secure|set ACADEMY

3801 FRANKLIN STREET
DENVER, CO 80205
800 445 0024

SSA Hacking 101 'Intro to Data Visualization as a HUNT tool'

Activity Objective

This Hacking 101 is to show how you can use visualization to spot deltas in datasets enabling you to detect threat actors or unauthorized activity or changes in your environment.

Required Resources

- Laptop
- A Computer Network that is providing DHCP
- A Computer network with an internet connection

References

https://en.wikipedia.org/wiki/Data_visualization

https://www.sas.com/en_us/insights/big-data/data-visualization.html

https://searchbusinessanalytics.techtarget.com/definition/data-visualization

https://en.wikipedia.org/wiki/Diff_utility

https://www.diffchecker.com/

https://online-image-comparison.com/

TERMS and CONCEPTS

- **Wetware**:
  - **Wetware** is a term drawn from the computer-related idea of hardware or software, but applied to biological life forms
- **Data Visualization**:
  - Data visualization or data visualization is viewed by many disciplines as a modern equivalent of visual communication. It involves the creation and study of the visual representation of data.
  - Effective visualization helps users analyze and reason about data and evidence. It makes complex data more accessible, understandable and usable. Users may have particular analytical tasks

3801 FRANKLIN STREET
DENVER, CO 80205
800 445 0024

- Cybersecurity Kill Chain:
    - The term kill chain was originally used as a military concept related to the structure of an attack; consisting of target identification, force dispatch to target, decision and order to attack the target, and finally the destruction of the target.
    - Conversely, the idea of "breaking" an opponent's kill chain is a method of defense or preemptive action.
    - More recently, Lockheed Martin adapted this concept to information security, using it as a method for modeling intrusions on a computer network.
    - This model has seen some adoption in the information security community.
    - However, acceptance is not universal, with critics pointing to what they believe are fundamental flaws in the model.
- Baseline or Normal:
    - Measurement is the assignment of a number to a characteristic of an object or event, which can be compared with other objects or events. The scope and application of measurement are dependent on the context and discipline.
    - A baseline is a benchmark that is used as a foundation for measuring or comparing current and past values. These values could be any measurable IT activity, Ip addresses, hostnames, URLS, rate, TCP sessions, session duration, etc etc.
    - Once you understand what is normal, you can measure for deviations from the norm.
    - This would be a analysts focus, to determine what caused the change, and do the security controls and measurements need to be remodeled to adapt to this.
- Delta or Change:
    - The upper-case letter Δ can be used to denote:
    - Change of any changeable quantity, in mathematics and the sciences.
    - Mathematicians are fond of Greek letters, and they use the capital letter delta, which looks like a triangle (Δ), to symbolize change. When it comes to a pair of numbers, delta signifies the difference between them. You arrive at this difference by using basic arithmetic and subtracting the smaller number from the larger one. In some cases, the numbers are in chronological order or some other ordered sequence, and you may have to subtract the larger one from the smaller one to preserve the order. This might result in a negative number.
- Dwell Time:
    - How do security professionals know they’re successful in protecting and defending their data? While we rely on several cybersecurity metrics to measure the effectiveness of our efforts, there’s one in particular that many of us here at Armor are passionate about. It’s called dwell time.
    - What is dwell time and why does it matter?
    - Dwell time is the duration a threat actor has undetected access in a network until it’s completely removed. Currently measured in days, the average dwell time varies depending on who you ask.

3801 FRANKLIN STREET
DENVER, CO 80205
800 445 0024

    - If anything, this is a quantifiable measurement of how good your cybersecurity program is.
- Zero Day Threat:
    - A zero-day threat is a threat that exploits an unknown computer security vulnerability. The term is derived from the age of the exploit, which takes place before or on the first (or "zeroth") day of a developer's awareness of the exploit or bug. This means that there is no known security fix because developers are oblivious to the vulnerability or threat.
    - Attackers exploit zero-day vulnerabilities through different vectors. Web browsers are the most common, due to their popularity. Attackers also send emails with attachments exploiting software attachment vulnerabilities.
    - A zero-day threat is also known as a zero-hour attack or day-zero attack.
- Why HUNT threats:
    - Reduce dwell time by expediting adversary detection and reducing investigation and forensic costs.
    - Evict adversaries with minimal business disruption.
    - Hunting enables an organization to identify, characterize, analyze, and remove advanced adversaries as early in the kill chain as possible, which can be facilitated with automated technologies that support the hunt.
    - Stopping adversaries early in the kill chain generally hinders them from reaching their ultimate target.
    - At a more technical level, there are additional benefits to hunting. Hunting can find attacks that can't readily be detected by passive defenses.
    - For example, hunting is effective at finding previously unknown attacks because it doesn't depend on already knowing the signs of a specific attack. Similarly, hunting can find attacks that don't use malware because it isn't specifically focused on malware-based attacks. Many passive defenses rely on prior knowledge of malware characteristics (i.e., signatures) and can't identify attacks that don't use malware.
    - Many passive defenses rely on prior knowledge of malware characteristics (i.e., signatures) and can't identify attacks that don't use malware.
- Automation and SAO (Security Automation and Orchestration)
    - Hunters benefit greatly from automating much of analysis. Advanced analytics enable hunters to resolve ambiguities and apparent conflicts in the collected data and explore high priority data more efficiently. Hunters can also take advantage of advanced analytics to fine-tune detection and refine remediation strategies.
    - Similarly, automated analysis enables hunters to focus on higher-order patterns and signals that are often lost when focusing on specific signatures. For instance, the data science methodologies are equally useful for hunting for Domain Name System (DNS) response errors, features, or other characteristics that deviate from the norm. In this regard, automated analytics help surface useful insights for the hunter, who then prioritizes and focuses on the most important anomalies.
- How to prepare your environment and obtain baselines, what should you gather baselines on?

    - Info IT Assets
    - Info IT Networks
    - Network Maps
    - What is Normal behavior
    - Deploying Sensors
    - Obtaining management prioritization
    - Conduct or obtain a Cyber Risk Assessment
- Its better to measure 10x and cut 1x instead of just turning on all the logs and collecting everything, it won't provide value and the quality and fidelity is low, this moves you farther away from your goals.
- Analytical Thinking and Approaches
    - Analytical thinking is a critical component of visual thinking that gives one the ability to solve problems quickly and effectively. It involves a methodical step-by-step approach to thinking that allows you to break down complex problems into single and manageable components.
    - Analytical thinking involves the process of gathering relevant information and identifying key issues related to this information. This type of thinking also requires you to compare sets of data from different sources; identify possible cause and effect patterns and draw appropriate conclusions from these datasets in order to arrive at appropriate solutions.
    - Analytical thinking can be broken down into three main steps:
        - Gather Information
            - Here you must gather all the necessary information that will be required to help you solve your problems. You also need to recognize whether you need to obtain more or higher quality information in order to collect all the relevant data you will need to arrive at an appropriate solution.
            - Gathering information requires that you ask appropriate questions of yourself and of others in order to gain the necessary insights that will enable you to make more effective decisions about the problems you are facing. However, you also need to consider the relevance of your sources and the means by which you will gather this information.
        - Identify Issues and Problems
            - When it comes to analytical thinking, it's important to develop your ability to recognize underlying issues or problems based on trends, associations and cause-effect relationships between datasets.
        - Analytical Thinking and Visual Thinking
            - Analytical thinking is very much integrated into the visual thinking framework, and especially into The Path. It's a part of the problem-solving process you will utilize as you work visually towards acquiring

3801 FRANKLIN STREET
DENVER, CO 80205
800 445 0024

the necessary insights that will help you achieve your goals and objectives.

- What is NORMAL
  - Every asset is different in terms of what's considered normal activity.
  - Each network is unique as well in terms of usage patterns, such as which assets communicate with each other, when this occurs, how they interact (e.g., network and application protocols), and how much information they pass back and forth.
  - A hunter who knows what's normal for an organization's assets and networks will be much better prepared to spot deviations from the norm. Pat is new to the company, so it's worth spending time talking with system administrators, incident responders, and other IT staff members to learn more about normal activity.
  - This can be as simple as knowing what the typical work days and hours are for people in different roles within the company (e.g., standard users, managers, developers, system administrators.)
  - This can also be complex, including gathering detailed information on which assets people in each role may access, and which applications are whitelisted (allowed) or blacklisted (prohibited) on the company's assets.
- History of Data Visualization
  - The concept of using pictures to understand data has been around for centuries, from maps and graphs in the 17th century to the invention of the pie chart in the early 1800s. Several decades later, one of the most cited examples of statistical graphics occurred when Charles Minard mapped Napoleon's invasion of Russia. The map depicted the size of the army as well as the path of Napoleon's retreat from Moscow – and tied that information to temperature and time scales for a more in-depth understanding of the event.
  - It's technology, however, that truly lit the fire under data visualization. Computers made it possible to process large amounts of data at lightning-fast speeds. Today, data visualization has become a rapidly evolving blend of science and art that is certain to change the corporate landscape over the next few years.
  - Data visualization: A wise investment in your big data future
  - With big data there's potential for great opportunity, but many retail banks are challenged when it comes to finding value in their big data investment. For example, how can they use big data to improve customer relationships? How – and to what extent – should they invest in big data?
  - In this Q&A with Simon Samuel, Head of Customer Value Modeling for a large bank in the UK, we examine these and other big data issues that confront retail bankers.
- Data Visualization
    - It makes complex data more accessible, understandable and usable. Users may have particular analytical tasks, such as making comparisons or understanding causality, and the design principle of the graphic (i.e., showing comparisons or showing causality) follows the task. Tables are generally used

3801 FRANKLIN STREET
DENVER, CO 80205
800 445 0024

where users will look up a specific measurement, while charts of various types are used to show patterns or relationships in the data for one or more variables.

- **Chronologies and timeline pitfall:** In using Timelines, analysts may assume, incorrectly, that events following earlier events were caused by the earlier events.
  - Also, the value of this technique may be reduced if the analyst lacks imagination in identifying contextual events that relate to the information in the **Chronology** or **Timeline**.
  - Heuer Jr., Richards J.. Structured Analytic Techniques for Intelligence Analysis (p. 58). SAGE Publications. Kindle Edition.
- **Matrices** are used to analyze the relationships between any two sets of variables or the interrelationships between a single set of variables. Among other things, they enable analysts to:
  - Compare one type of information with another.
  - Compare pieces of information of the same type.
  - Categorize information by type.
  - Identify patterns in the information.
  - Separate elements of a problem.
  - Heuer Jr., Richards J.. Structured Analytic Techniques for Intelligence Analysis (pp. 68-69). SAGE Publications. Kindle Edition.
- **Venn Analysis** is a visual technique that analysts can use to explore the logic of arguments. Venn diagrams consist of overlapping circles and are commonly used to teach set theory in mathematics. Heuer Jr., Richards J.. Structured Analytic Techniques for Intelligence Analysis (p. 72). SAGE Publications. Kindle Edition.
- **Venn Analysis** helps analysts determine if they have put like things in the right groups and correctly identified what belongs in each subset. The technique makes it easier to visualize arguments, often revealing flaws in reasoning or spurring analysts to examine their assumptions by making them explicit when constructing the diagram. Heuer Jr., Richards J. Structured Analytic Techniques for Intelligence Analysis (p. 73). SAGE Publications. Kindle Edition.
- **Network Analysis** has proved to be highly effective in helping analysts identify and understand patterns of organization, authority, communication, travel, financial transactions, or other interactions among people or groups that are not apparent from isolated pieces of information. Dependent on at least one good source of information Heuer Jr., Richards J. Structured Analytic Techniques for Intelligence Analysis (p. 79). SAGE Publications. Kindle Edition.
- **Mind Maps and Concept Maps** are visual representations of how an individual or a group thinks about a topic of interest. Such a diagram has two basic elements: the ideas that are judged relevant to whatever topic one is thinking about, and the lines that show and briefly describe the connections among these ideas. Heuer Jr., Richards J.. Structured Analytic Techniques for Intelligence Analysis (p. 86). SAGE Publications. Kindle Edition.

3801 FRANKLIN STREET
DENVER, CO 80205
800 445 0024

    - **Process Maps, including Gantt Charts**, are used by intelligence analysts to track, understand, and monitor the progress of activities of intelligence interest being undertaken by a foreign government, a criminal or terrorist group, or any other nonstate actor. Heuer Jr., Richards J. Structured Analytic Techniques for Intelligence Analysis (p. 93). SAGE Publications. Kindle Edition.
- There are 3 labs associated with this Hacking 101 – Data Visualization
    - DIFF using wetware
    - DIFF using words
    - DIFF using Data Visualization
- In computing, the **diff utility** is a data comparison tool that calculates and displays the differences between two files. Unlike edit distance notions used for other purposes, diff is line-oriented rather than character-oriented, but it is like Levenshtein distance in that it tries to determine the smallest set of deletions and insertions to create one file from the other.
  The diff command displays the changes made in a standard format, such that both humans and machines can understand the changes and apply them: given one file and the changes, the other file can be created.
- Typically, diff is used to show the changes between two versions of the same file. Modern implementations also support binary files. The output is called a "diff", or a patch, since the output can be applied with the Unix program patch. The output of similar file comparison utilities are also called a "diff"; like the use of the word "grep" for describing the act of searching, the word *diff* became a generic term for calculating data difference and the results thereof.

3801 FRANKLIN STREET
DENVER, CO 80205
800 445 0024