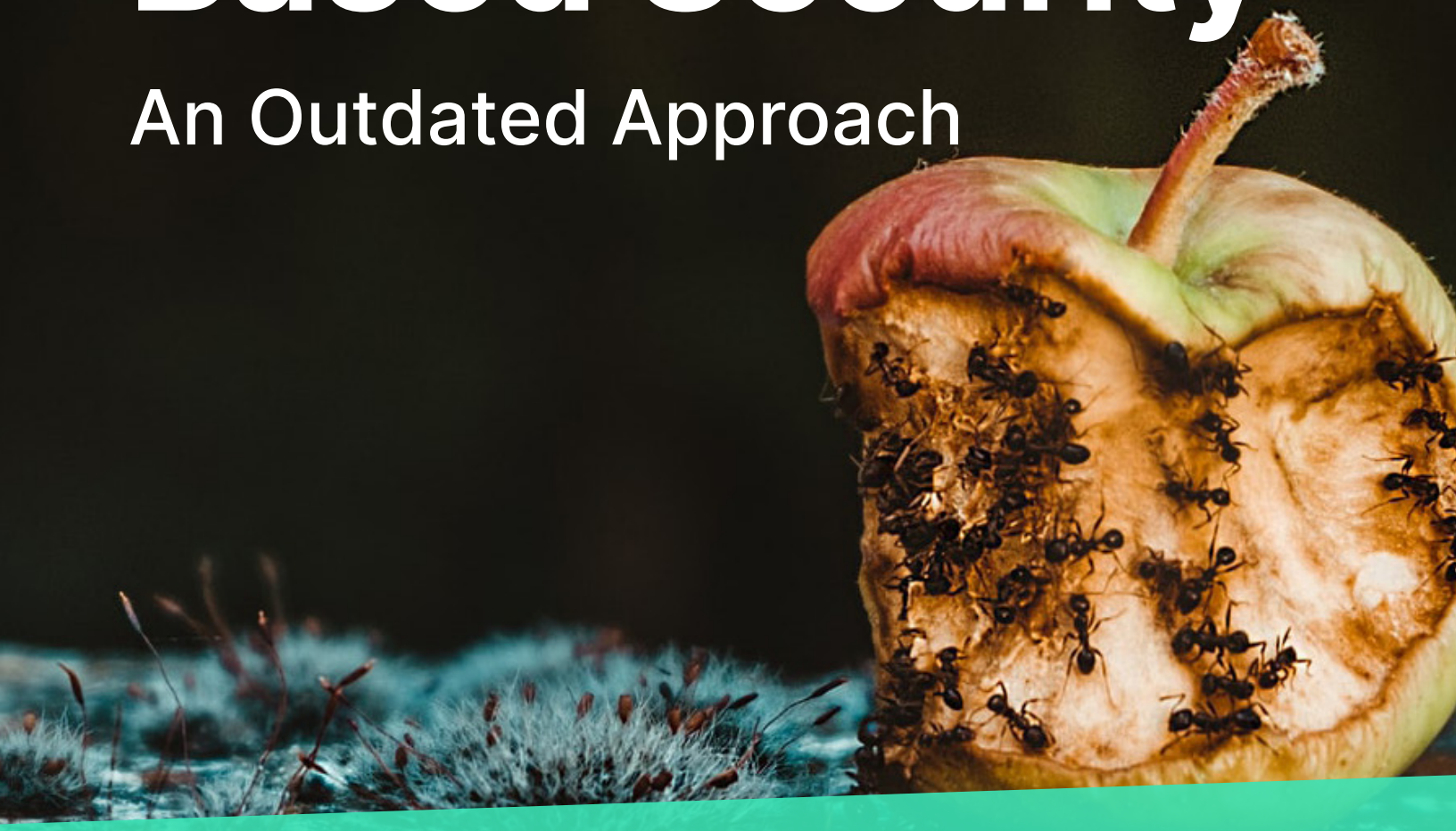
JupiterOne

Sampling Based Security

An Outdated Approach



If all it takes is one bad apple to spoil the entire bunch, should the owner of an apple orchard do a statistical sampling to look for the bad apple, or should they look at every apple in the orchard? Applying this to cybersecurity, is it smart to do a statistical sampling of your systems for compliance or should you do a deep analysis of every single one in order to find that one lurking risk?

What is Statistical Sampling

Statistical sampling is a type of data analysis that reviews a sample set of data to predict an outcome across an entire population. A basic example of statistical analysis is when manufacturing facilities conduct a quality assessment of the individual parts they are using in the production of a product. A car manufacturer generally doesn't double check every single bolt, nut, screw, or component that goes into a vehicle. Instead they look at a percentage of those components, and if the components they look at are built appropriately they assume that the rest of the population is as well. The number of samples that have to be tested is determined by the accuracy level you need in your results. If you've ever heard the term "confidence level", this comes from understanding the sample size in comparison to the population size.

Statistical Sampling or Population Sampling

Statistical sampling is chosen over complete population sampling when there is a limiting factor on the amount of resources available to execute the sampling. If the sampling process is backed with an unlimited number of resources, you can easily analyze every single item in the population. Statistical sampling becomes problematic when the problem you are looking for causes enough risk to necessitate a complete population analysis. Reconciling these two competing forces is how we choose from the two sampling models.

> **Population sampling becomes a requirement for detection of high risk problems.**

Sampling in Security

Statistical sampling is all around us. In the cyber security space we tend to see statistical sampling done in relation to compliance and security governance. It's impossible for a compliance officer to double check that every system and every risk is accounted for in an audit. For that to happen there would have to be an assumption of nearly infinite time and resources. We've become accustomed to our security audits being sample sets and not entire population analysis. However, this is a big issue. If the compromise of a single system can result in such a drastic impact on the business that you must find it at all costs, statistical sampling becomes useless. Population sampling becomes a requirement for detection of these high risk problems. To take it back to our original discussion on the apple orchard, if one issue spoils the entire bunch – we have to find that one issue.

It's Not a Matter of If, But a Matter of When

Security best practices also call for retesting after a change. Yet in today's digital and cyber operations, changes are constant. Combined with the above, this translates to the need for "continuous population analysis". If a full population analysis is already impossible, continuous population analysis becomes even more so. Security teams resort to the practices of detection and response because we have convinced ourselves that compromise is not a matter of if, but when.

Why Compromise isn't Mandatory

Today we live in a world where automation, APIs, and data analytics have grown in capability and are delivering significant advancements in security assessment capabilities. With modern enterprise infrastructure and security tooling being built with API first models, security auditors and teams can collect evidence from every system in their environment instead of looking at sample sets. Essentially, we have finally gotten to the point where security resources are automated enough that population sampling is within reach. Automation is the only way that we can possibly bring down the resource cost of actual security as opposed to sample set based analysis. Lucky for us, we now live in a world where it is possible to fully automate the collection and analysis of security and infrastructure data. With automation comes a lowering or elimination of cost, and with resource savings comes the ability to transition from sampling to population based security analysis.

JupiterOne is your security operations central command. Whether you're an established enterprise, or a budding startup, our cloud-native cyber asset management, visibility, and security platform empowers customers to centralize security operations, streamline processes, and drastically reduce enterprise risk.