**1.** **What security issue does the data speak to?**

PhishTank and Clean-mx provide data on phishing: from locations of compromised servers to URLs of phishing websites. According to Moore & Clayton (2007), phishing can be defined as follows: *“Phishing is the process of enticing people into visiting fraudulent websites and persuading them to enter identity information such as usernames, passwords, addresses, social security numbers, personal identification numbers (PINs) and anything else that can be made to appear to be plausible. This date is then used to impersonate the victim to empty their bank account, run fraudulent auctions, launder money, apply for credit cards, take out loans in their name, and so on.”*

The process of phishing often follows a typical pattern (Moore & Clayton, 2007). First, the victim receives an e-mail which resembles a legitimate e-mail from a bank or another trusted source. This e-mail typically contains a call to action, for example the victim should check a transaction of hundreds of euros. The user clicks on the link in the e-mail, which sends him to a fraudulent website that imitates the website of the bank or other source. If he fills in his data on the fraudulent website, this will be sent via webmail to the phishing criminals or stored in a plain text file on the website itself. The data is then sold to ‘cashiers’, who use it to empty the bank account or for a variety of other financial crimes. Often a ‘mule’, a middle man, is used to transfer and launder the money.

It is obvious that possessing such personal data opens many possibilities for financial crimes. But how large is the impact of that problem? According to Gartner (2007), about 3 billion dollars was lost by phishing attacks in 2007, with an average loss of $886 per incident. The most often attacked brands are PayPal and eBay.

Since phishing is one of the most lucrative cybercrimes, policy makers should develop a strategy to combat these criminal activities. This policy can be aimed at different stakeholders, such as:

* Internet Service Providers (ISPs): they should make sure that confirmed phishing URLs are blocked;
* E-mail providers: they should make it very difficult for phishing e-mails to pass through the spam filter;
* Banks and other institutions that can be spoofed: they should have clear safety policies and make it hard for phishing criminals to imitate their systems;
* Users of the services of these institutions: they should be aware of phishing tactics so they do not fall easily for the fake e-mails and websites.

In order to get more insight in the problem of phishing, to know what policies are most effective and efficient, and to come up with new, better policies, it is required to be able to measure various aspects of phishing. What aspects of phishing are crucial for the profitability and thus the existence of this crime? In this assignment, metrics will be proposed to measure these aspects.

First, ideal metrics on phishing for security decision makers are designed. In this first step the actual availability of data in practice is not taken into account yet –pretending all data to make good policy decisions is available. Unfortunately this is not the case at all. Therefore the metrics that exist in practice are observed. Lastly, the datasets of PhishTank.com and Clean-mx will be examined to assess what metrics can be derived from it.

**2. What would be the ideal metrics for security decision makers?**

The main purpose of metrics is to compare between alternatives. These metrics should provide useful information regarding security-related objectives.   
Rainer Böhme *[reference]* introduced a framework that covers the different components of decisions on security: ‘cost of security’, ‘security level’ (which together determine the security productivity curve) and ‘benefit of security’. By estimating the values of these different components, one can use them to enlighten the organisation by showing some type of progress. The organisation then can determine if action is needed or if there are changes to be made.

The estimations require a lot of data. Data regarding costs and benefits are usually internal data from a specific company. One has to know about the spendings of a company to determine the direct and indirect ‘costs of security’ and data about the results of these costs to determine the ‘benefit of security’. Since this kind of data is hard to get, this paper will focus on the ‘security level’ aspect. Usually data regarding the security level is easy to obtain. This data can contain information about for example the amount of attacks and success rates of these attacks.

Some metrics are better than others. The key to good security metrics is obtaining measurements that satisfy the following statements:

* Metrics should contain relevant information;
* Metrics should be reproducible;
* Metrics should be objective and unbiased;
* Over time, they should be able to measure some type of progression towards a goal.

Some metrics count items of a certain aspect in a region. For example, the amount of attacks in a country. These metrics by itself can be biased since the amount can be influenced by different factors. Therefore, to make these metrics more meaningful, they must be normalised. This way the influences of a certain factor are eliminated . For example, the amount of attacks per country can be heavily influenced by the amount of internet users per country. By dividing the amount of attacks by the amount of internet users, a measurement is derived which is unbiased towards the internet community of the country.

**Ideal metrics conducted from brainstorm and literature**

Green = we can do this with our data + some look-up data like GDP per country

Red = Maybe, we can try and see what comes out.

Blue = idea of Maaike, we’ll discuss it later. (Should we also state which metrics would be really great to have (like our top 2), but are not possible?)

| Metric | Explanation (include evaluation here already (I mean explain why this metric is so much better than existing ones)?) | Feasible or not | If not feasible, why? |
| --- | --- | --- | --- |
| Uptime of phishing website per country of origin | Shows for how long the website has been online differentiating from country to country. It could be useful to analyse the suppression response of the various countries in order to evaluate their law enforcement. | Feasible. |  |
| Uptime of phishing website per business sector targeted | Shows for how long the website has been online differentiating from country to country. Useful to see which sector pushes more for suppression. | Feasible, though hard. |  |
| Amount of phishing websites per business sector | Shows which sector is more/less targeted. | Feasible, though hard. | Targets should be categorized and clustered in business sectors manually, this would lead to some (still incomplete) information |
| Amount of phishing websites per business sector per country | Shows whether in different countries different sectors are targeted more/less. | Feasible if metric above is determined. |  |
| Effectiveness of phishing website during time | Meaning: what’s the trend (if there is one) in the number of people filling in the phishing websites with valuable information?  We calculate effectiveness as the number of victims divided by the total number of visitors of the page. | Feasible. Reference: Moore & Clayton, 2007. |  |
| Money lost per incident/country/target/time | Shows the amount of money lost due to phishing in the various dimensions. | Feasible/ unfeasible.  Reference: ”Measuring the cost of cybercrime”, 3.2 online banking fraud | It’s hard to determine the exact amount of money lost per incident. In some case the loss could be indirect. Also not every data is available.  It has been estimated. |
| Number of internet users per targeted/origin country | Useful to give a proper weight to country-based metrics. | Feasible. |  |
| GDP per targeted/origin country and presence of phishing (targeted /origin) | It could be interesting to study whether there is a correlation between a country’s wealth and attacks. | Feasible. |  |
| Distribution of attack goals: identity theft, money laundering, espionage etc. | Shows which goal is more less/popular in phishing attacks. | Not feasible. | Not all information is available.  Also multiple goals can be pursued with a single attack. |
| Distribution of attack goals per country (both as origin or target) | Shows whether different countries have different phishing problems.  E.g. countries with more modern economy could suffer more from home banking phishing rather than other kind of phishing. | Not feasible. | Not all information is available.  Also multiple goals can be pursued with a single attack. |
| Distribution of the lifespan of phishing websites during the days of the week | Shows whether the speed of taking down a phishing website depends on the day of the week it is reported, e.g. if the takedown is slower on the weekend. | Feasible.  Reference: Moore & Clayton, 2007 |  |
| Number of attacks per country/region, normalised by the number of local internet users | Shows where phishing (as origin and targeted country) happens more/less.  The metric should be normalised accordingly to the number of internet users of that country/region. | Feasible. |  |
| Number of attacks per specific target | Shows which company is targeted more. | Feasible. |  |
| Number and distribution of mules per country, normalised by the numbers of local internet users | Shows how many mules are present in the country. This metric should be normalised accordingly to the number of internet users of that country/region. | Not feasible. | The exact number of mules is really hard to measure. Data are scarce. It could be estimated. |
| Acceptance ratio of phishing emails through spam filters per e-mail client. | It would be useful to study how many phishing emails are blocked by the various e-mail clients (Gmail, Yahoo!, Hotmail etc.), in order to understand which one is better. | Partly feasible. | Data on how many emails are stopped by filters should be obtained in collaboration with e-mail providers. However, there are both false negatives and false positives which make the metric less reliable. |
| Acceptance ratio of phishing e-mails sent per criminal organisation (origin) | Shows how many of the phishing e-mails sent by a certain criminal organisation arrive to the potential victims. Maybe some organisations have different ratios of acceptance. | Not feasible. | It's hard to have exact data related to specific criminal organisations. |
| Average success rate per phishing email | Success rate: how many people fill in real data after receiving a certain mail out of how many people receive the emails | Not feasible. | People can have more email addresses and keeping track of every single mail is quite hard. We could get an idea from surveys though this is not precise enough for the ideal metric. Data will be available at the criminal side, but this is not accessible for obvious reasons. |
| Number of phishing emails received pro capita per country, per time unit | Shows how many emails a person receives on average per specific country in a determined period of time (a day, a week, a month, a year etc.). | Feasible, using for example data from ISPs and mail providers |  |
| Revenue per criminal organisation related to phishing activities. | Shows the profit of every single organisation gets from phishing related activities. | Not feasible. | It's hard to have exact data related to specific criminal organisations. |
| Amount of phishing sites per criminal organisation. | Shows how many online phishing websites an organisation has at a certain moment. | Not feasible. | It's hard to have exact data related to specific criminal organisations. |
| Amount of criminal phishing organisations per country. | Shows how many criminal organisations related to phishing exist in different countries. This value is useful both normalised to the number of Internet users of that country (or even the general population) and not normalised. | Not feasible. | It's hard to have exact data related to specific criminal organisations. |

**3. What are the metrics that exist in practice?**

In practice, most metrics are not as specific as in an ideal situation. They do not discriminate between business sectors, amount of internet users per area, and so on. It is obvious that the main goal of phishing information clearinghouses, like phishtank.com, is merely to expose phishing sites. Although sometimes it is possible to search for certain variables such as “target” and “country”, performance variables such as uptime are not kept track of consistently. It is therefore hard to retrieve an extensive database on which research can be conducted.

The variables, and therefore the conductable metrics, of phishing detection sites are based on incidents. They are hard to interpret without knowledge of underlying, sometimes hidden, causal relationships. The examples below are incomplete in that sense, since normalisation or underlying causalities are not taken into account.

Although no useful conclusions can be derived from these metrics, it does show that these phishing clearinghouses focus on measuring the prevalence and severity of phishing sites. One similar example of the use of such measures was found in literature (Doshi et.al., …).

PhishTank.com (2015) and Clean MX (2015) deliver data on the following statistics:

|  |  |  |
| --- | --- | --- |
| PhishTank | Clean MX | Doshi et.al. (…) |
| * Phishes Verified per Day * Phishes Submitted per Day and per Hour * Total Phishes Submitted / Valid / Invalid per month * Median Time to Verify (by the community) * Top 10 domains of valid phishes * Top 10 IP addresses of valid phishes * Top 10 Networks that host phishes * Popular Targets | * Phishes detected and closed per Day * Distribution of % phishing sites worldwide and per country * Nameserver distribution worldwide and per top 25 countries * Phishingsite distribution top 25 countries, per region | * Average Phishing URLs Per Day * Distribution of Obfuscation Types Used * Average Phishing Victims Per Day * Distribution of Phishing by Organization * Geographical Distribution of Phishing |

The Anti-Phishing Working Group (APWG, 2015) seems to do a somewhat better job in their quarterly reports. Some of their metrics actually do make distinctions on important factors such as industry sectors, or contribute to more insight when combined with other metrics.

|  |  |
| --- | --- |
| Not specific | More specific |
| * Number of unique phishing websites detected * Number of unique phishing e-mail reports (campaigns) received by APWG from consumers * Contain some form of target name in URL * No hostname; just IP address * Percentage of sites not using port 80 * Phishing by top-level domain * Malware infected countries (ranking between countries, and type of malware) * Measurement of detected malware | * Most targeted industry sectors *(discriminates between industrial sectors)* * Countries hosting phishing sites *(instead of looking at attacked countries)* * Phishing-based Trojans and Downloader’s Hosting Countries (by IP address) * Number of unique brands targeted by phishing campaigns *(instead of a total amount of attacks, this helps to provide insight in the distribution of attacks over brands).* |

Apart from the above examples of incomplete incident metrics, another phishing related type of metric is found in literature. This is about how well a technology is capable of recognizing phishing sites or e-mails. These metrics are based on controls and their performance and have some overlap with each other.

Performance metrics found in literature include:

* False positive rates (DeBarr et.al., 2013; Miyamoto et.al., 2008; Wardman, 2014);
* Area Under the receiver operating characteristic Curve (AUC) (DeBarr et.al., 2013; Miyamoto et.al., 2008) which measures accuracy;
* F(1) measure (DeBarr et.al., 2013; Miyamoto et.al., 2008) which can be seen as a weighed average of precision and recall;
* Precision, accuracy and recall (DeBarr et.al., 2013).

Searches in databases like Scopus and IEEE with search terms “phishing” and “metric\*”only deliver one page with results – which is not much. When “metric\*” is left out, suddenly over 5 pages of results come up. While it is never good to jump to conclusions, it does give food for thought regarding the status quo of experience and performance evaluation in phishing management.

*An interesting question that remains is how the submitted phishing sites were detected. Is it because of excellent technological systems, or because of high employee awareness?*

**REFERENCES**

* APWG (2013). *APWG Phishing Activities trend report 2013Q2.* Retrieved from http://docs.apwg.org/reports/apwg\_trends\_report\_q2\_2013.pdf on September 17, 2015.
* Clean MX (2015). *Phish Stats.* Retrieved on September 15 via <http://support.clean-mx.de/clean-mx/phishstats>
* DeBarr, D., Ramanathan, V., & Wechsler, H. (2013). Phishing Detection Using Traffic Behaviour, Spectral Clustering, and Random Forests. *Proceedings of* *ISI 2013, June 4-7, 2013, Seattle, Washington, USA.*
* Clayton, R. and Moore, T. (2007). Examining the impact of website take-down on phishing. *e-Crime '07, Proceedings of the ant-phishing working groups, 2007, New York, USA.*
* Gartner (2007). *Gartner Survey Shows Phishing Attacks Escalated in 2007; More than $3 Billion Lost to These Attacks.* Press release, December 17, 2007, Stamford, Connecticut, USA. Retrieved from: <http://www.gartner.com/newsroom/id/565125>
* Doshi et.al. (?)
* Miyamoto et.al. (2008)
* PhishTank.com (2015). *Stats*. Retrieved on September 15 via <http://www.phishtank.com/stats.php>
* Wardman (2014)

**Sources:**

<http://www.isaca.org/Journal/archives/2005/Volume-2/Documents/jpdf052-how-can-security.pdf>

<https://books.google.nl/books?hl=nl&lr=&id=Af8F00gTRN4C&oi=fnd&pg=PP3&dq=security+metrics&ots=NolTtr5LtP&sig=Ae2U4qubuqth_UreHJwLPY4SHkg#v=onepage&q=security%20metrics&f=false>