Employing Big Data for cyber Attacks Analysis

Ankit Srivastava[1], **K. Venkatesh**[2]

Abstract: ***Many enterprises have made substantial investments in security information and event management and log management technologies over the years to collect, manage and analyze logs. Advances in large-scale analytics enable a well-honed security program to use data to spot anomalies and analyze attacks. However, it is easy to be overcome by a deluge of indicators and warnings derived from this data. In this paper I am trying to explore the big data usage in the analysis of cyber-attacks.***

Keywords: Big Data, Security Investigation, Cyber-attacks, Cyber-Security

1. Introduction

The fundamental issue that associations face when battling digital assaults is that they don't recognize what to search for and think that it hard to translate all the information they get from their systems. Utilizing examination, associations can practice continuous checking of system and client practices, distinguishing suspicious action as it happens. The Organization can show different system, client, application and administration profiles to make discernment driven efforts to establish safety equipped for rapidly distinguishing irregularities and corresponding occasions showing a risk or attack:

1) **User:** authentication and access location, access date and time, user profiles, privileges, roles, travel and business itineraries, activity behaviors, normal working hours, typical data accessed, application usage
2) **Device:** type, software revision, security certificates, protocols
3) **Network:** locations, destinations, date and time, new and non-standard ports, code installation, log data, activity and bandwidth
4) **Customer:** customer database, credit/debit card numbers, purchase histories, authentication, addresses, personal data
5) **Content:** documents, files, email, application availability, intellectual property

Utilizing examination, associations can practice continuous checking of system and client practices, distinguishing suspicious action as it happens. Associations can show different system, client, application and administration profiles to make discernment driven efforts to establish safety equipped for rapidly distinguishing irregularities and corresponding occasions showing a risk or assault.

2. Utilizing Big Data with Analytics to Catch a Thief:

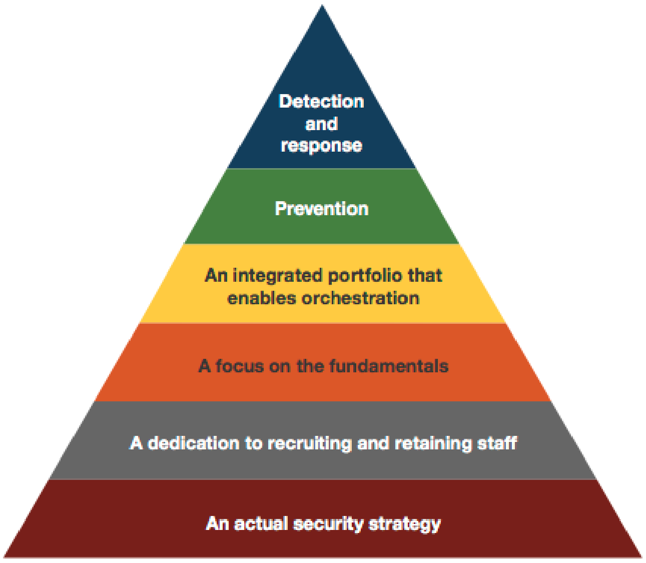
Utilizing investigation, associations can practice continuous observing of system and client practices, recognizing suspicious action as it happens. Associations can show different system, client, application and administration profiles to make discernment driven efforts to establish safety fit for rapidly recognizing peculiarities and associating occasions showing a danger or attack:

- Traffic anomalies to, from or between data warehouses.
- Suspicious activity in high value or sensitive resources of your data network.
- Suspicious user behaviors such as varied access times, levels, location, information queries and destinations.
- Newly installed software or different protocols used to access sensitive information.
- Identify ports used to aggregate traffic for external offload of data.
- Unauthorized or dated devices accessing a network
- Suspicious customer transactions
- Cyber-attack and criminal profiling

Examination can be exceedingly successful in distinguishing an attack not exactly in progress or prescribing an activity to counter an attack, in this way minimizing or killing misfortunes. Examination makes utilization of enormous information with opportune investigation of divergent occasions to obstruct both the smallest and biggest scale attacks. [1]

3. Cyber-Attacks and Big Data

There has always been a lively–and very relevant–debate in security circles about where and how to expend resources, be they cash, personnel or attention. Budgets are tight and we have not yet found a holistic and bullet-proof solution to stopping hackers and malware from getting into our systems.


Detection and response
Prevention
An integrated portfolio that enables orchestration
A focus on the fundamentals
A dedication to recruiting and retaining staff
An actual security strategy

Source: Forrester Research, Inc.

At Domain Tools, we are applying this vision to the area we know best: Domains and DNS data. By mapping the entire

Internet of domains and IP addresses, and collecting a lot of DNS related data in the process (as we have done over the last 12 years), we have built a huge database of “who is doing what on the Internet.” Through machine learning, trend analysis, and a pivot engine that draws connections between otherwise unconnected domains based on associated DNS data, we are able to track ‘bad actors’ in cyberspace often before they strike.By scoring areas and IP addresses on classifications of "disagreeableness," and making that accessible through our Cyber Threat Intelligence arrangements, we empower security experts to rapidly explore and triage cautions and suspicious activity to "partitioned the sign from the commotion." More proactively, we are working with industry accomplices including SIEM, IPS and interruption recognition suppliers to incorporate our information to consequently distinguish suspicious movement and piece awful traffic.

Yes, Big Data has the power to fundamentally change the game in cyber security. But it will take a combined effort of many solutions—thick firewalls and Intrusion Prevention Systems, powerful system and network analysis, and “outside the firewall” intelligence—to optimize the power of big data and put Security Ops teams back in the control seat.[2]

Cyber threats are constantly evolving - becoming more sophisticated, targeted and sustained. Traditional forms of security (Firewalls, IDS/IPS, Antivirus, Vulnerability Management etc.) are no longer enough to stop Advanced Persistent Threats. To protect yourself, your business and your reputation from the risk of cyber-attack, what’s needed now is a new way of monitoring activity with your network across all geographies and correlating it together so that it can be analyzed, interpreted and acted upon.

4. Profiling

Cyber-attack profiling plays a key role in Investigation. An informed endeavor to give particular data as to the sort of person who carried out a certain wrongdoing. A profile taking into account attributes examples or variables of uniqueness that recognizes certain people from the overall public. [3]

This area clarifies how Big Data is changing the investigation scene. Specifically, Big Data investigation can be utilized to enhance data security and situational mindfulness. For instance, Big Data investigation can be utilized to dissect budgetary exchanges, log records, and system activity to distinguish peculiarities and suspicious exercises, and to associate different wellsprings of data into a lucid perspective.

Data-driven information security dates back to bank fraud detection and anomaly-based intrusion detection systems. Fraud detection is one of the most visible uses for Big Data analytics. Credit card companies have conducted fraud detection for decades. However, the custom-built infrastructure to mine Big Data for fraud detection was not economical to adapt for other fraud detection uses. Off-the-shelf Big Data tools and techniques are now bringing attention to analytics for fraud detection in healthcare, insurance, and other fields. [4]

In the connection of data analytics for interruption recognition, the accompanying development is expected:

- First era: Intrusion location frameworks – Security planners understood the requirement for layered security (e.g,responsive security and break reaction) on the grounds that a framework with 100% defensive security is incomprehensible.
- Second era: Security information and event management (SIEM) – Managing cautions from diverse intrusion detection sensors and guidelines was a major test in big business settings. SIEM systems aggregateand filter alarms from many sources and present actionable information to security analysts.
- Third era: Big Data analytics in security (second era SIEM) – Big Data tools have the potential toprovide a significant advance in actionable security intelligence by reducing the time for correlating,consolidating, and contextualizing diverse security event information, and also for correlating long-termhistorical data for forensic purposes.

Examining logs, system bundles, and framework occasions for criminology and intrusion detection has customarily been a huge issue; nonetheless, customary advances neglect to give the devices to bolster long-term, extensive scale investigation for a few reasons:

1) Storing and retaining a large quantity of data was not economically feasible. As a result, most event logsand other recorded computer activity were deleted after a fixed retention period (e.g., 60 days).
2) Performing analytics and complex queries on large, structured data sets was inefficient because traditional tools did not leverage Big Data technologies.
3) Traditional tools were not designed to analyze and manage unstructured data. As a result, traditional tools had rigid, defined schemas. Big Data tools (e.g. Piglatin scripts and regular expressions) can query data in flexible formats.
4) Big Data systems use cluster computing infrastructures. As a result, the systems are more reliable and available, and provide guarantees that queries on the systems are processed to completion.

5. New Big Data technologies

As databases related to the Hadoop ecosystem and stream processing, areenabling the storage and analysis of large heterogeneous data sets at an unprecedented scale and speed**.**These technologies will transform security analytics by:

a) collecting data at a massive scale from many internal enterprise sources and external sources such as vulnerability databases;
b) performing deeper analytics on the data;
c) providing a consolidated view of security-related information;
d) achieving real-time analysis of streaming data.

It is essential to note that Big Data instruments still require system architects and examiners to have a profound information of their framework to appropriately configure the Big Data investigation tools. [4]

References

[1] Nazar Tymoshyk, “Beating Back Cyber Attacks with Big Data + Analytics”,October 28, 2014
[2] Jeff day,”Big Data and Cyber Threat Intelligence”,September 22, 2014
[3] Arun Warikoo,”Proposed Methodology for Cyber Criminal Profiling”,1–7, 2014.
[4] CSA (Cloud Security Alliance),”Big Data Analytics for Security Intelligence”,September 2013

Author Profile

Ankit Srivastava is an M.Tech Student in SRM University. His specialization is Information Security and Cyber Forensics.

K. Venkatesh is an Asst. Professor (Sr. G) in department of Information Technology at SRM University.His area of interest is Data Mining.