Article information

Article title

Dataset of Normal and Abnormal Behaviour of IoT devices.

Authors

Eniela Vela[a], Fehmi Jaafar[b](*), and Darine Ameyed[c].

Affiliations

[a] Gina Cody School of Engineering and Computer Science, Department of Computer Science and Software Engineering, Concordia University, 1455 Boul. de Maisonneuve Ouest, Montréal, QC H3G 1M8. Canada.

[b] Department of Computer Science and Mathematics, Quebec University at Chicoutimi, 555 Bd de l'Université, Chicoutimi, QC G7H 2B1, Canada.

[c] Synchromedia Laboratory- École de technologie supérieure, 1100 Notre-Dame St W, Montreal, Quebec H3C 1K3, Quebec, Canada.

Corresponding author's email address and Twitter handle

Fehmi.jaafar@uqac.ca

Keywords

Internet of Things.
Cybersecurity.
IoT malware.
Cyber-attacks.

Abstract

The Internet of Things (IoT) describes the set of connected devices that are exchanging data through the Internet. IoT devices include sensors, cyber physical component, robots, and other objects that are able to collect exchange and potentially analyze data. The increasing use of Internet of Things has brought about many new risks, making these devices more favorable to hackers. Discovering and detecting compromised IoT devices is essential to reduce IoT based cyber-attacks. Thus, we performed an experiment with six common commercially available IoT devices compromised with three different ways (i.e. malware infection, vulnerability exploit, and DDOS attacks). We collected the data related to these devices before and after being compromised. The collected data is consisting of packets exchange and energy consumption.

We collect the network traffic using the open-source packet analyzer Wireshark. We set it as a server with an active firewall to capture and store the network traffic of the analyzed IoT devices. We measure the energy consumption of the IoT devices using the energy *monitor detector Arduino Uno connected to the Allegro Current Transformers (CT) Sensor ACS712, which is considered as the most accurate CT Sensor for low current and voltage measurements. We measures the energy consumption of each device every almost one second. For each component, we collect the metrics in the attack state as well as in the normal state for a period of 30 minutes. Then, we compare the energy consumption of all the IoT devices when they are clean and when they are compromised. We did a statistical test P-value of Mann-Whitney U test and effect size to ensure that there is a significant difference of the network behavior and the energy consumption between clean and compromised IoT devices.*

Specifications table

Subject	Cybersecurity
Specific subject area	The detection of compromised IoT devices that may be involved in cyber-attacks based on their network behavior and energy consumption.
Type of data	The data consist on numerical and textual values that describe the energy consumption and the network behavior.
How the data were acquired	For each IoT device, we collect data related to the network traffic and energy consumption in a normal working condition and under malware attacks. Energy and network traffic were collected in different timeframe since the energy was captured in almost each second and the network data is present only when there is a traffic in the IoT device. We merged the dataset according to the energy timeframe to be used later in determining compromised IoT devices.
Data format	We collected 12 numerical and textual variables in the following order: energy consumption, number of packets, number of unique sources, number of unique destinations, median package length, mean package length and six of the recurrent protocols.

Description of data collection	For each device, we start capturing the network traffic and measuring the energy consumption before and immediately after compromising it. Concretely, we collected 30 minutes of the data of the network traffic using Wireshark and the energy consumption via the energy monitor. In fact, for each IoT device, we collect data related to the network traffic and energy consumption in a normal working condition and under cyber-attacks. Energy and network traffic were collected in different timeframe since the energy was captured in almost each second and the network data is present only when there is a traffic in the IoT device. We merged the datasets according to the energy timeframe to be used later in determining compromised IoT devices
Data source location	· City/Town/Region: Montreal, Quebec · Country: Canada
Data accessibility	All datasets are publically available in Mendeley Data at DOI: 10.17632/3bkdm53jph.1
Related research article	Jaafar, Fehmi, Darine Ameyed, Amine Barrak, and Mohamed Cheriet. "Identification of Compromised IoT Devices: Combined Approach Based on Energy Consumption and Network Traffic Analysis." In 2021 IEEE 21st International Conference on Software Quality, Reliability and Security (QRS), pp. 514-523. IEEE, 2021.

Value of the data

- These datasets can be used to analyze the impact of cyber-attacks in energy consumption of IoT devices.
- These datasets can be used to differentiate between normal and abnormal behavior of IoT devices using energy consumption data and network behavior.
- These datasets can be used to create machine-learning models that are able to detect compromised IoT devices.
- These datasets can be used to observe the deviation of network behavior of IoT devices depending on how they are compromised.
- These datasets can be used to help advance the understanding of IoT based cyber-attacks.
- These datasets can be used to analyze the impact of IoT malware.

Data description

The presented dataset includes a sequential dataset of energy consumption and network behavior for six different IoT devices, in normal state (clean devices) and abnormal state (compromised devices). Each abnormal state of IoT devices is triggered by attacking these devices with Mirai botnet or exploiting RouterSploit and UFONet to compromise them.

Devices	***Model***
1. Indoor Smart Home Camera	*Wyze Cam WYZEC*
2. Outdoor Camera	*HOSAFE Outdoor Wifi Camera*
3. Home Router (1)	*TP-Link ACS1750*
4. Home Router (2)	*D-Link AC1200*
5. Smart Home Hub	*Phillips Hue Smart Hub*
6. DVR Digital Video Recorder	*ANNKE DVR 8CH*

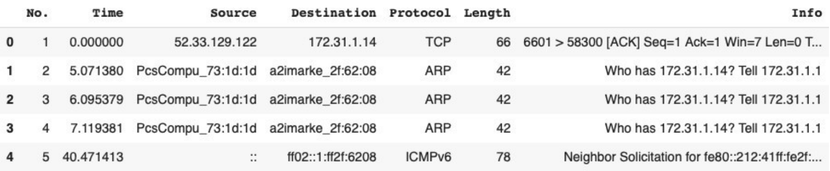
Table 1: Chosen IoT Devices Information

Indeed, the dataset can be divided into 2 major groups: Network Behavior and Energy Consumption of six different IoT devices, in Normal and Abnormal State.

The network behavior contains several attributes as shown in Figure 1. These attributes includes the timestamp saved in milliseconds. The Source and Destination, which describe the communication between the IoT devices. Both Source and Destinations are IP addresses, mostly IPv4 and some IPv6. The Protocol helps understand over which protocols these data packets are exchanged. The most common ones TCP, UDP and ARP in the normal network behavior. These used protocols varied a lot for the abnormal network behavior as shown in Figure 3. The length of the packet is an integer number. The Info is a string type that represents the log information exchanged between two devices.

Devices	***Model***
1. Wyze Cam WYZEC Indoor Camera	*5V*
2. HOSAFE Outdoor Wifi Camera	*12V*
3. TP-Link ACS1750	*12V*
4. D-Link AC1200	*9V*
5. Phillips Hue Smart Hub	*5V*
6. ANNKE DVR 8CH	*12V*

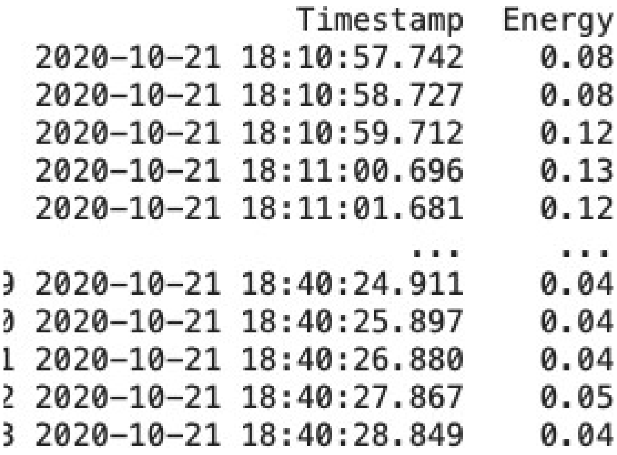
Table 2: IoT Devices Voltage Information



	No.	Time	Source	Destination	Protocol	Length	Info
0	1	0.000000	52.33.129.122	172.31.1.14	TCP	66	6601 > 58300 [ACK] Seq=1 Ack=1 Win=7 Len=0 T...
1	2	5.071380	PcsCompu_73:1d:1d	a2imarke_2f:62:08	ARP	42	Who has 172.31.1.14? Tell 172.31.1.1
2	3	6.095379	PcsCompu_73:1d:1d	a2imarke_2f:62:08	ARP	42	Who has 172.31.1.14? Tell 172.31.1.1
3	4	7.119381	PcsCompu_73:1d:1d	a2imarke_2f:62:08	ARP	42	Who has 172.31.1.14? Tell 172.31.1.1
4	5	40.471413	::	ff02::1:ff2f:6208	ICMPv6	78	Neighbor Solicitation for fe80::212:41ff:fe2f:...

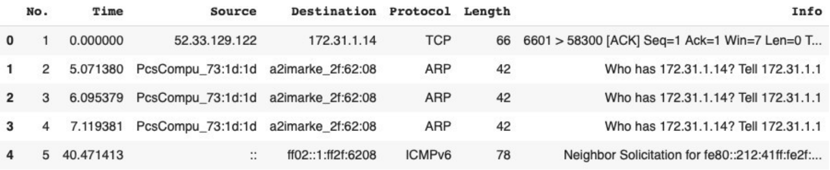
Figure 1: Normal Network Behaviour for DVR Digital Video Recorder

The energy consumption dataset consist of two main attributes as shown in Figure 2. The timestamp and the energy consumption in ampere. The voltage of all six IoT devices is presented in Table 2.



```
                Timestamp  Energy
  2020-10-21 18:10:57.742    0.08
  2020-10-21 18:10:58.727    0.08
  2020-10-21 18:10:59.712    0.12
  2020-10-21 18:11:00.696    0.13
  2020-10-21 18:11:01.681    0.12
                      ...     ...
9 2020-10-21 18:40:24.911    0.04
0 2020-10-21 18:40:25.897    0.04
1 2020-10-21 18:40:26.880    0.04
2 2020-10-21 18:40:27.867    0.05
3 2020-10-21 18:40:28.849    0.04
```

Figure 2: Normal Energy Consumption for DVR Digital Video Recorder



	No.	Time	Source	Destination	Protocol	Length	Info
0	1	0.000000	52.33.129.122	172.31.1.14	TCP	66	6601 > 58300 [ACK] Seq=1 Ack=1 Win=7 Len=0 T...
1	2	5.071380	PcsCompu_73:1d:1d	a2imarke_2f:62:08	ARP	42	Who has 172.31.1.14? Tell 172.31.1.1
2	3	6.095379	PcsCompu_73:1d:1d	a2imarke_2f:62:08	ARP	42	Who has 172.31.1.14? Tell 172.31.1.1
3	4	7.119381	PcsCompu_73:1d:1d	a2imarke_2f:62:08	ARP	42	Who has 172.31.1.14? Tell 172.31.1.1
4	5	40.471413	::	ff02::1:ff2f:6208	ICMPv6	78	Neighbor Solicitation for fe80::212:41ff:fe2f:...

Figure 3: Abnormal Network Behaviour for DVR Digital Video Recorder

Experimental design, materials and methods

In order to collect the datasets, we compromised the six IoT devices with three different ways; using the malware Mirai Botnet, using the vulnerability exploit tool RouterSploit, and using the DDOS execution tool UfoNet. To understand the anatomy of the all three attacks we referred to two papers [3] and [4]. Wireshark [5] captured the complete process of registering the data transmission in both abnormal and normal state for all IoT devices. The outcome reflects in csv files containing extracted features. The csv files are easy to use with various tools and programming libraries.

3.1. Normal Behaviour

Normal behavior of the IoT devices is considered the moment the IoT device is plugged in and connected to the internet. We observed an increase of energy consumption just the first 2 seconds when the IoT is plugged in and it is recorded in each devices. As per Network Behavior, from the dataset there is a smooth exchange of data packets.

3.2. Mirai Botnet

In order to generate data in abnormal state of IoT devices, we attack the analyzed devices with Mirai Botnet. Mirai is widely used in IoT based cyber attacks. It is a malware which infect mostly smart devices which run ARC processors and turn them into ’zombie’ devices or so called bots. These remotely controlled network bots to launch DDoS attacks.

3.3. RouterSploit

RouterSploit is an open source framework dedicated to scan and exploit known IoT vulnerabilities. The framework itself consist of many modules in which we used exploits, scanners and payloads.

We used scanners to scan through network and find all the connected devices with open ports. After we identified the IoT devices, we used exploits to break through the username and passwords. Once we had access to the IoT device, we used the payload module to exhaust the device.

3.4. UFONet

UFONet is a P2P and Cryptographic software, which allows performing DoS and DDoS attacks in different devices. The exploitation of the IoT device is done through Open Redirect vectors on a third-party website. After the exploitation, the payloads are sent from different websites, which act as botnet and abuse protocols in the third layer network.

Acknowledgments

An NSERC Discovery Grant and the Computer Research Institute of Montréal supported the work. We acknowledge the support of Ministère de l'Économie et de l'Innovation of Quebec, Canada. We gratefully thank Louis- David Perron, Khaireddine Mejri, and Amine Barrak for their generous effort during the empirical study.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Jaafar, Fehmi, Darine Ameyed, Amine Barrak, and Mohamed Cheriet. "Identification of Compromised IoT Devices: Combined Approach Based on Energy Consumption and Network Traffic Analysis." In 2021 IEEE 21st International Conference on Software Quality, Reliability and Security (QRS), pp. 514-523. IEEE, 2021.

[2] Jaafar, Fehmi (2022), "Dataset of Normal and Abnormal Behaviour of IoT devices", Mendeley Data, V1, doi: 10.17632/3bkdm53jph.1

[3] Antonakakis, Manos, Tim April, Michael Bailey, Matt Bernhard, Elie Bursztein, Jaime Cochran, Zakir Durumeric et al. "Understanding the mirai botnet." In 26th USENIX security symposium (USENIX Security 17), pp. 1093-1110. 2017.

[4] Hong, Bing-Kai, Jr-Wei Huang, Tao Ban, Ryoichi Isawa, Shin-Ming Cheng, Daisuke Inoue, and Koji Nakao. "Measurement study towards a unified firmware updating scheme for legacy IoT devices." In 2019 14th Asia Joint Conference on Information Security (AsiaJCIS), pp. 9-15. IEEE, 2019.

[5] Jain, Vinit. "Getting Familiar with Wireshark." In Wireshark Fundamentals, pp. 35-78. Apress, Berkeley, CA, 2022.