L08 Group Activity 2
Exploring Open-Source Intelligence (OSINT)
Part 2: Technique Sharing

IST 110

Group 05

Below are the brief overviews of an OSINT technique including key steps and potential use cases involved as explored by each group member in group 05.

Nicholas La Bella

The OSINT technique of **advanced searching** is a pretty simple but highly effective method that is used by security professionals, intelligence experts, and IT professionals to identify threats or acquire sensitive information. An advanced search function is called “Google Dorking”. The Google Dork queries can be used to look for specific types of information across the search engine. One example I found from a 2014 bulletin is as follows: “sensitive but unclassified” filetype:pdf site:publicintelligence.net. This query example tells Google to return any pdf document from the website publicintelligence.net that contains the words “sensitive but unclassified”.

As aforementioned, these queries can be used by many types of experts, but security professionals may utilize them more often than any other to look for a company’s vulnerabilities and then correct the issue to prevent leaks of sensitive of proprietary information.

References:

Imperva. (n.d.). *Open-Source Intelligence (OSINT).* https://www.imperva.com/learn/application-security/open-source-intelligence-osint/

Public Intelligence. (2014, August 25). *Feds Issue Bulletin on Google Dorking.* https://publicintelligence.net/feds-google-dorking/

David Leo Lenze

The technique I researched was passive recon. Unlike active recon, this technique does not directly interact with the system. Passive Recon is usually to gather OSINT intelligence. It can involve scraping publicly available websites, retrieving data from open-sourced APIs such as the Twitter API, or pulling data from deep web information sources. The data is then packaged and organized for consumption. For “penetration testers and security teams”, the goal of using ONSIT is to reveal public information about internal assets and other information accessible outside the organization. For example, useful information that can be revealed through OSINT

includes open ports, unpatched software, publicly available IT information such as device names, IP addresses and configurations, and other leaked information belonging to the organization. They can find the hole, trace it to the source and stop the leakage. The government is constantly using similar techniques for their own cyber security.

But if the bad guys can detect what is happening before they've got them then hackers can disappear without a trace. The only way to do it is by surprise. Targets are more likely to notice active scanning as an intrusion and detect any attempts to access any open ports.

Sources:

"Open-Source Intelligence (OSINT): Techniques & Tools: Imperva." *Learning Center*, 20 Sept. 2022, www.imperva.com/learn/application-security/open-source-intelligence-osint/.

"Open Source Intelligence (OSINT): Top Tools and Techniques: Upguard." *RSS*, www.upguard.com/blog/open-source-intelligence#:~:text=OSINT%20Techniques,-OSINT%20reconnaissance%20(recon&text=Passive%20recon%20involves%20gathering%20information,%2C%20Unix%2C%20and%20Linux%20systems. Accessed 21 Oct. 2023.

Sophia Kay Marvin

The technique I found was Active Collection. This technique looks for vulnerabilities in websites and servers to collect data. Active collection involves finding open ports that will let data out. The data collected is still public data, but it takes a bit more work to collect as it may be archived data or outdated. In simpler, more applicable terms, this could be something as simple as putting a friend request for a private social media account to find out more about them.

Sources

https://ntrepidcorp.com/managed-attribution/defining-active-vs-passive-osint/
https://www.imperva.com/learn/application-security/open-source-intelligence-osint/

David Reichart

The technique I decided to look more into was web scraping, which is a process of collecting mass amounts of data from the internet automatically. Though in addition to collecting the data, one must consider the dynamic nature of the web. The information scrapped today may be edited into something completely different tomorrow. As well, there's a near infinite amount of information to collect which calls for great levels of organization and adjustment of collection tools. So despite the lengths you can go to by programming a web crawling and scraping program with languages such as R and python, there's always a need for human intervention to ensure efficient data collection. The general steps involve include

analysis of a website/information source, crawling the website to retrieve the data using programmed scripts, and finally organization of collected data. (Korotov, Johnson, & Silva, 2020, pp. 540-541)

As for the use cases of such semi-automated data collection, there are many. Companies could keep track of pricing across the web to constantly re-price their own goods to stay competitive. Companies or organizations could collect opinions of their products and services without the need for traditional surveys. (harkiran78, 2023) Or - as is a hot topic in the world today - the massive amounts of data collected could be used to help train artificial intelligence models.

References

harkiran78. (2023, March 23). *What is Web Scraping and How to Use It?* Retrieved from GeeksForGeeks: https://www.geeksforgeeks.org/what-is-web-scraping-and-how-to-use-it/#

Korotov, V., Johnson, L., & Silva, L. (2020, December 10). Tutorial: Legality and Ethics of Web Scraping. *Communications of the Association for Information Systems, 47*, 539-563.

Pengfei Zhang

The technique I find that is important in OSINT is Social Media Analysis. Social media platforms have an abundance of information that can provide valuable insight into various aspects of human behavior and interests. Social media can be a starting point to gather information, like a gateway to find even more data that can then be analyzed to find trends, public opinions on events or current issues, influential individuals or groups of a particular community, geospatial analysis, verification, etc. There are many use cases that social media analysis can provide. One example is threat detection and monitoring. Analyzing social media posts to identify potential security threats and extremist activities can be particularly useful for intelligence agencies. Another example that is not related to cybersecurity or criminal offences is brand monitoring and competitive intelligence. Brand monitoring tracks and analyzes conversations about a product or brand on social media platforms. Competitive intelligence coincides with brand monitoring in that it brings insights into a company's strategies, market position, and also helps to refine your own business strategies.

References

IBM. (2022). What is social media analytics? Ibm. https://www.ibm.com/topics/social-media-analytics

Qualtrics. (2023). What is Social Media Analytics in 2022? Qualtrics. https://www.qualtrics.com/experience-management/research/social-media-analytics/