Differences in security performance of Internet Hosting Providers for the POODLE vulnerability

Delft University of Technology
WM0824TU Economics of Cybersecurity - Assignment Block 4

Group 16:
Marco Boom: 4393031
Maartje Flipsen: 4206444
Andrew Jiang: 4942795
Bryan van Wijk: 4363329

Monday October 22nd 2018

Attachments: https://github.com/bryanvanwijk/TUD_WM0824

1 Security issue

A server can become vulnerable as a consequence of outdated SSL certification, in particular this is the case for the SSL 3.0 version, commonly referred to as SSLv3. In this assignment, the security issue discussed is that of insecure servers due to SSLv3 certification that are therefore vulnerable to POODLE attacks. In a POODLE attack, the handshake protocol between client and server is compromised, as adversaries can exploit the vulnerabilities of SSLv3 at that point. Once the adversary is in, he can act as a man-in-the-middle. This allows him to look into and interfere with the data that is transferred over the insecure connection between client and server.

The data set explored in this assignment and previous assignments shows that many servers still use SSLv3. The data set is used to explore data on Internet Hosting Providers (IHP), which are parties that provide services and web servers to customers, i.e. individuals or organizations that own websites. IHPs are bound to legislation and regulations set by regulators. This can either be a governmental organisation, or a private organisation that shares best-practices and such.

The data set allows analyzing what percentage of servers is vulnerable per Autonomous System Number (ASN). We assume that each IHP runs one ASN, to say something about the number of vulnerable servers per IHS. It is also assumed that any entry in the last week of the data set is still vulnerable. These assumptions allow us to count the distinct IP addresses from the last week of the dataset and to group this by their corresponding ASN. The IP addresses count is then normalized by querying WHOIS databases for all the IP ranges belonging to a given ASN. Following, these ranges are aggregated to return the number of IPs per ASN. Figure 1 shows the resulting percentages of vulnerable IP addresses per ASN in a descending order. For readability purposes, 776 ASNs that has a vulnerable percentage of lower than 5% were filtered out of the figure, making for a total of 791 ASNs.

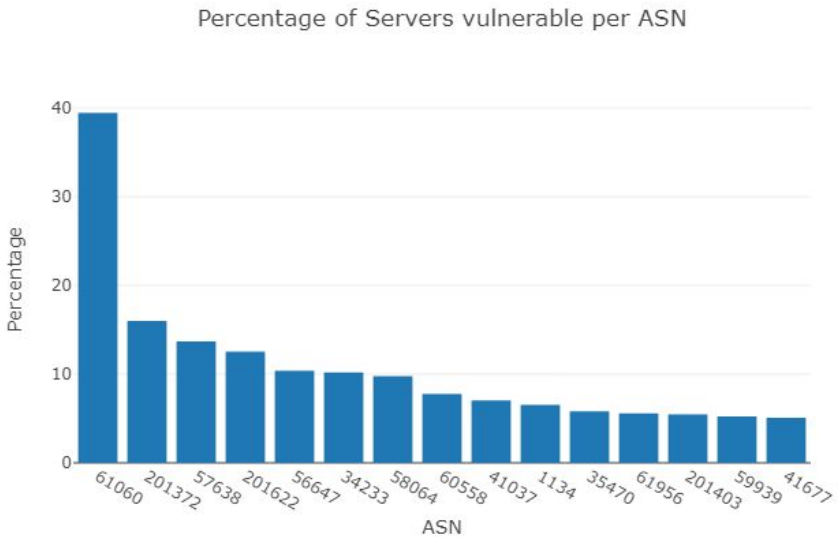

Figure 1: Percentage of servers vulnerable per ASN, for ASNs with vulnerability percentages >5%.

2 Actors and externalities

Actors involved in the security issue follow different strategies to mitigate its impact. Three important actors for the security issue are the Internet Hosting provider, the regulators, and the customers. In the next three sections, for each actor, one specific countermeasure and its externalities, i.e. effects for the other actors and society, are analyzed.

2.1 Problem owner: Internet Hosting Provider (IHP)

An IHP can update the configurations of the servers themselves enforcing their customers to use a newer SSL version. The incentive to use this countermeasure would be the prevented loss for when the communication with one of their servers is intercepted. A positive externality is a better competitive position for having better security and security support compared to other IHPs.

If the IHP deploys this strategy, this would not likely carry any costs for the regulators. The regulators will have benefit with this countermeasure, because it is a countermeasure that is easily checked and regulated. It will thereby prevent all attacks using the SSL vulnerability in an optimal way.

There is also a downside to this countermeasure, as the customers could experience costs due to this strategy. They could experience interoperability issues and their end-users are possibly unable to connect to the servers when they don't support a newer SSL version. Another cost the customers will experience is the server updating costs where they most likely will be charged for. On the other side, the customers will also experience benefits. From the customer's point of view, their risk is now transferred to the IHP. So, they benefit from the secure servers.

Attackers move to other IHPs. It becomes harder for attackers to decrypt the connection between end-users and the sites hosted at the IHP. Therefore, they move to sites hosted at IHPs with weaker security.

Thus, the negative externalities that accompany the countermeasure are the increased cost for the customers and interoperability issues for the end users. There are also positive externalities which are a better competitive position for the IHP and less risk for the customers since the IHP now takes care of the SSL version.

2.2 Regulators

Regulators can impose fines and regulations and enforce these to push the industry to mitigate the security issue in question.

If the regulators would deploy this strategy, this would result in increased costs for companies/organizations, and as such, more overhead in the cost of doing business, which might be detrimental to the economy. However, this would force the adoption of more secure

standards, thus reduce the risk of fraud and security incidents across the board, thereby benefiting the society as a whole.

Given the existence of a two-sided market relationship here, it is very difficult for the market to grow, and for dominant players to take actions in the interest of security due to the low risk of market displacement, as mentioned in the video lectures. Therefore, governments have an incentive to further grow and advance this market, and can do so by deploying the security measure as outlined.

Thus, the positive externality that accompanies the countermeasure is the reduced risk of fraud and security incidents for end users.

2.3 Customers (server owners)

Customers can update their server configuration to use only secure protocols. This would imply that the IHP enables the configuration by customers and it is agreed upon that customers are responsible for the configuration.

The deployment of this strategy by customers incurs several costs. First, for the customers themselves, costs are related to the wages paid for the activity of updating the configuration. Second, it's possible that some end-users do not support newer protocols and therefore they will be unable to continue visiting the customers' website. Due to this, the customer would suffer some loss of income. Another possible loss is the slower connection between customers' site and end-users due to the increased security. A positive externality is the end users benefitting from the secure connection especially when personal information or money is involved.

While there are costs for the customer associated with this strategy, it's benefit is that the risk of incidents drops. Therefore, there is a lower probability of leaking data. Depending on the business scale the customer operates on, the prevented losses are more significant. For example, a customer that operates a popular webshop could be more likely to benefit from a lower probability of leaked data than a webshop that has very little traffic. In cases like webshops, the customer gets into a better competitive position compared to webshops without good security. The customer could steal end-users from other webshops.

Thus, the negative externality that accompanies the countermeasure is the loss of customers due to a slower connection and end-users not supporting newer protocols. A positive externality is a better competitive position and reduced risk for website owners.

3 Factors influencing the metric

The metric displayed in Figure 1 shows the performance differences between the different ASNs or IHPs as we assumed an ASN to be equal to an IHP. Thus, the metric is an indicator of the IHP security performance.

Some IHPs might have a higher percentage of servers vulnerable for the poodle attack than others. Factors influencing the differences among IHP security performance are investigated in the next sections.

3.1 Factors explaining the variance in the metric

The variance in security performance among IHPs can be caused by several drivers. Firstly, the access policy on how customers of IHPs can access their servers might differ, making it for the customers harder or impossible to remotely access some configuration options. This could mean that they are simply unable to update the servers themselves and have to rely on the IHP to update the servers. Secondly, IHPs could be missing the capabilities to update the server configuration because knowledge is lacking. Or there is simply not enough capacity, i.e. not enough skilled employees, to update the servers. Thirdly, business culture can play an important role. Due to differences in business culture, IHPs can have different levels of risk tolerance. Thereby, some IHPs may simply accept the risk of using an insecure protocol, therefore scoring poorer on security performance. Fourth, even when the IHPs have security policies in place to only use the more secure SSL protocols, the quality between the security policies might differ. To ensure this rule is followed, you need to constantly monitor the SSL version used and take action when a lower SSL version is used on regular basis. Lastly, some IHPs might be more popular by customers with lower security demands which means there is less incentive to provide more secure servers with up to date protocols. Conversely, customers who demand more security could choose to merely use IHPs that have better security performance, or at least more controls in place.

To summarize, there are multiple facets influencing the variance in security performance of IHPs. Their respective impact is difficult to quantify, as not only the concepts are difficult to define quantitatively, but also data collection is a challenge. A possibility to still capture some of these factors is to look at the country that the ASN/IHP owner is located in. This is explored in the further sections.

3.2 ASN owner country

Investigating the impact of the country where the IHP is based, requires the collection of extra data, next to the dataset of vulnerable servers in the Netherlands. In the following calculations, the ASN lookup tool from Ultratools (2018) is used to retrieve the countries of the ASN owners. As mentioned before we assume an ASN to be equal to an IHP.

Table 1: Descriptive statistics for percentage of vulnerable servers of ASNs, grouped by origin country.

	asn_vul_servers_percentage			asn
country	**mean**	**std**	**var**	**count**
AE	0.442972	0.621966	0.386842	2
AI	0.065104	0	0	1
AR	0.016655	0	0	1
AT	0.01088	0.011595	0.000134	2
AU	0.002351	0.003704	0.000014	3
BE	0.130205	0.136447	0.018618	9
BG	0.325521	0	0	1
CA	0.06099	0.120979	0.014636	4
CH	0.160998	0.27691	0.076679	5
CN	0.066105	0.111907	0.012523	3
CO	0.000043	0	0	1
CW	0.127551	0	0	1
CY	0.000115	0	0	1
CZ	0.742282	1.049478	1.101405	2
DE	0.158906	0.500581	0.250581	24
DK	0.027982	0.039346	0.001548	2
EE	0.067409	0.088812	0.007888	2
EG	3.125	0	0	1
ES	0.038168	0.049603	0.00246	2
EU	0.533471	1.725272	2.976563	21
FI	0.104167	0	0	1
FR	0.037934	0.074611	0.005567	11
GB	0.246702	0.423025	0.178951	25

GR	0.000138	0	0	1
HK	0.031555	0.058252	0.003393	4
IL	0.071129	0.045329	0.002055	3
IM	0.097656	0	0	1
IN	0.702505	1.647167	2.713161	6
IR	0.002271	0	0	1
IT	0.039762	0.040687	0.001655	5
KH	0.00465	0	0	1
KW	0.000145	0	0	1
LK	0.000198	0	0	1
LT	5.240885	0	0	1
LU	0.976562	0	0	1
MD	1.206342	1.153598	1.330788	2
MT	0.390625	0	0	1
NL	1.013957	2.358408	5.56209	508
NO	0.072489	0.148887	0.022167	8
PH	0.000021	0	0	1
PK	0.000018	0	0	1
PT	0.169572	0.237787	0.056543	3
RU	0.216079	0.448632	0.20127	10
SE	0.383735	1.16966	1.368105	11
TR	0.047239	0.050143	0.002514	3
TW	0.113938	0.161116	0.025959	2
UA	1.022598	1.023907	1.048385	4
US	0.204655	0.701253	0.491755	82
ZA	0.000118	0	0	1

3.3 Analysis of variance

The impact of ASN owner country on the number of vulnerable servers is inspected by applying Analysis of Variance, in particular the one-way ANOVA. This analysis was chosen based on the measurement scales of the independent and dependent variables to be researched. The dependent variable, in this case, is the percentage of vulnerable servers for an ASN, which is a ratio scale. The independent variable is the ASN owner country, which is a variable of nominal scale (e.g. FR, NL, DE, etcetera). Because there are more than two tags on the nominal scale, the ANOVA applies.

In the analysis, the means of vulnerable servers for ASNs are compared for several independent comparison groups. By comparing all groups at once, two types of variance are researched: variance within comparison groups and variance between comparison groups. When the variance between is larger than the variance within, the statistically significant difference between the means of comparison groups becomes more clear (Lane, n.d.). The ratio of variance between groups and within groups is called the F-ratio.

One-way ANOVA has three important conditions to consider. First, that the dependent variable is normally distributed for all comparison groups. Second, there should be homogeneity of variances, which states that all comparison groups have the same variance. To be so, group sizes need to be equal. If this is not the case, the F-test within the ANOVA may be biased. Third, scores need to be independent (University of Alberta, n.d.). With regards to the first condition, the Central Limit Theorem should hold or each comparison group should have at least three samples in order for the Shapiro-Wilk test to be applied. With regards to the second assumption, Levene's test can be applied where the null hypothesis is that comparison groups do not have equal variance. The null hypothesis can be rejected when the resulting p-value is less than the chosen significance level of 0.05, thus fulfilling the homogeneity condition as samples for the comparison groups come from populations with similar variance. The third condition is assumed to be met, as data was collected with procedural IP range scans.

ANOVA is based on hypothesis testing, where the null hypothesis is that there is no difference in variance between and within comparison groups. In other words, the means of the comparison groups are roughly equal. Opposing, the alternative hypothesis is that the variance between groups is larger than the variance within groups, thus meaning that means are not equal and that there is a large enough difference in groups that it can explain the difference in security performance. The chosen significance level is $\alpha = 0.05$. The ANOVA is executed for the following hypotheses:

$$H_0: \mu_{AE} = \mu_{AI} = \ldots = \mu_{ZA}$$
$$H_1: \textit{Means are not equal}$$

3.4 Results

To begin the conditions needed to be met. As stated before, the independence condition is assumed to be met. With regards to the normality condition, the Central Limit Theorem does not hold as not all groups contain over 30 data points. In addition, the Shapiro-Wilk test could not be performed as certain comparison groups have less than three data points. Thus, the normality condition is not met. In fact, the normality condition is violated such that the one-way ANOVA method as described above is no longer reliable. Instead, the Kruskal-Wallis test for one-way analysis of variance will be applied (pytoLearn, n.d.). This test is similar to the one-way ANOVA, with the same null hypothesis, but it does not assume a particular distribution, and as such, its results still hold despite the failure to meet the prerequisites of ANOVA. A significant result of the Kruskal-Wallis test indicates, same as the one-way ANOVA, that the null hypothesis can be rejected. In that case, this means that at least one group is significantly different from at least another group. All calculations can be found in the appended Python notebook, also uploaded on the Github repository.

The Kruskal-Wallis calculations returned a Kruskal-Wallis H/test statistic of ca. 272, and a p value of ca. 4.25e-33. The Kruskal-Wallis H-statistic is to be compared to the critical Chi-square value, which can be determined using the p value and the degree of freedom, and locating the related cell in the Chi-squared distribution table. Should the critical Chi-square value be less than the H statistic, then the null hypothesis that the medians of the groups are equal can be rejected; otherwise, there does not exist enough evidence to suggest that the medians of the groups are unequal. In our case, the H statistic is far greater than the critical Chi-square value, which means that the null hypothesis can indeed be rejected.

The p value can be interpreted as follows: a value greater than 0.05 means that the data grouping does not lead to a statistically significant result, and vice versa for values less than 0.05. Considering that the p value is much less than 0.05, it is safe to say that the differences in the percentage of vulnerable servers of each ASN does have a statistically significant tie to the ASN's country of origin.

Table 2: Inputs and outputs for the Kruskal-Wallis test.

α	0.05
Kruskal-Wallis H	272.25677830133685
p	4.247475481702601e-33
Degree of Freedom (df)	49 - 1 = 48

3.4.1 Discussion on results

Since there are very few data points per group, the normality condition for the one-way ANOVA was not met. Alternatively, the Kruskal-Wallis test was applied. Because this test is non-parametric, i.e. free of distributions, it is less powerful than the one-way ANOVA. Therefore, though the Kruskal-Wallis test results in a significant difference for at least one of the comparison groups to at least one other, the results should not be considered fully reliable. Rather, more research is encouraged for statistical testing with more data points per comparison group, here the ASN owner countries. In particular, it would be interesting to see if the one-way ANOVA for more data points per group results in a similar conclusion. Additionally, a failed rejection of the one-way ANOVA would be a fruitful starting point for post hoc analyses to research which groups in particular differ, to make the explanation of differing variance in security performance for ASNs more meaningful.

4 Conclusion

Actors involved in the security issue follow different strategies to mitigate its impact. Three important actors for the security issue are the Internet Hosting provider, the regulators and the customers. The different countermeasures they follow are accompanied by different externalities. One of the negative externalities is the increased cost for the customers and interoperability issues for the end users. A positive externality is the reduced risk of fraud and security incidents for end users.

The security performance in server vulnerability by Internet Hosting Providers (IHPs) can vary. Several factors can explain this variance, e.g. access controls, capabilities and culture. As these factors are difficult to quantify, and data collection proves to be a challenge, the impact of IHP owner country is investigated as a proxy. Here it is assumed that an IHP is equal to an ASN. The ASN owner country is the country that the ASN is officially registered to. To determine the impact of the country of origin of the IHP on security performance, a statistical analysis using the Kruskal-Wallis test was conducted, which yielded a statistically significant result. This entails that the null hypothesis can be rejected and it can be concluded that at least one country significantly differs from at least one other country in security performance. This supports the theory posed in section 3.1, that the origin country of an IHP influences its security performance. However, this conclusion is debatable: while the Kruskal-Wallis test is not the strongest of tests, it was the best option given the low number of data points available for certain countries. For greater confidence, more data points should be collected for ASNs per country, such that there would be sufficient data to analyze for distributions on a group basis, and for the regular one-way ANOVA procedure to be employed.

5 References

Lane, D. (n.d.). One-Factor ANOVA (Between Subjects). Retrieved October 19, 2018, from http://onlinestatbook.com/2/analysis_of_variance/one-way.html

pytoLearn. (n.d.). One way ANOVA. Retrieved October 19, 2018, from http://pytolearn.csd.auth.gr/d1-hyptest/12/anova-one.html

Ultratools. (2018). ASN lookup tool August 2018 Overview. Retrieved October 13, 2018, from https://www.ultratools.com/tools/asnInfo