Economics of Cyber Security

DRAFT Assignment Block 2

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

September 16, 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 [4]. Security is something that is not coordinated and prioritized 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 are being recorded from the first of July and ends 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[5]. 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.

An example of a DDoS attack 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[3].

Another task for which botnets can be used are 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 and not sending them from one device.

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 no 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* a incident has happened, metrics in the prevented losses category try to incorporate the economic impact that a 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 an better overview of the total level of security, instead of only testing the level of security when dealing with a specific threat environment.

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:

- Number of packets per given time unit
- 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 *not* benign 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:

- Number of login attempts
  - Number of successful login attempts
  - Number of unsuccessful login attempts
- Frequency of attacks

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.

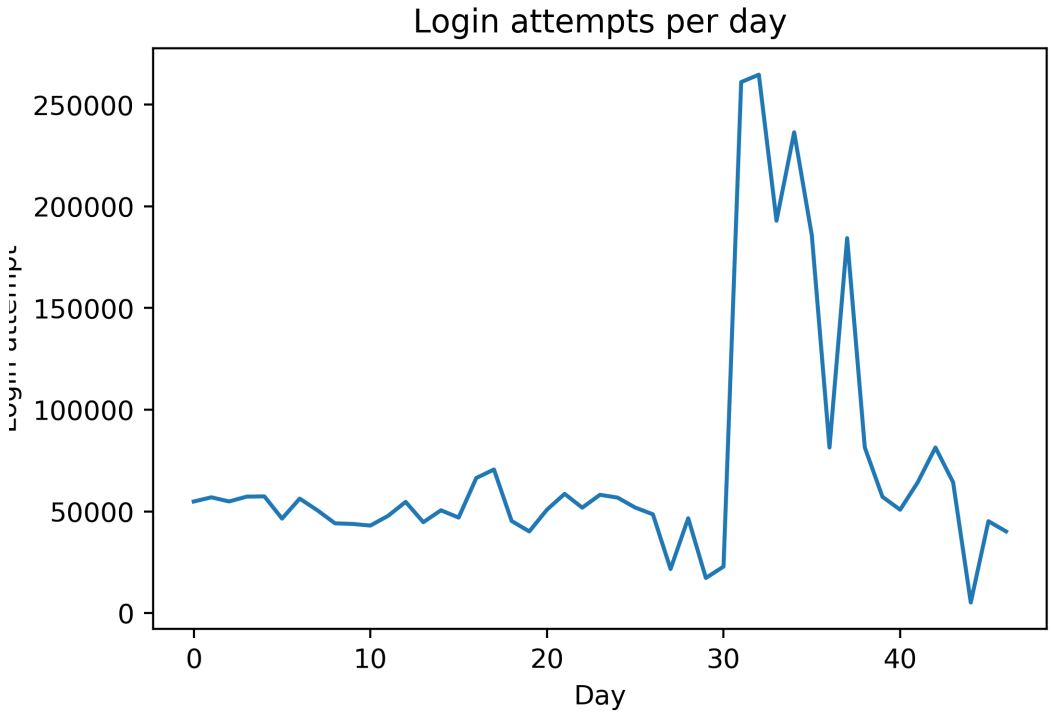
On the given data we can use some metrics described in section Existing Metrics, since we can still measure the most frequently occurring IP addresses and ports. Also the most frequently used ports can be measured. From the command list we can deduce the most frequently used/tried commands. Furthermore, 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.

6 Evaluation of Metrics

In this section we would like to display some graphs like:

- Most used source IPs
- Most used source ports
- Most used credentials
- Peak times/days of attacks



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] 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.

[4] 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.

[5] 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.