WM0824TU
Economics of Cyber Security

Block 4 Report - Group 13 - Phishing

Fiske Schijlen	(4947681)
Alexandros Korpas-Kamarianos	(4907167)
Joël Abrahams	(4443268)
Thanh-Dat Nguyen	(5025664)

{f.h.schijlen, a.korpas-kamarianos, j.s.abrahams, t.m.nguyen-3}
@student.tudelft.nl

October 19, 2019

Contents

1 Introduction

Previously, we analysed the variance in security performance based on the survival rate of IPs, and how different risk strategies of different actors affect this variance. Here, we proceed to analyse the underlying factors which influence this variance. This will allow us to uncover the underlying reason of the differences of security performance of different actors.

In section 2 we identify the mitigation strategies of the problem owners (organisations), end-users, and employees, and the costs and benefits these strategies provide. In section 3 we identify the type of actors which are visible in our metric, i.e. those that influence the up-time of phishing websites, and attempt to identify the factors which influence the variance in up-time.

2 Significant actors' mitigation strategies

In this section, we will identify some significant actors' most valuable mitigation strategies, as well as whether or not the execution of these strategies has any merit for the actors. More precisely, we will investigate how a (different) mitigation strategy that can be adopted by the problem owner, the end-users and/or the organisations' employees can benefit those three actors and what the adoption costs are. Furthermore, we will try to identify the incentives that drive the corresponding actor to adopt their specific strategy, as well as the externalities that occur as a result of this choice, namely the indirect consequences of this strategy to the rest of the actors.

2.1 Problem owner

Mitigation countermeasure

The problem owner refers to an organisation that suffers losses due to the consequences of phishing attacks on its employees and end-users. Therefore, the problem owner's most effective mitigation strategies focus on protecting those vulnerable parties. Specifically, **deploying a phishing awareness program for end-users** is an effective mitigation strategy[1] that reduces the success rate of phishing attacks by up to 15 percent.[2]

Cost-benefit analysis

The cost of this countermeasure comes from having to define a comprehensive summary of relevant phishing patterns, having to define mitigation steps to help end-users protect themselves against those patterns, and finally having to deploy the resulting awareness program through one or multiple platforms. The execution of these steps requires paying staff for their research and implementation work. Furthermore, the staff will not be able to work on other tasks while they are working on an awareness program, which results in more indirect losses.

The benefit is in the users' improved awareness, which reduces the success rate of phishing attacks. This means the average impact of each phishing attempt is mitigated. Since the mitigated losses scale with the original amount of losses, the countermeasure is less effective if the organisation's losses due to phishing were low to begin with.

Incentive

Assuming an organisation's goal is to maximise profit, it will naturally want to minimise losses as well. This means an organisation has an incentive to implement this countermeasure if its losses due to phishing are significantly high. To evaluate if these losses are high enough to warrant an awareness program, the organisation in question will have to weigh their implementation cost against the mitigated losses.
An awareness program incurs absolute costs, while the mitigated losses scale with the amount of original losses. This means the organisation will mostly have an incentive to implement this countermeasure if their losses due to phishing are exceedingly high.

Externalities

Employees are significantly affected by the presented mitigation strategy, as their workload is increased with extra working hours and in the meantime they need to complete their regular tasks. However, they are affected positively as well, as their awareness level is increased which could be very important for their personal life. Furthermore, the end-users are affected, as they don't have a direct incentive to follow the phishing awareness program and they might consider this action as unnecessary working hours or wasting of time.

2.2 End-user

End-users can be considered as those who use the service that is provided by the problem owner. Usually, this service requires the user to provide some personal information, such as address and nationality, that are protected through various regulations, e.g. GDPR. Moreover, in some cases, the end-user has given his bank account details in order to execute transactions. In case of a successful phishing attack, the attacker would have access to sensitive information of the end-user or even execute unauthorised transactions. Hence, the service provider is obligated to either pay fines for the data breach or refund the user for the unauthorised transactions.

Mitigation countermeasure

An effective countermeasure strategy for an end-user is to use the 2-factor authentication login service that is (or should be) provided by the organisation. Thus, the end-user is protected from asynchronous phishing attacks, namely attacks that aim to retrieve the credentials of the user and store them for later usage. However, the risk is not mitigated entirely, as there is a case that the attacker asks for the One-Time Password and executes the aforementioned attacks right away (synchronous attack).

Cost-benefit analysis

The presented countermeasure provides both costs and benefits to the involved actors. As previously, we will investigate those that are related to the examined actors, namely the problem owner, the end-user, and the employees.

- **End-user:** A significant cost of the end-users who use the mitigation countermeasure is the extra layer of inconvenience that is added when they want to login and/or complete a transaction. On the other hand, their assets are more protected by this countermeasure.
- **Problem owner:** Obviously, the implementation costs for a 2-factor authentication login service fall on the organisation. Furthermore, there is a change that some end-users won't be pleased by the aforementioned inconvenience of this measure and they might stop the provided service. This could result to a decrease of the organisation's revenues. However, an important benefit for the problem owner are the prevented losses that might be large in case of a successful phishing attack, as described above.
- **Employees:** The employees of the organisation are not directly positively or negatively by the countermeasure, but only indirectly, assuming that the presenting costs and benefits of the organisation affect them in any way.

Incentive

The only concern of the end-users when it comes to an unauthorised access to their account is the protection of their sensitive information and their bank account, when applicable. Hence, every action that aims to mitigate the risk of a phishing attack is driven by this incentive.

Externalities

The 2-factor authentication as a mitigation strategy adopted by the end-users affect indirectly other actors. When it comes to the problem owner, a downturn of the organisation's revenues is possible, as the aforementioned inconvenience of this countermeasure might discourage the end-users to use the provided service. As a result, the employees of the organisations are affected accordingly, as any loss of the organisation might lead to redundancies.

2.3 Employee

Employees of an organisation are often targeted by phishing. This attack aims to get access to the organisation through the mistake of an employee. For example employees can divulge credentials to the attackers which gives them access to potentialyl sensitive data.

Mitigation countermeasure

Employees can mitigate phishing by strictly following the security policies of the organisation where they work. Examples are reporting any phishing emails, not using company email address for anything other than work, notify when they fell for a phish, etc. The security policies depend on the organisation. Employees can also try to follow phishing awareness programs if they think that they are prone to such attacks.

Cost-benefit analysis

1. **Employees:** The benefits of these countermeasures are that the employees will be better in detecting phishing emails and therefore not fall for them. This allows them to not be a liability to the organisation. Employees that are likely to fall for phishing attacks are dangerous for the organisation, which can put their career in jeopardy. Following security policies has no cost for the employee. Phishing awareness programs are in general not paid by the employees.

2. **Problem Owner:** The main costs of these countermeasures fall on the organisation. Employees that are strictly following security policies generally are less productive, because no risky shortcuts can be taken. However, the small loss in productivity is highly beneficial to the organisation as they are less likely to incur losses from phishing attacks. These losses can come in various forms such as data breach, fines, loss of intellectual property, etc. To further avoid phishing attacks, organisations pay for phishing awareness programs for their employees.

3. **End-User:** The customer does not have any direct costs from the countermeasures. An indirect cost could be that the security investment of the organisation reflects on the price of the products and services that are offered. The benefits of the customer are that their data is less likely to be leaked.

Incentive

The major incentive for employees to counter phishing attacks is their career. An employee that is prone to phishing attacks is not well regarded in an organisation and industry. The organisation might want to let go of that employee because he poses a threat.

Externalities

The specific mitigation strategy barely affects the rest of the actors indirectly. However, it is noteworthy to mention that an indirect consequence of this mitigation strategy is that the employees are end-users in their personal life. Since their phishing awareness level is increased, they are benefited as they are more capable to protect their own assets (that are not related with their job) from phishing attacks.

3 Analysis of factors which influence the variance of phishing attacks per company

In the previous assignment we developed a metric which measures the average up-time of phishing names over time to measure their survival rate. This metric shows how effective different phishing attacks are when categorised by the entity they are impersonating (denoted by `virusname` in the clean-mx dataset). We would expect a logarithmic trend over time, with a large number of domains being shut down in a short amount of time, and a small amount of domains evading capture for a significant or large span of time.

3.1 Factors which cause variance in the metric

Assuming that the entities being impersonated are the ones reporting the domains, their performance is directly visible in our metric. In addition to this, phishing attacks are often reported to other concerned entities, such as phishing blacklists, enterprise protection, and anti-phishing communities [3]. Furthermore, ISPs receive pressure from both users of their services, who wish that these services do more to protect them, and from institutions targeted (either impersonated or directly targeted) by these attacks to shut down these domains from which attacks are staged [4].

Phishers in general seem to avoid traffic from areas where there are major organisations involved in internet security, as detection by a web host might lead to deactivation of the platform/domain where the phishing attacks are staged [3]. Phishers don't want anyone besides the victims to have access to a given phishing page or platform.

Furthemore, a 2014 study observed based on a 2014 dataset that the 9 countries with the most phishing domains are as follows: United States (76%), Czech Republic (9%), United Kingdom (5%), Germany (4%), France (2%), Canada (1%), Netherlands (1%), Australia (1%), and Ireland (1%). The United states hosted the largest amount of phishing domains, in part due to large registars/hosts such as Godaddy.com, Enom, and WildWestDomain.com. A 2018 dataset also showed that after the United States, in addition to the above countries, Hungary and Belarus were also found to host a large amount of phishing domains.
To summarise, phishers try to avoid targeting tech-dense locations, as large (security) tech companies are more likely to uncover their operations if they do. At the same time, however, due to the large amount of host providers in these locations, a lot of phishing domains are located in tech-dense areas such as California.

We wish to investigate whether domains are they also more likely be uncovered if they are hosted in such tech-dense areas.

Hypothesis	Variable
Phishing domains hosted in areas with a large amount of blacklisted IPs [3] have a low survival rate	Blacklisted/blocked IP's from from Cofense (formerly Phisme)

3.2 Data collection method

3.2.1 Data used

We can get the country of origin from the clean-mx dataset.
For getting the amount of blacklisted IP's per country, we make use of breakdown of the amount of IP's listed as infected by country by the CBL (Composite Blocking List) project[1].
CBL is a division of Spamhaus[2], which is an international organsiation based in Geneva, Switzerland, which originally was conceived to track email spammers and spam related activities. In fact, many ISPs use their their anti-spam lists to reduce the amount of spam that reaches their users[3].

3.2.2 Model

Using the above two data-sets, we can cross reference the location of where the domain is hosted against the amount of blacklisted IPs in that area. Thus, we end up with a dataset with 3 variables:

1. Survival rate (days). Also known as up-time in the dataset.
2. Country
3. Amount of blacklisted IPs in that country

3.2.3 Data preparation

We will map each phishing attempt's `country` field to a new `amount_blacklisted` field, which will store the amount of blacklisted IPs in the attempt's country of origin. For this mapping, we will use the list of blacklisted IPs per country [3] and normalise them by the number of internet users per country using the same dataset's `IPop%` column.

3.3 Methods

In order to accept or reject our hypothesis, we need to measure the correlation between phishing sources' *uptime* and *their originating area's amount of blacklisted IPs*. Since these variables are both numerical, we will masure the linear correlation between them by using the *Pearson correlation coefficient* (PCC) [5].

[1] https://www.abuseat.org/public/country.html

[2] https://www.spamhaus.org/

[3] https://www.spamhaus.org/organization/

We will also observe the relation between phishing sources' *survival rate* and their originating area's amount of blacklisted IPs by constructing a *Cox proportional hazards model.*

3.3.1 Pearson correlation

We will compute the *Pearson correlation coefficient* (PCC) between `uptime` and the amount of blacklisted IPs in an phishing source's country of origin. To do so, we calculated each source's `uptime` in days using our dataset and obtained each relevant country's amount of blacklisted IPs from the dataset provided by Spamhaus's *composite blocking list* (CBL) division.[6] Keeping our naming conventions in line with that of the provided dataset, we will henceforth let the `listings` field denote the amount of blacklisted IPs corresponding to a phishing source's country of origin.

The PCC between two sets of numerical values X and Y can be computed with the following equation:

$$\text{PCC}_{X,Y} = \frac{\text{covariance}(X, Y)}{\sigma_X \sigma_Y}$$

where σ_A is the standard deviation of set A.

In our case, we have columns $X =$ `uptime` and $Y =$ `listings`. To get a dataframe with these columns and a correct mapping for each phishing attempt, we executed an inner join on the `country` column of our dataset and CBL's blacklist dataset.

3.3.2 Cox proportional hazards model

The Cox proportional hazards model will be used to compute the correlation between the survival rate of a phishing domain and the amount of blacklisted IPs in the country of the phishing domain. To obtain the survival rate of the phishing domains, we used the uptime of the phishing domain and its state, with 0 denoting "alive" and 1 denoting "dead".

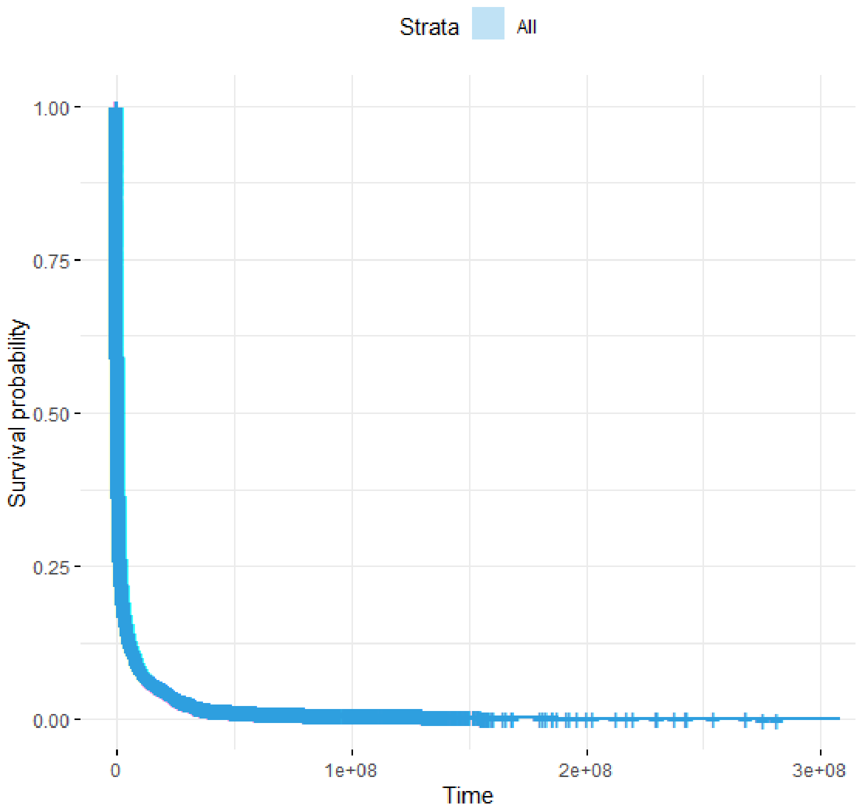


Figure 1: Survival rate of phishing domains with time in days.

To compute the correlation, additional data is needed. For each entry in the table, we added the normalized number of blacklisted IPs (IPop% [6]) depending on the country in which the phishing domain resides [7].
A Cox proportional hazards model is represented as a hazard function, which represents the risk that a phishing domain is taken down. It is computed using the formula:

$$h(t) = h_0(t) * exp(b_1x_1 + b_2x_2 + ... + b_px_p)$$

where t is the survival time, x_i is a covariate, b_i is the impact of the covariate x_i and $h_0(t)$ is the hazard function which can be computed by taking the complementary of the survival function (1 - survival function) [8].

3.4 Results

3.4.1 Pearson correlation

We used the `pandas` module in Python to determine $\text{PCC}_{\texttt{uptime,listings}}$ and obtained the heat map from figure 2.
The figure shows that $\text{PCC}_{\texttt{uptime,listings}} \approx -0.0039$, which implies that there is no significant correlation between a phishing source's up-time and its originating country's amount of blacklisted IPs. This does not align with our hypothesis that phishing domains

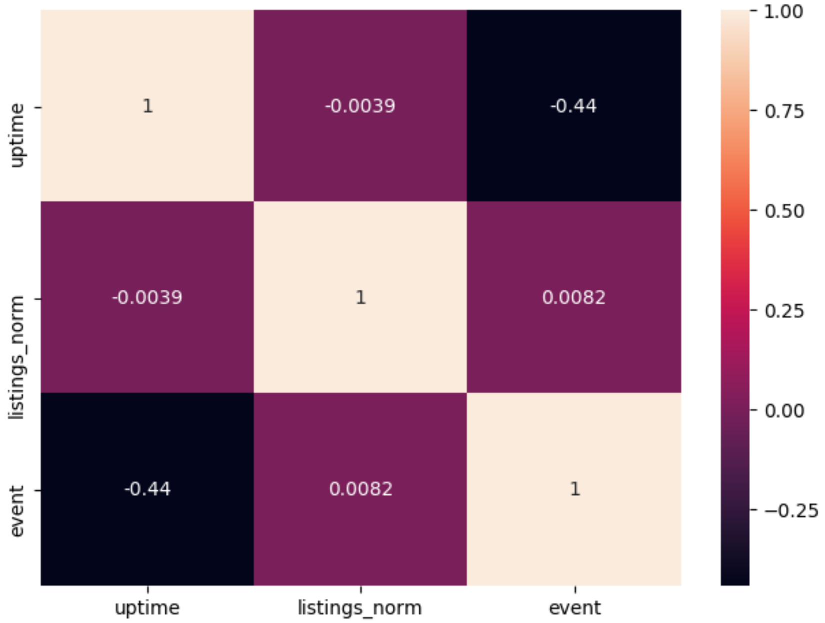
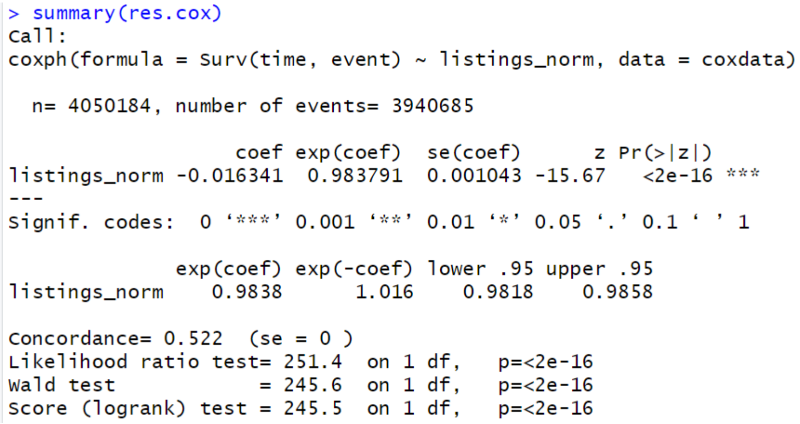


Figure 2: Heat map displaying the PCC between the `uptime` and `listings` columns

hosted in areas with a larger amount of blacklisted IPs have a lower survival rate. The figure also shows the trivial negative correlation between `uptime` and `event`, since `event` is 0 if the phishing source domain is still up as far as our dataset is concerned.

3.4.2 Cox proportional hazards model

The Cox model is a type of proportional hazards model that is typically used in statistical medical research to investigate the association between survival time of patients and one or more predictor variables. We can apply this model to our data, which allows us to investigate the association between survival time of phishing domains, their hosting location, and the amount of blacklisted IPs in that location.



```
> summary(res.cox)
Call:
coxph(formula = Surv(time, event) ~ listings_norm, data = coxdata)

  n= 4050184, number of events= 3940685

                   coef exp(coef)  se(coef)      z Pr(>|z|)
listings_norm -0.016341  0.983791  0.001043 -15.67   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

              exp(coef) exp(-coef) lower .95 upper .95
listings_norm    0.9838      1.016    0.9818    0.9858

Concordance= 0.522  (se = 0 )
Likelihood ratio test= 251.4  on 1 df,   p=<2e-16
Wald test            = 245.6  on 1 df,   p=<2e-16
Score (logrank) test = 245.5  on 1 df,   p=<2e-16
```

Figure 3: Summary of the COX model results

Our model's results can be seen in figure 3. They can be interpreted as follows:

Variable	Explanation	Interpretation
Statistical significance (z) (`coef`)	This corresponds to the ratio of each regression coefficient to its standard error. This evaluates whether the beta coefficient of a given statistic differs greatly from 0, which if it does means that it has an effect on the survival rate.	We obtained a coefficient of -0.016341 for the `listings_norm` variable. Due to it being close to 0, we can conclude that the amount of blacklisted IPs in a region does not influence the lifetime of a phishing domain significantly. The lifetime on average is only marginally smaller in regions with a large amount of blacklisted IPs, normalised for amount of internet users.
Regression coefficient (`coef`)	A positive sign means that the probability of being shut down is higher for domains with that variable present.	We obtained a regression coefficient of approximately -0.016 for the `listings_norm` variable. This implies that the survival rate *slightly* decreases per unit increase in the normalised amount of listings. In other words, the two variables are implied to be mostly uncorrelated. This aligns with our PCC results from figure 2.
Hazard ratios (`exp(coef)`)	Hazard ratios give the effect size of covariates.	The hazard ratio indicates how much the hazard is reduced depending on the number of blacklisted IPS. As there is no significant correlation, the hazard ratio does not provide any useful information.
Confidence intervals of the hazard ratios (`lower.95` and `upper .95`)	We also have an upper and lower 95% confidence interval for the hazard ratio.	The obtained confidence interval is so small that it has no significant implications for our results.
Global statistical significance of the model (`Pr(>|z|)`)	We also have the p-values for three alternative tests for overall significance of the model. These are the *likelihood-ratio test*, *Wald test*, and *score logrank statistics*. They are asymptotically equivalent, meaning that for large N they will give the same results.	Our global statistical significance was smaller than $2e^{-16}$, which means the null hypothesis is firmly rejected for this model and the results are meaningful.

4 Conclusion

We have analysed the factors that influence the *up-time* and *survival rate* of phishing sources in our dataset. In particular, we found that the victim's location can have a large influence on the attack's probability of success[4] and therefore also on the up-time of the source of the attack.
We investigated whether or not this claim holds if you distinguish phishing attacks by their source's country of origin and take each country's normalised amount of blacklisted IPs into account as an indication of security-related activity. However, our Pearson correlation and Cox model analyses showed that a phishing source's up-time and originating country's amount of blacklisted IPs are not significantly correlated. Using our Cox model, we found similar results for the correlation between a phishing source's survival rate and the corresponding country's amount of blacklisted IPs.

References

[1] J. Milletary and C. C. Center, "Technical trends in phishing attacks," *Retrieved December*, vol. 1, no. 2007, pp. 3–3, 2005.

[2] I. Dimov. (2019) Security awareness statistics. [Online]. Available: https://resources.infosecinstitute.com/category/enterprise/securityawareness/security-awareness-fundamentals/security-awareness-statistics/

[3] A. Oest, Y. Safei, A. Doupé, G. Ahn, B. Wardman, and G. Warner, "Inside a phisher's mind: Understanding the anti-phishing ecosystem through phishing kit analysis," in *2018 APWG Symposium on Electronic Crime Research (eCrime)*, May 2018, pp. 1–12.

[4] S. Gupta and P. Kumaraguru, "Emerging phishing trends and effectiveness of the anti-phishing landing page," in *2014 APWG Symposium on Electronic Crime Research (eCrime)*, Sep. 2014, pp. 36–47.

[5] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," in *Noise reduction in speech processing.* Springer, 2009, pp. 1–4.

[6] C. B. List. (2019) Cbl breakdown by country, highest by count. [Online]. Available: https://www.abuseat.org/public/country.html

[7] [Online]. Available: hhttp://www.sthda.com/english/wiki/survival-analysis-basics

[8] [Online]. Available: http://www.sthda.com/english/wiki/cox-proportional-hazards-model