelert is doing with pure math alone. Give it a whirl. It's worth a shot to see if behavioral analytics can give you a leg up on the sneaky bad guys.

— Roger A. Grimes — Security Adviser

Roger A. Grimes — Columnist
An InfoWorld security columnist since 2005, Roger Grimes holds more than 40 computer certifications and has authored eight books on computer security.