Universiteit of Twente

Economics of Security Course
AY 2018/2019

Internet of Things: Mirai

ASSIGNMENT 1

Authors:
Anna Prudnikova
Giovanni Maria Riva
Federico Casano
Sridhar Bangalore Venugopal

Contents

1 Introduction

IoT Mirai dataset shows the number of IoT devices infected across the different regions over a period of six months. The given dataset was examined with the use of specific tools, such as IBM SPSS, Tableau and Rstudio. The goal of the assignment is to analyze the given dataset and define the proper security metrics, that could be used by decision makers.

The report consists of six parts, including introduction, the main actor facing security issue, ideal and exciting in practise security metrics, the metrics from given dataset, graphs and conclusion.

To be able to further analyze the given dataset, we started with understanding of mirai malware nature, what security issues it raises and further proceeded with developing the possible security metrics. **Mirai** is a malware that infects IoT devices and turn them into remotely controlled bots. Every bot scans a an IPv4 address space linked to IoT devices addressing port 23 and 2323. Then, the infected machine attempts to log-in via a dictionary attack using a list of default passwords of IoT credentials.[3] Once infected, an attacker can mount a threat thought the controlled bots via a remote command and control server. The types of attacks can be various such as spam to a more dangerous ones such as DDoS that can lead to significant loses for the target company.

2 The main actor facing security issue

Security issues, raised by the dataset, could affect a number of actors, such as companies in different sectors of industry, individuals, governments and internet service providers (ISPs).

For our dataset, we decided to focus on one actor in particular: ISP as the most affected by threats of botnets. The actual dataset reports around 16200 IPs that are listening on port 7547 and around 8000 IPs on port 5555, (all the others listening on ports 23 and 2323). Mirai is known to attack mostly on port 7547, 5555 (from the point of view of ISP), where ISP links are affected and thus cause significant impact. These two ports are opened on many devices, even though they are supposed to be restricted [5].

This causes a security issue, since the infection of an IoT device by Mirai leads to severe attacks, such as DDoS, which causes a potential revenue losses, loss of customer trust and confidence, latency and bandwidth issue and even legal problems. At the same time, high network traffic can cause load balancing issues for ISP.

3 Ideal and existing in practise security metrics

3.1 Existing in practise metrics for measuring Botnets

The threat of botnets can be characterized by the following metrics:

Botnet size refers to the number of bots, x, that can be instructed to launch attacks (e.g., distributed denial-of-service attacks) at time t, denoted by y(t). Due to time zone difference, y(t) is often much smaller than the actual x as some of x is turned off during night time at time zones [4].

Network bandwith indicates the network bandwidth that a botnet can use to launch denial-of-service attacks [4].

Botnet Efficiency can be defined as the network diameter of the botnet network topology [4]. It measures a botnet's capability in communicating command-and-control messages and updating bot programs.

Botnet robustness measures the robustness of botnets under random or intelligent disruptions [4].

3.2 Ideal metrics for given dataset

The ideal metrics could be used by decision makers in ISPs (high management), but cannot be calculated in practice with given information. Mostly those decision makers can use this information to understand, if extra security controls are needed, how much money (if any at all) they need to invest to protect from botnets infections in their networks or to understand the efficiency of their existing implemented control against botnets. For example, an ISP could estimate the predicted money loss in case of botnet infections (proportional to the size of botnet), knowing the previous statistics on attacks (e.g.DDoS) on a specific company with an exact number of bots. Thud he can decide what is the proper amount of money they should invest in mitigating the botnets threat.
Based on the information, provided in publication by ENISA[6] we can identify the following ideal metrics for our case:

Direct financial damage

- financial loss, caused by equipment failure due to exceeding traffic and over utilization;
- hiring more qualified / skilled staff to mitigate / control botnets;

Second level financial damage

- penalties from breaking the service level agreement provider;
- network security liability coverage via theft, unauthorized access, or denial of service attack;

Number of clients that would be affected by botnet infections, who could complain about loss of bandwidth (based on business interruption for the clients from affected IOT devices);

Type of port that could cause business interruption for the companies (business-critical Internet link down for clients);

Amount of spam botnet can produce;

Number of devices that could be potentially affected by Mirai within the ISP Network, that leads to probability of being affected by Mirai overall (to understand if we need to invest money or no, e.g. if we don't have the type of devices that could be affected);

Number of DDoS attack botnet can perform;

Harvesting data (amount of data or types of data that was stolen).

4 The metrics from given dataset

4.1 The methodology to define security metrics

- To understand, what metrics we can obtain from the given dataset, we started with counting the infections over time, obtaining the infection rate. For example, the number of infected devices per day. Proceeding in this way, we can track the infection levels over time, to monitor the security level. At the same time, if the total amount of infected devices goes down, it does not necessarily mean that mitigation policy is effective, because this metric is highly influenced by an attacker's behavior.

- Then we proceeded with the question: "how could we control an attacker's behaviour?". To be able to understand if changes in the statistics are caused by an attacker behaviour or not, we can rank the networks (ISPs, countries etc), to understand, how well or poorly each network performs. Thus, the rank metric would tell us that some networks have more infections than others. But still it doesn't necessarily mean that their level of security is poorer, there are other factors involved that could influence the outcome / output. For instance, the more IoT devices in a country, the more infections. So we need to divide the number of infections by the number of IoT devices.

- To reach more accuracy we collected the number of IoT devices per country and worldwide. Unfortunately, this data is not freely open about all countries, so in part 4 of the report, we only have results on a number of countries. So, by dividing the number of infections by the number of IoT devices for that country, we can obtain more specific data. This data can be used to understand, if the infection percentage in the particular country is high (e.g. more than 50% of devices are infected) or not. If the percentage is high, it could be a trigger for an ISP to raise the awareness of network users about the danger of botnets, and persuade them to change default passwords or implement extra controls against botnet threat.

- We then analyzed ISPs. Focusing again on the number of IoT devices per country, we calculated the distribution of affected IoT devices over time (stochastic variable) in a particular ISP. Even though this is a stochastic metric, that doesn't take into consideration the behavior of an attacker, this might show that the measures put in place by the ISP helped to reduce the botnet infections (taken the specific ISP). In order to do so, we need to compare the distributions of botnet infections over time in different ISPs (countries) to compare whether the "drop" in the graph was at the same moment of time in different ISPs. That could mean that botnet was just abandoned by an attacker. At the same time, we evaluated the size of a botnet within a particular ISP knowing the size of the exact network of ISP. This is a more precise metric to give an understanding of the overall percentage of IP addresses that are affected by Mirai and thus could produce spam / DDoS attacks.

4.2 The definition of metrics from given dataset

The dataset speaks of metrics from the number of affected devices by Mirai.

To have a better understanding of security level it is better to use 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), the correlation of IP address to a particular country and a particular ISP.

Later on, to show the measurements of our metrics, we are going to use only the metrics, that could be useful to a decision maker. In our case, the decision maker is the high management of ISP.

	Metric	Definition	Use for decision maker	Comment
1	The overall number of IoT devices affected by Mirai	The number of IP addresses in the dataset	The metric is not precise, doesn't give any useful information to a decision maker	
2	The distribution of affected IoT devices over time	The distribution of number of IP addresses over time	This is a stochastic metric, doesn't' take into consideration the behavior of an attacker. Can't be useful for a decision maker	
3	The distribution of affected IoT devices over time (stochastic) in a particular ISP	The distribution of number of IP addresses within a particular ISP over time	This is a stochastic metric, doesn't' take into consideration the behavior of an attacker. Still can be used by a decision maker to understand if the measures put in place were effective or not.	Should be used in comparison to other ISP within a country to understand if it was influenced by attacker behaviour
4	The size of a botnet within a particular ISP	The percentage of affected IP addresses within a particular ISP amongst all the IP addresses of that network	This is a more precise metric for a decision maker to give understanding about the percentage of IP addresses that are affected by Mirai and thus could produce spam / DDoS attacks etc	Knowing the size of the exact network of ISP
5	The percentage of IoT devices affected by Mirai worldwide	The percentage of affected IP addresses given the number of overall IoT devices in the world	The metric is not useful for a decision maker, cannot influence any security decisions	
6	The percentage of IoT devices affected by Mirai in particular countries	The percentage of affected IP addresses given the number of overall IoT devices in a particular country	This metric helps ISP to understand if his country is highly affected by mirai (e.g. $>50\%$ of IoT devices are infected), if it is, ISP can raise the awareness of network users about the danger of botnets, persuade them to change default passwords (the percentage is the "trigger" and the proof).	Ex. in Netherlands there is a number of policies enforced to fight the botnets, the percentage of infection is already low, that is why extra measures are not necessary

5 Graphs

The *IoT-Mirai* dataset contains four columns, which we interpreted as followed:

- **Ip**: the network address of the device;
- **Port**: the port addressed by the malware attack;
- **Timestamp**: the time when the device has been detected as infected;

Figure 1 shows the distribution over time of the number of infected devices per ISP. We filtered the results to display those ISPs with *at least 50* infected devices over time.

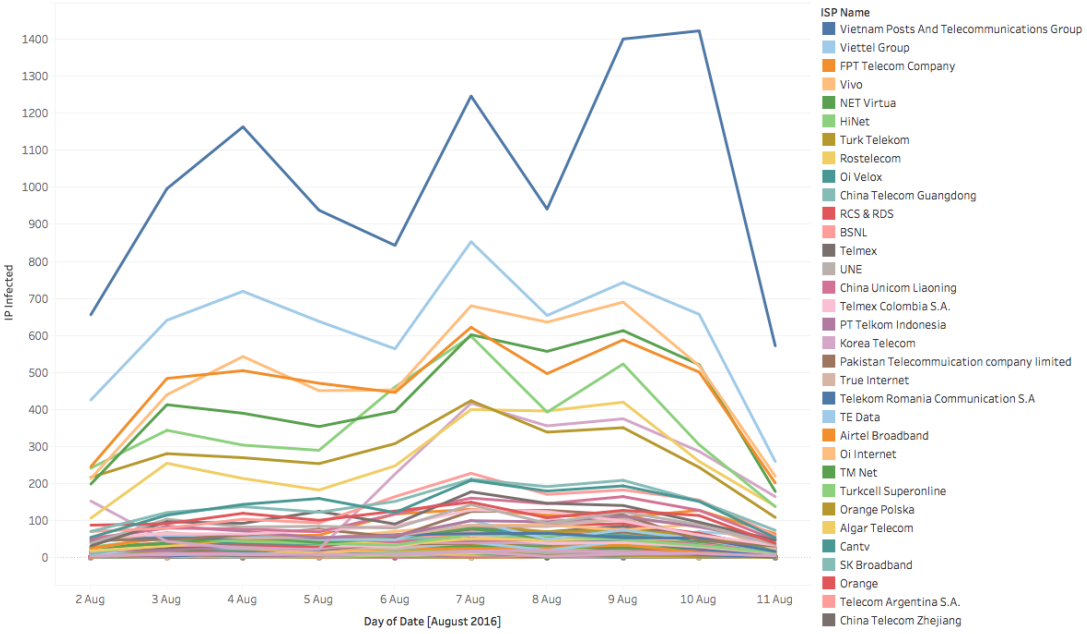

Figure 1: Distribution of infected IoT devices per ISP over days

Subsequently, we enriched our dataset by retrieving for each IP address, which country and which ISP it belongs to. It was unfeasible in a reasonable amount of time to query the whole dataset, therefore, our graphs are referring on a subset of 100 thousands records.

However, we can relate Graph 1 with Figure 2 where is possible to have an overview of the distribution of infected IoT devices per country. For example, Vietnam is one of the top countries for number of infections over months and every Vietnamese service provider is on the top list.

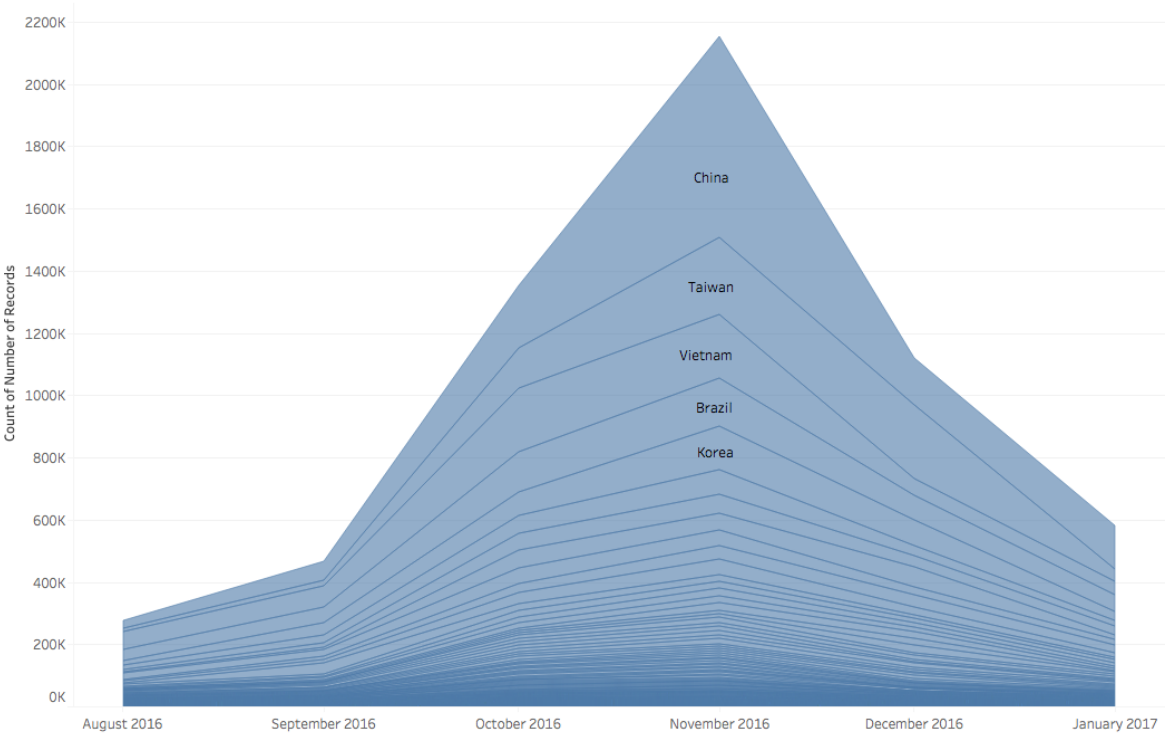

Figure 2: Distribution of infected IoT devices per contry over months

Figure 3 reports the percentage of infected systems per country. To compute the value we used two datasets. [1] contains the top ten countries with the most number of IoT devices per 100 persons. [2] lists the actual population for each country in the world. We joined the two datasets to estimate the total number of devices in the top 10 country from [1].

These countries are not among the top ones considering the original dataset given the fact we used [1] to retireve the number of IoT devices in total per country. Nonetheless, it is an insight for our validation of *metrics 6*.

Countries	Max. Devices in t..	Max. infected IP	Percentage Infected Ip ..
Sweden	2,204,836	17,123	0.78
United Kingdom	8,582,895	61,194	0.71
Portugal	1,667,582	10,769	0.65
France	11,812,882	49,159	0.42
Spain	9,267,834	31,869	0.34
Belgium	1,774,043	3,651	0.21
United States	81,104,075	152,275	0.19
Netherlands	4,231,815	4,244	0.10
Germany	18,523,680	5,644	0.03

Figure 3: Percentage of infected IoT devices over the total number of IoT devices.

Figure 4 shows the rank for the number of IoT devices infected per ISP. We further grouped every ISP for their relative country. We also filtered the data to display ISP with *at least 200* infected devices to enhance the readability of the image.

The graph provides an insight on the performance of each ISP. However, we could only retrieve the amount of infected devices per each service provider but not the actual size of every ISP network. This would have given another information which is the percentage of infected devices of the size of each company.

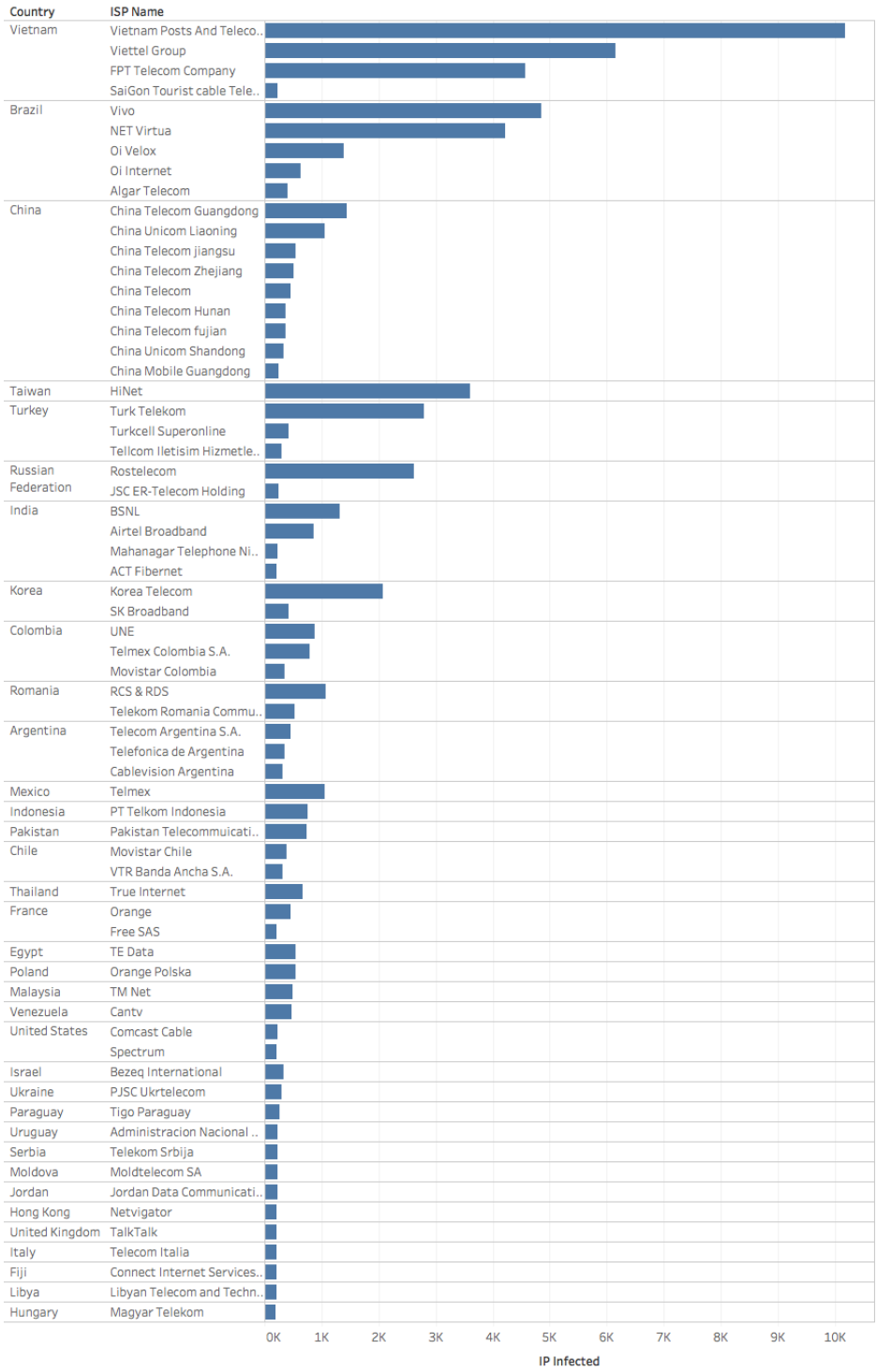

Figure 4: Number of infected IoT devices ranked by ISP and country.

6 Conclusion

As the first step of the assignment, we analyzed the provided dataset, identified the security issues that it speaks of, including the main actor, that could be affected by those security issues. The main actor, that was identified, is ISP and the security issues it might face, according to provided dataset, is the infection of IoT devices (within its network) by mirai malware, that could further lead to financial loses, network performance problems and legal issues.

Secondly, we analyzed the existing in practice security metrics, that could be used to further analyze the given dataset and that are already mentioned in literature [4],[6] and came up with ideal metrics that could be used by decision makers (in our case high management of ISPs). The metrics, that exist in practise, include botnet size, botnet efficiency, botnet robustness and network bandwidth. The ideal security metrics include direct and second level financial loses, the number of clients, that would be affected by botnet infection, the amount of spam, that botnet can produce, the number of DDoS botnet can perform, harvesting data.

Since the information provided within the given dataset is limited, it is not possible to evaluate the ideal metrics, that is why, as the final step, we developed the metrics, that could be calculated with existing data. Besides, we also used other data sources, to achieve better results, such as databases of IP addresses within countries and ISPs, the number of IoT devices worldwide and in particular countries [1][2].

We were able to evaluate three types of security metrics, that could be useful for decision makers: the size of a botnet within a particular ISP (knowing the size of ISP), the distribution of affected IoT devices over time in a particular ISP, the percentage of IoT devices affected by Mirai in particular countries. The size of a botnet within a particular ISP, knowing the size of ISP, can be used by a decision maker, to understand, if he needs to address the problem of botnet infections within his network by applying extra measures. The distribution of affected IoT devices over time in a particular ISP can be used by ISP as one of the signs, that the measures, that were taken to fight against botnet infections, were effective or not (should be analysed in comparison to other ISPs). The percentage of IoT devices, affected by Mirai in particular countries, could be used by decision makers to understand if its country of origin is highly affected by mirai malware (in comparison to other countries). Consequently it would be possible to decide if extra control measures against botnets are required.

References

[1] *Countries with the most IoT devices*, 2018 (accessed September 18, 2018). https://www.theatlas.com/charts/E1VUy4z0x.

[2] *The World Bank: total population*, 2018 (accessed September 18, 2018). https://data.worldbank.org/indicator/sp.pop.totl.

[3] Manos Antonakakis, Tim April, Michael Bailey, Matt Bernhard, Elie Bursztein, Jaime Cochran, Zakir Durumeric, J Alex Halderman, Luca Invernizzi, Michalis Kallitsis, et al. Understanding the mirai botnet. In *USENIX Security Symposium*, pages 1092–1110, 2017.

[4] David Dagon, Guofei Gu, Christopher P Lee, and Wenke Lee. A taxonomy of botnet structures. In *Computer Security Applications Conference, 2007. ACSAC 2007. Twenty-Third Annual*, pages 325–339. IEEE, 2007.

[5] Constantinos Kolias, Georgios Kambourakis, Angelos Stavrou, and Jeffrey Voas. Ddos in the iot: Mirai and other botnets. *Computer*, 50(7):80–84, 2017.

[6] Daniel Plohmann, Elmar Gerhards-Padilla, and Felix Leder. Botnets: measurement, detection, disinfection and defence. In *ENISA workshop on*. Mar, 2011.