This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2020.2999327, IEEE Internet of Things Journal

# Characterizing DNS Behaviors of Internet-of-Things in Edge Networks

Kuai Xu, *Senior Member, IEEE*, Feng Wang, *Member, IEEE*, Sergio Jimenez, Andrew Lamontagne, John Cummings, and Mitchell Hoikka

***Abstract*—The recent spate of cyber attacks and security threats towards Internet-of-things (IoT) systems in smart cities, smart homes, and industry 4.0 calls for effective techniques to understand if, when, who, what IoT systems are exploited and compromised by Internet attackers. Towards this end, this paper attempts to study DNS behavioral patterns of IoT systems in edge networks as a first step of characterizing their communication patterns and their interactions with IoT users, cloud servers, and other IoT or non-IoT devices in the same edge networks. Specifically, we analyze the temporal-spatial patterns of DNS behaviors of a variety of IoT systems in two dozens of edge networks, and develop a simple yet effective bloom filter mechanism for detecting anomalous traffic patterns based on unusual DNS queries and answers. To the best of our knowledge, this paper is the first effort to systematically measure and monitor IoT network traffic from DNS perspective for providing the security of heterogeneous IoT systems and ensuring IoT user privacy.**

***Index Terms*—IoT network traffic, security and privacy, smart cities, smart homes.**

## I. Introduction

The recent decade has witnessed the explosive growth and deployment of IoT systems in smart homes, smart cities, and industry 4.0. These IoT systems bring a wide spectrum of innovative and disruptive applications and services such as smart energy, remote healthcare, and transportation safety. However, the recent spate of cyber attacks towards IoT systems such as Mirai botnet [1] has exploited the weakness and vulnerabilities of these massive and distributed systems including weak password configuration, out-of-date and unpatched firmwares, softwares and operating systems as well as insufficient security monitoring and management.

As a first step for securing billions of IoT systems in millions of edge networks, this paper proposes and develops a built-in DNS traffic monitoring systems on programmable edge routers for characterizing how IoT systems communicate with remote cloud servers from DNS query and response activities which capture when and what on the data communication of IoT systems. The availability of DNS monitoring system allows us to systematically examine DNS traffic of IoT systems and correlate with network traffic flows, and ultimately exploit the benefits of these DNS traffic patterns for IoT security monitoring and anomaly detection.

Kuai Xu, Feng Wang, Sergio Jimenez, Andrew Lamontagne, John Cummings, and Mitchell Hoikka are with Arizona State University, USA.



Our temporal analysis on DNS traffic on two different levels of DNS names: fully qualified domain name (FQDN) and effective second-level domain (e2LD) has discovered that heterogeneous IoT systems have distinctive DNS queries patterns due to the difference in behavioral activities and functionalities. In addition, our convergence analysis of DNS behaviors reveals that the numbers of FQDNs and e2LDs for all IoT systems tend to remain stable after a certain observation time period, which suggests that these systems typically communicate with a fixed set of hosts and domains on the cloud.

Based on the insights gained from characterizing and analyzing DNS traffic, we explore the applications of DNS-based traffic analysis for the security monitoring of IoT systems in distributed edge networks as well as for for anomaly detection of IoT systems via a simple yet effective bloom filter data structure. Specifically, we leverage if and when DNS queries and response occur for any IoT network traffic activities to gain an in-depth and multi-dimensional understanding of remote cloud servers with which IoT systems have communicated. In addition, to detect new domain names being queried by IoT systems, we develop a Bloom-based probabilistic data structure in the built-on DNS monitoring system in programmable edge routers to effectively monitor and capture emerging and novel FQDNs and e2LDs sent by IoT systems.

The contributions of this paper are summarized as follows:

- We develop a lightweight DNS traffic monitoring system in edge networks to capture DNS traffic of IoT systems for behavioral analysis and modeling.
- We systematically analysis DNS traffic of IoT systems from a range of perspectives, and discover a number of important findings with critical implications in IoT security and privacy.
- We exploit the benefits of the behavioral analysis of modeling of IoT systems in IoT security monitoring and anomaly detection in real-world edge networks.

The remainder of this paper is organized as follows. Section II briefly describes the research background of this paper and introduces our DNS behavioral monitoring system which captures DNS traffic and network flows in real-world edge networks for analysis. Section III is devoted to the analysis of DNS traffic data as well as network flows for gaining a deep understanding of DNS behaviors of various IoT systems. In Section IV, we shed light on a variety of applications of DNS behavioral analysis in security monitoring and anomaly

detection. Section V discusses existing studies in the research area. Finally, Section VI concludes this paper and outlines our future work.

## II. Background and DNS Traffic Monitoring System

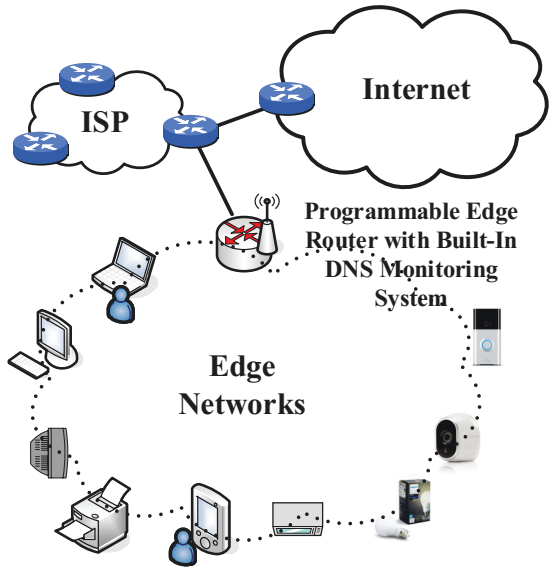



Fig. 1: An edge network example with the proposed DNS traffic monitoring system for securing IoT devices.

The last decade has witnessed the explosive growth and deployment of IoT systems in smart homes, smart cities, and smart campus. The Internet-connected IoT systems have introduced a variety of disruptive applications and services such as home automation, remote healthcare, and smart energy. However, the recent spate of cyber attacks towards vulnerable IoT systems calls for effective techniques for securing billons of IoT systems and protecting the privacy of IoT users [2], [3].

As IoT systems often communicate with cloud servers for normal services and functions, analyzing IoT traffic has recently become a natural approach for security monitoring and behavioral analysis. Similar to computers, tablets and smart phones, both IPv4 and IPv6 based IoT systems rely on DNS service for resolving hostnames to IP addresses to connect and communicate with cloud servers or other Internet end-systems. As a first step of characterizing IoT network traffic, analyzing their DNS traffic behaviors offer a unique perspective to understand behavioral patterns and activities of IoT systems.

For DNS traffic, this study mainly focuses on *DNS queries* from IoT systems in edge networks as well as *DNS replies* from DNS servers. Our empirical observations show that IoT systems typically send DNS queries to three types of DNS servers: edge routers, public DNS servers, and manufactory DNS servers. The edge routers such as home routers often serve as the default DNS routers for IoT systems as well as other Internet-connected devices, while the latter two cases, i.e., public DNS servers, and manufactory DNS servers, are preconfigured by IoT manufacturers or configured during the initial setup stage.

The combination of simple DNS queries and replies provide a rich set of behavioral insights on IoT systems and their interactions with cloud servers. In particularly, these DNS traffic reveal when IoT systems attempt to communication with cloud servers, and how frequent of such communications. In addition, correlating with the underlying network flows or IP packets, we can also explore if DNS traffic is or is not triggered before any communications.

Towards this end, this paper develop a lightweight DNS and network traffic monitoring system on programmable edge routers in edge networks for automatically capturing, analyzing, and making sense of DNS traffic of IoT systems. The edge networks in this study refer to single-home or multi-homed access networks that directly support the IPv4 or IPv6 Internet connections to end users in smart cities, smart homes, small business offices, and enterprise networks. Unlike transit, core or backbone networks run by Internet service providers, edge networks do not provide Internet transit services for other networks. As illustrated in Figure 1, our built-on DNS monitoring system on programmable edge routers is able to fully capture all non-encrypted DNS queries originating from or non-encrypted DNS replies destined to IoT systems in edge networks. More importantly, a unique strength of our edge router based IoT security monitoring platform is robust in the coming age of DNS encryption, since edge routers are often configured as the default DNS servers of IoT devices thus gaining the full visibility of DNS queries and responses from or towards IoT devices in the same edge networks. In the rest of this paper, we will first systematically examine DNS traffic of IoT systems and subsequently exploit the benefits of these DNS traffic patterns for IoT security monitoring and anomaly detection.

## III. DNS Behavioral Analysis

In this section, we present our preliminary results from making sense of DNS logs. In particularly, we analyze DNS behaviors from temporal perspective, and conduct convergence analysis for discovering the stability of DNS query patterns of IoT systems in edge networks.

### A. Temporal Analysis of DNS Behaviors

We first start to study when IoT systems send DNS queries over time. Fig. 2 shows the number of DNS queries in every 5 minutes during a 24-hour time window for six different IoT systems including Amazon Echo plus, IP camera, Amazon Echo dot, Philip Hue smart light, LG smart tv, and Samsung SmartThing hub, respectively. As shown in Fig. 2, the temporal patterns of IoT systems can be generally classified into three different types: always-on, periodical, and random. The Amazon Echo plus, Amazon Echo dot, and Samsung SmartThing hub exhibit always-on traffic activities, the Philip Hue smart light bulb shows periodical traffic activities, while

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2020.2999327, IEEE Internet of Things Journal

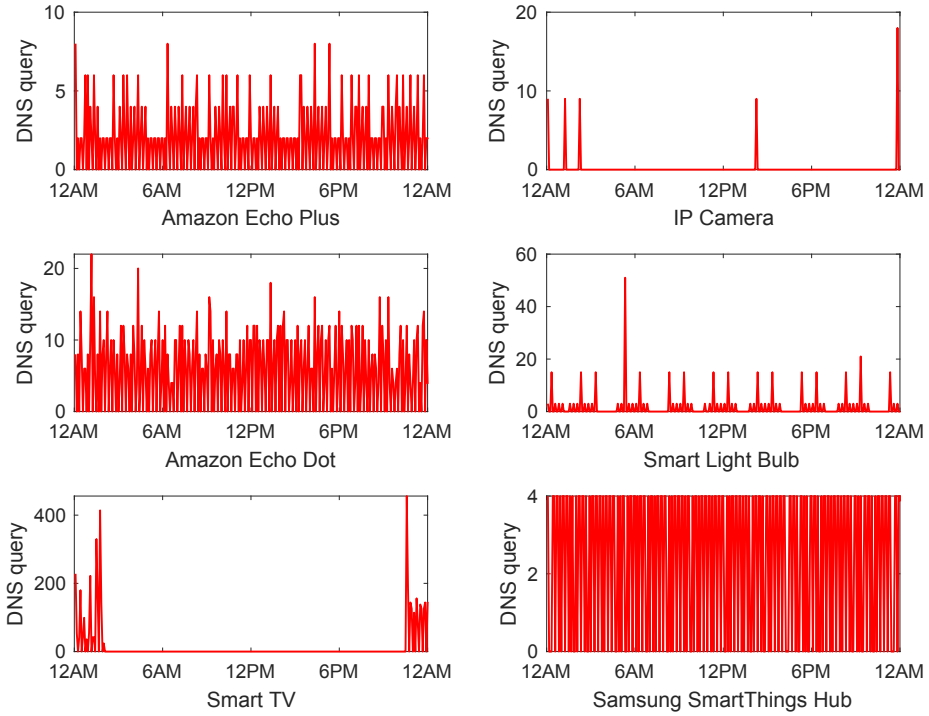
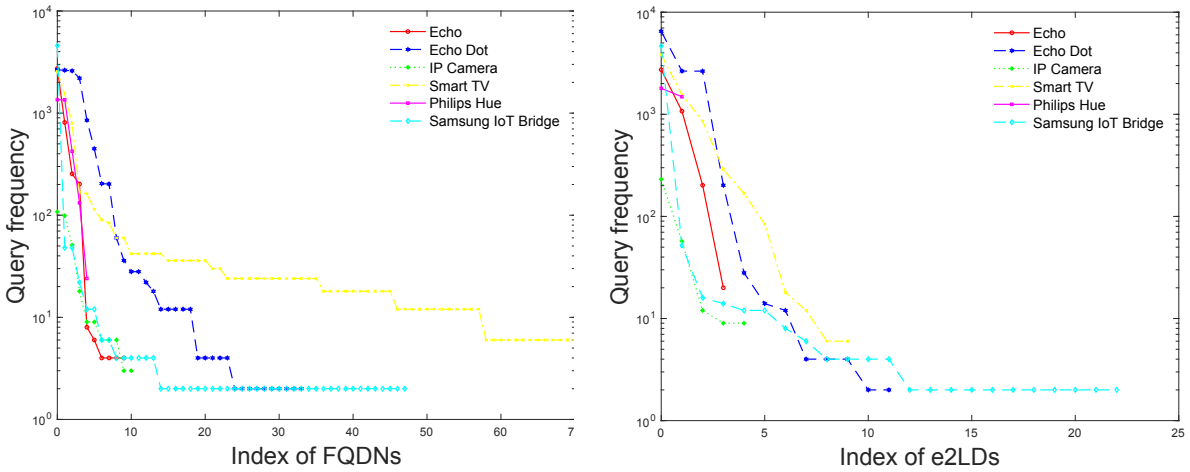



Fig. 2: Temporal observation of DNS queries over time for a variety of IoT devices in edge networks.




(a) FNDNs (b) e2LDs

Fig. 3: The query frequency of FQDNs and e2LDs by a variety of IoT systems.

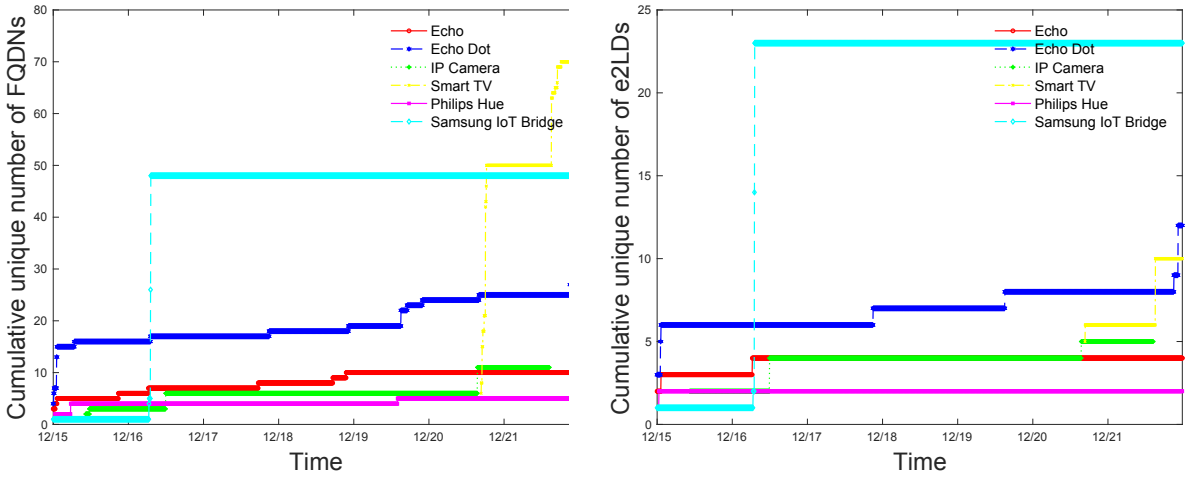
the IP camera and the LG smart TV have random traffic activities.

The diverse patterns of DNS queries from IoT systems motivate us to explore what hosts or domains are being queried by IoT systems and how frequent these hosts or domains are being queried by these systems. To understand the domains and hosts being queried by IoT systems, we focus on two different levels of DNS names: fully qualified domain name (FQDN) and effective second-level domain (e2LD).

A FQDN such as `developer.twitter.com`, also referred to as an absolute domain name, is a complete domain name of a host on the Internet in the form of [hostname].[domain].[tld], while an e2LD such as `twitter.com` or `bbc.co.uk` captures the real domain ownership and is a common DNS hierarchy for aggregating FQDNs in the DNS research literature.

Figures 3[a-b] illustrate the query frequencies of FQDNs and e2LDs by six different IoT systems including Amazon Echo dot, IP camera, smart TV, Philips Hue smart bulb, and Samsung SmartThings bridge, respectively. As shown in Figures 3[a-b], the query frequencies for six IoT systems exhibit similar patterns: a few FQDNs and e2LDs are queried extensively by IoT systems. Table I presents the most frequent queried FQDNs and e2LDs by six different IoT systems over

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2020.2999327, IEEE Internet of Things Journal




(a) FNDNs (b) e2LDs

Fig. 4: The convergence of unique and cumulative FQDNs and e2LDs over time by a variety of IoT systems.

one-week time-span. However, the majority of FQDNs and e2LDs are occasionally queried by these systems. It is not surprising to observe that the general query pattern for all of these IoT systems is to communicate with the FQDNs and e2LDs associated with the manufacturers or their content distribution or cloud service providers such as Amazon CloudFront and Amazon Web Services.

### *B. Convergence Analysis of DNS Behaviors*

The convergence analysis of DNS behaviors is centered on the growth and persistence of FQDNs and e2LDs over time. In other words, we are interested in understanding if the numbers of FQDNs and e2LDs remain stable or keep growing during our observation time window. Figures 4[a-b] show the convergence patterns of unique and cumulative FQDNs and e2LDs, respectively. For each of these six IoT systems, the numbers of FQDNs and e2LDs remain stable after a certain observation time period. These observations suggest that IoT systems typically communicate with a set of fixed hosts and domains on the Internet.

We believe that such insights are critical and valuable in anomaly detection of IoT behaviors. For example, a new hostname or domain queried by any IoT systems is worth investigating since communicating with such emerging and novel hosts indicates unusual activity of IoT systems which are popular targets by cyber attackers due to various vulnerabilities and weaknesses.

## IV. Applications

In this section, we explore the applications of DNS traffic analysis for the security monitoring of IoT systems in edge networks as well as anomaly detection for DNS behavior based on a simple yet effective bloom filter cache digest.

### *A. Security Monitoring of IoT Network Traffic via DNS Perspectives*

DNS traffic analysis plays an important role in security monitoring of IoT systems. Communicating with cloud servers on the Internet is the dominant activity between IoT systems and end systems on the Internet. Characterizing the DNS traffic can reveal if and when DNS query and response happen for such communication.

Table II shows the distribution of cloud servers seen in DNS vs. not seen in DNS for six different IoT systems in edge networks. For all these systems, we observe every system communicates with some cloud servers without resolving the IP addresses of these servers. For example, among 281 unique cloud servers the smart TV communicated with, there are no DNS traffic observed for 9 (3%) cloud servers. Considering the possible DNS caching feature enabled on IoT devices, this study searches for DNS activities for network applications from the previous 12-hour time window until the timestamp on which the application traffic was started. To fully understand these servers without DNS queries and response, we continue to further characterize the remote ports, network prefixes, and BGP domains, also referred to as autonomous system numbers (ASN) associated with cloud servers with which IoT systems have communicated.

As shown in Table III, our analysis on 383 unique cloud servers without DNS traffic activity shows all the communications between six IoT systems and cloud servers are centered on a small set of application ports including 53/UDP (DNS), 80/TCP (HTTP), 123/UDP (NTP), 443/TCP (HTTPS), 4070/TCP (Spotify), 33434/UDP (Alexa Media Streaming), and ICMP echo request. We conjecture that the similar observations will hold for IoT systems with dedicated or simple functionality such as smart plugs, smart thermostats, and smart locks. However, IoT devices with a rich set of features such as smart watches Google Home could communicate with cloud servers for a variety of functionalities. Our in-depth analysis shows the following data communications for these services during the traffic collection time period:

- For ICMP traffic, we have observed Amazon Echo and SmartThings Hub send ICMP echo requests, i.e, the ping networking utility, to Google public DNS servers 8.8.8.8 and 8.8.4.4 for ensuring the availability of DNS services.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2020.2999327, IEEE Internet of Things Journal

| IoT device | FQDNs (top 5) | Frequency | e2LDs (top 5) | Frequency |
|---|---|---|---|---|
| Amazon Echo | d3p8zr0ffa9t17.cloudfront.net | 2734 | cloudfront.net | 2734 |
| FQDNs: 10 | device-metrics-us.amazon.com | 810 | amazon.com | 1074 |
| e2LDs: 4 | dcape-na.amazon.com | 254 | amazonalexa.com | 202 |
| | api.amazonalexa.com | 202 | spotify.com | 20 |
| | guc3-ap-b-zljs.ap.spotify.com | 8 | | |
| IP Camera | mcs.arlo.com | 108 | arlo.com | 231 |
| FQDNs: 11 | ntp03.arlo.com | 99 | amazonaws.com | 57 |
| e2LDs: 5 | mcs.west-1.elb.amazonaws.com | 51 | arloxcld.com | 12 |
| | updates.arlo.com | 18 | akamaiedge.net | 9 |
| | arlo-device.messaging.arlo.com | 6 | edgekey.net | 9 |
| Echo Dot | spectrum.s3.amazonaws.com | 2650 | amazon.com | 6502 |
| FQDNs: 34 | kindle-time.amazon.com | 2630 | amazonaws.com | 2650 |
| e2LDs: 10 | d3p8.cloudfront.net | 2594 | cloudfront.net | 2646 |
| | ntp-g7g.amazon.com | 2198 | amazonalexa.com | 202 |
| | api.amazonalexa.com | 202 | amazoncrl.com | 28 |
| Smart Light | dcp.dc1.philips.com | 1353 | meethue.com | 1797 |
| FQDNs: 5 | www2.meethue.com | 1350 | philips.com | 1485 |
| e2LDs: 2 | diagnostics.meethue.com | 423 | | |
| | www.ecdinterface.philips.com | 132 | | |
| | ws.meethue.com | 24 | | |
| Smart TV | api-global.netflix.com | 2559 | netflix.com | 3837 |
| FQDNs: 70 | art-s.nflximg.net | 90 | nflxvideo.net | 858 |
| e2LDs: 9 | us.rdx2.lgtvsdp.com | 84 | nflximg.net | 288 |
| | ipv4-phx001-ix.nflxvideo.net | 42 | lgtvsdp.com | 84 |
| | us.info.lgsmartad.com | 6 | lgsmartad.com | 6 |
| SmartThings Hub | fw-update2.smartthings.com | 4576 | smartthings.com | 4654 |
| FQDNs: 51 | www.google.com | 48 | googleapis.com | 12 |
| e2LDs: 20 | dc.connect.smartthings.com | 40 | google.com | 52 |
| | api.smartthings.com | 12 | samsungcloud.com | 8 |
| | www.amazon.com | 2 | samsungiotcloud.com | 6 |

TABLE I: The most frequent queried FQDNs and e2LDs by six different IoT systems over one-week time-span.

| IoT device | Total Cloud Servers | Server seen in DNS | Server not seen in DNS |
|---|---|---|---|
| Amazon Echo | 65 | 38 (58%) | 27 (42%) |
| IP Camera | 19 | 12 (63%) | 7 (37%) |
| Echo Dot | 1406 | 1079 (77%) | 327 (23%) |
| Smart Light | 20 | 16 (80%) | 4 (20%) |
| Smart TV | 281 | 272 (97%)) | 9 (3%) |
| SmartThings Hub | 69 | 60 (87%) | 9 (13%) |

TABLE II: The distribution of cloud servers seen in DNS vs. not seen in DNS for different IoT systems in edge networks.

| Port/Protocol | Echo | Camera | Dot | Smart Light | Smart TV | SmartThings Hub | Note |
|---|---|---|---|---|---|---|---|
| ICMP | 1 | | | | | 2 | Ping |
| 53/UDP | 2 | 3 | 1 | | 1 | 2 | DNS |
| 80/TCP | 1 | | 319 | | 1 | 4 | HTTP |
| 123/UDP | 18 | | 2 | 4 | | | NTP |
| 443/TCP | 5 | 4 | 4 | | 7 | 3 | HTTPS |
| 4070/TCP | | | 1 | | | | Spotify |
| 33434/UDP | 1 | | | | | | Alexa Streaming |
| Total dstIP | **28** | 7 | 327 | 4 | 9 | **11** | |
| Unique dstIP | **27** | 7 | 327 | 4 | 9 | **9** | |

TABLE III: The remote application ports associated with cloud servers without DNS traffic activity for six IoT systems.

- For port 53/UDP DNS traffic, Amazon Echo, IP camera, Amazon Echo Dot, LG smart TV, and Samsung SmartThing hub have communicated with a variety of Google public DNS servers and OpenDNS DNS servers.
- For port 80/TCP HTTP traffic, Amazon Echo and Amazon Echo dot communicate with 1 and 319 Amazon cloud servers respectively, LG smart TV communicates 1 Akamai content distribution server, while Samsung Smartthing hub communicates with 3 Samsung cloud servers and 1 Amazon cloud server.
- For port 123/UDP NTP traffic, we have observed 18 different servers from 12 ASNs including Amazon NTP servers as well as unexpected NTP servers such as `ntp3.junkemailfilter.com`. The high and diverse number of servers on NTP ports warrants further security analysis, which is part of our ongoing effort for deep packet investigation on such *unusual* traffic behaviors. Similarly, we also have noticed two *unexpected* servers hosted by two cloud server providers: Linode and NetActuate. Unlike Amazon Echo and Amazon Echo dot, Philip Hue smart light *normally* communicates with 4 Google NTP servers, .i.e., `time[1-4].google.com` for network time synchronization.
- For port 443/TCP HTTPS traffic, we have observed Amazon Echo communicates with 3 Amazon cloud servers and 1 Google server, IP camera communicates with 2 Amazon servers and 2 servers in Cox Networks, Amazon Echo Dot communicates with 4 Amazon servers, LG smart TV communicates with 6 Amazon cloud servers and 1 Netflix server, and Samsung SmartThings communicates with 1 Fastly cloud server and 2 Samsung servers.
- For port 4070/TCP Spotify traffic, we observe Amazon Echo Dot communicates with 1 Spotify server hosted on Google cloud.
- For 33434/UDP Amazon media streaming traffic, we observe Amazon Echo communicates with 1 Amazon streaming server.
- It is interesting to note that we have observed Amazon Echo and SmartThings Hub engage DNS traffic and ICMP traffic on different destination ports with the same one `dstIP` and the same two `dstIPs`, respectively. Our follow-up and in-depth investigation reveals that first case is due to 8.8.8.8 while the second case is due to 8.8.8.8 and 8.8.4.4. Both of these IP addresses are Google public IPv4 DNS servers that are widely configured in millions of IoT systems in edge networks. This observation explains why the number unique `dstIPs` is less than the total number of `dstIPs` in Table III for Amazon Echo and SmartThings Hub.

### *B. Bloom-filter based Anomaly Detection of DNS Behavior*

Our analysis on the DNS query frequency on FQDNs and e2LDs has discovered that all FQDNs and e2LDs observed during the one-week time period are queried by IoT systems for two times or more. This insight leads us to develop a simple yet effective technique based on Bloom filter to monitor and detect emerging and novel FQDNs and e2LDs sent by IoT systems. Monitoring such emerging domain names is critical for detecting anomalous behavior of IoT systems due to a wide range of cyber attacks that exploit the weakness and vulnerability of weakly protected IoT systems.

Figure 5 illustrates the Bloom filter structure developed in the built-on DNS monitoring system in programmable edge routers for detecting new domain names being queried by IoT systems. For each IoT system in edge networks, we build a Bloom filter with $m$ bits which chooses $k$ uniformly distributed hash functions which independently calculates $k$ hash values for each domain queried by the IoT system.

The memory-efficient Bloom filter is essentially a probabilistic data structure to test if a newly observed domain is always included in the filter via running the same $k$ hash functions and comparing the hash results with the existing $m$-bit filter. The false negative for the Bloom filter data structure in testing the existence of a previously observed domain $d$ is zero, since all the bits mapped by the $k$ hash functions on the domain $d$ must be set to 1 during the initial mapping stage on the domain $d$. However, the Bloom filter could introduce false positives by confirming the existence of a never-seen domain $d'$ since all the bits of mapping the domain $d'$ might be set to 1 due to other previously observed domains rather than $d'$ itself.

Once a total of $n$ domains are mapped by $k$ hash functions to the $m$-bit Bloom filter, the probability, $p$, of a given bit being 0 is $p = (1 - \frac{1}{m})^{kn}$. Thus, the probability of reporting a false positive, $fp$ of a never-seen domain $d'$ is calculated as $fp = (1-p)^k = (1 - (1 - \frac{1}{m})^{kn})^k \approx (1 - e^{kn/m})^k$.

A previously observed domain name suggests the IoT system repeats the data communication with the remote system on the Internet, while a *newly* observed name indicates a novel or unusual communication that requires further analysis on the follow-up traffic activities. Therefore the Bloom filter in the built-in DNS monitoring system serves as an anomaly detection module for capturing the emerging domain names that IoT systems attempt to communicate with. Part of our future work is to correlate with network traffic flows for determining if the domain names are associated with benign or malicious activities.

## V. Related Work

The explosive growth of IoT systems deployment in smart homes, smart cities and other real-world environments has successfully introduced a variety of innovation and disruptive applications and services such as voice assistants, self-driving vehicles, and delivery drones [4]. However, protecting the security of these IoT systems and ensuring user privacy have met significant challenges due to heterogeneous hardwares and softwares, lack of IoT transparency, as well as insufficient resources on IoT systems [5], [6], [7], [8]. As a first step for addressing the security and privacy challenges, many recent studies have devoted to measuring, monitoring, analyzing, and fingperprinting IoT network traffic and application behaviors [9].

Due to the wide deployment of IoT systems in smart homes, a number of research studies have focused on the security evaluation, threat assessment, and network traffic monitoring [10],

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2020.2999327, IEEE Internet of Things Journal

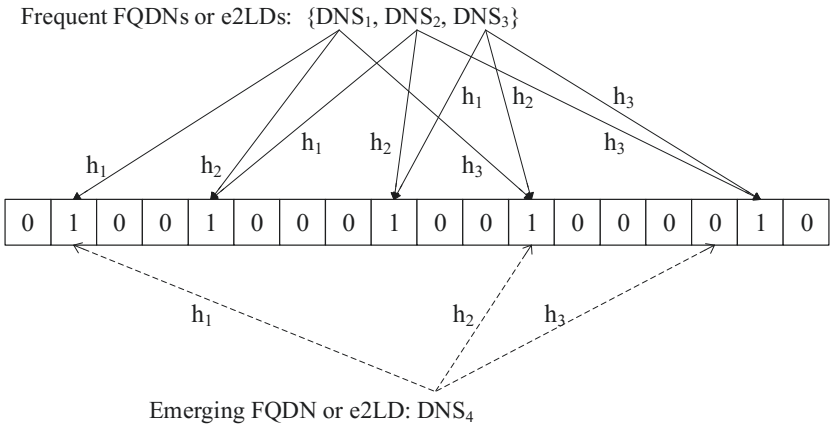



Fig. 5: Detecting new domain names being queried by IoT systems via Bloom filters.

[11], [12], [4] of home-based IoT systems. For example, the study in [10] evaluates a broad range of attack techniques, mitigations, and stakeholders for 45 different home-based IoT devices, while the survey in [11] explores the secure challenges and threats of smart home technologies such as activity recognition, health monitoring, and automation. Similarly, the HoMonit system [12] explores side-channel inference capabilities to evaluate 181 smart home applications on the Samsung SmartThings platform and discovers 60 malicious applications.

Fingerprinting network traffic of IoT systems for security monitoring has recently demonstrated an effective approach for characterizing, understanding, and modeling behavioral patterns of IoT systems [13], [14], [15], [16]. For example, the research in [13] develops a measurement framework to build multidimensional behavioral profiles for heterogeneous IoT devices in edge networks, while the study in [14] extracts various features from IoT network traffic for effectively and accurately identify IoT systems. In parallel, in [15] the researchers monitor traffic patterns of the work-life cycles of IoT devices, and extracts behavioral features from multiple layers in the IoT protocol stacks for device spoofing detection and device fingerprinting, while the IoT SENTINEL tool [16] automatically identifies the types of IoT devices based on 23 network traffic features, and subsequently develops mitigation strategies for minimizing the potential threats from attackers and securing data communications of IoT devices.

Analyzing DNS traffic for system and network security is a widely used approach in enterprise networks and wireless networks [17], [18], [19], [20], [21]. For example, [17] applies the combination of bipartite graphs, one-mode projections, behavioral modeling and graph embedding to characterize DNS traffic patterns for detecting malicious domains which act as advanced persistent threat command and control communication channels or host phishing Web sites, while the research in [18] analyzes the co-clusters from DNS failure graphs in a large campus network for identifying novel traffic anomalies. In addition, the surveys [19], [20] independently present overviews of DNS-based botnet detection techniques and systematically classifies these techniques into distinctive categories based on the underlying methodologies, while the study in [21] explores the correlation of DNS queries with outgoing Internet connections to rapidly detect worm propagation within enterprise networks. Different from these prior studies, this paper focuses on measuring and monitoring the network traffic of IoT systems from DNS perspective for providing the security of IoT systems and ensuring the privacy of IoT users.

## VI. Conclusions and Future Work

The recent spate of cyber attacks towards IoT devices in edge networks have highlighted the importance of effectively managing, monitoring, and securing billions of such vulnerable systems on the Internet. Towards this end, this paper develops a DNS traffic monitoring system for exploring DNS traffic analysis to monitor and mitigate security attacks towards these systems. The available DNS traffic data allows us to characterize DNS behaviors from the temporal perspective and perform convergence analysis for understanding the stability and persistence of DNS query patterns of heterogenous IoT systems in distributed edge networks. Based on the critical insights gained from DNS traffic analysis, this paper further explore the applications of DNS traffic pattern analysis for the security monitoring of IoT systems in edge networks as well as anomaly detection for DNS behavior based on a simple yet effective bloom filter cache digest. Our future efforts are centered on i) monitoring and analyzing DNS traffic of IoT systems from distributed edge networks to identify and mitigate coordinated cyberattacks or emerging exploits towards IoT systems, and ii) developing a prototype system of DNS traffic monitoring and analysis for detecting and mitigating security threats and malicious attack towards millions of Internet-connected IoT device in edge networks.

## Acknowledgment

This research was supported in part by NSF grant 1816995. The information reported here does not reflect the position or the policy of the funding agency.

## REFERENCES


[1] M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. Halderman, L. Invernizzi, M. Kallitsis, D. Kumar, C. Lever, Z. Ma, J. Mason, D. Menscher, C. Seaman, N. Sullivan, K. Thomas, and Y. Zhou, "Understanding the Mirai Botnet," in *Proceedings of USENIX Security Symposium*, August 2017.

[2] G. Ho, D. Leung, P. Mishra, A. Hosseini, D. Song, and D. Wagner, "Smart Locks: Lessons for Securing Commodity Internet of Things Devices," in *Proceedings of ACM on Asia Conference on Computer and Communications Security (ASIACCS)*, May 2016.

[3] J. Obermaier and M. Hutle, "Analyzing the Security and Privacy of Cloud-based Video Surveillance Systems," in *Proceedings of ACM International Workshop on IoT Privacy, Trust, and Security (IoTPTS)*, May 2016.

[4] D. Kumar, K. Shen, B. Case, D. Garg, G. Alperovich, D. Kuznetsov, R. Gupta, and Z. Durumeric, "All Things Considered: An Analysis of IoT Devices on Home Networks," in *Proc. of the USENIX Security*, 2019.

[5] D.-T. Nguyen, C. Song, Z. Qian, S.-V. Krishnamurthy, E.-J.-M. Colbert, and P. McDaniel, "IotSan: Fortifying the Safety of IoT Systems," in *Proc. of CoNEXT*, 2018.

[6] E. Ronen, A. Shamir, A.-O. Weingarten, and C. O'Flynn, "IoT Goes Nuclear: Creating a ZigBee Chain Reaction," in *Proceedings of IEEE Symposium on Security and Privacy (S&P)*, May 2017.

[7] S. Soltan, P. Mittal, and H.-V. Poor, "Blackiot: Iot botnet of high wattage devices can disrupt the power grid," in *Proc of USENIX Security*, 2018.

[8] A. Wang, A. Mohaisen, and S. Chen, "XLF: A Cross-layer Framework to Secure the Internet of Things (IoT)," in *Proc. IEEE ICDCS*, 2019.

[9] Q. Wang, W. Hassan, A. Bates, and C. Gunter, "Fear and Logging in the Internet of Things," in *Proceedings of Network and Distributed System Security Symposium (NDSS)*, February 2018.

[10] O. Alrawi, C. Lever, M. Antonakakis, and F. Monrose, "SoK: Security Evaluation of Home-Based IoT Deployments," in *Proceedings of IEEE Symposium on Security and Privacy (S&P)*, May 2019.

[11] J. Dahmen, D. Cook, X. Wang, and H. Wang, "Smart Secure Homes: A Survey of Smart Home Technologies that Sense, Assess, and Respond to Security Threats," *Journal of Reliable Intelligent Environments*, vol. 3, no. 2, pp. 83–98, 2017.

[12] W. Zhang, Y. Meng, Y. Liu, X. Zhang, Y. Zhang, and H. Zhu, "HoMonit: Monitoring Smart Home Apps from Encrypted Traffic," in *Proceedings of ACM SIGSAC Conference on Computer and Communications Security (CCS)*, October 2018.

[13] K. Xu, Y. Wan, G. Xue, and F. Wang, "Multidimensional Behavioral Profiling of Internet-of-Things in Edge Networks," in *Proc. of IEEE/ACM IWQoS*, 2019.

[14] B. Bezawada, M. Bachani, J. Peterson, H. Shirazi, I. Ray, and I. Ray, "Behavioral Fingerprinting of IoT Devices," in *Proceedings of ACM CCS Workshop on Attacks and Solutions in Hardware Security (ASHES)*, October 2018.

[15] T. Gu and P. Mohapatra, "BF-IoT: Securing the IoT Networks via Fingerprinting-Based Device Authentication," in *Proceedings of IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, October 2018.

[16] M. Miettinen, S. Marchal, I. Hafeez, N. Asokan, A.-R. Sadeghi, and S. Tarkoma, "IoT SENTINEL: Automated Device-Type Identification for Security Enforcement in IoT," in *Proceedings of IEEE International Conference on Distributed Computing Systems (ICDCS)*, July 2017.

[17] K. Lei, Q. Fu, J. Ni, F. Wang, M. Yang, and K. Xu, "Detecting Malicious Domains with Behavioral Modeling and Graph Embedding," in *Proceedings of IEEE International Conference on Distributed Computing Systems (ICDCS)*, July 2019.

[18] N. Jiang, J. Cao, Y. Jin, L. Li, and Z.-L. Zhang, "Identifying Suspicious activities through DNS Failure Graph Analysis," in *Proceedings of IEEE International Conference on Network Protocols*, October 2010.

[19] M. Singh, M. Singh, and S. Kaur, "Issues and Challenges in DNS based Botnet Detection: A Survey," *Computers & Security*, vol. 86, pp. 28–52, September 2019.

[20] K. Alieyan, A. ALmomani, A. Manasrah, and M. Kadhum, "A Survey of Botnet Detection based on DNS," *Neural Computing and Applications*, vol. 28, no. 7, p. 15411558, July 2017.

[21] D. Whyte, E. Kranakis, and P. C. van Oorschot, "DNS-based Detection of Scanning Worms in an Enterprise Network," in *Proceedings of Network and Distributed System Security Symposium (NDSS)*, February 2005.