Spamhaus

Security Metrics

Felix W. Dekker, Nick Ho Sam Sooi, Jorai Rijsdijk, Jakob S. Kok

Group number
6

GitHub
https://github.com/Erackron/EoC/

Introduction

Spamhaus is a not-for-profit organisation which collects data regarding several cybersecurity topics such as spam, botnets and malware. For this assignment, one of these datasets will be analysed in order to assist security decision makers. The Spamhaus dataset contains a list of machines which have been flagged as part of a botnet. For each machine, it also includes the corresponding IPv4 address block, location, domain name, timestamp, Autonomous System Number (ASN), and the reason for listing. Therefore, the dataset is able to provide insight into relevant threats with regard to the various types of botnets. Consequently, the goal of this assignment is to take the first steps towards the analysis of this dataset in order to assist actors responsible for making security-related decisions.
First of all, the relevant security issue will be discussed. The security issue describes who are the most prominent attackers, which threats will be focused on and which assets can be harmed if the attackers are successful. Then, the ideal metrics for measuring cybersecurity for the security issue will be presented. These metrics show how the security level would be measured in an ideal situation without any restrictions on the availability and quality of data. Furthermore, the metrics which are used in practice by governmental bodies are discussed. Finally, the metrics that can be used to measure the security level from the Spamhaus dataset are introduced. The value of the metrics will be calculated from the dataset and used to generate an advice to the security decision makers in a later stage of the research process.

Security issue

For this assignment, the perspective of the cybersecurity agency of a single country will be taken. This agency has been given the task to analyse the general level of security in the country in order to assess the threat level to its core resources. In particular, the focus lies on threats to assets owned by its residents and the companies rather than the government agencies themselves.

Furthermore, the assignment specifically targets threats caused by botnets, since the Spamhaus dataset mainly contains data regarding infections of this type. The agency would like to investigate how botnets affect the security level of the country and investigate how the situation can be improved, if necessary.

Firstly, the *attacker* is a person or organisation that tries to create botnets. The motivation of the attacker is not significant, though many botnets are created to be leased to others to deliver a payload to compromised systems. However, the entity responsible for spreading the botnet itself will be considered as the attacker in this analysis.

The main *threat* being addressed in this analysis is the infection of a machine by a botnet. If a threat is realised, this may result in damage to the assets, which are described below.

The *assets* that may be harmed by the threat are the availability of computation power and the integrity of business-critical data. If machines are infected by a botnet, they will spend time serving the master of the botnet rather than performing their regular tasks. Furthermore, business-critical data could be attained, altered, or made unavailable by the attacker. This harms the confidentiality and integrity of the business-critical data, since the attacker is not authorised to access the data and is able to make invalid changes to the data.

Ideal metrics

In this section, the ideal metrics with regard to measuring the security level of the country as a result of botnet threats will be described. These metrics are ideal in the sense that they are not limited by availability and correctness of the data they rely on.
The metrics presented here have been classified in two ways; by security level concept and by implementation type. First, the metrics will be discussed in their relevant context. After that, the metrics will be placed in the theoretical framework, shown in Table 1.

IM1: The amount of money spent on cyber security relative to the IT budget. This often gives a fair indication of the security level of a business.

IM2: The percentage of companies that have antivirus installed. This metric also provides an indication of how well the general business in a country is protected against known threats, since most antiviruses normally protects against these threats.

IM3: The percentage of companies that have up-to-date patches. Patching your systems to the latest configuration is one of the most efficient security mitigation actions that can be done, since it provides security to newly emerged threats. Obtaining this percentage will provide a country with insights of it general security level.

IM4: The total number of vulnerabilities discovered during penetration testing per company. If this value is normalised by dividing it by the number of companies in a country, it will provide an approximation of how vulnerable the average company in a country is.

IM5: The percentage of servers being vulnerable to malware infections. This metric points out how vulnerable the servers in a country are towards malware.

IM6: The number of infected machines per million internet subscribers. This ratio correlates to the average awareness of users in a country.

IM7: The number of networks or autonomous systems infected over time, per botnet. This describes the likelihood of being affected, and points out how much awareness is needed regarding that specific virus.

IM8: Ranking autonomous systems by infection percentage. Measuring the value of this metric helps other autonomous systems raise better defences and deeper inspections towards packets incoming from that autonomous system and thus mitigate the risk of virus spreading from that autonomous system.

IM9: The amount of money lost relative to the value of the IT sector. As stated earlier the amount of money invested in security often reflects the security level. This applies to the money that is being lost as well. If the amount of money lost relative to the It sector is low, this could indicate a high level of security, even though there are many factors such as influencing this, such as the attack behaviour.

	Controls	**Vulnerabilities**	**Incidents**	**Losses**
Organizational	IM1			
Procedural	IM2	IM4		
Technical	IM3	IM5	IM6 IM7 IM8 IM9	IM10

Table 1: Ideal metrics

Metrics in practice

The metrics in practice are being used by real-world actors to measure the security level with regard to the threat of machines being compromised by an attack making use of botnets. Since the focus of this assignment lies with governmental actors, the metrics that are discussed are relevant for this type of actor to assess the security level with regard to the threat imposed by botnets.

The metrics in practice have been based upon three publications by governmental bodies regarding measuring cyber threats. The first publication that has been used is the Cyber Security Assessment of the Netherlands in 2017. In this document, the most relevant cybersecurity threats for the Netherlands are presented (Ministry of Justice and Safety, 2017). Secondly, a report published by the European Union Agency for Network and Information Security has provided a comprehensive analysis of methodologies to assess threats imposed by botnets and secure information systems against these threats (ENISA, 2011). Thirdly, the Information-Technology Promotion agency of japan has published a paper which gives an in depth description of threats imposed by botnets and provides measures to mitigate them (IPA, 2006). The metrics that have been derived from these reports are discussed and placed in the framework below.

MP1: The percentage of companies that have up-to-date patches. A lack of patching in companies causes a higher risk for the machines of the companies to be compromised since older versions of software often contain vulnerabilities which are known to adversaries.

MP2: Percentage of companies that have an Intrusion Detection System(IDS) in place. An IDS is an automatic security system which notifies the administrator of a system when an adversary is attempting to compromise the system by means of malicious activities or security policy violations.

MP3: Percentage of companies that have a specific security policy for threats imposed by botnets. By implementing a specialised security policy, companies can better protect their assets against these threats. For example, a security policy can contain recovery procedures for incidents.

MP4: Percentage of companies that have specifically trained personnel to deal with threats imposed by botnets. Companies with specifically trained personnel to manage these threats are better able to identify vulnerabilities and recover after incidents occur.

MP5: Percentage of companies that invest in security awareness of their employees. Awareness of employees is an important factor in improving the resilience against cyber attacks. Many incidents are caused by a lack of awareness, for example in the case of a phishing attack where an employee downloads malicious content.

	Controls	Vulnerabilities	Incidents	Losses
Organizational	MP4 MP5			
Procedural	MP3			
Technical	MP1 MP2			

Table 2: Metrics in practice

The analysis displayed above shows that the metrics used by governmental bodies primarily focus on the controls which can be put in place to mitigate the risks imposed by botnets. There is a lack in metrics for the measurement of the security level with regard to vulnerabilities, incidents and losses. Therefore, research should focus more on the development of metrics that enable security decision makers to measure metrics of these types. The Spamhaus dataset contains data that would enable such an analysis. In the following section, metrics will be developed to measure incidents with regard to the botnet threat.

Metric design

Finally, it is key to define metrics which can be used to generate advice to security decision makers at governmental agencies by analysis of the Spamhaus dataset. The data allow for the analysis of emergence processes and locations of specific types of botnets. Because of the inherent properties of the dataset, all metrics in this section would be classified as “technical” and “incidents”.
The following metrics will be used to analyse the Spamhaus dataset:

DM1: *Rank countries by infected machines per million internet users*
In order to assess the security level of the country with regard to incidents, countries will be ranked by the number of infected machines per Internet user. By doing so, the density of infected machines per country can be assessed. This gives insight into how one country compares to other countries. If one’s own country is ranked highly, more attention should be given to cybersecurity in general.

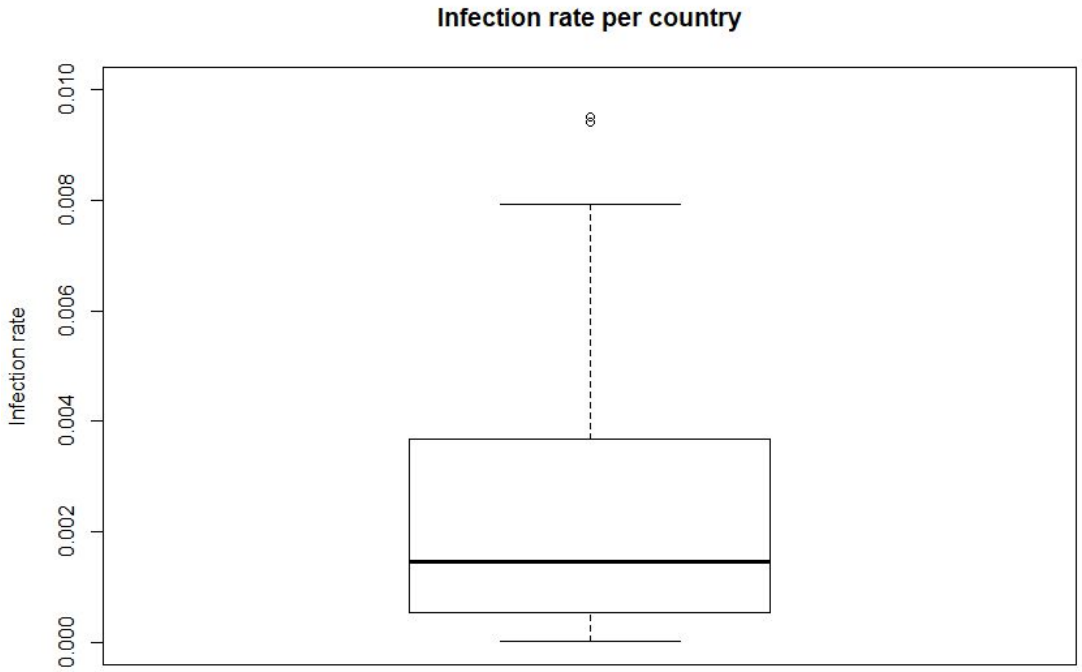
DM2: *Infection rate of an Autonomous System (AS) based on size*
This metric is calculated by dividing the number of infected machines in an AS by the number of IP addresses assigned to that specific AS. By looking at the AS’s of a specific country or region, it may reveal correlations between the size of an AS and its risk. On one hand, it could show that larger ASes have more threats because they are more exposed and therefore more attractive targets for attackers; but on the other hand they may also have more professionals and experience available to protect their systems.
As such, this metric provides a governmental agency with information on what types of AS are most at risk, and will thus help in deciding where to spend money to improve security.

DM3: *Infection rate per bot type per country*
The relative amount of infections per bot type gives insight into which botnets are currently the most dangerous, and indirectly tells something about where the current vulnerabilities are. For example, if the bots with the highest infection rates are all spread through email, then more should be invested in raising awareness on the opening of spam.

Metric evaluation

Country infection rate ranking

The data on the amount of Internet users per country is available in the repo as `countries.csv`. This data was taken from https://www.internetworldstats.com/list2.htm.



From the boxplot we can observe that 75% of the of the countries have an infection rate of less than 0.4%, and that almost all of the countries have an infection rate less than 0.8%, expect for a few outliers. The median is somewhere around 0.15%. These observations are as expected.

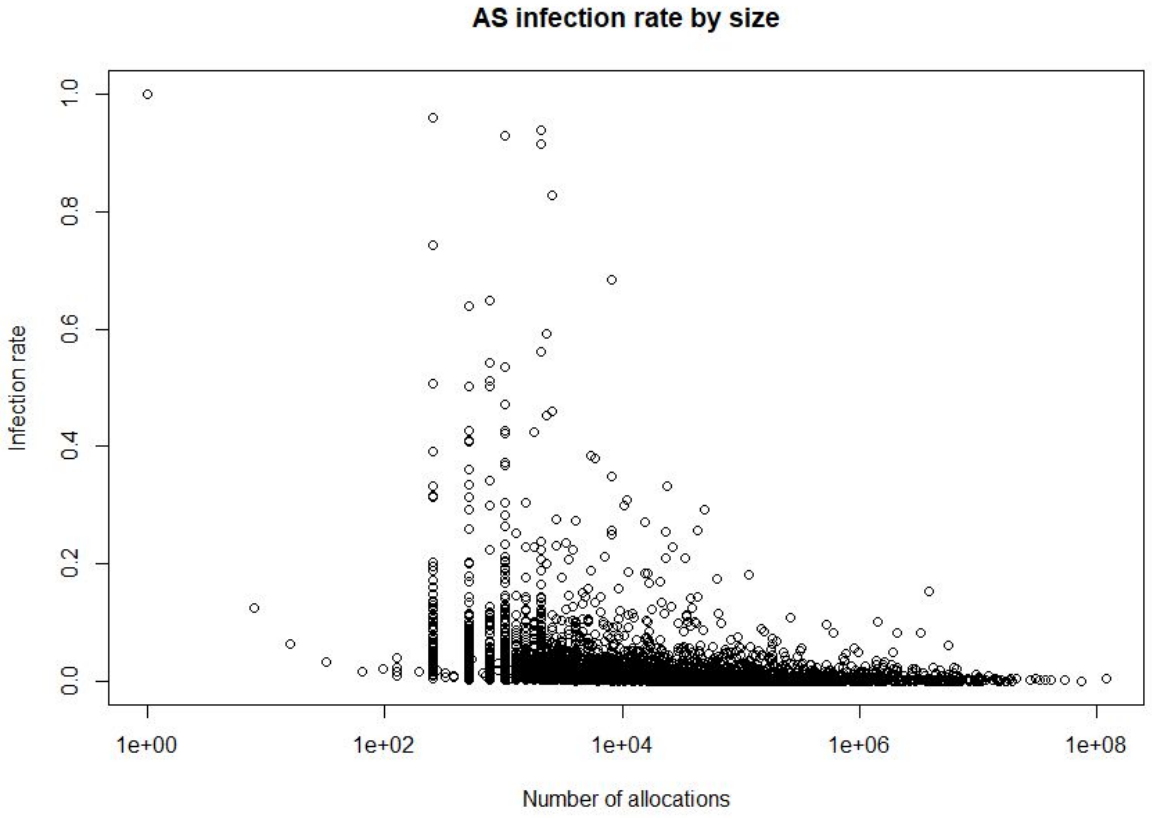
Country	**Internet users**	**Infection rate**
Montenegro	36,922	64.2%
Dominica	43,335	6.6%
Turks and Caicos Is.	14,760	2.1%
Virgin Islands, British	14,620	1.8%
Macedonia	1,280,132	1.7%
Guyana	295,200	1.7%
Vietnam	64,000,000	1.5%
Uruguay	2,017,280	1,5%
Belarus	5,204,685	1.0%
Romania	10,812,784	1.0%

Table 3: Top 10 countries with highest density of infected machines

We can see a trend that countries with very few internet users might have higher infection rates. These countries are mostly developing countries. These anomalies may be the result

of inaccurate measurements, or because there are disproportionately many non-resident infections.
The metric also shows that most of the infected machines actually comes from Vietnam, India and China. However, since these countries have millions of internet users, the infection rate obviously is not as high.

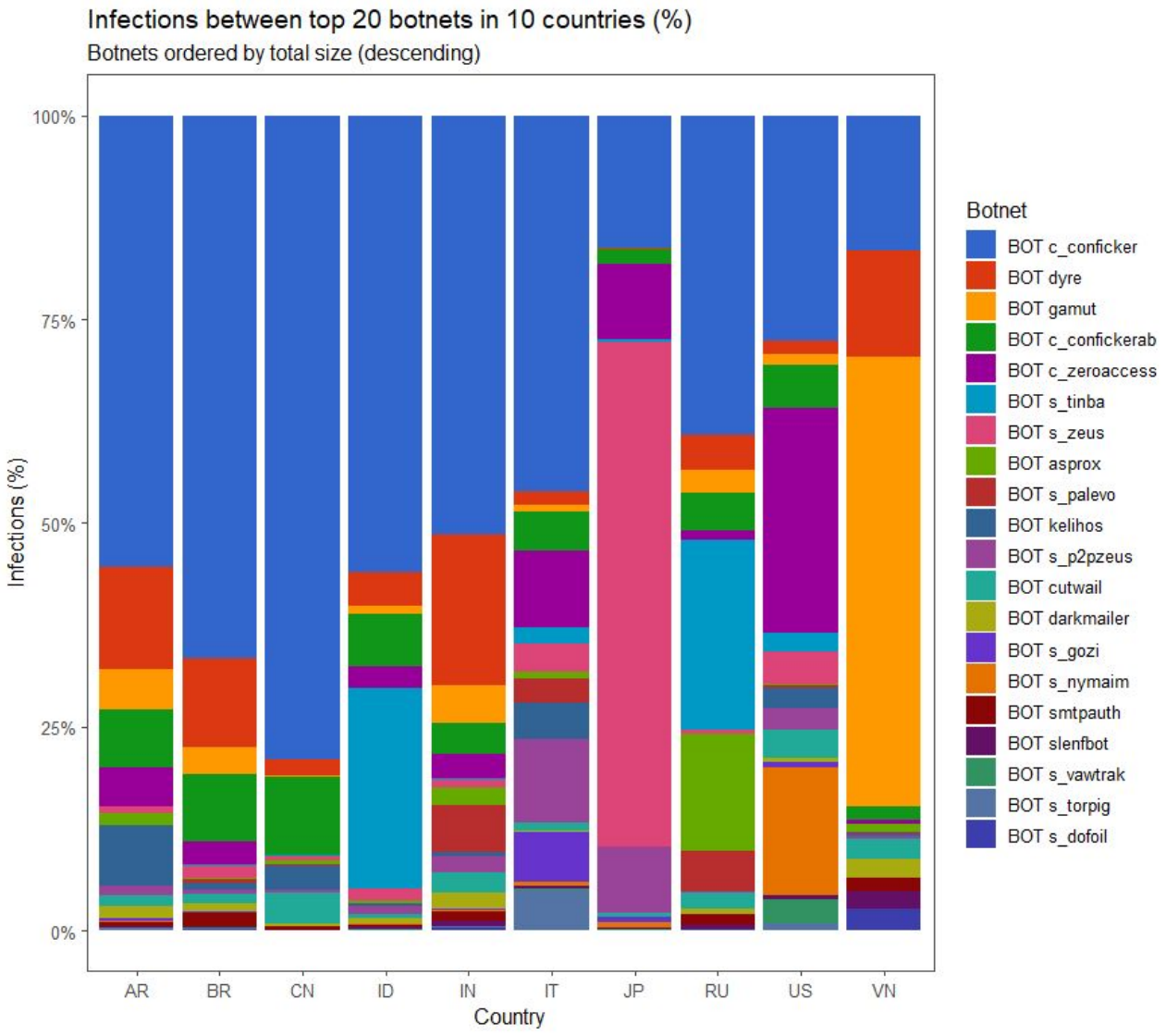
AS infection rate by size



The graph seems to show a correlation between the size of an AS and its infection rate. In particular, as the number of IP addresses available to an AS increases approaches several hundreds of millions, the infection rate approaches 0. This would seem to indicate that the operators of larger autonomous systems are better at protecting their systems against botnet infections, which might have a variety of causes.

However, because we divide by the size of the allocated IP blocks rather than by the actual number of machines in an AS, the decrease may also be the result of a decreasing usage rate of available IP addresses as the size of the AS increases, or it may be the result of both. Because we do not have data on the usage rate of available IP addresses in ASes, this metric is inconclusive.

Bot type infection rate per country



This graph shows the division between botnet infections within the countries with the highest overall infections. By limiting ourselves to just the 20 botnets with the highest overall infection rate, we can already distinguish different security threat landscapes between countries.

In Japan, there is a much higher than average activity of the ZeuS botnet, while Vietnam has a disproportionately high percentage of infections due to the Gamut botnet. Both ZeuS and Gamut spread by opening and (accidentally) downloading email attachments, and Japan and Vietnam should therefore increase efforts to inform people of these risks.

References

Ministry of Justice and Safety. (2017). Cyber Security Assessment Netherlands 2017: Digital resilience is lagging behind the increasing threat. Retrieved from https://www.ncsc.nl/english/current-topics/Cyber+Security+Assessment+Netherlands/cyber-security-assessment-netherlands-2017.html

ENISA. (2011). Botnets: Detection, Measurement, Disinfection & Defence. Retrieved from https://www.enisa.europa.eu/publications/botnets-measurement-detection-disinfection-and-defence

IPA. (2006). Countermeasures against bots. Retrieved from https://www.ipa.go.jp/security/english/virus/antivirus/pdf/Bot_measures_eng.pdf

IWS (2018). Internet world stats (internet penetration rates and population statistics. Retrieved from https://www.internetworldstats.com/list2.htm