Universiteit of Twente

Economics of Security Course
AY 2018/2019

Internet of Things: Mirai

ASSIGNMENT 1

Authors:
Anna Prudnikova
Giovanni Maria Riva
Federico Casano
Sridar Bangalore

Contents

1 Introduction

2 Context and Scoping

2.1 What security issue does the data speak to?

IOT_Mirai dataset shows the number of IOT device scanned across different region over a period of six months. We have used IBM SPSS, Tableau and Rstudio to examine the dataset. Mirai is a malware that infect IoT devices turns into remotely controlled bots. They scan IPv4 address space of IOT devices that run port 23 and 2323, attempt to log-in using a universal default password or using dictionary based of IoT credentials and thus gain control of these devices. If an IOT devices are compromised then this could be a stochastic indicator, due to random nature of losses and these are events are driven by unknown attacker. Attacker can gain control over the device and use them as a bot to mount the attack on organization that would lead to impact of losing the asset. There is reduction in scanning the devices not only because of security measures but change in attacker behavior.

2.2 What would be the ideal metrics for security decision makers?

We saw many metrics are based on controls and vulnerabilities, since there are vulnerabilities organization can put controls in place but the threat may not be completely reduced. In IOT case, incident metric could be best metric to measure by decision makers, since they can show how existing controls are performing. Incident data can show how many devices are infected and thus determine the incident rate and analyze these incidents over time. Turning these metrics into ranking order can show how much weak is the devices. By having strong security policies incident rate could be controlled. Alternatively, all three metrics (control, vulnerabilities & incidents) can be aggregated and determine most variance in these metrics.

2.3 What are the metrics that exist in practice?

The dataset speaks of metrics from the number of affected devices by mirai. The basic metrics from this type of data is the number of affected devices (incidents) changing over time. In theory, this type of metrics should show if the security was improved over time, but in practice this type of metrics doesn't include the other factors, for example, the behavior of an attacker .

3 Methodology and Metrics

To have a better understanding of security level it is better to use the multiple sources of data and then aggregate them to receive the final metric. For our case, we can use statistics on IoT devices (the number of devices worldwide or per country). That is why in practice we can also use another types of metrics, as followed:

- **Number of affected devices changing over time.** This is a stochastic metric, that doesn't' take into consideration the behavior of an attacker;
- **Number of affected devices per country.** This is a more precise metric to try to understand, which countries have better security levels, still not efficient, because doesn't take into consideration the number of IoT devices or users in the specific country;
- **Number of affected devices per country taking into consideration the number of people in the country.** The more specific deterministic metric, that shows the rank of security level in the country;
- **Number of affected devices per country taking into consideration the number of IoT devices in the country.** The more specific deterministic metric, that shows the rank of security level in the country;
- **The overall percentage of IoT devices affected by mirai taking into consideration the number of IoT devices in the world.** This statistic is not precise, because not all IoT devices can be actually affected by mirai, but we only have statistics on overall bumber of IoT devices;
- **The percentage of IoT devices in 1 country affected by mirai taking into consideration the number of IoT devices in this country.** This statistic is not precise, because not all IoT devices can be actually affected by mirai, but we only have statistics on overall number of IoT devices.

Procedure - We started considering "What metrics could you extract from this data?" where this data are the information extracted from the sinkhole (dataset provided for the assignment). The sinkhole allows us to define how many machines (many of them or even all) are part of that specific botnet.

1- We started counting the infections obtaining the incident rate. For example the number of infected machines we see everyday. Proceeding in this way we can track this incident over times to see whether our security is improving. Anyway if the total amount of infected machines goes down, it does not necessarily means that our mitigation policy is valuable, the attacker behaviour influences highly this metric. Incident data is stochastic.

2. "How could we control for attacker behaviour?" We need to rank the networks, understanding how well or poorly each network does. Thus the rank metric would tell us that some countries have more infections than others. But "is their security worse?". Not necessarily, there are other factors involved that could influence the outcome/output. For instance, the more internet users in a country, the more infections. So we need to divide the number of infections by the number of users.

3. To reach more accuracy we planned to search how many IoT devices are online per country. In this way dividing again, but this time, the result (point 2) by the number of IoT devices for that country, we get a more specific data.

4. One further step could be analyze the ISP and vendors of IoT products worldwide. With the first analysis we would like to understand which ISPs implement Anti-DDoS security measures. As we studied during the Economics of Security course, ISPs are one of the main actors that could prevent attacks and improve overall cyber security worldwide. The goal of the latter analysis is figure out the policy adopted by the vendors. We would like to define the level of asymmetric information between customers and vendors, trying to deal with the problem of default credentials, main misconfiguration exploited by the Mirai malware.

4 Graphs

Each dataset's row contains the date and ip address of the infected machine. For evry month, we list the sum of the number of infected machines at the grouped by month as shown in Figure 1.

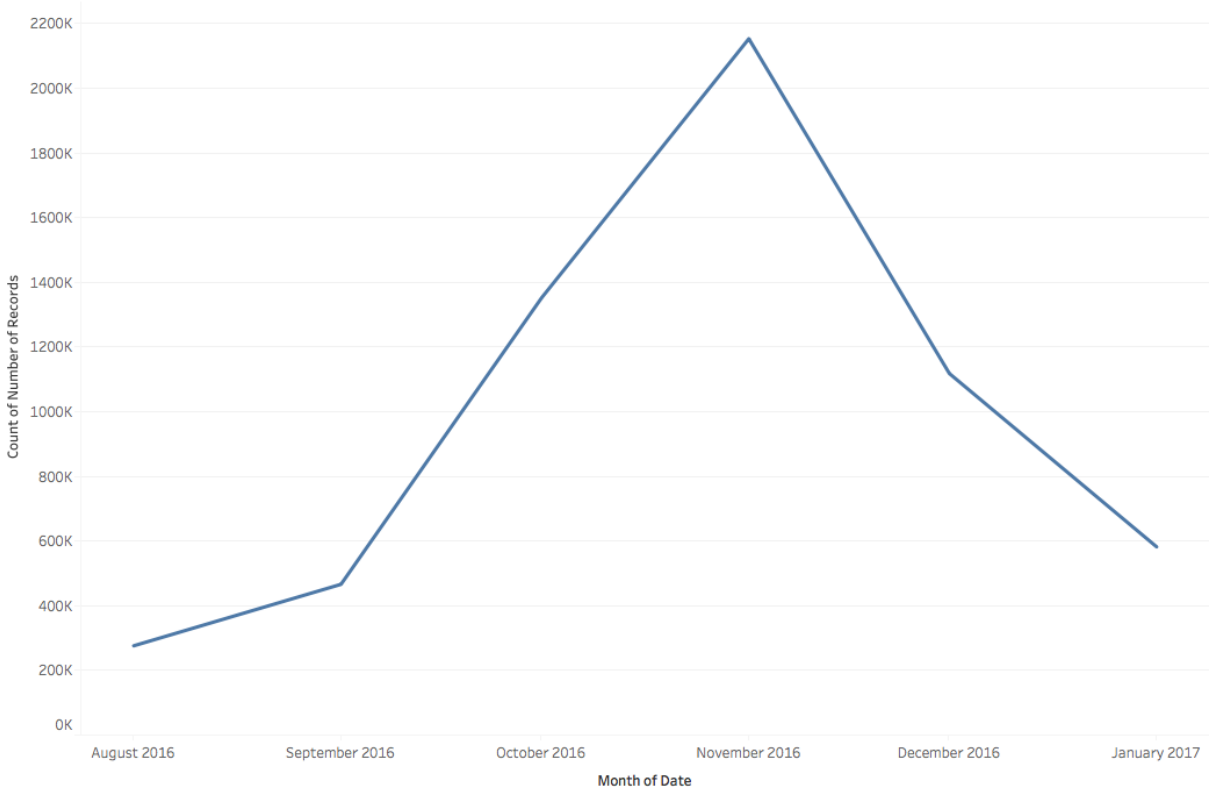

Figure 1: Number of total infected machines over months

In following graph we represents what the total number of infected machines over months. In this case, figures 2 and 3 show the sum of infected machines for each country in order to have an overview on the distribution of the malware.

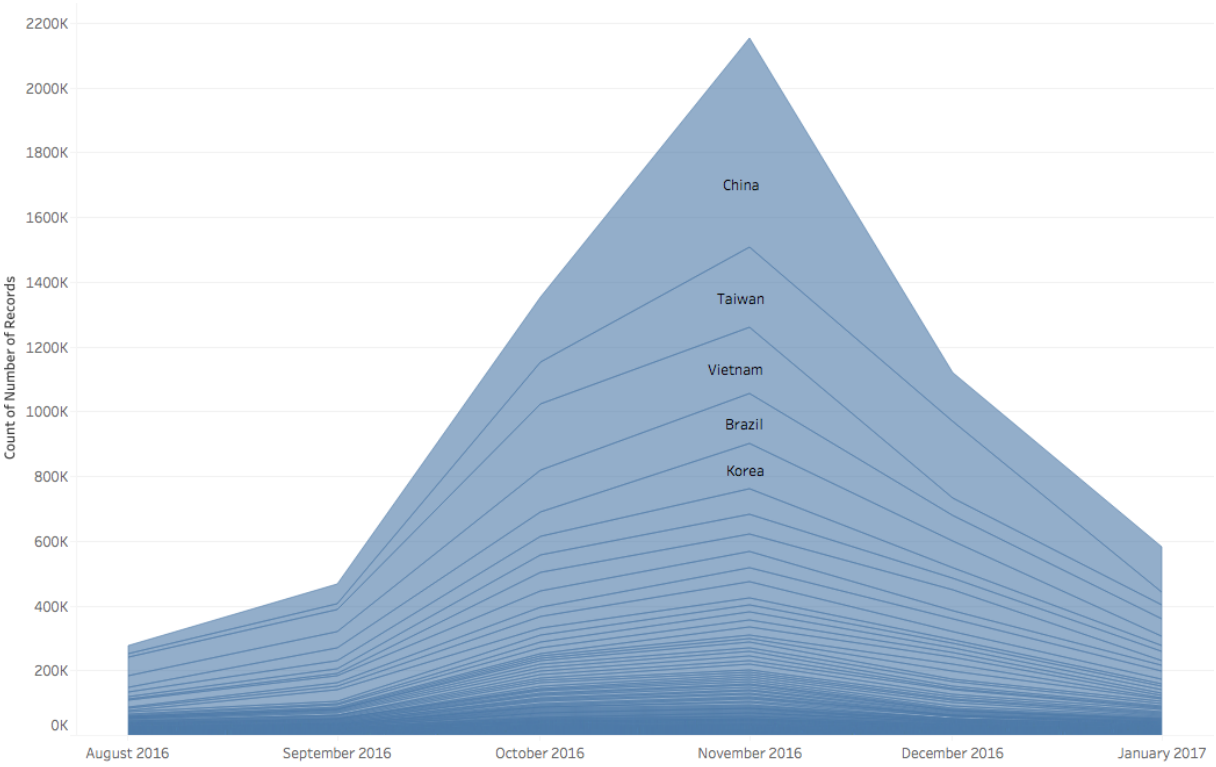

Figure 2: Number of total infected machines over months per country

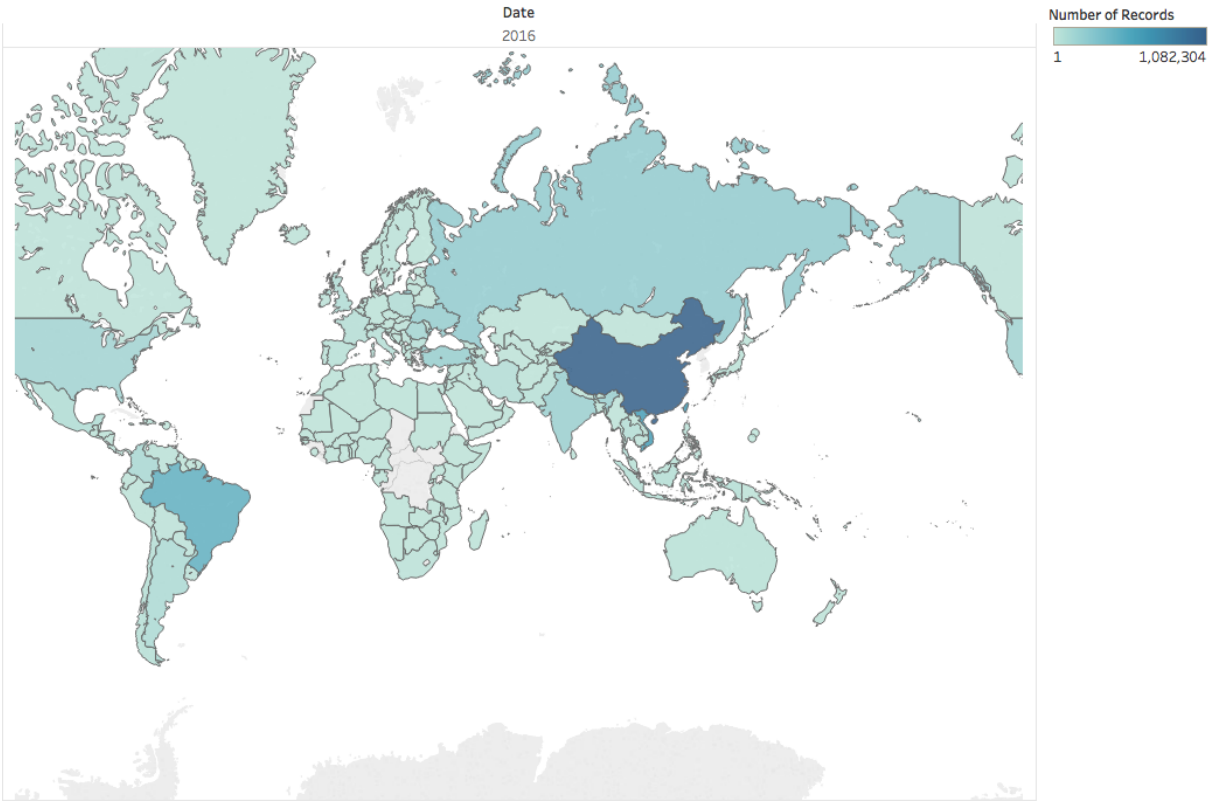

Figure 3: Number of total infected machines over months per country . A map view of the data

5 Conclusion

References