Big Button: A Situational Awareness and Cyber Defense Solution

STEVE PYCHA
Forcepoint
UNITED STATES

spycha@forcepoint.com

ARTICLE I. ABSTRACT

Millions of cyber events occur throughout our IT systems every day. These events; such as, logons, send/receiving emails, internet activities, application updates/installs; or file sharing are logged and recorded. Most are expected and offer no threat to the system. Others are considered suspicious and their detection becomes paramount when a hostile actor (whether internal or external) is introduced into your system. Quick detection prevents the hijacking of your system and ensures that intellectual property, or regulatory compliant data, is not compromised. Most organizations have anti-virus software, firewalls, blacklists and other detection systems in place for exposing the malicious behavior. This paper discusses the North Atlantic Treaty Organization Multinational Cyber Defense Capability (NATO MN CD2) project. As part of this effort, we discuss how Big Button Analytics (BBA) can facilitate situational awareness and mission readiness enabling rapid response to known and unknown threats.

1.0 DATA WAREHOUSING

The NATO MN CD2 data warehouse incorporates data from various cyber data sets into a single authoritative repository. This data includes incident tickets; IP addresses; hostnames; MAC addresses; email addresses; source and destination information; application, service, device, and user accounts; URLs; watch lists; vulnerabilities, and so on.

Traditionally, the logging of cyber data is often held in disparate data sources within network appliances or various silos. This represents a major analytical challenge. Overcoming this challenge requires (1) the identification of entity types across these unrelated sources and (2) the resolution of any inconsistencies that prevent comprehensive analysis.

A comprehensive data model was developed for the purpose of bridging these sources. Data is ingested into the warehouse using a fully automated process that is managed by our data warehouse. The first step in the BBA approach is to store the data from these various sources in a single data warehouse. Then, through a proprietary ETL (Extract, Transform, Load) process, data from these various sources are imported into a standardized format and structured according to a pre-defined schedule that provides a near-real time view of the data. This warehouse is managed by us; however, the original data maintains its security, access and integrity to eliminate any ownership issues.

1.1 Big Button Analytics

The data sources are now accessible and standardized. Your analysts, mission commanders, and other approved personnel can utilize BBA to provide user-friendly, streamlined interfaces that are custom tailored to NATO MN

CD2 specifications and use cases. Users can quickly search for entities, create dossier reports, view data through dashboards, perform basic pattern analysis, and create alerts for when data changes.

The Search capabilities allow the analyst to perform information retrieval against various data sources. Network activity related to a certain application, host machine, IP address, timestamp, Common Vulnerability and Exposure (CVE), organization, Advanced Persistent Threat (APT) or event/incident can all be discovered through an ad-hoc search interface. Prior to executing any query, the analyst can choose which sources will be searched. Results are returned in a user-friendly tabular format that offers sorting, filtering and downloading capabilities.

Typically, an analyst's investigation does not end with the results from one search. Results need to be cross referenced with other data sources and additional details from the same source may be required. The BBA tool expects such actions to be taken. With a few clicks of the mouse, the Drill Down feature allows for retrieval of additional detail from the same or other sources.

The presentation of this data is important; therefore, results are returned in several intuitive formats (including data grids). Data sets that involve several fields can be challenging to view and reviewing data from multiple data sets can add another layer of complexity. BBA includes a feature called Dossier Reports that displays the data from one or multiple sources in a more readable format. As an example, a report created on an infected host machine may contain detailed information from multiple sources. When these sources are combined into a single report, the analyst has a comprehensive view of (1) the host machine; (2) recently ran applications; (3) suspicious network activities; (4) affected services; and (5) a list of vulnerabilities.

Reactive queries (where a specific entity is targeted from the outset) are part of an analyst's responsibility. In addition, there is a requirement to be proactive in the investigation where a hypothesis based on an experience is observed. For such cases, the BBA tool provides Dashboards. These Dashboards are critical for NATO MN DC2 command staff to quickly review aggregate data across missions. Data within a dashboard can be displayed in various formats; such as, a table, a bar/line graph, or a pie chart. For example, a dashboard may contain (1) a table that details the most recent suspicious activity by site and/or malware type; (2) a bar graph depicting suspicious incidents by type; (3) resolved issues by year in a pie chart; and (4) a line graph showing events by organizations over time.

Visualization of data is another approach included in Forcepoint's framework. BBA analyzes patterns using Link and Geospatial analytical methods. Through link analysis, relationships between various entities can easily be discovered. Then, using latitude/longitude coordinates these entities can be plotted on a map. These visualizations assist in the detection of suspicious behaviors and provide another form of data presentation.

New cyber events are occurring every second; however, it's not practical to expect analysts to query the data on a continual basis. BBA offers a more efficient approach by employing Alerts to search for known targets. They can be scheduled to run at specific times or intervals, and look for changes or specific conditions in the data. When a condition is met during the execution, an Alert notifies specific users through the BBA interface or email.

1.1.1 Worm Incident

John, an employee, encounters an issue installing a software application on his computer. The answer cannot be found in the software installation guide and he decides to search for the answer on the Internet. Starting with a search engine, he enters a keyword/question and the search returns numerous results. John then scans the results to find the links that will provide the best solution. As the he clicks on the links, he discovers that some of these sites are accessible and others are not. The websites he cannot reach are being block by a firewall inside his network. As intended, the firewall is using the security rules that have been set in place by the IT Network Administrators. This policy prevents employees (such as John) from visiting any untrustworthy web sites that are known to cause harm to a system.

John then focuses his attentions on the websites that he can access. As with most search results, some links are more relevant than others. Through trial and error, he visits the accessible websites until he finds a suitable answer. During this process, he inadvertently clicks on a website that links to a software download. Fortunately, John is immediately prompted by his anti-virus software that this download contains malicious code and has been quarantined. Like the firewall, the anti-virus software operates as expected and uses the rules set in place by the IT department.

Eventually, John finds a website that will satisfy his question. He closes his web browser and successfully completes the installation of the software. John continues on with other daily activities that include sharing files with other machines and reading/replying to emails. Unbeknownst to John, the firewall and anti-virus software was not completely effective. While John was browsing the Internet, he clicked on link that surreptitiously downloaded malware to his computer. The website was not on any blacklist and the malicious program was not detected by his antivirus. John's computer is now infected.

This downloaded program was not "just a virus". It was computer worm (which can quickly infect an entire network of computers). Jane, another employee, uses a computer that is targeted by this worm. Fortunately, Jane keeps her anti-virus software more up-to-date than John. The worm is detected and is safely isolated. It can not to cause any harm to her machine. Following proper procedure, Jane fills out an incident ticket with her IT department. She provides information; such as:

- Date that malware was detected/isolated,
- Host machine name,
- IP Address,
- Facility/Site, and
- Virus type

The IT Network Administrator subsequently sends an email to all employees informing them of the worm and instructs everyone to update their anti-virus definitions. John sees this email and updates his anti-virus software. The worm is quarantined on John's machine and he follows up with an incident ticket.

1.1.1.1 Incident Analysis

This is just one example of the many threats a network can face. As an analyst, it necessary to find the "what" and "where" of a situation while identifying "how" and "why" the situation occurred. Knowing what systems resources were affected and where they are located is important; however, it is equally important to find out why they were attacked and how they were disturbed. Unfortunately, the "how" and "why" is typically not found in just one set of data.

Let's assume our analyst is unaware of this worm incident. Through the use of the BBA dashboard, the analyst is able to view a graph that visualizes a spike in recent incident tickets related to a computer worm. This spike represents the incident tickets that were submitted by several employees, including Jane and John, in relation to the computer worm event.

Our analyst then searches the Incident database for any tickets involving computer worm events. The results come back in the form of a data grid/table. The analyst can see the host machine names, their IP address and the facility where they are located. The analyst notices that a majority of the machines that were recently affected are located at the same facility. By drilling down on these affected host machines, the analyst can view related information located in other data sets. In this case, the following data sources are being cross-referenced:

- Device
- Software
- Common Vulnerabilities
- IP Traffic – source and destination IP address with timestamp.

In addition to the sources listed above, our analyst re-queries the Incidents database to search for other events involving these host machines. Nothing related to the worm activity is determined; therefore, the analyst turns his attention to results from these new sets of data.

The Device data set will contain information about the host machines themselves (similar to an inventory). It includes the name of the machine along with the MAC address, IP address(s), subnet mask, type of machine, and manufacture model. In this scenario, the analyst checks to see if the affected machines are associated with a certain subnet or model of computer (e.g. Mac Book or Dell). Our analyst finds a correlation and recommends that similar model machines in the same subnet be checked for this worm as well.

The Software data set catalogs recently installed applications for each machine. Information being collected includes the software's name, version and publisher. With this data set, our analyst can check for a connection between affected machines and the application being installed. It is possible that a certain versions of an application, or several applications from the same software vendor, are related to the computer worm. In this case, our analyst finds a similar application and recalls the same application being installed on his own machine. The install was part of a system wide software update and our analyst feels fairly certain that this application was approved and not related.

As a safeguard, our analyst checks the Common Vulnerabilities data source for any known security threats or weaknesses with this recently installed software update. A CVE (Common Vulnerabilities and Exposures) is listed for the application but pertains to an earlier version of the software.

Next, our analyst examines the Network traffic. The IP traffic data set is a log of all the interactions between internal IP address (identified as the Source) and the other IP addresses.(identified as the Destination). Destination IPs can be both internal to the network and external. For example, traffic may include a recently visited website, a file being shared from one machine to another, or an email being received/sent. Our analyst compares the IP addresses of the affected machines with the source IP's in this data set, and then views the destinations IPs with which these affect machines have communicated.

In this case, the IP traffic source has been integrated with geospatial data. This referential source associates an IP address to latitude and longitude coordinates. With these geo-values, the physical location of both internal/external Destination and Source IPs are observed on a map. Our analyst finds that the location of most of the IP traffic is as expected. The Source IPs are plotted on the map where the facility of the affected machines is situated and the internal Destination IPs are displayed on the map where the organization's other facilities are located. In addition, most of the external Destination IP's seemed to be inconspicuous as well. Most are located in Western Europe or North America; however, the traffic for one Source IP depicts communication with a Destination found in a hostile country that has known ties to cyber-terrorism. This Source IP is associated with the host machine name of John's computer. Alas, the origination of where and how this worm was introduced inside the network becomes apparent.

1.1.1.2 Conclusion

Our analyst concludes the investigation by submitting a dossier report detailing his data findings and a list of recommended actions. The details of this report include the following:

Data:

- Incident Tickets submitted related to this Worm
- Known host machines impacted by this worm
- Possible machines affected by this worm
- Suspected IP address of the origination of this worm
- Machine suspected of first downloading the worm.

Actions:

- Check all machines (known, and probable) for …
- Anti-virus software is updated
- This malicious code is isolated
- Add suspected IP address/URL to a blacklist ensuring that the firewall blocks any future communication with this harmful site
- Identify and interview the owner of the machine discussing his activity that introduced this worm into the environment.

Our systems are constantly under threats from outside sources and security measures are not always 100% successful. To detect these threats and provide a rapid response, it is important to integrate and analyze data from several IT sources. An analyst, who is often the last line of defense, has a comprehensive view of these threats and can quickly identify what machines were affected, where they are located, why the system was attacked, and how it was disturbed.