TIME TO REACT

Mike MacIntyre *explains why organisations need to take a data science approach to cyber security*

In revolutions, those that fail to adapt quickly tend to get left behind. Security is in the perilous position of being left behind by an accelerating and evolving digital revolution, just when it might be needed the most. This has big implications for the people at the helm of security and is further compounded by the need to consider the actions and intent of dynamic and intelligent adversaries. How can a Chief Information Security Officer (CISO) and their team keep pace? What's more, after years of screaming to be heard by the business or having given up trying, security now has their attention, but just how many CISOs are ready and able to make their case and tell their story?

For today's CISO to be effective they need to take a business-first approach powered by solid, data-driven decision-making capabilities. It's inconceivable, given the rapidly evolving threat landscape, a competitive business market and an increasingly distant (cloud) and heterogeneous technology environment, for a CISO to fly blind without timely and accurate instrumentation to help them navigate challenges.

It's vital that every third-party contract is reviewed for liability protection purposes

Security has traditionally been viewed as a back-office function that slows business innovation rather than enabling it. Now is the time to dismantle this stereotype and align the security mission and objectives with those of the business. This does not come naturally to many in the function, as a CISO often comes from a technical background (usually within the same organisation) and can find it challenging when they step into the business arena. More than ever they are being held directly accountable for their decisions (and are more present at board level due to the general increase in security awareness of those stakeholders) so it's vital that they can both pitch their needs in a way that the business understands (in line with its drivers) and back this up later with evidence of progress. However, to do this they must move their decision making from gut instinct to an evidence-based existence.

GETTING DATA RIGHT

For CISOs that believe they have already embraced a data-driven future, the information they receive is either meaningful but not timely, or it is timely but not meaningful. This is because the content is too technical and siloed. For example, the preparation for board reports or risk committee meetings can take weeks to produce (due to data extraction, curation and narration), so by the time it goes in front of those stakeholders it's out of date and has consumed large quantities of the team's time, which would otherwise have been spent on security! Conversely, the data that's available at short notice may be the total number of vulnerabilities, which has no place in a board report. Reporting isn't the only use case for data science. Corporate entities are awash with data from a multitude of sources so there is no shortage of opportunity to dive in. But where to start?

A COMPREHENSIVE INVENTORY POWERED BY AN AUTOMATED DATA PIPELINE IS ACHIEVABLE

Unsurprisingly, the market has seen a glut of security data analytics products that use maths to quickly identify the bad guys that are inevitably in your network and increase Security Operations Centre efficiency all through the use of Machine Learning or Artificial Intelligence. It's a problem that is ripe for data analysis innovation, but the realities are more intricate than many realise. For a start, AI is mostly marketing hype applied to a small subset of machine learning techniques, so don't be fooled by how a product is branded. At best, the algorithms embedded in these products perform highly specialised analysis in a single field and have been trained on large volumes of data. This is a far cry from general AI, which is a system that can perform any generalised task and answer questions across multiple domains. That's ok so long as you know what you are buying and are happy with the effort required to make it operational and know how to make it effective within your organisation.

Here is an example of some of those considerations. Reading the small print of a typical product you will find that they detect anomalies – not threats – which are just deviations from some measure of normal. What does normal mean? The model wasn't trained on your network, so it might not find your anomalies. And even if it does, are those anomalies truly a risk to the business? Are they the alerts you should be focusing your time on? Don't get me wrong, these approaches have a lot of merits and if applied correctly will undoubtedly be an effective tool in the defensive arsenal.

Data science should not just be confined to the 'detect' space. It's easy to understand how rapid identification of threats is appealing, but it takes a slightly defeatist view that you can't prevent attackers from breaching your defences. However, there are many opportunities for CISOs to use data to be more proactive in preventing threats from taking hold. For example, using data to raise the general cyber hygiene of an organisation is an underserved use case. An automated metrics and measurement programme can tell you a lot about how well your control infrastructure is deployed, configured and managed. Many organisations currently use point-in-time assessments, conducted manually or via questionnaires to assess their control status. No modern organisation can genuinely believe this is a sufficient fidelity of measurement to feel confident that a control infrastructure is operating as expected or needed (particularly considering regulatory reporting needs).

Another example, linked to the cyber hygiene use case is the creation, management and maintenance of an asset inventory. Critical to understanding what you need to defend, many organisations have given up on keeping such a system up to date. However, there is no reason that this has to be done manually. Whether it's with network discovery or scanning tools or modern CMDB applications, endpoint agents or even some superset of all of the above data sources, a comprehensive inventory powered by an automated data pipeline is achievable.

A particular benefit in these preventative data analysis use cases is that as well as delivering insight into where defences might be weak and at risk from attack, it can drive down the noise that plagues some of the detect solutions and can increase the effectiveness of the AI algorithms, as well as traditional rule-based detection. Given that the attacker signals are generally pretty weak in the data, this can have tremendous benefits.

STAYING RELEVANT

Regardless of the objectives of a data science program or product, a common theme is still needed to make it relevant to the business. Data science will only be successful if you can instrument, capture, move, combine and analyse data that covers the technical infrastructure (eg Endpoints, Network, Authentication, Vulnerabilities), the business context (eg Asset management, CMDB, application architecture, business processes), the identities that interact with these systems and processes (eg HR data, Identity Management) and threat intelligence that's pertinent to your business or sector.

Tackling business relevance leads the CISO to an unusual opportunity. Many organisations attempt a homogenous, 'one-size-fits-all' security solution.

▶

However, with security control performance measurement, contextualised to give each geographical region, business unit or product line their personalised view of security, the informed CISO can ask each group to set their own risk appetite, which can now be successfully tracked and monitored. This elevates them to an agile business operator where they can divert sparse resources to the most relevant (risk averse) parts of the business.

This holistic visibility is challenging to acquire, but with modern data analytics products it's by no means unattainable. There are many decisions to take along the way as to how to achieve this data utopia. What tools or mechanisms should I use to capture the events or data? What are the events of interest? Does my network have the bandwidth to move this data? Where am I moving this data to – cloud or data centre? How often do I want or need to analyse the data? How will the insight be consumed? And by who?

BUILD VERSUS BUY

Finally, do you have the team and resources to build such an analysis capability? Data scientists are not commonly found in security functions and those assigned to the business are often distracted by revenue generating problems. If you don't have the team, how can you meaningfully evaluate the vendors that might help? The decision to build versus buy is personal to the organisation, but some key questions to ask that can point you in the right direction are: Do you trust a vendor to deliver on all your requirements? Can you live with 80 percent? What is the total cost of ownership of a build versus buy solution? (Note that people usually underestimate the recurring support and maintenance costs of a build solution).

The scale of setting up such a program can put many people off, not least those that have been in the post for some years and may not be ready or willing to expose the painful realities that have built up on their watch. However, you have the opportunity to start small and build out the capability incrementally. Whether it's a threat detection use case or a preventative cyber hygiene measurement program, having clearly scoped use cases, evaluating the data that is readily available and seeing what value can be demonstrated quickly is a great starting point upon which you can build. Taking two years to develop a polished data collection program is the wrong first move to data-driven decision making. A great starting place is to look at how you can automate the repetitive tasks that your team spend a disproportionate amount of time sorting out, thus freeing them up to focus on the other problems you never thought you'd get to.

CISOS NEED TO TAKE A BUSINESS-FIRST APPROACH POWERED BY DATA-DRIVEN DECISIONS

Having accurate and up-to-date data at your fingertips means you can stop security becoming blockers to the business by being more agile and relevant and bring security into the digital revolution. The future for security is data driven. The journey to get there is fraught with uncertainty, but should you succeed the rewards could be plentiful. Good luck ●

Mike MacIntyre is Chief Scientist and VP of Product at Panaseer. He joined the company in 2015 and is responsible for researching, experimenting and applying computational and analytical techniques to derive new insights in cyber security. His team's mission is to make cyber security more data driven.

AI is mostly marketing hype applied to a small subset of machine learning techniques