

Network Traffic Characteristics of IoT Devices in Smart Homes

Md Mainuddin, Zhenhai Duan
Department of Computer Science
Florida State University
Tallahassee, FL 32306, USA
{mainuddi, duan}@cs.fsu.edu

Yingfei Dong
Department of Electrical Engineering
University of Hawaii
Honolulu, HI 96822 USA
yingfei@hawaii.edu

Abstract—Understanding network traffic characteristics of IoT devices plays a critical role in improving both the performance and security of IoT devices, including IoT device identification, classification, and anomaly detection. Although a number of existing research efforts have developed machine-learning based algorithms to help address the challenges in improving the security of IoT devices, none of them have provided detailed studies on the network traffic characteristics of IoT devices. In this paper we collect and analyze the network traffic generated in a typical smart homes environment consisting of a set of common IoT (and non-IoT) devices. We analyze the network traffic characteristics of IoT devices from three complementary aspects: remote network servers and port numbers that IoT devices connect to, flow-level traffic characteristics such as flow duration, and packet-level traffic characteristics such as packet inter-arrival time. Our study provides critical insights into the operational and behavioral characteristics of IoT devices, which can help develop more effective security and performance algorithms for IoT devices.

Index Terms—Internet of Things, IoT Devices, Smart Homes, IoT Security

I. INTRODUCTION

Internet of Things (IoT) devices have been increasingly adopted and deployed in diverse environments including both homes and enterprises to provide a variety of data collection, monitoring, and control functionalities. For example, the number of IoT devices surpassed the number of non-IoT devices in 2020 for the first time, and it was estimated that there would be more than 30 billions of IoT devices in 2025, which would be three times more than non-IoT devices at that time [1].

On the other hand, the unprecedented growth of IoT devices poses a significant security threat on the Internet. Typical IoT devices possess low processing capabilities, limited memory and storage, as well as minimal network protocol support, to prolong the battery life and to lower the prices of IoT devices in an increasingly competitive IoT market [2], [3]. As a consequence, IoT devices in general lack strong security measures to protect themselves from security attacks. Furthermore, security patches might not be updated regularly due to irregular software release or a lack of awareness and expertise of the IoT device users [4]. These factors ultimately keep a huge number of IoT devices vulnerable on the Internet [5].

IoT devices and systems are deployed in public, private, and corporate spaces, and a significant amount of information is

being collected from these environments to make intelligent decisions. This huge amount of information, along with the vulnerabilities of the IoT devices, attracts the attackers who wish to exploit the IoT systems. As an example, in 2016 the massive DDoS attack on Dyn data centers made its DNS service unreachable to its users [6]. The attack was performed using Mirai botnets on IoT devices, and it took down more than 1200 domains that Dyn was supporting, including Amazon, Twitter, Shopify, and many others.

A thorough investigation of network traffic characteristics of IoT devices is critical in our understanding of the network operation and behavior of IoT devices, and in developing effective IoT device identification, classification, and anomaly detection mechanisms. Given the significance of the IoT device security issue, a number of studies have developed machine-learning based algorithms to identify and classify IoT devices, and to detect compromised IoT devices, relying on various network traffic characteristics (or features) of IoT devices. However, none of them have provided comprehensive studies on the network traffic characteristics of IoT devices.

In this paper we investigate the network traffic characteristics of IoT devices (and non-IoT devices) in a typical smart home environment consisting of 10 common smart home IoT devices and 5 non-IoT devices. In this paper we adopt a commonly accepted definition of IoT devices. An IoT device performs a specific functionality, and can operate autonomously without directly human controls. After the network traffic is collected for the devices, we perform a thorough investigation of the network traffic characteristics from three complementary aspects: remote network servers and port numbers that IoT devices connect to; flow-level traffic characteristics such as flow duration, and packet-level characteristics such as packet inter-arrival time.

Based on our investigation of the network traffic characteristics of both IoT and non-IoT devices, we make a few promising observations that can provide insights into developing effective IoT security mechanisms, including IoT device identification, classification, and anomaly detection. In particular, IoT devices behave very differently from non-IoT devices in a number of aspects, for example, in terms of number of network domains they communicate with, and in terms of both control traffic and data traffic. Second, IoT

devices are also more consistent and stable in certain network traffic that they generate. We will report the details of our studies in Section IV.

The remainder of the paper is organized as follows. In Section II, we discuss related work. In Section III we describe the set-up of the smart home environment, and tools for data pre-processing. We perform the detailed study of the network traffic characteristics of IoT devices in Section IV. We summarize our study and discuss the future work in Section V.

II. RELATED WORK

In this section we discuss a number of research efforts that are most relevant to our work. We first note that the existing works focused on developing machine-learning based mechanisms for IoT identification, classification, and anomaly detection. Although they also explored network traffic characteristics (or features) in the development of their machine learning based algorithms, none of them have provided a thorough study on network traffic characteristics of IoT devices, which is the objective of our work.

Mazhar and Shafiq [7] performed a study on IoT traffic analysis in a home network environment, based on flow-level summary information obtained through a home gateway management company. They made a few interesting observations based on this flow-level summary data. For example, they found that the majority of IoT traffics were related to communications with Google Cloud and amazon AWS. In addition, they also noted that most of IoT devices used hard-coded Google DNS servers for various reasons. However, the researchers only had access to the flow-level summary data; and as a consequence, they cannot perform packet level studies of IoT traffic.

Meidan *et al.* used IoT traffic data to detect and white list IoT devices connected to the current network [8]. They used selected features from TCP sessions information to train their machine learning model. Doshi *et al.* presented a model to distinguish normal and attack IoT traffic by using machine learning techniques on different network traffic features [9]. They used simulated IoT traffic data for their study.

Shahid *et al.* developed a few machine learning algorithms to recognize IoT devices using the sizes and inter-arrival times of the first N packets of TCP flows of IoT devices [10]. Their study used only four IoT devices with no non-IoT devices. Sivanathan *et al.* developed a machine-learning algorithm to classify IoT devices based on statistical attributes of their network activities [11], including features such as activity cycles, remote servers and ports, signaling patterns, and cipher suites. The authors investigated the characteristics of these features. Our work complements their studies and provide a more thorough investigation of the network traffic characteristics of IoT devices.

Miettinen *et al.* worked on IoT device fingerprinting and identification that utilized the network traffic during the setup phase [12]. Ammar *et al.* presented a different machine-learning based approach of identifying IoT devices [13]. Their

work used service discovery and DHCP protocols to collect information about devices during their setup phases.

III. NETWORK SETUP AND TRAFFIC COLLECTION

A. Devices and Network Setup

Figure 1 illustrates the setup of the network testbed in a home environment, from which we collect the network traffic used in this study. This network comprises a variety of 10 IoT devices as well as a number of non-IoT devices (see Table I), which reflects a typical network setup in a smart home environment. All devices are connected to a Netgear AC3200 Wi-Fi router. The Netgear router is connected to a Technicolor DOCSIS 3 cable modem, which is also a Wi-Fi router, to connect the home network to the Internet. The Netgear router and the Technicolor router assign IP addresses in two different private IP address ranges.

This home network contains three pairs of security cameras. A pair of Logitech Circle-2 cameras are placed at different outdoor locations; two indoor cameras (Wyze and Eufy Indoor) are placed at the same location at the front porch; and another pair of indoor cameras (Eufy pan/tilt and LittleElf) are placed inside the living room. This combination of similar cameras at different locations and different cameras at the same location helps us to observe the traffic behavior of these security cameras from different angles. This home network also contains two smart plugs from two different manufacturers, a smart bulb, and an HP printer. Two laptops, two smart phones and an iPad, representing non-IoT devices, are also connected with the Netgear Wi-Fi router.

All devices are assigned with static IP addresses on the Netgear router so that we can keep track of their traffic with IP addresses in addition to their MAC addresses. Table I shows these IoT and non-IoT devices used in this study, including their corresponding MAC and IP addresses (assigned by the Netgear router). In the table, the column "Device" indicates the manufacturer and type of the device. For each device, we also provide an "Alias Name", which is used to refer to the device throughout the paper to simplify our exposition. We also note that none of the IoT and non-IoT devices have been intentionally shut down for the studies reported in this paper.

The Netgear Wi-Fi router is flushed with OpenWrt firmware [14] to ease the collection of network traffic on the home network. A few software packages are installed on the router to help capture network traffic from all the connected devices. A Unix cron job runs the tcpdump tool on the router, which captures the traffic from all the devices and saves the captured traffic to an attached 512GB USB flash drive. The cron job runs tcpdump for each device separately and restarts the tcpdump processes at midnight. In this way we obtain a separate trace file every 24 hours for each device while keeping the file sizes within a reasonable limit.

All security cameras have motion detection enabled. When any motion is detected, a camera will upload the footage to the remote server and send a notification to the companion mobile app. Also, all the cameras have a live view feature,

Table I: Connected Devices in Smart Home Network

Device	Alias Name	Category	MAC Address	IP Address
Logitech Circle-2	Cam1	IoT	44:73:D6:0C:36:AD	192.168.1.158
	Cam2		44:73:D6:09:BD:C9	192.168.1.186
Wyze Cam	Cam3		2C:AA:8E:95:F3:18	192.168.1.228
Eufy Indoor Cam	Cam4		8C:85:80:38:98:AF	192.168.1.182
Eufy Pan and Tilt	Cam5		8C:85:80:3A:12:B4	192.168.1.131
LittleElf Cam	Cam6		0C:8C:24:61:50:29	192.168.1.168
Epicka Smart Plug	Plug1		DC:4F:22:0E:C6:36	192.168.1.127
Amazon Smart Plug	Plug2		F8:54:B8:25:AA:C9	192.168.1.204
Smart Bulb	Bulb		84:0D:8E:7F:4B:B4	192.168.1.207
HP Envy Printer	Printer		94:57:A5:0C:5B:66	192.168.1.248
HP Elitebook	Laptop1	non-IoT	AC:FD:CE:01:7C:9B	192.168.1.105
HP ZBook	Laptop2		CC:2F:71:3B:0E:DE	192.168.1.247
Apple iPhone X	iPhone		34:08:BC:DE:E9:7E	192.168.1.203
Apple iPad	iPad		E8:8D:28:14:82:30	192.168.1.125
Samsung S20	Android		16:05:DD:78:5F:20	192.168.1.215

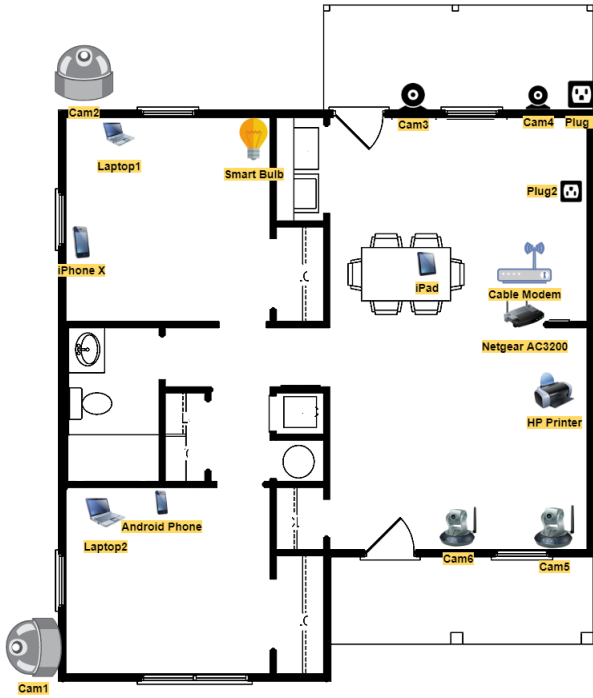


Figure 1: Smart home network setup

which enables a user to view the live video through the mobile app.

We note that we sometimes directly connect our Android phone to the Technicolor router (instead of the Netgear Wi-Fi router) in order to examine the communication behavior of IoT devices when an IoT device and the corresponding companion mobile phone are on different LANs. In the following we highlight some observations we made on the

behaviors of different IoT devices to illustrate the diversity and complexity of the IoT device operations. At a high level, the behavior of the same type of IoT devices can behave very differently depending on the manufacturers. In addition, they can also operate differently depending on the location of the corresponding companion mobile apps.

For example, Logitech cameras (Cam1 and Cam2) always send the data to their corresponding remote servers whenever an event happens (for example, a motion is detected), or when we watch the live stream through the mobile app. In contrast, Wyze (Cam3), Eufy (Cam4 and Cam5), and LittleElf (Cam6) cameras send the traffic to their corresponding remote servers when there is an event, but they may send the traffic directly to the mobile device (where mobile app is installed) for a live stream under certain scenarios, for example, when the mobile device is connected locally. Similar behavior is observed for smart plugs and bulb. When the mobile phone is connected on the local network, the app sends the control instruction to the IoT device directly; otherwise, the control instruction is delivered from the app to the remote server, from where it is in turn delivered to the device.

B. Data Pre-processing

The USB flash drive has different folders for each device and daily traffic is stored in separate pcap files in corresponding folders. We download those pcap files from the router to a laptop for data processing. In our study we use two tools—CICFlowMeter [15] and tshark [16]—and a number of home-made Python scripts to process the collected data traces. In particular, we use CICFlowMeter to extract the flow-level information of the data traces (and also the packet inter-arrival times in a flow), and we use tshark to extract the other network traffic characteristics.

We updated the source code of CICFlowMeter to reflect a more commonly used definition of TCP and UDP flows. A

TCP or UDP flow is defined by a 5-tuple of source IP address, source port number, destination IP address, destination port number, and protocol. In the original CICFlowMeter, a flow is identified by a pair of SYN and FIN packets or a *duration* timeout threshold for TCP, and a duration timeout threshold for UDP. For example, with a duration timeout threshold of 120 seconds, CICFlowMeter will conclude a flow when the next packet is received 120 seconds *after the flow starts*. As a consequence, the maximum flow duration reported in CICFlowMeter is 120 seconds. Instead of a duration timeout threshold, we update it to a *packet inter-arrival* timeout threshold. That is, we only conclude a flow when a packet inter-arrival time is greater than the specified timeout threshold (if we have not received FIN packet for TCP flows). We use a packet inter-arrival timeout threshold of 120 seconds for extracting flows in this study [17]. Both TCP and UDP flows are bidirectional, containing both outgoing and incoming packets.

IV. NETWORK TRAFFIC CHARACTERISTICS

In this section we analyze the network traffic characteristics of the IoT devices based on the data collected on the home network, and comparing with those of non-IoT devices. We investigate the characteristics from three complementary perspectives: the properties related to the remote servers that IoT devices connect to, the flow-level traffic properties, and the packet-level traffic properties. Although our data collection spans more than two months, our investigation will be focused on the data trace collected on one day, which is somewhat randomly chosen to be a Monday. For a number of studies, we also investigate the network traffic characteristics of data traffic in a week, containing the one-day mentioned above. The results obtained are representative of the general network traffic characteristics.

A. Remote Servers and Services

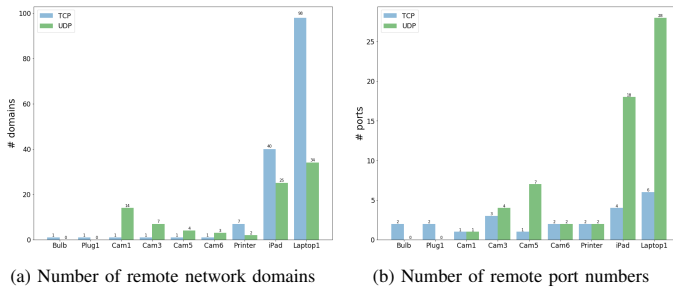


Figure 2: Number of remote network domains and port numbers.

In this subsection we will study the behavior of IoT devices in terms of both remote servers and services an IoT device communicates with. For the remote servers, we will aggregate at the network domain level instead of individual server machines, given that it is common for an IoT device to communicate with multiple servers in the same domain for performance, reliability, or other purposes. For this reason, we use the terms *remote server* and *remote network domain*

interchangeably. The remote services are represented by the corresponding TCP or UDP port numbers. We will first study the number of remote network domains and port numbers that a device communicate with in one day, and then we study the corresponding behavior on a daily basis for one week. To make the figures more legible, in the figures we only show the results of seven IoT devices and two non-IoT devices. We omit the results of other devices. They have similar results as the ones reported here.

1) *Remote Servers and Services in One Day*: Fig. 2a shows the number of remote network domains that the devices communicate with in the chosen day. We report the results for the TCP and UDP traffic separately to better illustrate the communication patterns of the devices. As we can see from the figure, IoT devices only communicate with a small number of remote network domains. For example, most of the IoT devices only communicate with one network domain for TCP traffic, except the Printer, which communicates with seven remote network domains for TCP. The IoT devices in general communicate with a slightly higher number of remote network domains for UDP traffic. The UDP is normally used for control traffic such as DNS and NTP to support the autonomous operations of the IoT devices. However, we note that in general, the number of UDP network domains is still a relatively small number. In contrast, the non-IoT devices communicate with a much higher number of remote network domains for TCP traffic. Of course, these are normally affected by the specific usage of the device by the users. Non-IoT devices also have a smaller number of remote network domains for UDP traffic. This is understandable as UDP is mostly used for control traffic such as DNS and NTP etc. Users of these devices may not initiate any UDP traffic.

These observations are consistent with what we have expected. As IoT devices are connected to their corresponding remote server for their status update and other operational activities, they regularly send traffic to their remote servers and maintain longer TCP sessions (see Subsection IV-B2 for more studies on flow-level characteristics of IoT devices). Smart bulb and plug1 do not have any external UDP connections while cameras maintain some UDP connections with different time servers.

Figure 2b shows the number of remote port numbers that devices communicate with. In general, IoT devices communicate with a small number of remote port numbers, except Cam5. We inspected the traffic of Cam 5 manually, and it revealed the following observation. Cam5 was connected very often for a live view from iPhone, which resulted in communications with a number of different UDP ports. However, it still holds that IoT devices in general communicate with a limited number of remote port numbers. In contrast, non-IoT devices in general communicate with a larger number of remote port numbers. The specifics of the remote port numbers of non-IoT devices may depend on the specific usage of these devices by the users. For example, the non-IoT devices iPad and Laptop1 communicated with various port numbers on websites such as Youtube and Facebook.

Table II: Remote Servers and Port Numbers of IoT Devices

Device	TCP Ports	UDP Ports (Up to 4 most frequent)	TCP Servers	UDP Servers (Up to 4 most frequent)
Cam1	443	123	amazonaws.com	mattnordhoff.net, chaosfire.net internap.com, latt.net
Cam2	443	123	amazonaws.com	ardoin.me, cloudflare.com comcastbusiness.net, ctyme.com
Cam3	8443, 443, 8883	53, 1001, 10240, 123	amazonaws.com	dns.google, homeassurednow.com ip-147-135-36.us, nist.gov
Cam4	443	32100, 53, 8006, 123	amazonaws.com	amazonaws.com, dns.google 224.0.0.251, flashdance.cx
Cam5	443	32100, 53, 26807, 25249	amazonaws.com	amazonaws.com, dns.google 114dns.com, comcast.net
Cam6	19000, 19001	53, 123	amazonaws.com	dns.google, nist.gov colorado.edu
Plug1	80, 1883	-	amazonaws.com	-
Plug2	443	123	amazonaws.com	amazonaws.com
Bulb	443, 8886	-	amazonaws.com	-
Printer	5222, 443	9930, 5353	google.com, 1e100.net	amazonaws.com, 224.0.0.251
Some common ports: 443 = HTTPS, 8443 = SSL, 80 = HTTP, 1883/8883 = MQTT, 5222 = XMPP, 123 = NTP, 53 = DNS, 5353 = mDNS				

Table II shows the most frequently used remote network domains and port numbers of TCP and UDP traffic for IoT devices. We see that IoT devices mainly use port 443 for their TCP traffic, which is an official port number for HTTPS. IoT devices rarely use port number 80 for the HTTP traffic. Some IoT devices use port numbers other than 443 for HTTPS traffic. For instance, Cam6 uses port numbers 19000 and 19001 to communicate with their server instead of the more commonly used 443 port number. The remote network domains and port numbers of non-IoT devices are much more diverse compared to IoT devices and unpredictable, because their behavior is largely determined by the users who use these devices.

We also note from Table II that most of the IoT devices use Google DNS server as their DNS provider, which is also supported by other studies [7], while some others use the local router for this purpose (not shown in table). On the other hand, all laptops, iPhone, and iPad use the local router as their DNS provider. The Android phone uses the router as well as Google DNS as its DNS provider.

2) *Daily Remote Network Domains and Port Numbers in One Week*: Fig. 3 shows the number of remote network domains for a representative IoT (Cam1) and non-IoT device (Laptop1). In the figure, in addition to showing the total number of remote network domains in a day, we also show the number of new remote network domains, that is, the network domains that we have not seen in the previous days in the chosen week. From the figure we can see that, the IoT device Cam1 only communicates with one or two remote network domains (which is amazonaws.com, and additionally cloudfront.net, which is an integrated Content Delivery Network

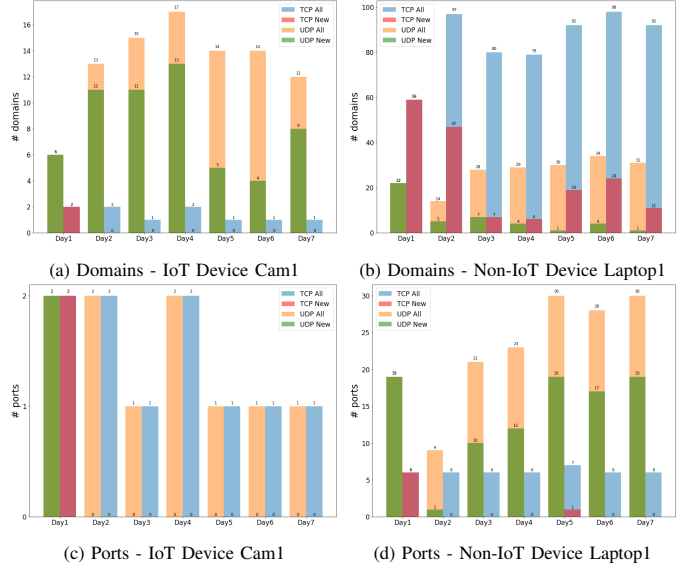


Figure 3: Daily Remote Network Domains and Ports in One Week.

service with amazon aws [18]) throughout the week for TCP traffic. All IoT devices show similar behavior for TCP traffic. In contrast, for UDP traffic, Cam1 has communicated with a changing number of remote network domains throughout the week, with new remote network domains added each day. Most of the UDP traffic is related to NTP and DNS queries. Unlike TCP traffic, there are slightly different behaviors for UDP traffic among IoT devices. In particular, the IoT devices Plug1 and Bulb do not communicate with NTP or DNS servers.

Compared to the IoT devices, non-IoT devices behave quite

differently. From the figure, we can see that Laptop1 communicates with a large number of remote network domains for both TCP and UDP traffic. In addition, it also communicates with new remote network domains each day. This is a reasonable observation, as the behavior of a non-IoT device is largely determined by the users who use this device. Users may access different remote servers and applications at will. On the other hand, the behavior of IoT devices is more likely to be pre-programmed and thus less likely to access random remote network domains and applications.

Figure 3c shows the daily number of remote port numbers that Cam1 communicates with. From the figure we can see that, Cam1 communicates with a relatively small and stable set of remote port numbers. In particular, even though there are fluctuation in terms of number of remote port numbers, there are no new port numbers observed after the first day. In contrast, as we can see in Figure 3d that Laptop1 almost always communicates with some new remote port numbers for UDP traffic (and also one day for TCP traffic). Again, this observation can be similarly explained by the fact that, IoT devices are more likely pre-programmed with a fix set of applications, and consequently they only communicate with a relatively small and stable set of remote port numbers.

B. Flow-Level Traffic Characteristics

In this subsection we investigate the *flow-level* traffic characteristics of IoT devices, including number of flows, flow duration, flow size, among others. These provide us with insights into the behaviors of IoT devices at the granularity of TCP and UDP flows.

1) *Number of Flows and Flow Duration:* We first study the number of flows and flow duration of TCP and UDP flows for the deployed devices on the representative day (for 24 hours). The duration of a flow is defined as the time difference between the last packet and the first packet of the flow.

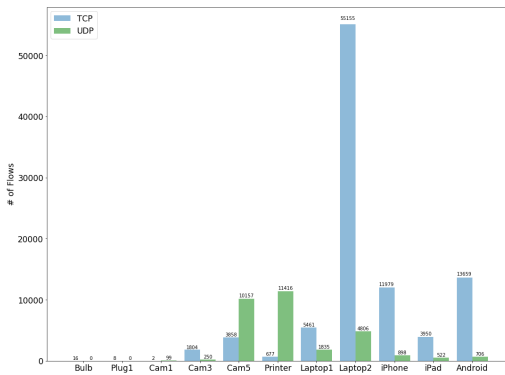


Figure 4: Number of TCP and UDP Flows.

Figure 4 shows the number of flows of the representative IoT and non-IoT devices. From the figure we first note that non-IoT devices have more TCP flows than IoT devices. For example, the non-IoT device iPad generated 3950 TCP flows, the lowest among the non-IoT devices, which is still higher than the highest number of TCP flows generated by

IoT devices (3858 generated by Cam5). As we will see when we study the flow durations, this could be caused by the fact that IoT devices tend to generate long-lasting TCP flows. The situation of UDP flows is quite different. Some of the IoT devices do not generate UDP flows at all (including Bulb and Plug1); however, some other IoT devices have generated a large number of UDP flows (including Cam5 and Printer). After examining the traffic trace, we note that Cam5 has generated frequent UDP flows to remote port numbers such as 32100 for its operational functionalities. Printer has some mDNS flows (with destination IP addresses 224.0.0.251 and 224.0.0.252, and port number 5353), while all other UDP flows are to amazonaws servers with port 9930.

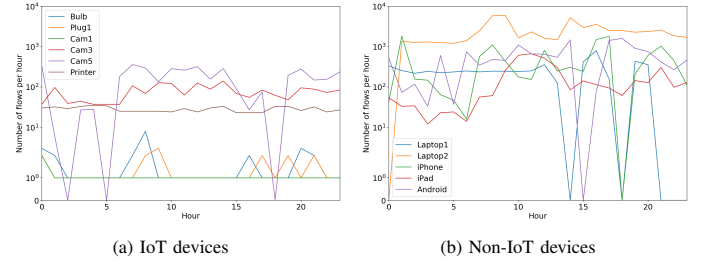


Figure 5: Hourly number of TCP flows.

Figures 5a and 5b show the hourly number of TCP flows for representative IoT and non-IoT devices, respectively. From the figures, we note that most of IoT devices have relatively stable number of hourly number of TCP flows, with some fluctuations, except Cam5. Overall, the hourly number of TCP flows of Cam5 is also stable; however, it has a number of instances where it has zero TCP flows. We manually examined the data trace for Cam5 and other IoT devices around the times when these instances occurred (for both TCP and UDP). Our conclusion is that, this was likely related to the specific application running on Cam5 for video delivery, which may crash or malfunction from time to time.

Non-IoT devices also generate relatively stable hourly number of TCP flows, though with a relatively larger fluctuation (and also larger number of TCP flows). A number of non-IoT devices also generate zero number of TCP flows during a number of hourly time intervals. Based on our observation, this could be caused by either taking a device outside the home or accidentally turning off a device.

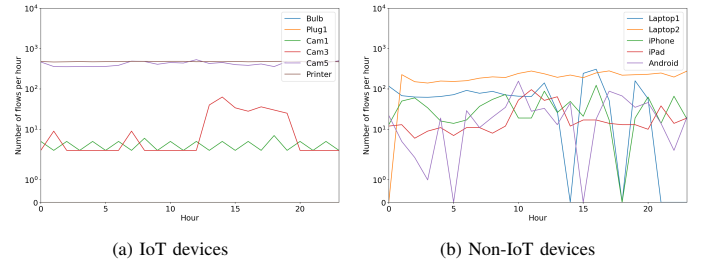


Figure 6: Hourly number of UDP flows.

Figures 6a and 6b show the hourly number of UDP flows of representative IoT and non-IoT devices, respectively. From

the figures we can see that, the hourly number of UDP flows of IoT devices are very stable, except Cam3, which generate larger number of UDP flows from the noon time to the early night time. This could be caused by a larger amount of data traffic to be delivered during that time interval.

The hourly number of UDP flows of non-IoT devices has higher fluctuation than those of IoT devices. In addition, except one instance around 5AM for Android, all the instances with zero hourly number of UDP flows are coincident with those of TCP flows of the same device. As we have discussed above, these should be caused by us either taking a device outside the home or accidentally turning it off.

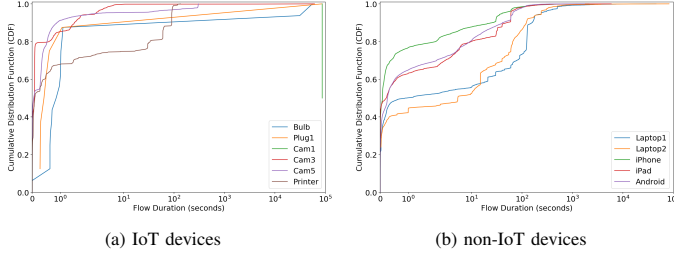


Figure 7: TCP flow duration.

Figure 7a shows the TCP flow durations of the representative IoT devices. We first note that most IoT devices tend to generate either long-lasting flows or very short flows. For example, Cam1 only has two TCP flows and both of these flows last for close to 24 hours (the length of our one-day data trace). Some flows of a few other IoT devices also last for close to 24 hours for the same reason. These devices use long-lasting TCP flows to communicate with their corresponding servers. On the other hand, most TCP flows of IoT devices have short durations, except Printer. For example, about 85% of TCP flows of most IoT devices have a duration less than 1 second. Printer behaves slightly different from the other IoT devices; about 30% of its TCP flows have a duration between 1 second and 100 seconds.

Figure 7b shows the TCP flow duration of the non-IoT devices. We note first that, most non-IoT devices do not have long-lasting TCP flows (except Laptop2). The majority of them last less than 1000 seconds (about 16 minutes). In addition, their flow durations are more spread out between 0 second and 1000 seconds. Like IoT devices, non-IoT devices also have a large portion of short TCP flows (for example, with flow duration less than 1 second). However, the percentage of short TCP flows of non-IoT devices is in general lower than that of IoT devices.

Figures 8a and 8b show the UDP flow durations of IoT and non-IoT devices, respectively. From the figure we can see that, IoT and non-IoT devices have different behaviors in terms of UDP flow durations. First we note that the majority of UDP flows of IoT devices have very short UDP flow durations (close to 0 seconds). This is caused by the fact that IoT device only periodically communicate with certain remote UDP servers such as NTP servers. In contrast, UDP flows of non-IoT are in general much longer. For example,

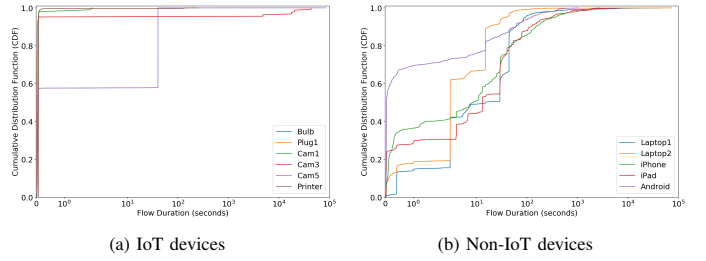


Figure 8: UDP flow duration.

a large portion of UDP flows of all non-IoT devices have a duration between 1 second and 1,000 seconds. When a non-IoT device communicates with a remote UDP server, they tend to exchange traffic continuously, and therefore, resulting in longer UDP flows.

Cam5 in Figure 8a behaves differently from other (IoT) devices. It has a concentrated flow duration of around 40 seconds for about 40% of its UDP flows. Our manual examination of the data trace shows that, Cam5 uses remote port 32100 to communicate with amazonaws servers to send video traffic. A large number of these flows have similar duration (around 40 seconds). We have similarly inspected the UDP traffic of Laptop2 (Figure 8b), which shows that it uses SSDP (with destination IP address 239.255.255.250, and port 1900) to facilitate UPnP (Universal Plug and Play). Most of the SSDP flows a duration around 3 seconds. We see another jump around 15 seconds, which are contributed by flows to a google service (google safe browsing).

2) *Flow Size*: In this subsection we study the size of TCP and UDP flows in terms of both the number of packets and the amount of traffic (in bytes). We note that the amount of traffic in a packet is the number of bytes in the payload of a TCP or UDP packet; and therefore, the amount of traffic in a flow is the sum of the payload of all packets in the flow.

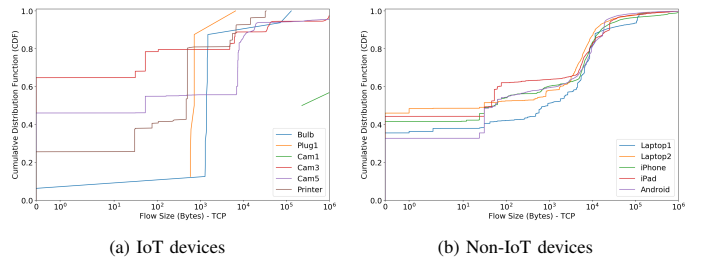


Figure 9: TCP flow size (bytes).

Figures 9a and 9b show the flow sizes of TCP flows of IoT and non-IoT devices (in bytes), respectively. We note that all non-IoT devices have similar behavior in terms of flow sizes, for example, about 70% of TCP flows of all non-IoT devices have a flow size less than or equal to 10^4 bytes. In contrast, IoT devices have more diverse behaviors in terms of flow sizes. For example, both TCP flows of Cam1 are large flows. One generates 229,892 bytes of traffic, while another generates more than 10 GB of traffic (not shown in the figure). We note that Cam1 is positioned at the front

porch facing a reasonably busy street, and as a consequence, Cam1 generates large volume of traffic. In addition, both TCP flows of Cam1 are long-lasting flows, which also contributes to the observation of large flow sizes. The other IoT devices in general generate smaller TCP flows compared to non-IoT devices.

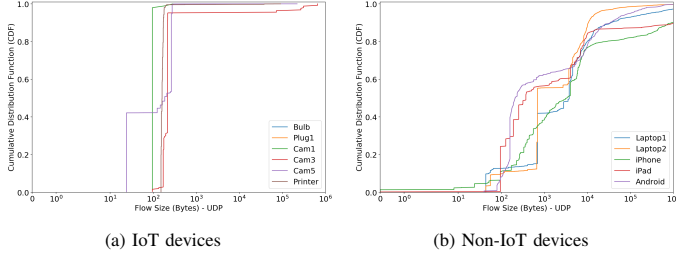


Figure 10: UDP flow size (bytes).

Figures 10a and 10b show the sizes of UDP flows of IoT and non-IoT devices (in bytes), respectively. We have similar observations with that of TCP flow sizes. In particular, UDP flows of non-IoT devices behave similarly in terms of flow sizes. They follow a similar shape and they in general generate more traffic than IoT flows. In contrast, most UDP flows of IoT devices have small sizes; for example, about 90% of all UDP flows of IoT devices have a flow size less than 200 bytes.

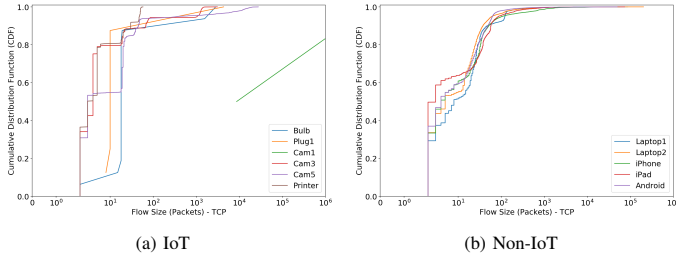


Figure 11: TCP flow size (packets).

Figures 11a and 11b show the TCP flow sizes in terms of number of packets for IoT and non-IoT devices, respectively. From the figures we can see that in general non-IoT devices generate larger flows. For example, about 80% of TCP flows of IoT devices generate less than 20 packets. On the other hand, only about 60% of TCP flows of non-IoT devices generate less than 20 packets. Cam1 is an exception; both of its flows generate large number of packets (8681 and 10893101, respectively). Given that they are both long-lasting flows, this observation is not a surprise.

Figures 12a and 12b show the UDP flow sizes in terms of packets for IoT and non-IoT devices, respectively. As we can see in the figures, IoT and non-IoT devices differ greater in UDP flow sizes than in TCP flow sizes. For example, above 90% of UDP flows of most IoT devices generate 20 packets; on the other hand, only less than 30% of UDP flows of most non-IoT devices generate that amount of packets. In general, UDP flows of non-IoT devices generate greater amount of packets than IoT devices. This can be similarly understood based on different usages of UDP traffic in IoT and non-IoT

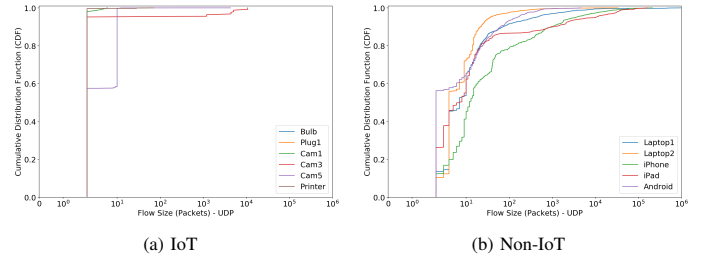


Figure 12: UDP flow size (packets).

devices, in particular, many IoT devices use NTP (on top of UDP) to synchronize their clocks.

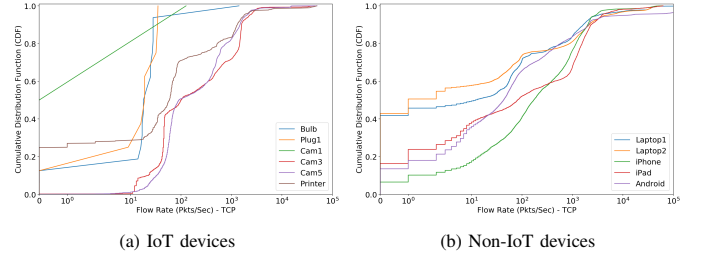


Figure 13: TCP flow rate (packets/second).

3) *Flow Rate*: Figures 13a and 13b show the TCP flow rates (in packets per second) of IoT and non-IoT devices, respectively. From the figures we can see that, the majority of TCP flows of IoT devices have a flow rate between 10 and 3000 packets/second. In contrast, TCP flow rates of non-IoT devices are more spread out, with majority of them ranging from 1 to 10,000 packets/second. We also note that both IoT and non-IoT devices have a large portion of TCP flows with a flow rate close to zero.

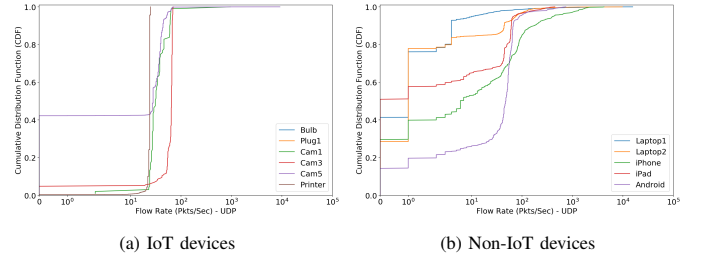


Figure 14: UDP flow rates (packets/second).

We can make some similar observations on UDP flow rates of both IoT and non-IoT devices, as shown in Figures 14a and 14b. In particular, majority of UDP flows of IoT devices have a concentrated flow rate range between 20 and 90 packets/second. On the other hand, the UDP flows of non-IoT devices have a wider range of flow rates, ranging from 1 to 1000 packets/second. Non-IoT devices have a large portion of UDP flows with a flow rate close to zero. The majority of IoT devices have a low percentage of flows with a flow rate close to zero, with an exception of Cam5, which has above 40% of UDP flows with a close to zero flow rate.

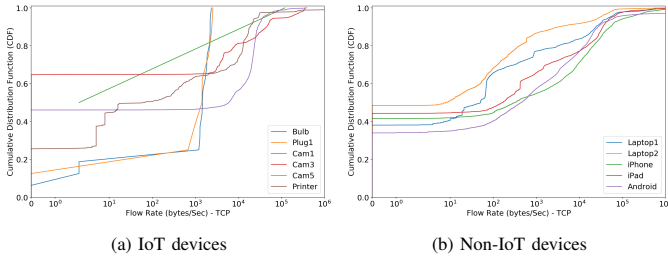


Figure 15: TCP flow rates (bytes/second).

Figures 15a and 15b show the TCP flow rates (in bytes/second) for IoT and non-IoT devices, respectively. From the figures we can again observe that non-IoT devices have a similar behavior in terms of UDP flow rates; their CDFs show similar shapes. In contrast, IoT devices have more diverse TCP flow rates in terms of bytes/second.

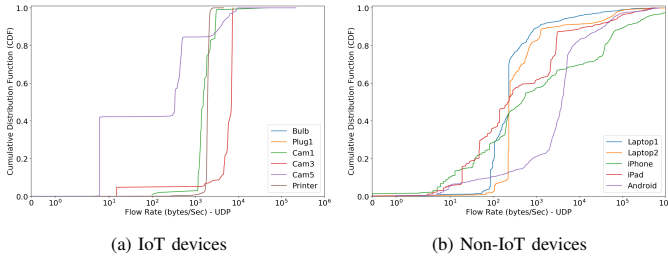


Figure 16: UDP flow rates (bytes/second).

Figures 16a and 16b show the UDP flow rates (in bytes/second) for IoT and non-IoT devices, respectively. From the figure we can see that the majority of UDP flows of most of the IoT devices have flow rates between 1,000 and 10,000 bytes/second. Cam5 is an exception, with a few smaller UDP flow rate ranges. The UDP flows of non-IoT devices have more diverse flow rates, ranging from 10 to 10^6 bytes/second.

4) *Number of Flows*: In this subsection we study the number of flows in each time interval of 120 seconds. We note that, a flow may be counted multiple times if it spans multiple time intervals.

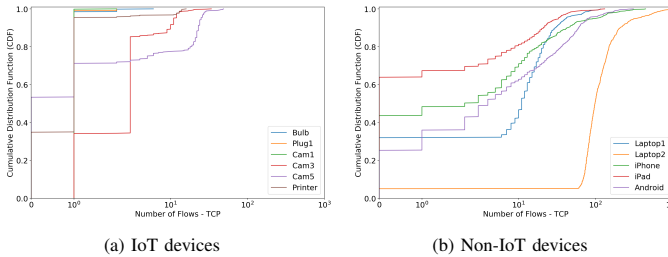


Figure 17: Number of TCP flows.

Figures 17a and 17b show the distribution of number of TCP flows in each time interval for IoT and non-IoT devices, respectively. From the figures we can see that, for all the IoT devices, a large portion of time intervals have no TCP flows at all. For example, Printer has no TCP flows in about 33% of time intervals, and Plug1 has only one TCP flow in about

99% of time intervals. Moreover, no time interval has more than 50 TCP flows among all these IoT devices. In contrast, although non-IoT devices have a large portion of time intervals without any TCP flows, in general, the percentage of such time intervals is lower. In addition, the number of TCP flows in the time intervals is more diverse, ranging from 0 to 1,039 TCP flows.

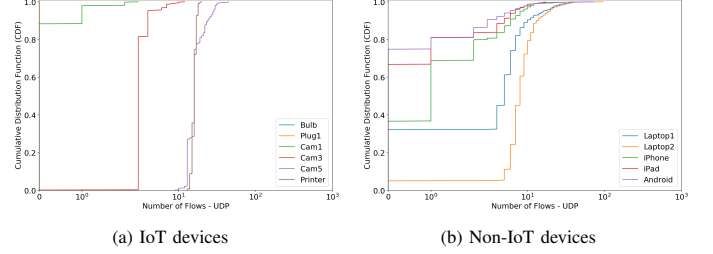


Figure 18: Number of UDP flows.

IoT and non-IoT devices also behave quite differently in terms of the number of UDP flows, as shown in Figures 18a and 18b, respectively. First we note that Cam1 and Cam3 have smaller number UDP flows, with close to 80% and 90% of time intervals without UDP flows at all. In addition, they do not have more than 45 UDP flows in any of time intervals. Cam5 and Printer have relatively larger number of UDP flows, with majority of them having about 16 UDP flows. Overall, they also do not have more than 80 UDP flows in any of the time intervals. Again, non-IoT devices have more diverse number of UDP flows, ranging from 0 to about 100 UDP flows.

C. Packet Level Characteristics

In this subsection we study the packet-level characteristics of IoT devices.

1) *IATs of Packets in a Flow*: First we study the inter-arrival times (IATs) of outgoing packets of both TCP and UDP flows of IoT devices and contrast them with those of non-IoT devices. Note that although flows are bidirectional, in this study, we only investigate the IATs of consecutive outgoing packets of a flow.

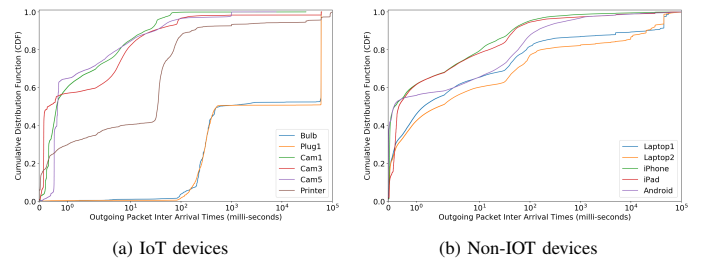


Figure 19: IATs of outgoing packets of TCP flows.

Figures 19a and 19b show the IATs of outgoing packets of TCP flows of IoT and non-IoT devices, respectively. From the figures we first note that the longest IATs of both IoT and non-IoT devices are close to 120 seconds. This could be caused by the timeout threshold that we have used in extracting flows.

We point out that, theoretically speaking, an IAT longer than 120 seconds is possible, given that the flow timeout threshold is applied on bidirectional packets (instead of one way traffic). We also note that both IoT and non-IoT devices have a large portion of small IAT values. For example, all the security cameras (IoT devices) and phones and ipad (non-IoT devices) have around 60% of IATs smaller than 1ms. Two notable exceptions are Bulb and Plug1; they have in general larger IATs. We also note that, in general, IoT devices behave more differently from each other than non-IoT devices in terms of IATs of outgoing packets of TCP flows.

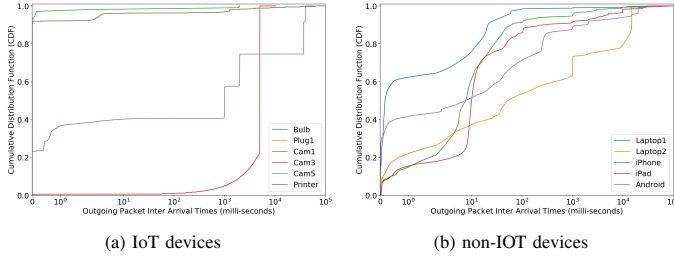


Figure 20: IATs of outgoing packets of UDP flows.

Figures 20a and 20b show the IATs of outgoing packets of UDP flows. From the figures we can similarly observe that, non-IoT devices behave closer to each other than IoT devices. Moreover, while IATs of IoT devices tend to concentrate at a few small ranges, the IATs of non-IoT tend to be more spread out.

2) *Hourly Number of Bytes*: In this subsection we study the hourly amount of all TCP and UDP traffic in terms of number of bytes. We note that the amount of traffic only consider the payload of a TCP or UDP packet.

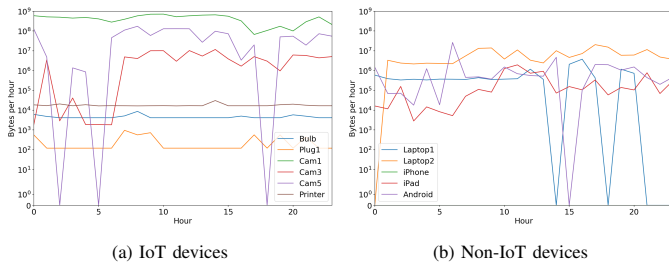


Figure 21: Hourly number of bytes of TCP traffic.

Figures 21a and 21b show the hourly number of bytes of TCP traffic of IoT and non-IoT devices, respectively. From the figures we can see that, some of the IoT devices (Bulb, Plug1, and Printer) have very stable amount of hourly TCP traffic. On the other hand, the security cameras, in particular, Cam3 and Cam5, have more fluctuations in the amount of TCP traffic. Cam5 have a few hours without any TCP traffic, for the potential reason that we have discussed previously. In general, non-IoT devices generate TCP traffic with more fluctuations.

The difference in terms of fluctuations in the number of bytes in UDP traffic is more notable between IoT and non-IoT devices (see Figures 22a and 22b). While all the IoT devices

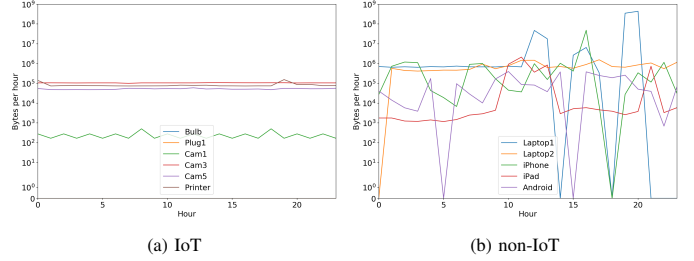


Figure 22: Hourly number of bytes of UDP traffic.

generate relatively stable amount of UDP traffic, the amount of UDP traffic of non-IoT devices has more fluctuations.

3) *Hourly Number of Packets*: Figures 23 and 24 show the hourly amount of traffic in terms of the number of TCP and UDP packets, respectively. From the figures, we can make similar observations as we have seen in the hourly number of bytes (these two characteristics are closely related). In particular, we emphasize the notable difference between the UDP traffic of IoT and non-IoT devices as seen in Figures 24a and 24b.

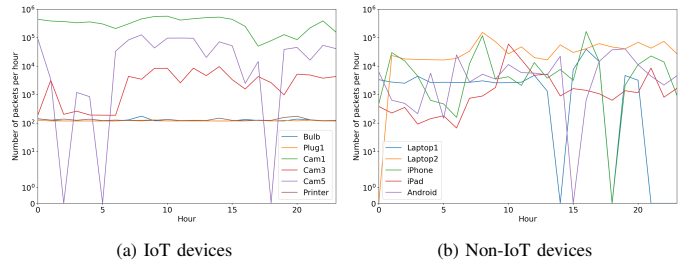


Figure 23: Hourly number of packets of TCP traffic.

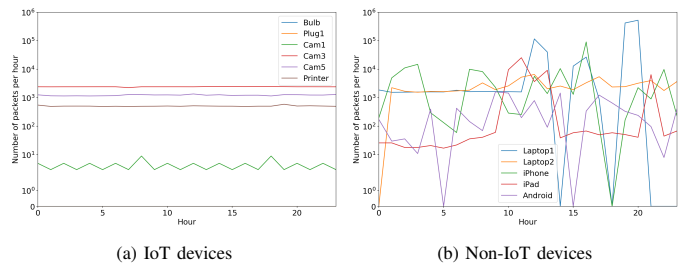


Figure 24: Hourly number of packets of UDP traffic.

V. SUMMARY AND FUTURE WORK

In this paper we analyzed the network traffic characteristics of IoT devices collected on a typical smart home environment consisting of a set of common IoT (and non-IoT) devices. We analyzed the network traffic characteristics of IoT devices from three complementary aspects: remote network servers, flow-level traffic characteristics, and packet-level traffic characteristics. Our study provided critical insights into the operational and behavioral characteristics of IoT devices, which could have important implications on developing effective security and performance algorithms for IoT devices. In our future

work, we plan to study the network traffic characteristics of additional IoT devices, and furthermore, we will apply the observations and insights made in the paper to develop effective algorithms to identify and classing IoT devices and to detect compromised IoT devices.

REFERENCES

- [1] Knud Lasse Lueth. State of the iot 2020: 12 billion iot connections, surpassing non-iot for the first time. <https://bit.ly/3jp5zGI>.
- [2] Mocana. 5 key challenges in securing resource-constrained iot devices. <https://bit.ly/3oItX76>.
- [3] Pedro Abreu. Why manufacturers make insecure iot devices and how you can protect them. <https://bit.ly/3roRWtY>.
- [4] Bruce Schneier. The internet of things is wildly insecure — and often unpatchable. <https://bit.ly/3cH26lu>.
- [5] Noy Hadar, Shachar Siboni, and Yuval Elovici. A lightweight vulnerability mitigation framework for iot devices. In *Proceedings of the 2017 Workshop on Internet of Things Security and Privacy*, pages 71–75, 2017.
- [6] Tim Greene. How the dyn ddos attack unfolded. <https://www.networkworld.com/article/3134057/how-the-dyn-ddos-attack-unfolded.html>.
- [7] M Hammad Mazhar and Zubair Shafiq. Characterizing smart home iot traffic in the wild. In *2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI)*, pages 203–215. IEEE, 2020.
- [8] Yair Meidan, Michael Bohadana, Asaf Shabtai, Martin Ochoa, Nils Ole Tippenhauer, Juan Davis Guarnizo, and Yuval Elovici. Detection of unauthorized iot devices using machine learning techniques. *arXiv preprint arXiv:1709.04647*, 2017.
- [9] Rohan Doshi, Noah Aphthorpe, and Nick Feamster. Machine learning ddos detection for consumer internet of things devices. In *2018 IEEE Security and Privacy Workshops (SPW)*, pages 29–35. IEEE, 2018.
- [10] M. R. Shahid, G. Blanc, Z. Zhang, and H. Debar. Iot devices recognition through network traffic analysis. In *2018 IEEE International Conference on Big Data (Big Data)*, pages 5187–5192, 2018.
- [11] Arunan Sivanathan, Hassan Habibi Gharakheili, Franco Loi, Adam Radford, Chamith Wijenayake, Arun Vishwanath, and Vijay Sivaraman. Classifying iot devices in smart environments using network traffic characteristics. *IEEE Transactions on Mobile Computing*, 18(8):1745–1759, 2018.
- [12] Markus Miettinen, Samuel Marchal, Ibbad Hafeez, N Asokan, Ahmad-Reza Sadeghi, and Sasu Tarkoma. Iot sentinel: Automated device-type identification for security enforcement in iot. In *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, pages 2177–2184. IEEE, 2017.
- [13] Nesrine Ammar, Ludovic Noirie, and Sébastien Tixeuil. Network-protocol-based iot device identification. In *2019 Fourth International Conference on Fog and Mobile Edge Computing (FMEC)*, pages 204–209. IEEE, 2019.
- [14] Openwrt for netgear r8000. <https://openwrt.org/toh/netgear/r8000>.
- [15] Cicflowmeter. <https://github.com/ahlashkari/CICFlowMeter>.
- [16] tshark. <https://www.wireshark.org/docs/man-pages/tshark.html>.
- [17] J. Cai, Z. Zhang, P. Zhang, and X. Song. An adaptive timeout strategy for profiling udp flows. In *2010 First International Conference on Networking and Computing*, pages 44–48, 2010.
- [18] Aws-cloudfront. <https://aws.amazon.com/cloudfront>.