Economics of Cyber Security

Block 2 Report

Fiske Schijlen (4947681)
Alexandros Korpas-Kamarianos (4907167)
Joël Abrahams (443268)
Thanh-Dat Nguyen (5025664)
{f.h.schijlen, a.korpas-kamarianos,j.s.abrahams, t.m.nguyen-3}@student.tudelft.nl

September 16, 2019

Todo list

Introduction

General intro about the assignment content

Outline of the upcoming chapters

Security Issue

Phishing attacks can be examined from two different perspectives as a security threat. First, such an attack could cause direct damages to an organisation, as their members are constantly possible victims of an attack that would provide access to a non-trusted party to sensitive data or infrastructures. On the other hand, a phishing attack might affect an organisation indirectly an organisation by hurting its reputation. Thus, to describe the security issue, we need first to define precisely what is security under these two viewpoints. **(author?)** [1] uses two specifications to define the security under a certain context: *security for whom* and *security for which values*.

Actual numbers and citations to support the questions

As a direct security issue

- security for whom: individuals/end-users
- security of which values: credentials, sensitive information –¿ confidentiality, integrity

As an indirect security issue

- security for whom: organizations
- security of which values: economical damage, reputation

Ideal Metrics for Security Decision Makers

In general, add more citations (arguments need strong motivation)

Metrics can aid security decision makers during their decision making process. Below, we discuss some metrics which we believe would be ideal for security decision makers, regardless of their actual feasibility. It should be noted that we attempt to frame these metrics in terms of organisations in general, say for instance government organisations, and not just in terms of corporations specifically.

Security gains from security measures are often intangible in from [2], which, when compounded by the actual cost of security, which is tangible, can make it difficult to justify the extra cost and potential burden placed on users by implementing said security measures. Therefore, metrics can make it easier to measure the hypothetical gains made by implementing a certain security measures, making it easier to justify its existence.

In this context, phishing attack can target either an individual or an organisation. When an individual is targeted, an attacker might for instance want to retrieve the log-in details of a user. When an organisation is targeted, an employee is target to for instance retrieve their employee credentials.

Detection and Prevention Metrics, Measuring the Risk

Detection metrics can be used to detect how eminent an attack might be, or how likely that the organisation will be attacked in the (near) future. This is akin to the *likelihood* portion of the *risk* = *likelihood* × *consequence* relation.

Add more general metrics?

Improve wording

Locations known for frequent origin of attacks — Certain locations or servers are often known to be frequent origins of attacks. Requests made from such locations could raise red flags, indicating that the request is more likely to be malicious in intent.

No. of attacks relative to organisation size — It stands to reason that organisations that are larger in size are more attractive to malicious users, as from their perspective, a successful breach is more likely to reap larger rewards. Accurately knowing the relationship between one's size, and the amount of attacks, can make it easier to assess risk.

Types of attacks over time — Following trends in the types of attacks can indicate whether a request made is malicious in intent.

Employees awareness of phishing attacks — Employee awareness of phishing can reduce the likeliness of a successful attack, as reckless behaviour can often circumvent many safety mechanisms put in place.

Effectiveness of counter measures — This is a very general metric, and would be specific to counter measures implemented. Ideally, we want to be able to see directly how effective they are, and how many successful attacks would take place without the tool as opposed to with.

Proportion of emails that are malicious —

Amount of money lost as metric

Measuring the Consequences

In also useful to measure the potential consequence of a successful attack. This indicates how important it is to actually prevent said attack, indicating how many resources should be poured into preventing it. This is akin to the *consequence* portion of the *risk* = *likelihood* × *consequence* relation.

Add more general metrics?

Rewrite below slightly for phishing specifically

Effect on brand image — A data leak, which may be the result of leaked employee credentials, can be brought to the public's attention, which can results in (lasting) damage to a brand's image and reputation. A damage to reputation could, in the case of a corporation, lead to decreased sales, potentially causing lasting financial damage long after the attack itself. Therefore, a method of measuring the effect on a brand's image would give a more accurate image of the actual cost of a security breach.

No of users lost as metric

Legal ramifications — Depending on the type of data leaked, there can be legal ramifications for leaking said data. Whereas leaking a set of usernames to the public might damage reputation as it could give notice to perhaps sloppy security practices, leaking credit card or bank information might lead to lawsuits to compensate for potential financial damages inflicted upon users of a system.

What are the metrics that exist in practice?

click-through rate: This metric measures the number of times users are falling for phishing emails. The rate can also be for a particular phishing email. This metric is mostly used for phishing awareness programs, in which users are faced to simulated phishing emails. It allows to track whether the users are becoming more aware of this attack and hence the effectiveness of the program. There are flaws to this metric. The metric does not take into account how elaborate and how well designed the phishing email is. As such, the click-through rate might increase leading to wrong interpretation that the program is failing. Another flaw is when the user base does not stay constant. New users probably will lead to a increase in the click-through rate. [3, 4, 5]

click-through rate per user This is a variant of the above mentioned metric that focuses on one single individual. It reflects the user's resilience to phishing emails. This metric allows to identify users that are prone to phishing attacks. [4, 6]

average detection/report time This metric measures the time until a phishing email is reported by a user or detected by the IT security team. It allows to track the response time to phishing emails. [7]

number of reported phishing emails reported This metric measures how many phishing emails are reported. This is particularly useful in phishing awareness programs where the number of simulated phishing emails are known. [7]

number of phishing domains The metric measures the number of reported and detected phishing domains in a Top-level domain (TLD). [8]

average up-time of phishing site Tracking the average up-time of detected phishing site by TLDs shows how effective phishing detection is. [8]

average traffic TLDs can track traffic of known phishing sites to estimate their effectiveness, [8]

phishing domains per ISP Phishing domains are not equally distributed in ISPs. Some host a larger number than the average. Such a metric can be a good tool to identify traffic that should be more closely monitored or even blacklisted. [8]

elaborate further

email?

Metrics that are applicable to the dataset

Relevant data attributes

In practice, you can only make use of metrics which the dataset allows you to produce. Our phishing dataset contains several attributes that can be used for this purpose:

- `firsttime` and `lasttime`, which can be used to determine the threat's uptime and whether it has been recently active;
- `virusname`, which actually refers to the entity or organisation as which the phisher is acting and can be used to find out which organisations are most commonly used in phishing procedures;
- `recent`, which shows whether the used phishing link is still up or not;
- `ip`, `domain`, and `email`, which can be used to determine commonly occurring phishing sources;
- `country`, which can be used to determine which countries have the most active phishing attempts.

Practical metrics based on the dataset

Our dataset describes several properties of phishing sources, but it does not contain source- or target-specific information that could determine the effectiveness of the phishing attempts. Specifically, we have no information regarding click rates, detection times, and user reports, which rules out the practical metrics from [3], [4], [6], and [5]. We can, however, use the `firsttime` and `lasttime` attributes to estimate the mean uptime per TLD, country, IP, or domain. This can then be used to determine the threat level of certain phishing threats.

Metrics that can be defined using the dataset

Using our dataset, we can measure phishing sources' threat levels by comparing their uptime, recent activity, and frequency. We can then use these threat levels to gather phishing domains and IP addresses to blacklist. We believe this approach to be effective because blacklisting is one of the most popular anti-phishing solutions.[9]

With threat level estimation in mind, we define the following metrics *for each occurring phishing source in the dataset*, where we treat a row's IP address as a phishing source:

- Source IP occurrence frequency;
- Amount of days the phishing source has been active;
- Source country occurrence frequency;

- Source domain occurrence frequency;
- Source email occurrence frequency.

Include a metric that makes use of the domain name servers

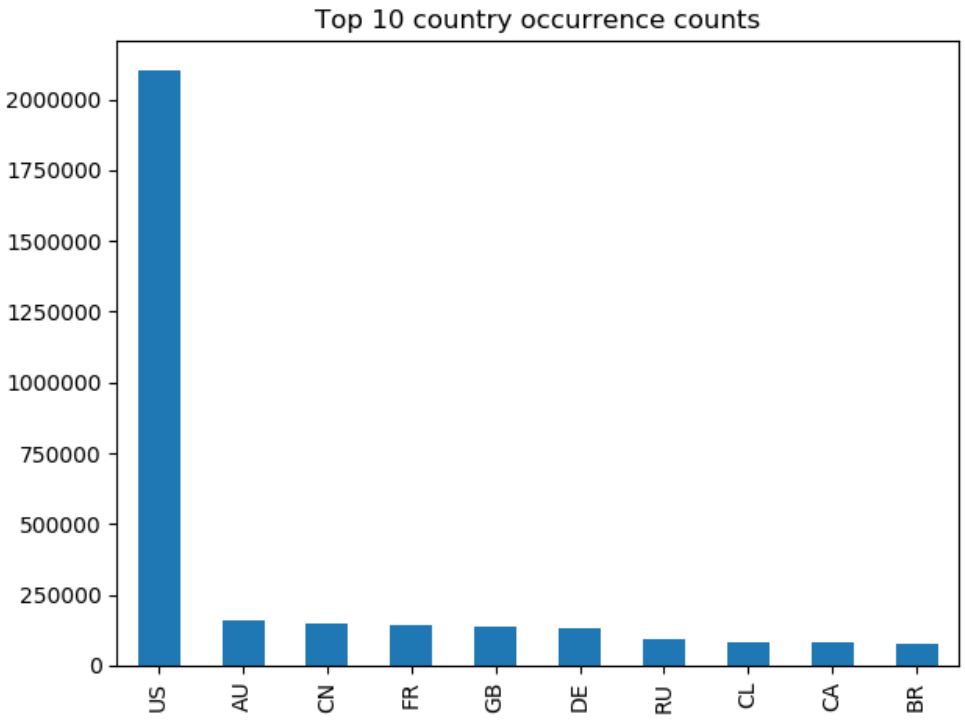
Attempt to find a more security-related metric

Evaluation of Metrics

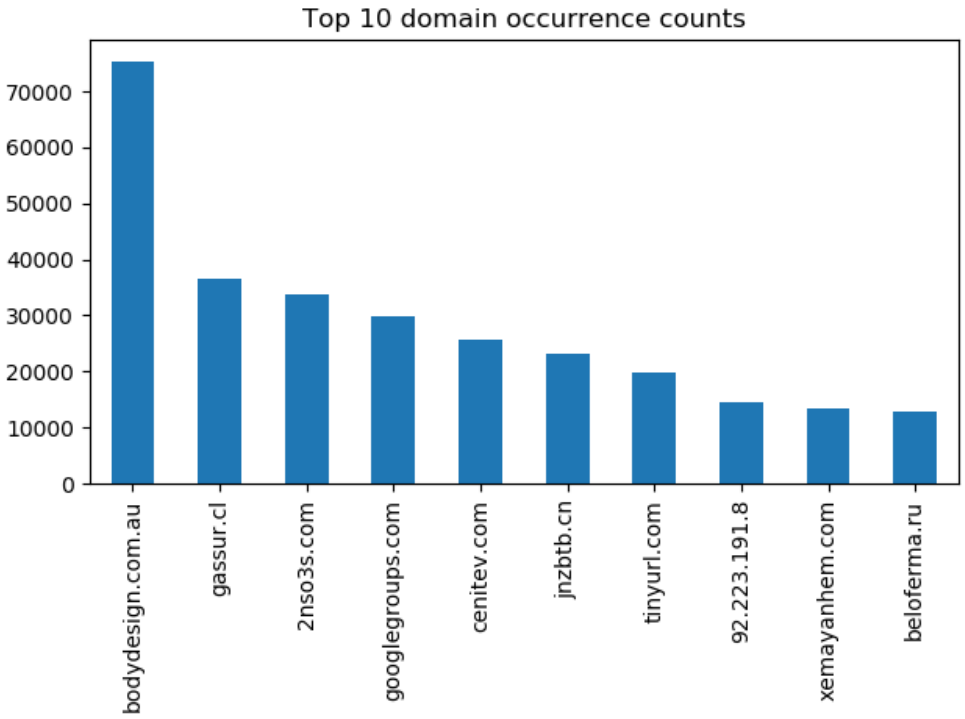
Elaborate

Maybe use top50 instead of top10

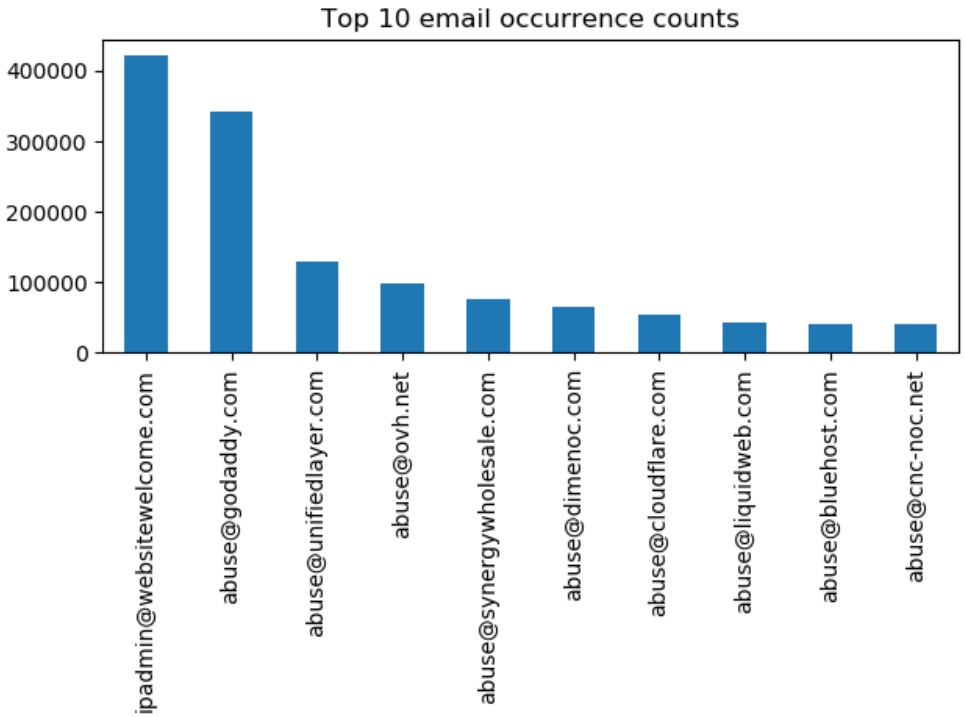
Combine similar conclusions



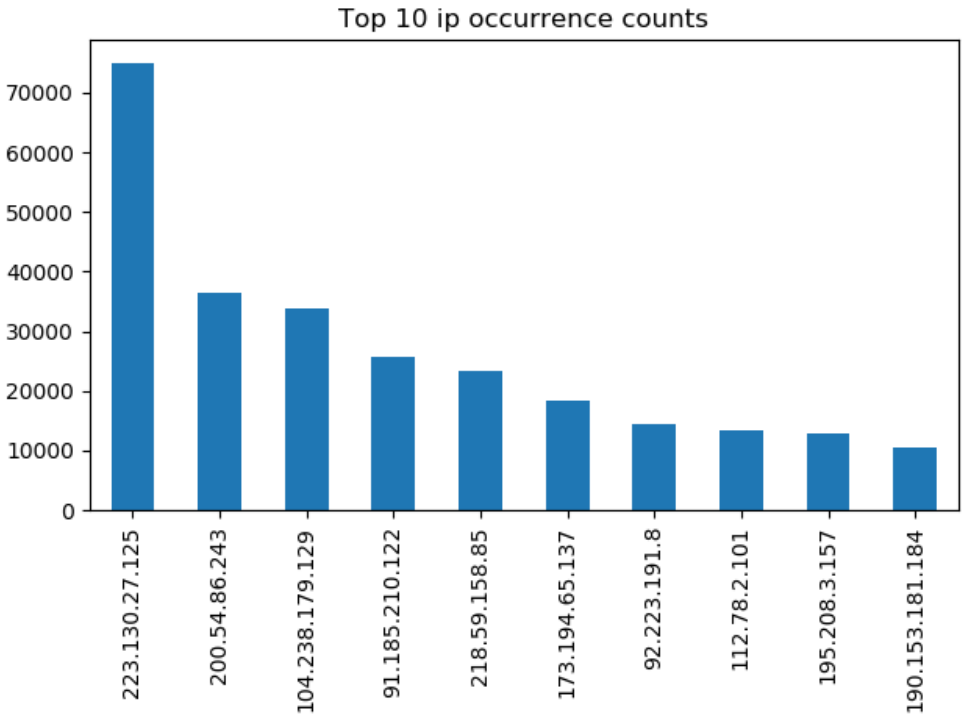
- US is signifficantly high
- Blocking traffic from a specific country is not ususally effective. Only for specific organisations that don't have usually traffic from US.



- Block high occurrences urls



- Block email that are connected with phishing emails



- Block malicious IPs

References

[1] D. A. Baldwin, "The concept of security," *Review of International Studies*, vol. 23, pp. 5 – 26, 01 1997.

[2] R. West, "The psychology of security," *Communications of the ACM*, vol. 51, no. 4, p. 34, 2008.

[3] H. Siadati, S. Palka, A. Siegel, and D. McCoy, "Measuring the effectiveness of embedded phishing exercises," 2017.

[4] CybeReady. Assessing your phishing risks — what metrics should you rely on?n. [Online]. Available: https://cybeready.com/assessing-your-phishing-risks-what-metrics-should-you-rely-on

[5] K. R. (2018) The trouble with phishing. [Online]. Available: https://www.ncsc.gov.uk/blog-post/trouble-phishing

[6] CybeReady. Is click rate the right way to measure phishing simulation training? [Online]. Available: https://cybeready.com/phishing-simulation-training-right-metrics

[7] R. Fahey. Anti-phishing: Measuring phishing awareness training effectiveness. [Online]. Available: https://resources.

infosecinstitute.com/category/enterprise/phishing/phishing-countermeasures/anti-phishing-measuring-phishing-awareness-training-effectiveness/

[8] M. Korczynski, S. Tajalizadehkhoob, A. Noroozian, M. Wullinkz, C. Hesselmanz, and M. van Eeten, "Reputation metrics design to improve intermediary incentives for security of tlds," 2017.

[9] N. Abdelhamid, A. Ayesh, and F. Thabtah, "Phishing detection based associative classification data mining," *Expert Systems with Applications*, vol. 41, no. 13, pp. 5948 – 5959, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417414001481