EoS - Security Metrics

Noël Keijzer (s1602349)
Tom Leemreize (s1745506)
Joost Prins (s1723545)

September 2019

1 Introduction

With the rapid growth in digitalisation of the world, the importance of digital security has grown significantly. With the rise in importance of digital security, a need for measuring the digital security of an individual or entity has grown as well. Measuring digital security starts with having data about your security. Thus it is very important to keep logs of your digital activity. In this assignment we are tasked to analyze such data and draw conclusions from them. Our group analyzes the malware droppers data set that is provided by the course.

In this document, the analysis will be divided into different sections to make sure it is correct and easy to follow. First, in section 2 malware droppers and its security issue is defined. Second, section 3 gives an overview of the ideal metrics regarding malware droppers that were discovered during a literature review. Third, section 4 discusses the actual metrics that are retrieved from the given data set and which variables can be estimated from these metrics. Finally, section 5 concludes the research.

2 Malware droppers

According to Naval et al. malware has become a major threat to the cyberspace [10]. Contemporary anti-malware solutions are often able to detect malware before they can cause damage to systems, however by utilising malware droppers, these infections can still take place, regardless of anti-malware solutions [13]. Malware droppers are a piece of software that is used to infect devices with the threat which the author wants to spread [15]. Malware droppers can deliver ransomware, a remote administration tool (RAT) or any other piece of malware. Malware can cause damage to the reputation of companies as well as resulting in financial losses [14, 7, 3]. This directly implies that malware droppers are

a threat as well, as they can deliver a malware payload onto a target machine without detection.

The security issue that malware droppers speak to is that of malware infection itself. As mentioned above, malware infection is a problem that can have major consequences for companies and individuals. In order to mitigate this problem a company can implement both technical as well as organizational measures. An example of a technical measure could be the deployment of anti-malware solutions, or a blacklist for unknown domains. Organizational measures include training employees to recognise suspicious e-mails, which could potentially include links to malicious programs, such as malware droppers.

3 Ideal Metrics

In this section the ideal metrics are described. The ideal metrics for malware droppers are distilled from a literature study that was done on the topic of security metrics in combination with malware droppers. An overview of the ideal metrics found in the literature study is given along with further explanation of these metrics. A summary of the literature study is given in Table 1.

Miani *et al.* looked at different metrics that could be retrieved from IDPS's and found that metrics 1, 2, and 3 from Table 1 were the most telling for the security of an entity [9].

In [1] the results of 46 papers are combined to produce the best metrics to identify the security of an entity. These resulting metrics are metric 4, 5, and 6 in Table 1.

Ahmed *et al.* listed several cybersecurity metrics for use in healthcare IT systems, however these metrics could be applied to other IT systems as well [2]. Metrics 7 and 11 could be used to spot data exfiltration, which is especially important with laws such as the GDPR. GDPR violations could result in both financial loss due to fines, and damage of the company's reputation. Depending on the organization, metric 8 could also provide a lot of insight into malware droppers. Organizations which do not have servers outside their geographical area should not connect to IP addresses in those areas. If this is the case, it could indicate a security hazard.

In [11], Skopik *et al.* described several metrics, including the financial loss caused by the incident and the time until the incident was discovered, which provide insight into the severity of an incident.

Takamura *et al.* did a study into the definition of security of NASA mission operation centers (MOCS). In a NASA MOC, security is of utmost importance because it is the critical element in NASA's space missions [12]. Takamura *et al.* Describe Metric 16, 17, 18 of Table 1 as the most important metrics to have in such an environment.

Metrics 19 and 20 are related to the domain names and registrars used for the malware droppers. Hao *et al.* found that 46% of spam domains come from two registrars, so this information could prove vital in blocking domains used for malware droppers, and thus preventing infection [6].

Metric Nr.	Metrics	Study
1	Number of (distinct) attackers per week	[9]
2	Number of (distinct) objectives per week	
3	Number of (distinct) signatures per week	
4	Current number of untrusted network connections	[1]
5	Total number and number of unsuccessful connection attempts to authenticate devices communicating via untrusted networks per week	
6	Number of access retries associated with a source address recorded by each access point	
7	Volume of outbound traffic	[2]
8	Volume and number of IP addresses connecting to your network from outside of your geographical area	
9	Number of simultaneous logins by the same user from different locations that has not been detected by the network security tools	
10	Number of unknown accounts with elevated privileges found on compromised systems	
11	Number of hosts communicating with external networks on non-standard ports	
12	Damage of attack (in euros)	[11]
13	Incident date	
14	Time until discovery	
15	Time until recovery	
16	Number of new vulnerabilities in the past thirty days	[12]
17	Number of open vulnerabilities in the past thirty days	
18	Number of closed vulnerabilities in the past thirty days	
19	Number of malicious domains per registrar	[8]
20	Number of malicious domains per TLD	

Table 1: Literature overview of ideal metrics

4 Practical Metrics

This section describes the metrics that can be gathered from the provided data set. First, the data set will be explained. Then, the different metrics from section 3 that are applicable to this data set are measured. Finally, visualizations will be given of these metrics.

The data that is provided is an access log containing when infected computers

download the payload of the dropper. The access log is in the format of date-time stamps along with URLs. A snippet of the data is given in Listing 1.

Listing 1: Snippet of malware dropper data

```
,2013-08-24 07:42:04,http://www.intro2seo.com/-.php
,2013-11-05 13:30:09,http://www.007museum.com/movie.htm
,2013-11-24 00:00:15,http://solariumibg.com/novi.html
,2013-11-24 00:30:04,http://paw.compnet.com.pl/h1l.html
,2013-11-24 01:00:09,http://www.doucetpol.net/taille.html
```

The metrics mentioned in the previous chapter are metrics that may not be applicable to the data provided. There are however a lot of practical metrics that can be used to quantify the metrics mentioned in the previous chapter. In this chapter we will be going over each metric applicable to the data, and how one can use practical measurements to quantify it. Additionally, we will provide metrics of our own.

Many of the metrics described in section 3 are not applicable to the data as not enough information is available. Metrics 1 through 3, for example, cannot be applied to the data as we cannot tell what was targeted or by whom the attacks were executed. Other metrics, such as metric 8, cannot be executed as no data is available of the targets themselves, so nothing can be said about the geographical location of the target. Because of the lack of available data, most of the metrics can unfortunately not be applied.

Even though no data was available on the attackers, metric 1 can still be applied when looking at the different URLs. This also holds for metric 16, where a new vulnerability can be seen as a new URL in the logs. By making these small modifications, some metrics found in literature can still be applied on our limited dataset. Additionally, metrics 13 and 20 were also applied on the dataset. For metric 20 some extra processing of the dataset had to be done, as the TLD for each domain had to be separated and compared to the list of total domains for that TLD. Aside from the metrics of the literature, we also looked at the country to which the IP address of URL belongs, as this might indicate that some countries host more servers serving malware. This information can in turn be used in blacklists or other technical countermeasures.

After processing the data set and extracting IP addresses, Top Level Domains (TLDs) and countries of origin it was possible to analyze a few of the metrics described above. In Figure 1, one can see the distribution of the TLDs used by droppers between 2014 and 2016. This distribution was made by summing up the amount of unique domains in the dataset and grouping them by TLD. It was then normalized by dividing each TLD by its sum of total assigned domains[4]. This results in a dataset that shows the percentage of malicious users on a certain domain.

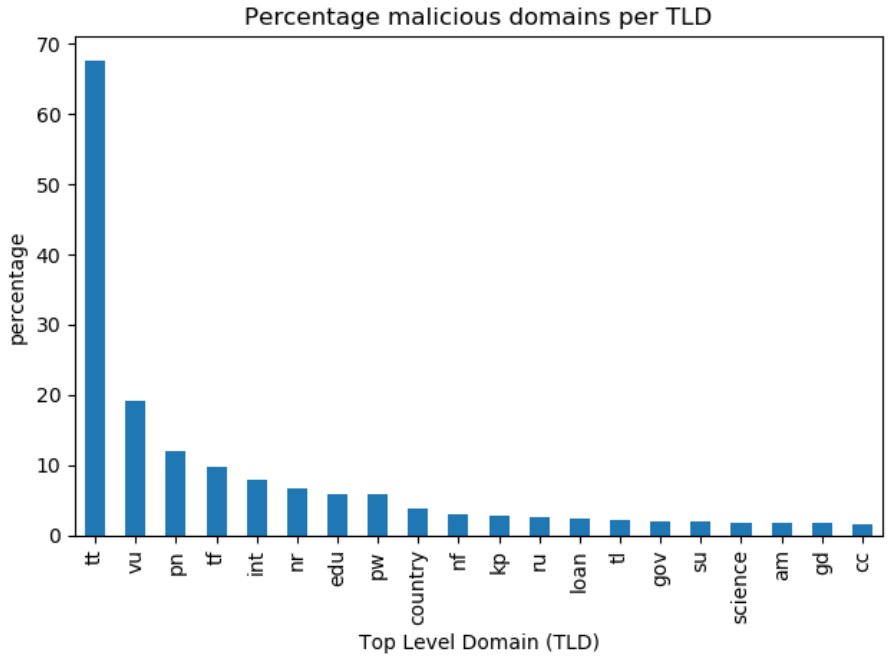

Figure 1: Percentage of malicious domains per TLD

As one can see from Figure 1 there are some TLDs for which the percentage of malicious domains is very large. For example for the .tt domain, malicious domains from the dataset used in this assignment make up almost 70% of the domain space. From this we can conclude that there are several TLDs that pose a particularly high risk to users, which should be avoided. This information can be used when informing employees of companies about malware. When giving a training one could mention that links from these particular TLDs often contain malware and thus that the employees should take extra care when clicking links from these domains. When taking more precautions a company could opt to block these TLDs altogether.

In the histogram in Figure 2 one can see the percentage of unique IP addresses used by malware droppers per country relative to IP address space of that country[5]. From the histogram one can see that compared to other countries Hong Kong has a reasonably high percentage of dropper domains. However this is still only 0.08%. Thus, it is not really feasible to mark Hong Kong domains as potentially dangerous. We learn from this that there is not one single country that is preferred by malicious actors to host their malware, instead it is more of a global epidemic.

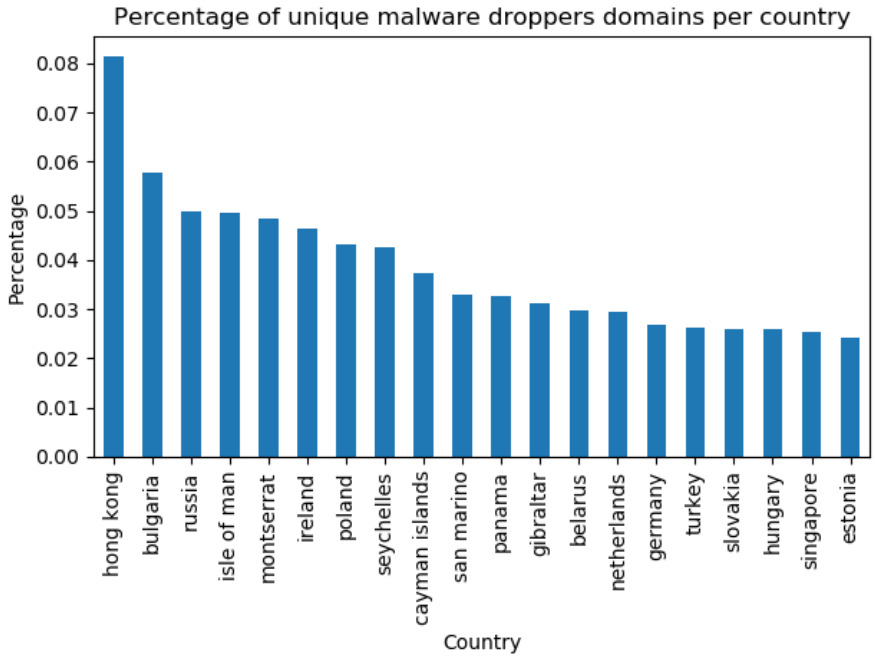

Figure 2: Percentage of malicious domains per country

In order to get better insight into when most malware infections occur, a plot was made showing the amount of clicks aggregated per hour of the day. This plot can be seen in Figure 3. In this plot it can be seen that the most amount of infections actually occur during the evening and at night. Thus, it is more likely that an employee of a company clicks on a malicious link during the night than during work hours. Using this data one can better inform employees that they need to pay attention, especially as the day goes on (and when they work at home during night time).

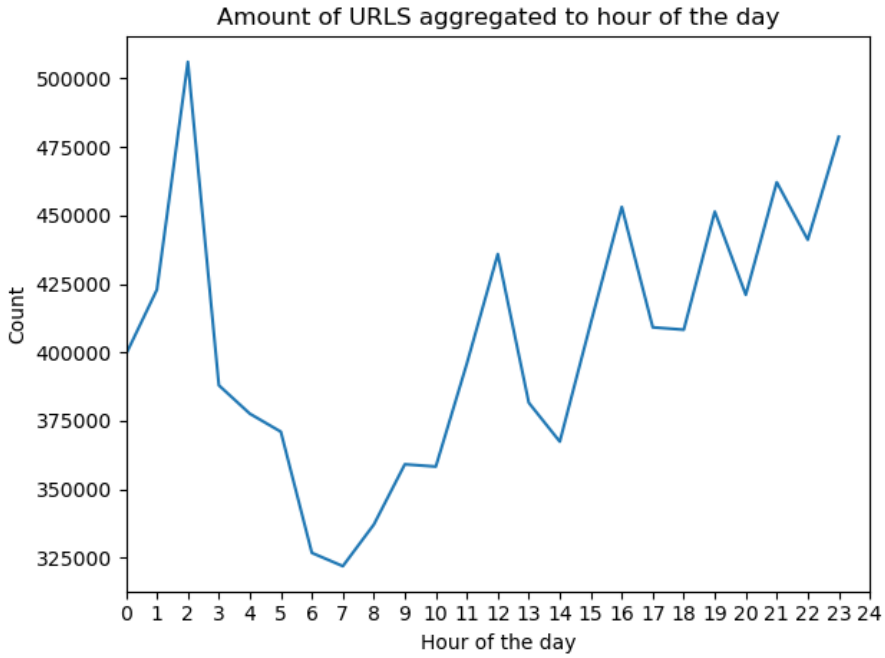

Figure 3: Amount of malware dropped per hour of the day

5 Conclusions

This article reports about the research that was done on malware droppers. The research shows that malware droppers are the source of a lot of malicious behavior. From a literature study we have learned which ideal metrics that one can use.

The literature has shown several metrics security decision makers could use in order to find the areas that need more attention. These ideal metrics can, however, not always be applied, depending on the data that is available. For the malware droppers data set very few of the ideal metrics could be applied. Those that could be applied did however prove useful, as they have indicated several areas with increased risk of malware droppers, such as the Top Level Domain of a website.

From the data set we have been able to derive some key metrics that can be used to make decisions about how to deal with the security issue. Firstly, there are several Top Level Domains that consist for a large part of malware droppers. Thus, these domains should be avoided or blocked whenever possible. Second,

the malware droppers are spread quite evenly across countries. Thus, there is not one single country that needs to address the issue, instead it is a global issue. Finally, we have learned that the amount of malicious clicks (and thus the likelihood that someone clicks a malicious link) increases as the day goes on and is highest during the night time. This data can be used to educate users on when and what to pay attention to and should attribute to enabling people to make safe use of their computers.

References

[1] "2011 Future of Instrumentation International Workshop, FIIW 2011 - Proceedings". In: *2011 Future of Instrumentation International Workshop, FIIW 2011 - Proceedings*. 2011.

[2] Y. Ahmed, S. Naqvi, and M. Josephs. "Cybersecurity Metrics for Enhanced Protection of Healthcare IT Systems". English. In: *International Symposium on Medical Information and Communication Technology, ISMICT*. Vol. 2019-May. 2019. URL: www.scopus.com.

[3] *Cybersecurity Ventures Official Annual Cybercrime Report*. [Online; accessed 21. Sep. 2019]. Sept. 2019. URL: https://cybersecurityventures.com/hackerpocalypse-cybercrime-report-2016/.

[4] *Domain Count Statistics for TLDs*. [Online; accessed 21. Sep. 2019]. Sept. 2019. URL: http://research.domaintools.com/statistics/tld-counts/.

[5] *DomainTools Internet Statistics - IP Addresses*. [Online; accessed 21. Sep. 2019]. Sept. 2019. URL: http://research.domaintools.com/statistics/ip-addresses/.

[6] Shuang Hao et al. "Understanding the Domain Registration Behavior of Spammers". In: *Proceedings of the 2013 Conference on Internet Measurement Conference*. IMC '13. Barcelona, Spain: ACM, 2013, pp. 63–76. ISBN: 978-1-4503-1953-9. DOI: 10.1145/2504730.2504753. URL: http://doi.acm.org/10.1145/2504730.2504753.

[7] *Is cybercrime the greatest threat to every company in the world?* [Online; accessed 21. Sep. 2019]. Sept. 2019. URL: https://www.csoonline.com/article/3210912/is-cybercrime-the-greatest-threat-to-every-company-in-the-world.html.

[8] M. Korczyński et al. "Reputation Metrics Design to Improve Intermediary Incentives for Security of TLDs". English. In: *Proceedings - 2nd IEEE European Symposium on Security and Privacy, EuroS and P 2017*. Cited By :6. 2017, pp. 579–594. URL: www.scopus.com.

[9] R. S. Miani et al. "A Practical Experience on Evaluating Intrusion Prevention System Event Data as Indicators of Security Issues". In: *Proceedings of the IEEE Symposium on Reliable Distributed Systems*. Vol. 2016-January. 2016, pp. 296–305.

[10] S. Naval et al. “Employing Program Semantics for Malware Detection”. In: *IEEE Transactions on Information Forensics and Security* 10.12 (2015), pp. 2591–2604.

[11] F. Skopik et al. “Establishing national cyber situational awareness through incident information clustering”. In: *2015 International Conference on Cyber Situational Awareness, Data Analytics and Assessment, CyberSA 2015*. 2015.

[12] E. Takamura et al. “Information security considerations for protecting NASA mission operations centers (MOCs)”. In: *IEEE Aerospace Conference Proceedings*. 2015.

[13] *Trojan.Dropper | Symantec*. [Online; accessed 21. Sep. 2019]. Sept. 2019. URL: https://www.symantec.com/security-center/writeup/2002-082718-3007-99.

[14] *”WannaCry” ransomware attack losses could reach $4 billion*. [Online; accessed 21. Sep. 2019]. Sept. 2019. URL: https://www.cbsnews.com/news/wannacry-ransomware-attacks-wannacry-virus-losses/.

[15] *What is dropper? - Definition from WhatIs.com*. [Online; accessed 21. Sep. 2019]. Sept. 2019. URL: https://whatis.techtarget.com/definition/dropper.