TELNET SCAN

#Ivan Lukman s1967754, #Metin A. Açıkalın s1984853,
#Valentine Legoy s2178192

1. **What security issue does the data speak to?**

   "Telnet is a client-server protocol, based on a reliable connection-oriented transport. Typically, this protocol is used to establish a connection to Transmission Control Protocol (TCP) port number 23, where a Telnet server application (telnetd) is listening. "

   In provided csv file, telnet scan results from various ip addresses in The Netherlands is included. The details of information types that are recorded are timestamp, ip, protocol, hostname, tag, asn, geo, region, city, naics, sic and banners.

   The data speaks about a specific time interval:
   [Mon, 01-01-2018 15:09:36 GMT ; Thu, 30-08-2018 19:01:53 GMT]

   The important issue in this dataset is leakage of information about critical systems in the banners section. To illustrate, in record 2018-01-01 15:11:14 Preventative Protection, Detection, Investigation, & Resolution system's "brand" was leaked and in record 2018-01-01 15:15:04 the system versions' of server and running application was provided before successful login. Those informations could lead to exploitations according to vulnerabilities of brands or system versions.

2. **What would be the ideal metrics for security decision makers?**

   We can identify three types of security decisions makers level: organizational (local), governmental (national), and both (global). Telnet scans' related issues are usually more of a concern for the decision makers at an organizational level as governmental level decision makers would mostly need metrics in case they want to create policies.

At the organizational level, multiple possible ideal metrics can be listed based on the four key metrics defining their level of security:

- Controls
  - Restrictions on the port numbers
  - Restrictions on the origin of the scan
  - How often and how many systems have been tested
- Vulnerabilities
  - What kinds of informations are accessible through telnet banners and how those informations can be used to exploit the system
- Incidents
  - Number of attacks which were facilitated by the information retrieved from a telnet scan.
- Losses
  - The cost of the damages resulting from the attacks relying on the information accessible through telnet scan

For the policymakers, the previous metrics could be interesting on a more global level (not for a precise organization). In addition, it would be interesting for them to know how popular these scans are compared to other techniques, such as SSH and DNS, and who are usually the targets of these scans.

3. What are the metrics that exist in practice?

Information found from:

- https://duo.com/decipher/mapping-the-internet-whos-who-part-three
- https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=935044
- https://www.jstage.jst.go.jp/article/ipsjjip/24/3/24_522/_pdf/-char/ja

No	Metrics	Definition	Decision-making level
1	Default Credentials Usage	Attempts to brute-force proxies	Local
2	Message that indicates the command shell	$,#,: etc. gives lots of information about system used	Local

3	Increase in IoT Devices	Increase in average scanning sources per day	Global

According to https://www.usenix.org/system/files/conference/usenixsecurity14/sec14-paper-durumeric.pdf, it is clearly seen that telnet scans are not commonly used nowadays. In table 1 of the paper which indicates Commonly targeted services for small scans (targeting <10% of the IPv4 address space) telnet scans form 2.8% of total scan types. It was also really hard to find security metrics of telnet scans in real life situations.

4. A definition of the metrics you can design from the dataset

No.	Metrics	Definition	Decision-making level
1	System information	Amount of information about the system (for example, system version, operating system, etc.)	Local
2	Number of scans per city	The number of telnet scans identified per city	Local
3	Number of scans per region	The number of telnet scans identified per region	Local
4	Number of scans per hostname	The number of telnet scans identified based on hostname	Local
5	Number of scans per IP address	The number of telnet scans identified based on IP address	Local
6	Number of scans which failed	The number of telnet scans which refused access to the person performing the scan (in some cases, the reason why is included)	Global/local
7	Types of informations which could be used to	The different information which could be used to	Global

	perform an attack	perform attacks regrouped by type of information	
8	Evolution of the protection against telnet scan for each IP	Most IPs are target of several scans during the time period these data were gathered. In some case, we'll be able to see different information in the banner over time.	Local
9	Change in banner messages for each IP	Change of banner messages might not be always related with increase of security and can show changes in company network	Local

5. **An evaluation of the the metrics you have defined. This should include graphical representations of the metrics (e.g., histograms, scatter plots, time series, bar charts).**

   -In final Report-

6. **Methodology**

   - Since the CSV file contains more than one million rows, for convenience and speed, it can be opened using RStudio.
   - Analyze ip, hostname, region, city, and the most importantly "banners", for example, finding the number of rejected telnet connections based on region.