Final Deliverable 3 - Actors and Security Strategies

Group 2 - TelnetScans

Ivan Lukman - S1967754

Metin A. Açıkalın - S1984853

Valentine Legoy - S2178192

Economics of Security

University of Twente

October 22, 2018

Contents

1 Introduction **1**

2 Countermeasures of actors involved in the security issue **1**
2.a Problem Owner 1
2.a.1 Example of possible countermeasure 1
2.a.2 Distribution of Costs and Benefits 2
2.a.3 Actor's incentives 3
2.a.4 Externalities of the countermeasure 4
2.b Government 5
2.b.1 Example of possible countermeasure 5
2.b.2 Distribution of costs and benefits 5
2.b.3 Actor's incentives 6
2.b.4 Externalities of the countermeasure 6
2.c Non-profit Organizations 6
2.c.1 Example of possible countermeasure 7
2.c.2 Distribution of costs and benefits 7
2.c.3 Actor's incentives 7
2.c.4 Externalities of the countermeasure 8

3 Explanation of the metrics influencing the security performances **8**
3.a Security Metrics 1: Usage of SSH on port 23 9
3.b Security Metrics 2: Rejection of Telnet sessions 10
3.c Security Metrics 3: Authentication banners 11
3.d Security Metrics 4: System information leakage from the banners 13

4 Influences on the security performance of owners of services using Telnet **15**
4.a Identification of different factors causing variance in security performance 15
4.b Statistical analysis to explore the impact of the factors on the metric 16
4.b.1 Type of industry 16
4.b.2 Location (City) 18

5 Limitations and Conclusion **19**

1 Introduction

In the two previous assignments, the security issue that the telnet protocol represents has been studied through related metrics and security investments. The whole project has been written based on a dataset from ShadowServer giving various information about IP address scanned on their port 23 or 2323, from the location of the host and the industry of the owner, to the banner grabbed in the scanning process.

The following report will give first details about the security strategies the three different actors can adopt: the network administrators of the companies owning the services using Telnet, governments and non profit-organizations studying this issue. For each of them, an example of a countermeasure, the details of the costs and benefits, as well the incentive and the externalities, they might bring, will be explained.

The second part of this paper will be dedicated to the different factors which might impact the security performances of the companies owning services running on Telnet. After reminding the different metrics linked to the security performances, the factors impacting the use of SSH instead of Telnet will be described and two of them will be evaluated to confirm the relationship between the two.

2 Countermeasures of actors involved in the security issue

2.a Problem Owner

As mentioned in the previous report, the problem owner could be referred to Chief of Information Security Officers (CISOs) and individual network administrators. For this report, the problem owner is focused on the CISOs, because the cyber risk within a company where a CISO is working, is relatively much more significant compared to the risk facing the individual environment.

2.a.1 Example of possible countermeasure

One of the effective and direct ways to improve cybersecurity, especially to mitigate the security issue raised by leaving telnet ports accessible from anywhere in the internet, is by creating internal regulations for the Information Technology (IT) staff to limit access to the systems as narrow as possible, for example by only opening the ports to whitelisted IP addresses. Creating

the policy should result in direct modification of the whitelisted access and closure of the openly-accessible Telnet ports. In addition, whenever a Telnet port is opened, there should be a policy to periodically evaluate whether the port is still in use or should be closed.

By having a higher standard of security, a company can be more resistant to cyber threats. To illustrate, by taking the highest frequency of appearance of domain names in the dataset (kpn.net), one of the largest mobile telecommunication company, KPN, can be used as a reference. Although measuring the success of this countermeasure seems tricky from the dataset, the possibility is still there. This is because, if a Telnet port is accessible from an IP address at a specific time, then it does not appear again after that time, some possibilities might occur, such as:

- The change of system configuration, including changing the IP address, closing Telnet port, or shutting down the system.
- The second option would be, the organization that owns the openly-accessible system through Telnet asks ShadowServer to exclude their systems from Telnet-Scan results.
- In the worst-case scenario, as also mentioned in [1], attackers might gain access into the system, and then close the port for all other attackers. This means that, these attackers could also configure the device such that the port is not scannable by ShadowServer.

For example, a server with IP address '145.129.217.87' last appeared at 2018-04-01. This does not necessarily mean that KPN must have closed the telnet port. However, this could be possible.

In this report, to mitigate the security issue as mentioned in the previous submission, one of the concrete countermeasures is by directly restricting or closing the openly-accessible Telnet port through internal company regulations. Since the access to the system is more restricted, the security issues could be mitigated, which will make the company less fragile to malicious attacks that exploit Telnet vulnerabilities.

2.a.2 Distribution of Costs and Benefits

In brief, the distribution of the cost of creating and implementing the regulation (modifying the access through Telnet port of a system) can be estimated as listed in Table 1. On the other hand, the benefits can be estimated as listed in Table 2.

Cost	Ratio
Creating Access Policies	5-10%
Direct Closure of Telnet Ports	70-80%
Periodic Evaluation of Opened Ports	10-25%

Table 1: Cost Distribution for Problem Owners.

Benefits	Value
Improvement in reputation	Reputation
Private files are stored more securely	The estimated price of each asset kept private
Saving of work hours and labour price (compared to if an incident happens)	Estimated time, human resources, and price of conducting repair and maintenance of the compromised systems.

Table 2: Benefit Distribution for Problem Owners.

Since the types of values of the benefits are different from one another, it is quite difficult to estimate the distribution of the benefits. Also, the values of some benefits, such as the price of private files, could vary in a wide range.

2.a.3 Actor's incentives

By definition, "incentive" is referred to a thing that gives motivations and encouragement to do something. This could be in a form of money, such as prizes or bonuses, or other types of incentives as well, such as appreciation or more-challenging duties, as also mentioned in [2]. In this report, incentives will be referred mostly as the monetary rewards to motivate the actors to countermeasure the security issues. Besides, incentives could be delivered within a company, or from external parties.

Since KPN was privatised in the 1990s, the company received no funds from the government. However, there are several ways where the CISO of KPN could still get some incentives, such as:

- Internal incentives.
  Profit sharing. When KPN receive annual profits, some percentage of this could be used to fund the training and assessment program.

- External incentives.
  Certifications and awards received from external parties, such as privacywaarborg [3]. Receiving this award, to some extent, could motivate the CISOs to maintain and improve the cybersecurity of KPN. Also, since some of this kind of awards and certifications will be held annually, these CISOs will be also motivated to win the upcoming editions. This means that the appreciation and the money received from winning the awards could be used to fund the training and assessment program.

2.a.4 Externalities of the countermeasure

As explained in the video lecture, there are two types of externalities, namely the positive (create benefits to third parties) and the negative externalities (create harm to external companies). In general, having lots of openly-accessible Telnet ports could bring both sides of externalities, because:

- Positive Externalities.
  Educational or research institutions in the field of cybersecurity could be benefitted from this security issue. By deploying a large number of systems that can be accessed from anywhere in the world using Telnet ports, the CISOs must have also predicted and calculated the possible risks and the worst-case scenarios. Taking the other perspective, a large number of openly-accessible devices also means that educational institutions could extract useful informations to further improve the security of these devices in the real-world scenarios.

- Negative Externalities.
  Having a lot of systems that can be accessed openly from all over the world means that the chance of hackers taking over these devices are higher. Once they gain access to these machines, this could lead to cybercrime acts, such as using the machines as Distributed Denial-of-Service (DDoS) botnets. This means other companies are in danger of being DDoS victims as their services could be inaccessible from legitimate users.

Considering the possible positive and negative externalities around the security issue, the openly accessible Telnet ports tend to bring negative externalities, compared to the positive ones.

2.b Government

As mentioned in the previous assignment, "governments" are among thes actors with a positive impact on the security issue. They usually work with agencies, such as NIST (National Institute of Standards and Technology) and ENISA (European Union Agency for Network and Information Security), created in order to improve network and information security, by evaluating security issues and developing appropriate guidelines. However, these are mostly the law makers and enforcements side of the government who we'll be able to create effective countermeasure for the security issue that the use of Telnet represents.

2.b.1 Example of possible countermeasure

The general countermeasure that governments could take in this situation is to create a set of law on the protection of the ports of a system. This law could forbid the use of the telnet protocol over the Internet, except if telnet sessions use encrypted channels and are limited to specific authorized devices. Such a directive would rely on already existing official recommendations such as in [4] and [5]. It would however force the use of this good practice.

If the creation of a law to counter the Telnet issue seems excessive, it is important to remember that the Mirai attack, which used Telnet to propagate, ended up representing a large loss for some companies [6]. Moreover, the use of other secure protocols, such as SSH, instead of Telnet has been on the rise for the last few years [7]. Thus, such a law would simply reinforce a current phenomenon to a serious issue.

2.b.2 Distribution of costs and benefits

The costs of the creation of this law by government mainly lie mainly in the monetary costs and time consumption that the creation of a law represent, as well as the verification of its application. The respect of this law would represent a cost also for the companies owning systems with services accessible via Telnet. They would have to pay for the transitioning to a more secure system. The non-respect of this law would, however, represents a much larger cost. At the same time, because it would result in a fine, but also to potential lawsuits, for instance, if the law was not respected and an attack exploited this issue, or a loss of reputation. Additionally, all companies would have costs related to this countermeasure. Each of them would have to perform verifications, generally from an external organization, that they are complying with the law.

The creation of a law, most likely, ensure the enforcement of the security countermeasure which will benefit the government, but also the companies following this countermeasure and the clients of these companies. For the government, the benefits lie in the costs it would suffer of, if the attacks, the law is preventing from, were to happen. The main benefits are economic, as a country where all organizations follow good cybersecurity standards will not suffer from the economic costs of potential attacks. A country with advanced security policy has overall a better branding. A country which seems to be more targeted of cyber attacks is less attractive to companies, which will be more likely to set up themselves in a country where cybersecurity is promoted. Companies would benefit, as they would be more secure and less prone to be attacked, thus avoiding all the costs it would represent. And as a result, the client of these companies would have access to more secure services from them.

2.b.3 Actor's incentives

Unless an attack related to telnet scan targets the government, this actor is not directly impacted. However, the fact that some of the companies or private organization might be attacked will have an economic impact on the whole country. Thus it should encourage governments to create such a law.

The incentive for companies to follow this law should be mainly the overall security benefit it represents. However, the threat of a fine and the resulting costs the non-compliance with a law might represents is also an incentive for the companies to follow this countermeasure.

2.b.4 Externalities of the countermeasure

The creation of this law produces mostly positive externalities, as they are mostly created to improve the security of the organizations. If its application represents a cost, it cannot be seen as a negative externality as the benefits is supposed to be superior to it.

The main negative externality, however, might be an unlikely but potential need for a company to use Telnet on the internet without devices limitations or use encrypted channels.

2.c Non-profit Organizations

The telnet scan CSV file was created by Shadowserver, which is a non-profit organisation working to understand and help put a stop to high stakes cybercrime in the information age [8]. The organisation is trying to supply

timely and relevant information to the security field [9]. In their goals, they indicate that Shadowserver wants to work closely with ISPs, Hosting and DNS providers in the identification and mitigation of botnets and malware propagation.

2.c.1 Example of possible countermeasure

As those kinds of organisations work by volunteering basis, all the participants of those organisations are engaged with the goals of their organisations. In telnet scan, they are already working to provide information to problem owners to inform that they have unsecured telnet ports accessible from the internet. They can move forward to increase the impact of providing security and make a "whois" search on the IP addresses they found to gather information about the owner of the address and send mail about the discovery to inform the problem owners.

2.c.2 Distribution of costs and benefits

As shadow server is a non-profit organisation [10], there will be no costs of deployment for any problem owner or any actor. Only costs will be the labour of volunteer workers of the organisation and the technical tools they already bought to make scans and store the scanned data daily. By deploying this method, each problem owner's system administrators or CISOs will recognise if they are not aware of the situation. On the other hand, if they know the situation and they are not doing anything with it, this leakage of information might push them to do something with it.

2.c.3 Actor's incentives

No company works for losing money and more or less each company tries to mitigate the risks of losing money. The incentives can be grouped section as follows.

- Internal Incentives.
  In the data that was analysed, there were huge companies such as KPN and Infopact. They have lots of confidential information about their clients. Risking them by leaving a trapdoor to their internal network can be costly in case of an exploit. Thus they have internal incentives to fix openly accessible telnet port issue.

- External Incentives.
  When an IP is listed as it has an openly accessible telnet port, this can

cause loss of reputation. Especially in the data that was analysed, the majority of problem owners were companies related with technology. Loss of reputation by a security-related issue might create significant impacts on such companies. Thus, they will be interested to fix it if they see a mail from Shadowserver about the issue.

2.c.4 Externalities of the countermeasure

- Positive Externalities.
  As Shadowserver is based on voluntariness, it is an excellent place for a security enthusiast to improve his/her skills. As those kinds of security issues keep existing, it will always create an area of improvement for security specialists.

- Negative Externalities.
  Disclosure of IP addresses of openly accessible telnet ports may lead more attackers to use this information against problem owners and choose them as one of their targets. However, any IP can be removed from scanning list of shadow server if the owner of the IP sends an e-mail to request his/her IP address to put into the whitelist. By this way, scan excludes that IP address.

3 Explanation of the metrics influencing the security performances

In order to make the analysis of the factors influencing the security performance more understandable, four metrics, extracted from the dataset and related to the security performances of the security issue owner, will be explained in the following part.

First of all, for all of the metrics that were discussed, the top 10 hostnames based on the dataset where the security metrics can be applied is listed. Then, the total number of scans based on each hostname rank is calculated from the whole dataset. To normalize the security performance, from the top 10 list for each security metrics, the number of host names that fulfil the security metrics is divided by the total number of scans during the time period of the dataset. Specifically for each security metrics, this can be seen as follows.

3.a Security Metrics 1: Usage of SSH on port 23

In Figure 1, the data shows the top 10 domain names that was scanned most in the entire dataset. The percentages of ssh usage on port 23, which can be seen in sshPortNb column, are all 0% except kpn.net at 0.0065% which is smaller than even 1%. Percentages are calculated as the ratio between banners containing information that SSH is running on port 23 and the total number of scans for each domain name. This reflects the fact that domain names which scanned the most stayed vulnerable because no improvement was done on security performance regarding migrating port to run SSH.

domainname	totalScanNb	sshPortNb
routit.net	438413	0.000000000
kpn.net	413474	0.006563895
ziggozakelijk.nl	282611	0.000000000
ziggo.nl	245564	0.000000000
xenosite.net	213835	0.000000000
solcon.nl	153053	0.000000000
chello.nl	96416	0.000000000
xs4all.nl	90049	0.000000000
infopact.nl	85135	0.000000000
prioritytelecom.net	74903	0.000000000

Figure 1: Percentages (sshPortNb) of most scanned domain names which run SSH on port 23.

As the domain names that were scanned the most during the entire scan period results were unsatisfying, it was also conducted another analysis in Figure 2. This time the dataset was first ordered by the domain names that had the maximum number of banners that indicates SSH is running on port 23 and then calculated the ratio between the number of total scans on the domain names and the number of banners which indicates SSH is running on port 23 to show the security performances. The results were opposite of the results of the previous analysis. This time, top 10 SSH running domain names on port 23, the ratio explained above was 100%. This shows that the companies who are aware of the security issues show a stable security performance.

domainname	totalScanNb	sshPortNb
zeus.pm	1595	1
phpwebhosting.com	1027	1
dnazone.net	975	1
growingminds.nl	910	1
achos.nl	694	1
hostje.nl	618	1
miamo.nl	618	1
coloprovider.nl	603	1
aliensonacid.nl	429	1
pure-isp.eu	421	1

Figure 2: Percentages (sshPortNb) of highest SSH usage per domain names.

The differences between the top 10 domain names regarding SSH usage on port 23 might be because of each company's decision of running SSH in their prefered port.

3.b Security Metrics 2: Rejection of Telnet sessions

domainname	totalScanNb	rejectScanNb	ratio
routit.net	438413	*NA*	*NA*
kpn.net	413474	38645	0.093464160
ziggozakelijk.nl	282611	*NA*	*NA*
ziggo.nl	245564	*NA*	*NA*
xenosite.net	213835	*NA*	*NA*
solcon.nl	153053	*NA*	*NA*
chello.nl	96416	*NA*	*NA*
xs4all.nl	90049	*NA*	*NA*
infopact.nl	85135	*NA*	*NA*
priorityтelecom.net	74903	*NA*	*NA*

Figure 3: Top 10 most scanned domain names which reject the telnet port connection.

In Figure 3, data shows the rejection of connections in the top 10 most scanned domain names. It can be clearly seen from the table that, other than

kpn.net which has nearly 1% of rejected connections according to banners, rest has never rejected any connection. This shows, in the entire scan period, no improvements on the rejection of connection was done.

To be able to further explain the data also for this metric, another analysis was conducted. This time the top 10 domain names which have the most rejection ratio. To be able to obtain this data, some specific keywords were searched in the banner. Afterwards, the ratio between the total number of scan and the scans which have rejection keywords was calculated. The results which can be seen in Figure 4, show that there are 6 hostnames in top 10 hostnames which were scanned significant amount of times and gave more than 78% of rejection. This shows their security performances are improved over time or was already in the highest position for the ones who had 100% rejection.

domainname	totalScanNb	rejectScanNb	ratio
kpn.net	413474	38645	0.093464160
by.pcextreme	2286	1794	0.784776903
tele2.net	1763	1763	1.000000000
weserve.nl	36636	1548	0.042253521
as197156.net	2323	987	0.424881619
nova-ict.nl	977	977	1.000000000
solidbe.nl	819	819	1.000000000
vqbn.com	784	783	0.998724490
colocenter.nl	4153	623	0.150012039
clientshostname.com	578	578	1.000000000

Figure 4: Top 10 domain names that have the most telnet connection rejections.

The differences between the top 10 domain names in terms of rejection, was because of each company's security approaches.

3.c Security Metrics 3: Authentication banners

The result of evaluating this metric can be seen at the Figure 5. The table is sorted by the number of records that requires authentication (TotalAuthRequired) in a decreasing order. In this table, column TotalAuthRequired shows the number of records from the dataset that requires authentication, based on each domain name. TotalScans shows the total number of scans based on the domain names from the dataset. AuthScansRation is computed by dividing TotalAuthRequired with TotalScans per domain names.

NoAuthScansRatio is the opposite of AuthScansRatio and computed as (1 - AuthScansRatio).

Domain Names	TotalAuthRequired	TotalScans	AuthScansRatio	NoAuthScansRatio
kpn.net	349672	413474	0.8456928	0.15430716
xenosite.net	185334	213835	0.8667150	0.13328501
weserve.nl	28270	36636	0.7716454	0.22835462
as51050.net	6630	6988	0.9487693	0.05123068
xsdsl.net	5689	6319	0.9003007	0.09969932
3winfra.com	5373	7372	0.7288388	0.27116115
i3d.net	5292	24062	0.2199318	0.78006816
blackgate.nl	4772	5723	0.8338284	0.16617159
host-palace.com	4051	4720	0.8582627	0.14173729
onweerswolk.nl	3937	4025	0.9781366	0.02186335

Figure 5: Top 10 domain names that require authentication.

Visually, this statistics can be visualized in two line charts as shown in Figure 6 and 7. Figure 6 shows the comparison of TotalAuthRequired and TotalScans, and Figure 7 displays the comparison of AuthScansRatio and NoAuthScansRatio, of the top 10 domain names.

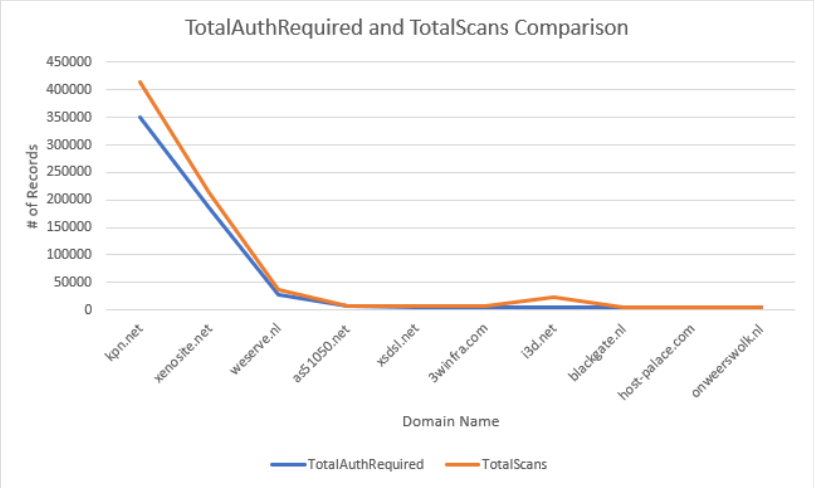

Figure 6: TotalAuthRequired and TotalScans Comparison.

The higher value of NoAuthScansRatio means that to some extent, the security of that system is better because based on the dataset, these records that contain no authentication requirements are stating that the telnet connections were rejected.

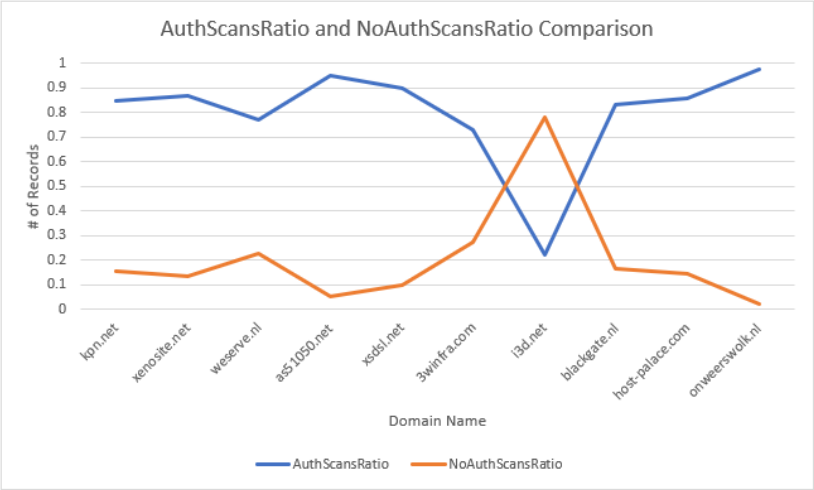

Figure 7: AuthScansRatio and NoAuthScansRatio Comparison.

The variance within the top 10 domain names that requires authentication is mostly caused by the different limit of telnet sessions each company allows. Even within an organization, the telnet limitation could differ due to different purposes of each system. For instance, this is visible for "leaseweb" company, whereby there are different total number of authentication in the telnet banners for the hostnames.

3.d Security Metrics 4: System information leakage from the banners

The result of evaluating these metrics can be seen at the Figure 8. The table is sorted by the number of records that leaks the system information (TotalLeakingSysInfo) in a decreasing order. In this table, column TotalLeakingSysInfo shows the number of records from the dataset that reveals the system information, based on each domain name. TotalScans shows the total number of scans based on the domain names from the dataset. LeakageScansRatio is computed by dividing TotalLeakingSysInfo with TotalScans per domain names.

Visually, this statistics can be visualized in two line charts as shown in Figure 9 and 10. Figure 9 shows the comparison of TotalLeakingSysInfo and TotalScans, and Figure 10 displays the comparison of LeakageScansRatio, of the top 10 domain names.

Domain Names	TotalLeakingSysInfo	TotalScans	LeakageScansRatio
kpn.net	38391	413474	0.09284985
i3d.net	17171	24062	0.71361483
weserve.nl	10474	36636	0.28589366
datafiber.nl	2242	2896	0.77417127
cloudvps.com	2065	2099	0.98380181
xs4all.net	2012	2012	1.00000000
heberjahiz.com	1881	1881	1.00000000
zeus.pm	1595	1595	1.00000000
dhcp	1567	1567	1.00000000
ip-eend.nl	1559	1559	1.00000000

Figure 8: Top 10 domain names that leaks system information.

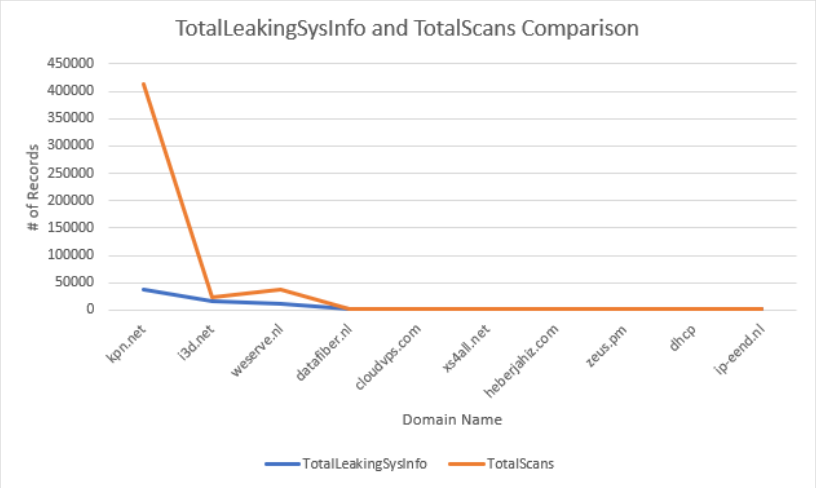

Figure 9: TotalAuthRequired and TotalScans Comparison.

To be able to mitigate the security issues, the value of TotalLeakingSysInfo should be minimized.

One of the factors that make the top 10 domain names leaks the system information differently is the different devices that each company use. Taking the example from the dataset, the domain name "kpn.net" has 24287 unique banners. Many of those reveal system information, such as Wandy and MikroTik version numbers. This fact also applies to other domain names.

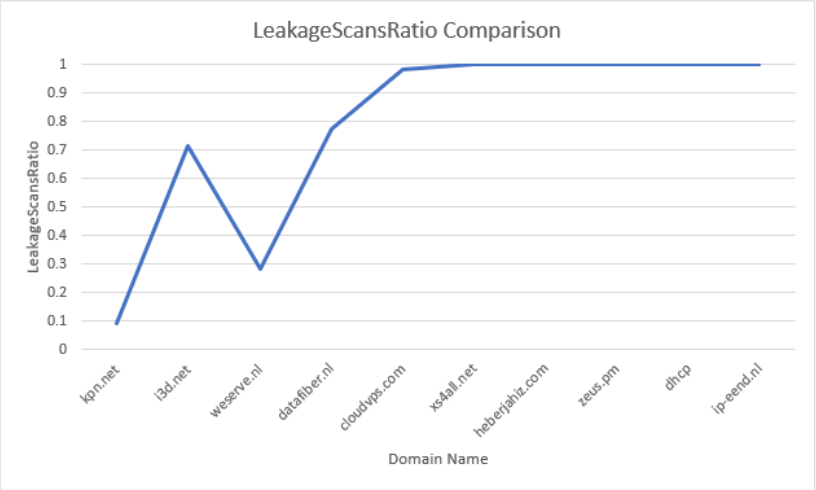

Figure 10: AuthScansRatio and NoAuthScansRatio Comparison.

4 Influences on the security performance of owners of services using Telnet

Usage of SSH on port 23 by the owners of systems using Telnet over the Internet will be the focus of the statistical evaluation part of this paper.

4.a Identification of different factors causing variance in security performance

Based on the previously analyzed metric, there are variances in the result. From the dataset and other external sources, several factors lead to this variance, namely:

- Industry of the owner of the openly-accessible telnet ports (based on NAICS number).
  The industry, to which belong the owner of the system having openly-accessible telnet ports, could be one factor that causes the variance in the metric, as different industries have different perspective on cyber security, especially when discussing about securing the openly-accessible systems through telnet ports. In addition, some industries have been more targeted by cyberattack than others, in the past, which might give them more incentives to using countermeasures. Based on the dataset, the type of industries can be determined based on the NAICS code.

- Location (based on region or city).

Another factor that could cause the variance in the metric is the location of the systems. Some of cities where the systems from the dataset are located could be described as "cybersecurity-aware". For example, some cities in the Netherlands have more frequent cyber security events or conferences. Another example would be that some cities might be populated with more people who are feeling concerned about cybersecurity, have universities which propose cybersecurity classes, or have companies specialized in cybersecurity. In addition, the security metric could be influence by the GDP (Gross Domestic Product) per region or city. Regions with lower GDP may tend to buy cheaper devices, sometimes at a cost of the security measures implemented in the devices.

- Characteristics of the company.
  Organizations with larger size or better reputation, to some extent, means that more cybersecurity graduates or professionals are attracted to apply for a job there. This means that the company will also be able to employ the workers of a higher quality and filter out the applicants that have less knowledge in cybersecurity. Thus, the characteristics of the company could be one cause of the variance in the metric.

4.b Statistical analysis to explore the impact of the factors on the metric

The following analysis will explore the impact of two factors previously mentioned: the type of industry to which the owner belongs and the security awareness of the city where the telnet-open system is located.

4.b.1 Type of industry

Different industries show different behaviours regarding using SSH or Telnet on port 23. This creates a variance in the security metrics defined. This variance will be explained by using Pearson's χ^2. The chi-square test applies to many situations in which experimental frequencies are compared to theoretical frequencies based on a hypothesis [11]. In this case, it was compared five different industries, namely, Electronic Computer Manufacturing; Wired and wireless telecommunications carriers (except satellite); Internet Service Providers; Data Processing, Hosting, and Related Services and Computer Systems Design Services to see if there is a correlation between type of industry and usage of SSH or Telnet on port 23. The hypothesis is as follows:

H_0 = Type of service used and type of industry are independent.

H_1 = Type of service used and type of industry are not independent.

To be able to calculate chi-square, the data that was used is shown in Table 3.

Type of Company/Used Service	Electronic Computer Manufacturing	Wired and wireless telecommunications carriers (except satellite)	Internet Service Providers	Data Processing, Hosting, and Related Services	Computer Systems Design Services	TOTAL
SSH	10859	10	4	311	120	11304
Telnet	473	909	456	10176	177	12191
TOTAL	11332	919	460	10487	297	23495

Table 3: SSH and Telnet usages of each company.

According to the table, Pearson's χ^2 can be calculated by using the following formula

$$\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{5}\frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

The result is $\chi^2_{4,0.01} = 13.28 < \chi^2 = 20132.72$ which shows H_0 is rejected and type of service and type of industry have relations.

χ^2 says that there is a significant relationship between variables, but it does not say just how significant and important this is, for this purpose Cramer's V is used. Cramer's V is a way of calculating correlation in tables which have more than 2x2 rows and columns. It is used as post-test to determine the strengths of the association after χ^2 has determined significance. Cramer's V varies between 0 and 1. Close to 0 it shows little association

between variables. Close to 1, it indicates a strong association. For this example, Carmer's V can be calculated by using the following formula:

$$\text{Cramer's}V = \sqrt{\frac{\chi^2/n}{\min\{I-1, J-1\}}}, 0 \leq V \leq 1$$

The result is 0.65, which means that the association is strong for the type of industry and type of service they are using.

4.b.2 Location (City)

The second factor analyzed in this report is based on locations. Based on the dataset, the location of the systems could play a role in creating variances in the security metrics. For this, Pearson's χ^2 could also be used to determine the correlation between the city of the device and the level of security (the ratio of SSH usage over Telnet on port 23).

One of the factors that can be raised from the dataset is, whether the ratio of SSH usage has any dependency on the frequency of a city organizing cybersecurity conferences or events. Putting "cybersecurity" into the search field and "The Netherlands" in the location field shows the calendar of cyber security events in the Netherlands, as shown in [12]. To evaluate the correlation, the cities in the Netherlands that appear in the dataset is categorized into "more frequent" and "less frequent". When a city hosts more than 5 events, it will fall into the first category, otherwise, it will be categorized as "less frequent".

- "More frequent" cities captured from Eventbrite [12] are Amsterdam, Rotterdam, The Hague, Eindhoven, Groningen, Utrecht, Leiden, Delft, Deventer, Nijmegen, and Roden.
- Other cities will be labeled as "less frequent".

To compute Pearson's χ^2, the hypothesis are as follows. The null hypothesis (H_0) will be "Type of service used and the city is independent". The opposite hypothesis (H_1) will be "Type of service used and the city are not independent".

The 2x2 table for this measurement can be seen at Table 4. Computing in a similar way of the previous computations, the Pearson's χ^2 outputs P-value of 132049.2332. Taking 0.05 as the significance level and the degree of freedom is $(2-1)\cdot(2-1) = 1$, the $\chi^2_{2,0.05}$ is 5.991. Since the computed P-value is much larger than $\chi^2_{2,0.05}$, the null hypothesis is rejected. This means that the city and the type of used service (based on the frequency of cyber security events) have a correlation.

Cities / Used Service	More Frequent	Less Frequent	TOTAL
SSH	184238	70688	254926
Telnet	1764439	3083416	4847855
TOTAL	1948677	3154104	5102781

Table 4: 2x2 Table for Cities.

5 Limitations and Conclusion

In the statistical analysis, NAICS numbers are used to create the type of industries. Dataset had 128,220 distinct IP addresses which leaked the information either SSH or Telnet service on port 23 and 2323. 104639 of IP address' records didn't have any NAICS number which forms 81% of the dataset. This inevitably brought a significant limitation on the data that was worked on during the statistical analysis part.

To conclude, in the paper, for the chosen three different actors, countermeasures that can be taken to mitigate the security issue are visible in the dataset was discussed. Furthermore, the variance caused by different factors in the dataset was interpreted. For interpretation of these correlations, Pearson's χ^2 and Cramer's V were used.

References

[1] It threat evolution q2 2018. statistics. `https://securelist.com/it-threat-evolution-q2-2018-statistics/87170/`. Accessed: October 22, 2018.

[2] Incentives types: Financial and non-financial incentives – explained! `http://www.yourarticlelibrary.com/hrm/incentives/incentives-types-financial-and-non-financial-incentives-explained/35360`. Accessed: October 22, 2018.

[3] Kpn - privacy waarborg. `https://privacywaarborg.nl/bedrijven/koninklijke-kpn-b-v/`. Accessed: October 22, 2018.

[4] Keith Stouffer, Joe Falco, and Karen Scarfone. Guide to industrial control systems (ics) security. *NIST special publication*, 800(82):16–16, 2011.

[5] Cyber security and resilience of smart cars. https://wiki.unece.org/download/attachments/42041673/TFCS-03-09e\%20ENISA\%20Cyber\%20Security\%20and\%20Resilience\%20of\%20smart\%20cars.pdf?api=v2. Accessed: October 22, 2018.

[6] Mirai botnet cost you $13.50 per infected thing, say boffins. https://www.theregister.co.uk/2018/05/09/berkeley_boffins_infect_things_with_mirai_in_a_good_cause/. Accessed: October 22, 2018.

[7] National exposure index. https://www.rapid7.com/globalassets/_pdfs/research/rapid7-national-exposure-index-2018.pdf. Accessed: October 22, 2018.

[8] Shadowserver foundation - main - homepage. https://www.shadowserver.org/wiki/. Accessed: October 22, 2018.

[9] Shadowserver foundation - shadowserver - mission. https://www.shadowserver.org/wiki/pmwiki.php/Shadowserver/Mission. Accessed: October 22, 2018.

[10] The shadowserver foundation | crunchbase. https://www.crunchbase.com/organization/the-shadowserver-foundation#section-overview. Accessed: October 22, 2018.

[11] Ronald J Tallarida and Rodney B Murray. Chi-square test. In *Manual of Pharmacologic Calculations*, pages 140–142. Springer, 1987.

[12] Amsterdam, netherlands cyber security events | eventbrite. https://www.eventbrite.com/d/netherlands--amsterdam/cyber-security/?page=1. Accessed: October 22, 2018.