Final Deliverable 1 - Security Metrics

Group 2 - TelnetScans

Ivan Lukman - S1967754

Metin A. Açıkalın - S1984853

Valentine Legoy - S2178192

Economics of Security

University of Twente

September 24, 2018

1 Introduction

The Request For Comments (RFC) 854 define the TELNET connection as "a Transmission Control Protocol (TCP) connection used to transmit data with interspersed TELNET control information" [1]. The Telnet protocol allows the communication in-between terminals or processes, but mainly it allows "a standard method of interfacing terminal devices and terminal-oriented processes to each other" [1].

Even though there are many similar services available today such as SSH and Rlogin, Telnet is still an important service in providing remote access. However, a standard way of ensuring privacy and integrity of Telnet session has been lacking [2]. Telnet service is still one of the internet vulnerabilities that creates a backdoor for attackers [3].

Although most Telnet ports are protected with authentication, such as username and password combination, unencrypted remote terminals often display the running system's name, model and sometimes even versions before the login prompt. This leads to leakage of important information about critical systems.

Besides this, allowing telnet port to be accessible from the internet makes the vulnerabilities even easier to exploit. The main purpose of this study is to dig in to the security of open and reachable telnet ports from the internet by defining suitable security metrics for them. Further in this paper, some unique examples of vulnerabilities are going to be discussed as well.

1.a Data Used in Assessment

Telnet scan data which obtained from Shadow Server Organization was used to conduct this study. According to organization, the aim of this scan is "to identify openly accessible systems that have the Telnet service running and report them back to the network owners for notification" [4]. Unfortunately, since the data can also be accessed publicly through the internet, and the fact that people can also conduct a similar telnet scans around the world, hackers can misuse these data.

The details of information types that are recorded between dates [Mon, 01-01-2018 15:09:36 GMT ; Thu, 30-08-2018 19:01:53 GMT] from various IP addresses in the Netherlands are listed and explained as follows:

- *Timestamp*: Date and time of each scan.
- *IP*: Internet Protocol Addresses of scanned TELNET ports.

- *Protocol*: Protocol used in scanning. For the data in this study, it only includes TCP.
- *Hostname*: Hostname of the machine that was reached.
- *Tag*: Tag of the search. For the data we have, it is either telnet or telnet-alt.
- *Asn*: Autonomous System Number of the scanned IP's. ASN is an identifier for a collection of IP networks and routers under the control of one entity.
- *Port*: The port used in the scan. For the data we have it is either 23 or 2323.
- *Geo*: Geographical code of the country where the search was done. For the data we have it is only NL.
- *Region*: Region of which the IP address belongs to.
- *City*: Name of the city where the IP address belongs to.
- *Naics*: North American Industry Classification System number.
- *Sic*: Standard Industrial Classification number.
- *Banners*: Welcome messages of the CLIs (Command Line Interface) of successful TELNET connections.

Obtaining valuable information from the recorded data will be discussed in Methodology And Results section of this study.

2 Metrics

2.a Related Work

In practice and in literature, several metrics have been defined to evaluate the level of security of the studied issue. These existing metrics will be discussed in this section.

If any official regulations could not be found, most organizations using Telnet services have regulations and policies to make their system more secure [5]. The use of Threshold Random Walk (TRW), which helps to identify malicious remote hosts doing scans, is also frequent [5].

Organizations, such as rapid7 [6] or Shadowserver [7], also scan the Internet in search of open telnet ports. Their report offers information such as the number of devices with Telnet accessible over the world [4] [8], but also the percentage of Telnet device used, compared to the use of SSH per country [9]. The scans done by Carna Botnet could also identify the numbers of devices which, additionally to be accessible remotely via telnet, were still using the default credentials [10].

Other metrics also made available by studies about IoT or the Internet in general. The authors of the paper "IoTPOT: Analysing the Rise of IoT Compromises" [11] analyzed results from NICTER [12], which is a Japanese darkweb monitoring system, to find the importance of Telnet scans and the types and proportions of each attacking hosts. A. Cui and S. J. Stolfo describes in "A Quantitative Analysis of the Insecurity of Embedded Network Devices: Results of a Wide-Area Scan" describes the most vulnerable devices by geographical location, functional category, and brand, based on a scan to access Telnet and HTTP ports [13].

2.b Ideal Metrics for Data Used

Being able to access services through Internet using Telnet is an important issue. Therefore it is necessary for any security decision maker to have access to reliable metrics, which can help them to know more about the security issue. In the studied case, several levels of security decision makers can be identified:

- Local, which includes sysadmins and organization decision makers.
- National, regarding governments willing to install policies for their whole country.
- Global, which includes both local and national level as well as being international, for example, the Internet Society (ISOC).

For each of these types of decision makers, it is defined ideal metrics, which are part of at least of the four key metrics defining the level of security: control, vulnerability, incident, loss [14].

On a global level, it would be ideal to know the number of systems with telnet services accessible. Having the numbers per region of the world, but also the proportion compared to the use of SSH. This would help to define how to adapt the regulations and the education on the issue for these regions. It is also important at this level, to be able to know which attacks have been helped by the use of Telnet, and what were the resulting losses

of these attacks. By comparing them with the same kind of information for attacks not using Telnet, it would help to define how big of an issue this is and if solving it is a priority compared to other issues.

All these previous metrics would also be precious at a governmental level to adapt their regulations by comparing themselves to other countries, which have of Telnet a lesser security issue.

At a local level, the ideal metrics are more numerous. It would be interesting for them to know:

- Which and how many systems have services accessible through telnet.
- Which and how many systems have preventive measures deployed, and which are those. For instance, whether the machines are deployed with authentication mechanisms, such as username-password combination. Another example is, whether the machines' accessibility is limited in terms of physical, local network or remote access. Furthermore, are there any limitations on the number of login attempts the service allows or the number of telnet sessions which can be active at the same time and how long can a connection lasts in an idle state.
- What is the number of services that are making the system vulnerable by evaluating which of them uses default settings, such as default login credentials. Moreover, evaluating how many of them have banners giving information about the system which could be exploited by attackers.
- How many systems have been attacked so far due to Telnet services being accessible from the internet, but also what were the losses resulting from these attacks.

3 Methodology and Results

3.a Methodology

After obtaining the data from ShadowServer, opening the csv file using several options have been experimented, namely Microsoft Excel, CSView, and RStudio. Unfortunately, the data was too large to be opened with Excel. In addition, CSView only provides access to view the data and does not allow filtering or analyzing options. Therefore, RStudio was the most suitable option to conduct this study. To be able to analyze all data, readr library of R needed to be downloaded, otherwise RStudio was only showing the first few hundred thousands line of the data.

The first thing to focus on after the complete data has been successfully loaded is identifying what security issues this dataset was containing. Just by directly looking at the data, several rows pose security problem, which is leaking the critical information of the system in the telnet banner. One of the example is "|MikroTik v6.11|Login:". Since the firmware version of the MikroTik device is quite old, bugs exploitations are available in the internet. This means everyone from internet can try attacking this device.

Further analysis was conducted about this dataset, about why this data was published in the first place. According to data provider ShadowServer's idea, this is because there are still a lot of devices that reply to Telnet scans. This means, the devices are open for configurations remotely through the internet. Therefore, organizations like ShadowServer exist and they attempt to make the Internet more secure.

Based on the two main problems of this dataset, several metrics can be defined into the dataset, as can be seen at Table 1.

Since the goal of the metrics were clear, R scripts were written to be executed using the complete data. To speed up the testing phase, we first selected the first 10000 rows of the data, because executing some minor changes in R scripts was very expensive in terms of resources and time, in all three computers we used. After the query gave the desired output, it was then applied into the complete dataset. The results of the security metrics will be discussed in the next section.

No.	Metrics	Definition	Decision-making Level
1	The number of telnet sessions require authentication during the entire scan period.	This metrics quantifies how many rows from the telnetscan data shows authentication process, for instance banners that contain “password” or “login”.	Local
2	The number of rows rejected the telnet attempts.	This metrics measures the number of records declining telnet session requests. For instance, error messages that are displayed when the session limit is reached or when the telnet session is initiated using a wrong port number.	Local
3	The number of telnet sessions that publish the system information.	The number of rows from the dataset that contains the system information in the banner is calculated using this method.	Global
4	The number of telnet sessions that release the whitelisted IP addresses.	Based on “System administrator is connecting from \textit{IP address}” in the banners, the IP addresses that are allowed to access a system can be identified.	Global
5	The number of telnet sessions that are accessible, grouped by each type of industry.	The main goal of this metric is finding which industry is most openly accessible, based on SIC and NAICS.	Global

Table 1: Table of Applicable Metrics.

3.b Results

The overall result of executing the aforementioned metrics can be seen at Figure 1 below. The R scripts can be found at our GitHub repository here [15]. Out of 5102781 telnet records, the results of the security metrics are as follows.

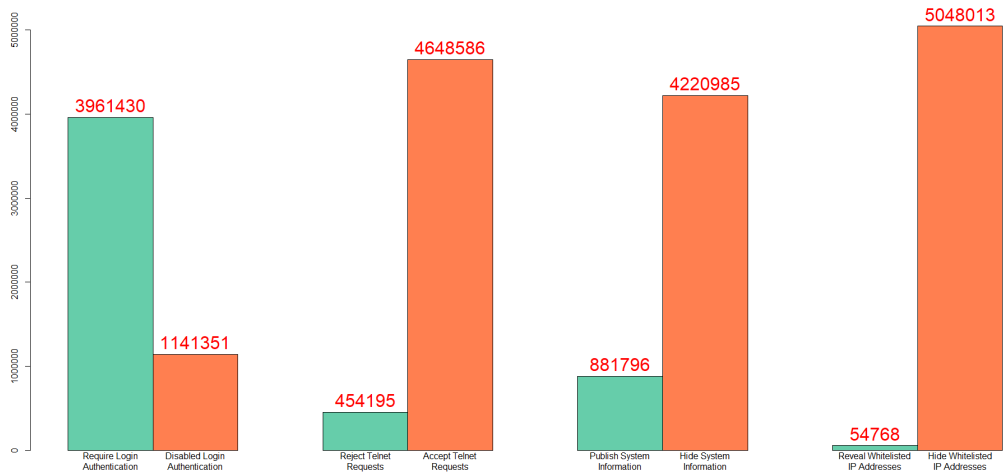

Figure 1: Bar Chart of Metrics Results.

1. The number of telnet sessions require authentication during the entire scan period.
   3961430 telnet sessions require authentication. Script auth.txt contains the R script used to capture this value. To capture the number of telnet sessions that enable authentication, telnet banners can be used, since in most of the login phase cases, telnet banners will contain "account" or "password".

2. The number of rows rejected the telnet attempts.
   454195 telnet requests were rejected. Script rejected.txt contains the R script used to capture this value. To capture the number of telnet data that rejected session requests, the banners can also be used. For instance, telnet banners that show "too many active sessions" shows that telnet communication was rejected.

3. The number of telnet sessions that publish the system information.
   881796 telnet sessions published system informations. Script sysinfo.txt contains the R script used to capture this value. To capture the number

of telnet data that reveals the system information, the regular expression used in the R script turns to be more complicated. For example, just by looking for "SSH", does not mean that the corresponding device does actually have SSH service running.

4. The number of telnet sessions that release the whitelisted IP addresses. 54768 telnet sessions published whitelisted IP addresses. Script ipaddress.txt contains the R script used to capture this value. To capture the number of telnet data that shows the whitelisted IP addresses, the regular expression in the R script can be written to look for "connecting" and IP addresses.

5. The number of telnet sessions that are accessible, grouped by each types of industry.
   The result for this metric can be seen at Figure 2, which shows the

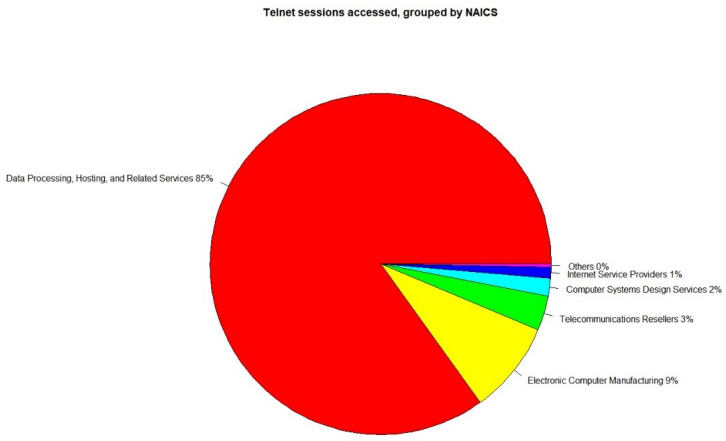

Figure 2: Telnet Sessions Accessed, Grouped by NAICS.

NAICS result, and Figure 3, which shows the SIC result, below. In both case, we ignored the session for which the NAICS, or SIC respectively, was not given. We grouped the NAICS, or SIC, having a number of scans inferior to 10,000 in the "Others" category. Based on NAICS and SIC, the number of accessible IP addresses using Telnet is dominated by Data Processing, Hosting, and related services, being the large majority of the overall telnet records, based on NAICS. On the other hand, based on SIC, Internet Service is in majority, while Data Processing service

is second. Combined, these services represent the same proportion as the one found with the separation made by NAICS.

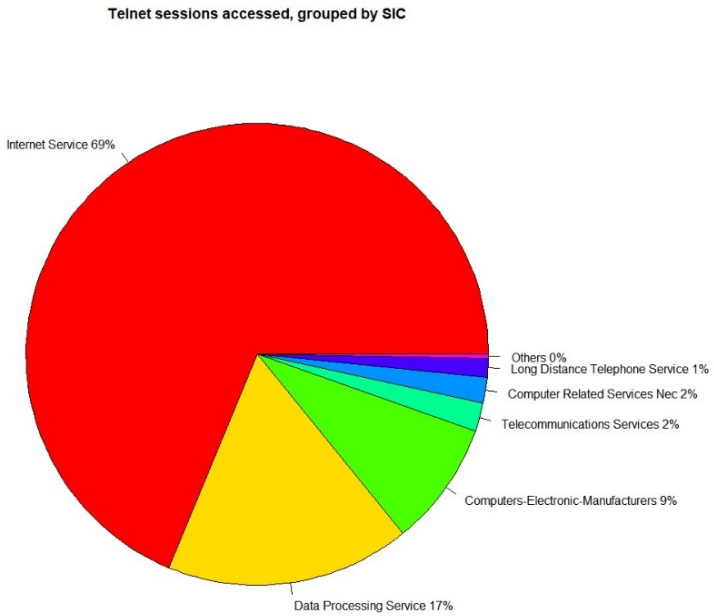

Figure 3: Telnet Sessions Accessed, Grouped by SIC.

4 Conclusion and Future Work

To conclude, in this research, it is successfully identified the security problems which Telnetscan report contains. These problems concern the fact that a lot of devices published their system information into the Telnet login banner, and that these devices were accessible remotely through the internet through an open Telnet port. Based on the security threats, security metrics were identified, which discusses about the ideal, the existing, and the applicable security metrics using the Telnetscan report. Then, the results of executing the applicable security metrics were shown.

However, due to limitation of time for the study, R query might not be the optimal when analyzing the dataset. One of the cause is the data contains too much varieties in the Telnet banner, and it was quite hard to provide the analysis with a high accuracy. Therefore, as the future work, it is hoped

that the queries could be improved to provide a deeper insight of the data. For example, analyzing unique device ID, such as "VMG8324-B10A", which corresponds to a ZyXEL wireless router.

Based on the security metrics we have identified, we hope that this will help decision makers to improve their devices security, to illustrate, in their security investment and management. By understanding the threats and decision makers, it can be taken a proper security countermeasures to fix the security issues. For example, Chief of Information Security Officers (CISOs) can train their SysAdmins to replace the usage of Telnet with a more secure channel, such as SSH.

References

[1] Telnet protocol specification. `https://tools.ietf.org/html/rfc854`. Accessed: September 24, 2018.

[2] Transport layer security protocol in telnet. `https://ieeexplore.ieee.org/document/1274255/`. Accessed: September 24, 2018.

[3] Haslina Binti Mahmood. Transport layer security protocol in telnet. In *Communications, 2003. APCC 2003. The 9th Asia-Pacific Conference on*, volume 3, pages 1033–1037. IEEE, 2003.

[4] The shadowserver foundation: Telnet scanning project. `https://telnetscan.shadowserver.org/`. Accessed: September 24, 2018.

[5] Hwanjo Heo and Seungwon Shin. Who is knocking on the telnet port: A large-scale empirical study of network scanning. In *Proceedings of the 2018 on Asia Conference on Computer and Communications Security*, pages 625–636. ACM, 2018.

[6] Accelerate security, vuln management, compliance | rapid7. `https://www.rapid7.com/`. Accessed: September 24, 2018.

[7] Shadowserver foundation. `https://www.shadowserver.org/wiki/`. Accessed: September 24, 2018.

[8] Bob Rudis. National exposure index.

[9] 2017 national exposure index map | rapid7. `https://www.rapid7.com/data/national-exposure/2017.html`. Accessed: September 24, 2018.

[10] Internet census 2012. https://web.archive.org/web/20151013010243/http://internetcensus2012.bitbucket.org/paper.html. Accessed: September 24, 2018.

[11] Yin Minn Pa Pa, Shogo Suzuki, Katsunari Yoshioka, Tsutomu Matsumoto, Takahiro Kasama, and Christian Rossow. Iotpot: analysing the rise of iot compromises. *EMU*, 9:1, 2015.

[12] Masashi Eto, Daisuke Inoue, Jungsuk Song, Junji Nakazato, Kazuhiro Ohtaka, and Koji Nakao. Nicter: A large-scale network incident analysis system: Case studies for understanding threat landscape. In *Proceedings of the First Workshop on Building Analysis Datasets and Gathering Experience Returns for Security*, pages 37–45. ACM, 2011.

[13] Ang Cui and Salvatore J Stolfo. A quantitative analysis of the insecurity of embedded network devices: results of a wide-area scan. In *Proceedings of the 26th Annual Computer Security Applications Conference*, pages 97–106. ACM, 2010.

[14] Rainer Böhme. Security metrics and security investment models. In *International Workshop on Security*, pages 10–24. Springer, 2010.

[15] Github - metinacikalin/economics-of-security. https://github.com/metinacikalin/Economics-Of-Security. Accessed: September 24, 2018.