Poster: Analyzing Endpoints in the Internet of Things Malware

Jinchun Choi[◇‡⋆], Afsah Anwar[◇⋆], Hisham Alasmary[◇],
Jeffrey Spaulding[†], DaeHun Nyang[‡] and Aziz Mohaisen[◇]

[◇]University of Central Florida [†]Niagara University [‡]Inha University [⋆]Equal contributors
{jc.choi, afsahanwar, hisham}@knights.ucf.edu, jspaulding@niagara.edu
nyang@inha.ac.kr, amohaisen@gmail.com

***Abstract*—The lack of security measures in the Internet of Things (IoT) devices and their persistent online connectivity give adversaries an opportunity to target them or abuse them as intermediary targets for larger attacks such as distributed denial-of-service (DDoS) campaigns. In this paper, we analyze IoT malware with a focus on endpoints to understand the affinity between the dropzones and their target IP addresses, and to understand the different patterns among them. Towards this goal, we reverse-engineer 2,423 IoT malware samples to obtain IP addresses. We further augment additional information about the endpoints from Internet-wide scanners, including Shodan and Censys. We then perform a deep data-driven analysis of the dropzones and their target IP addresses and further examine the attack surface of the target device space.**

I. Introduction

With the number of seamlessly connected and online IoT devices soaring into the 10's of billions [1], potential adversaries set such devices on target via malicious codes. Such codes or malware not only infect the devices themselves but also receive updates from their Command and Control (C2)/dropzones around the world. Forming a network, these devices have the potential to launch attacks on other targets resulting in distributed denial-of-service (DDoS) attacks [2].

Reckoning that the malware sources, C2 servers, the intermediary targets, and the victim must be connected to the Internet for the attacks to happen, studying these endpoints is important. This work attempts to understand such endpoints to decipher such patterns. In particular, we extract endpoints from IoT malware samples and perform a data-driven analysis to understand geographical affinities and their exposure to risk.

Our work is important given its insight into understanding the Indicators of Compromise (IoCs), and the behavioral aspects necessary for threat hunting. Specifically, we make the following contributions: 1) IP Centric Analysis: We investigate the target IP addresses among different dropzone IP addresses. Additionally, we analyze the locations of dropzones and their target IP addresses. Moreover, we analyze the risk associated with the IP addresses through insights gained from Shodan [3]. 2) Network Centric Analysis: For the masked target endpoints, we examine the entire network and study the network devices and their exposure to risk.

II. Dataset Creation

Our primary dataset contains a total of 2,423 IoT malware samples collected from IoTPOT [4]. We design a tool to reverse-engineer each of the malware samples and extract the IP addresses from the malware code base. In total, we extract a total of 106,428 target IP addresses, resulting in 2,211 *unique target* IP addresses associated with 973 malware samples. We also find a total of 2,407 IP addresses, resulting in 877 *unique dropzone* (source of attack)/C2 IP addresses corresponding to 2,318 malware samples. We notice that only 40% of the 2,423 malware samples contain target IP addresses, and a total of 95.66% of 2,423 malware samples contain dropzone IP addresses. Additionally, we augment the IP addresses with other details including their: country, city, Autonomous System Number (ASN), and location, using online DNS and IP lookup tools like the *UltraTools* [5]. Furthermore, we utilize Shodan [3] to obtain the vulnerable endpoints.

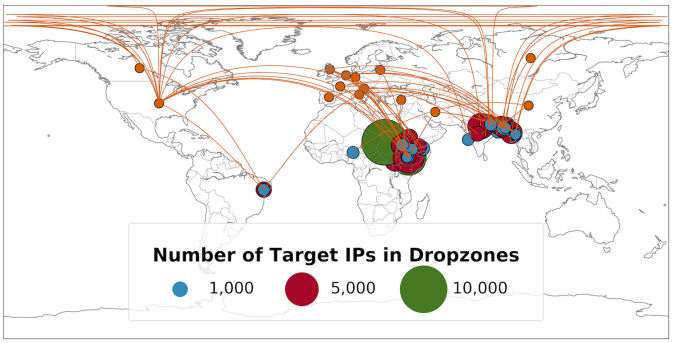


Fig. 1. Attack trends between dropzones and target IPs corresponding to dropzones having over 500 target IPs. The orange circle represents dropzones, and blue, red, and green circles stand for target areas.

III. IP Centric Analysis

A. Dropzone-Target Inter-relationship

We examine the affinity between the dropzone and the target IP addresses and find that ≈77% of the unique target IPs receive less than 10 attacks, while we see that one unique target IP received 72 attacks. We also find a dropzone IP associated with one malware targeting 1,265 target IP addresses, which is significantly larger than the average (121).

Shared targets between dropzones. To inspect the shared targets between dropzone IP addresses, we group the dropzone IP addresses and capture the common targets among the dropzones. There were 2,199 cases (12.11%) with 100% overlap between dropzones. Overall, we found 6,451 cases

TABLE I. TOP 5 NUMBER OF TARGET AND DROPZONE IPS BY COUNTRY. Countries include: United States (US), Netherlands (NL), France (FR), United Kingdom (GB), Italy (IT), Vietnam (VN), Brazil (BR), China (CN), India (IN), and Pakistan (PK).

Country	# Dropzones	%Total	Country	# Targets	%Total
US	1,041	43.25	VN	26,290	24.70
NL	278	11.55	BR	20,572	19.33
FR	188	7.81	CN	15,799	14.84
GB	183	7.60	IN	5,598	5.26
IT	177	7.35	PK	5,076	4.77

(35.53%), with >80% overlap, and 886 cases (4.88%) with <10% overlap. If the target IP addresses between different dropzones are identical, it is possible that the attacker obtained the same targets through similar vulnerability analysis (*e.g.* Shodan) or shared the target list from other attackers through underground communities.

B. Geographical analysis

In this section, we focus on the distribution of the distances between the dropzones and their target IPs. To visualize the flow of attacks in a holistic sense, we plot circular areas whose sizes are proportional to the number of targets placed according to their *average* position (not the exact position) on a world map with geodesic lines originating from various dropzone locations. Fig. 1 shows the country-level perspective of the relationship. We observe a large concentration of target areas focused in Central Asia.

The distribution of dropzone and target IPs by country is shown in Table I. We notice a large distribution of dropzones from US pointing to targets in Asian countries such as Vietnam. Additionally, China and Brazil are target of attacks originating from European countries. Imperva Incapsula states that Vietnam (12.8%), Brazil (11.8%) and China (8.8%) were the countries with the most-infected devices (from the Mirai botnet) [6].

C. Network Penetration Analysis

This section focuses on analyzing attributes gathered from Shodan and Censys [7], namely, active ports, vulnerabilities.

Active Ports. For each dropzone and target IP address, we use information gathered from Shodan and Censys the list of active ports. We extracted 5,745 active ports from 716 of 877 dropzone IPs and 1,114 active ports from 129 of 189 non-masked target IPs. Each port number is typically associated with a service, such as port 80 for HTTP traffic. While we observe common services like SSH (port 22), HTTP (port 80), and HTTPS (port 443), we point out to the usage of Network Time Protocol (NTP) on port 123. NTP is UDP-based and can be prone to "IP spoofing" for DDoS attacks [8]. This attack is also emphasized in [9], since the attacker can amplify attack packet size 1,000 larger by exploiting NTP.

Vulnerabilities. We then examine the susceptibility of the IP addresses; in particular, we determine the vulnerabilities present in the IP addresses. We gather the Common Vulnerabilities and Exposures (CVE) identifier, maintained by MITRE [10]. In our analysis, the second-most common CVE among the dropzones: CVE-2014-1692. The NVD [11] reports the severity of this vulnerability as "high" since it might allow remote attackers to cause a Denial of Service (DoS) through memory corruption due to uninitialized data structures from the `hash_buffer` function in OpenSSH.

IV. NETWORK CENTRIC ANALYSIS

We observe the malware targeting multiple IoT devices for propagation. In this regard, they often mask the endpoints on target. For such endpoints, we analyze their CIDR network. In total, we inspect 27 unique /24, 435 unique /16, and 125 unique /8 IP addresses. Towards this, we evaluate and analyze these networks to investigate their exposure to risk. In particular, we examine the devices on these networks and looked at the services being used.

The corresponding 100,793,403 active IP addresses are then clustered by their device type. Considering that open ports lead to increased security risks, we look for ports that are necessary for a device to operate uninterrupted. Taking a conservative approach, we suggest that if a port is being used by less than 10% of devices in a given device type, it should be closed to reduce its exposure to risk. Our results show the susceptibility of high-wattage IoT devices, such as heating, ventilation, and air conditioning (HVAC), power distribution units (PDU), etc., can be abused by the attackers to launch large-scale coordinated attacks. Additionally, the high presence of such susceptible devices lays the foundation for attacks as demonstrated by Soltan *et al.* [12].

V. CONCLUDING REMARKS

In this paper, we analyze the ≈78.2% of total responsive public IPv4 endpoints among dropzones and their targets as extracted from IoT malware and spread across the globe from diverse perspectives. Additionally, we augment our analysis results by leveraging the use of IoT search engines like Shodan or Censys.

ACKNOWLEDGMENT

This work was supported by NSF CNS-1809000, NRF-2016K1A1A2912757, and collaborative seed grant from the Florida Cybersecurity Center (FC2).

REFERENCES

[1] P. Middleton, "Forecast Analysis: Internet of Things–Endpoints, Worldwide, 2016 Update." [Online]. Available: http://gtnr.it/2oRo4aN

[2] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, "DDoS in the IoT: Mirai and other botnets," *Computer*, vol. 50, no. 7, pp. 80–84, 2017.

[3] "Shodan landing page," 2018. [Online]. Available: www.shodan.io

[4] Y. M. P. Pa, S. Suzuki, K. Yoshioka, T. Matsumoto, T. Kasama, and C. Rossow, "IoTPOT: A novel honeypot for revealing current IoT threats," *Journal of Information Processing*, vol. 24, no. 3, pp. 522–533, 2016.

[5] "Ultratools free IP tools," 2018. [Online]. Available: https://bit.ly/2v2cLk4

[6] B. Herzberg, D. Bekerman, and I. Zeifman, "Breaking down mirai: An IoT DDoS botnet analysis," Oct 2016. [Online]. Available: https://bit.ly/2dQbvYo

[7] "Censys landing page," 2018. [Online]. Available: https://censys.io/

[8] J. Graham-Cumming, "Understanding and Mitigating NTP-based DDoS Attacks," Jan 2014. [Online]. Available: https://bit.ly/2ifu8pa

[9] "UDP Flood Attacks." [Online]. Available: https://bit.ly/2NTNNu6

[10] "About CVE." [Online]. Available: https://bit.ly/1FxfOi2

[11] "National vulnerability database," 2018. [Online]. Available: https://nvd.nist.gov

[12] S. Soltan, P. Mittal, and H. V. Poor, "BlackIoT: IoT botnet of high wattage devices can disrupt the power grid," in *Proceedings of the 27th USENIX Security Symposium*, 2018, pp. 15–32.

Analyzing Endpoints in the Internet of Things Malware

Jinchun Choi* jc.choi@knights.ucf.edu
Afsah Anwar* afsahanwar@knights.ucf.edu
Hisham Alasmary hisham@knights.ucf.edu
Jeffrey Spaulding jspaulding@niagara.edu
DaeHun Nyang nyang@inha.ac.kr
Aziz Mohaisen mohaisen@ucf.edu
* Equal contributors

INTRODUCTION

Malware and Endpoints

- 10's of billions seamlessly connected IoT devices.
- Adversaries use malicious codes or malware to infect the IoT devices, or use them for propagation, or launch attack.
- Forming a network of devices disguised as botnets, these devices have the potential to launch DDoS attacks.
- Since the botnets and the victims have to be connected to Internet for success of attack, it is important to analyze the endpoints.

Aim Analyze the dropzones and targets to understand the affinities, their exposure to risk, and examine network devices.

Idea To do so, we divide the analysis into 2 parts and perform a deep data driven analysis: IP & Network Centric.

Dataset Creation Primary dataset of 2,423 IoT malware. We reverse engineer the malware to extract dropzone and target IP addresses. Finally, we augment the IP addresses with diverse information, like, country, city, location, ports, vulnerabilities, etc.

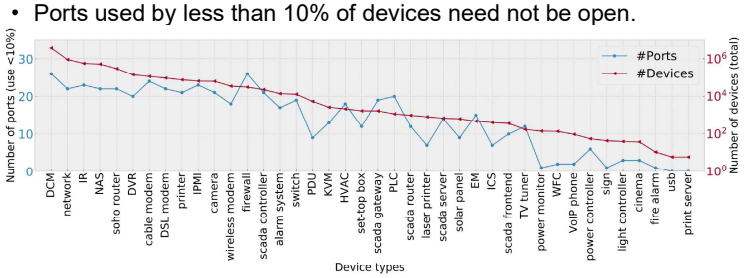
Network Centric

- **Mapping IP addresses from Masked IP (Network address)**
  - Attackers uses masking to hide the targets.
  - We analyze the CIDR IP addres--ses to examine exposure to risk.

Composition of Target IPs for masked and not-masked networks. "In Total" means the total number of target IPs, "In Unique" means the composition of non-duplicated target IPs.

Address	In Total	%	In Unique	%
/24	137	0.13%	27	1.22%
/16	104,369	98.07%	1,869	84.53%
/8	776	0.73%	126	5.70%
Not-masked	1,146	1.08%	189	8.55%
Total	106,428	100.00%	2,211	100.00%

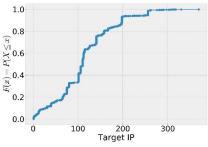
- **Open ports - Devices relation**
  - We determine ports required.
  - Ports used by less than 10% of devices need not be open.



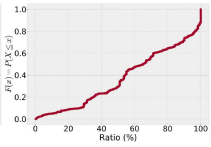
Total number of devices and the number of ports that used less than 10% of devices.

IP Centric

- **Dropzone-Target Inter-relationship**
  - Shared target between dropzones.
    - Target list gathered after similar vulnerability analysis.
    - Target list shared among attackers.



CDF graph of the number of overlapped target IPs.



Ratio of number of overlapped target IPs to the number of union target IPs in dropzones.

- **Geographical Analysis**
  - US has a large distribution of dz.
  - VN, BR, and CN are the most attacked by European dz.
  - **Trend of the Attack**
    - Dropzones are spread, targets are focused in Central Asia.

Top 5 number of target and dropzone IPs by country. Countries include: United States (US), Netherlands (NL), France (FR), United Kingdom (GB), Italy (IT), Vietnam (VN), Brazil (BR), China (CN), India (IN), and Pakistan (PK).

Country	# Dropzones	%Total	Country	# Targets	%Total
US	1,041	43.25%	VN	26,290	24.70%
NL	278	11.55%	BR	20,572	19.33%
FR	188	7.81%	CN	15,799	14.84%
GB	183	7.60%	IN	5,598	5.26%
IT	177	7.35%	PK	5,076	4.77%

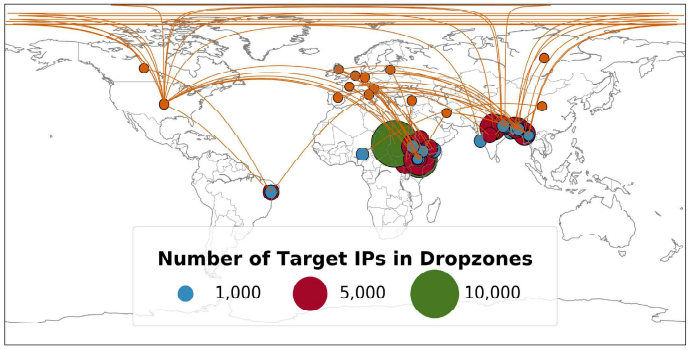
- **Network Penetration Analysis**
  - **Active Ports**
    - Most common ports are active.
    - 111, 123 ports can be utilized by adversaries (UDP flood, DDoS).
  - **Common Vulnerabilities**
    - CVE-2017-15906: medium severity in NVD (exhaust disk attack)
    - CVE-2014-1692: high severity in NVD (DoS attack)

Top 10 active ports in target IPs. "%" column means percentage of active ports in total matched IPs.

Port	Count	%	Service	Description
80	111	17.85%	HTTP	World Wide Web HTTP
22	106	17.04%	SSH	The Secure Shell (SSH) Protocol
443	67	10.77%	HTTPS	HTTP protocol over TLS/SSL
21	51	8.20%	FTP	File Transfer Protocol [Control]
25	49	7.88%	SMTP	Simple Mail Transfer
3306	40	6.43%	MySQL	MySQL database system
53	29	4.66%	DNS	Domain Name Server
8080	29	4.66%	HTTP-alt	HTTP Alternate (see port 80)
111	28	4.50%	SunRPC	SUN Remote Procedure Call
123	26	4.18%	NTP	Network Time Protocol

Top 5 Vulnerabilities by the number of dropzone IPs.

Vulnerability	#IP	#Malware
CVE-2017-15906	203	448
CVE-2014-1692	142	320
CVE-2016-0777	142	325
CVE-2012-0814	140	307
CVE-2011-4327	140	307



Trend of the attack between dropzone and target IP addresses

Result & Conclusions

- We reverse-engineer the IoT malware and extract the endpoints.
- We analyze approximately 78.2% of responsive IPv4 endpoints.
- US has largest distribution of dropzones that target Central Asia.
- Port 111, 123 (NTP) can be used by adversaries.
- CVE-2017-15906 and CVE-2014-1692 are common vulnerabilities.
- Presence of large number of susceptible high-wattage devices exhibits the possibility of large scale coordinated attacks.
- The results present the relationship between the dropzones and target and susceptibility of the endpoints – these along with the behavior of endpoints can be used to help augment threat hunting.

NDSS Symposium,
San Diego, California
24 – 27 February 2019

This work was supported by NSF CNS-1809000, NRF-2016K1A1A2912757, and collaborative seed grant from the Florida Cybersecurity Center (FC2).