Macroscopic Analysis of IoT Botnets

Hatem A. Almazarqi*, Mathew Woodyard†, Troy Mursch‡, Dimitrios Pezaros*, Angelos K. Marnerides*
*School of Computing Science, University of Glasgow, Glasgow, Scotland, UK
h.almazarqi.1@research.gla.ac.uk, [dimitrios.pezaros, angelos.marnerides]@glasgow.ac.uk
† Okta, Inc., San Francisco CA, USA
mathew.woodyard@okta.com
‡ Bad Packets LLC, Chicago IL, USA
troy@badpackets.net

***Abstract*—The adoption of the IoT by modern sociotechnical systems in synergy with the rapid deployment of insecure IoT devices and services has transformed the cyber-threat landscape. Thus, the vast majority of cyberattacks are underpinned by the orchestration of compromised IoT devices that are globally distributed and controlled through carefully designed IoT botnets. Contrary to conventional belief, cybersecurity vectors instrumented by such botnets are not always uniformly distributed across Internet Autonomous Systems (ASes). By virtue of network structural characteristics imposed by each individual Autonomous System (AS) as well as the diversity in terms of AS-level cybersecurity policies, the spatiotemporal manifestation of IoT botnets differs. In this work, we provide a novel measurement study that empirically quantifies AS tolerance of IoT botnet propagation in the global IPv4 Internet. We assess and correlate measurements gathered by globally distributed honeypots, Internet regional registries and IP blacklists for a 15-month period and observe more than 3.2M malicious events triggered by IoT botnets spanning 9.5K ASes. Our work demonstrates that ASes connected to a low number of providers are prone to embrace a high portion of malicious activities. Hence, we provide evidence on concentrated botnet activities and determine the effectiveness of widely used IP blacklists. In general, this study contributes towards empowering knowledge on large-scale cyber-attacks as being crucial for the composition of next generation data-driven cybersecurity defence applications.**

***Index Terms*—Autonomous Systems, IoT botnets, cybersecurity, Internet measurements**

I. Introduction

Cyber-criminals and organised hacking groups managing large-scale IoT botnets strive for the adequate and efficient maintenance of their networked resources. In order to achieve this it is necessary for their resources such as malware downloaders and command and control (C&C) services to be hosted on tolerant Internet Autonomous Systems (ASes) that employ lax security policies. As highlighted in [1], several ASes mapped to particular Internet geographical regions have a disproportionately high number of hostile hosts compared to others. In parallel, the vast majority of global botnet activity underpinning a range of Advanced Persistent Threats (APTs) in various sectors (e.g., energy, manufacturing, defense) is predominantly caused by Mirai or Mirai-like variants where critical botnet assets are mostly hosted in ASes residing in Asian countries [2], [3].

Undoubtedly, the weak implementation of AS-level security practices plays a crucial role on the prevalence of malicious activity targeting core socio-technical systems (e.g., finance) and critical infrastructures (e.g., nuclear, utilities) [4]. In fact, certain ASes may be considered as "bad harvest" and implicitly offer incentives and flexibility to attackers when deploying large-scale attacks [2], [5]. As discussed in various studies [3]–[6], the diversity of AS-level security policies in synergy with the minimal enforcement of such policies due to political and monetary constraints requires mechanisms that rely on consistent measurement studies such as to capture the emerging properties of large-scale threats. Hence, relating the influence of AS tolerance over IoT botnet deployments is of high significance for equipping next generation defense mechanisms with up to date knowledge [5], [6].

The majority of AS-level measurement studies examined security best practices in ASes with the use of metrics distilled by third-party abuse data (e.g., [7], [8]) or looked at AS structural characteristics from a business perspective where security practises were peripherally discussed (e.g., [6], [9]). Moreover, work investigating IoT botnet activity placed greater emphasis on providing an overview of vulnerabilities exploited by specific malware variants [2]–[5]. However, to the best of our knowledge, most of past and recent studies fail to determine the influence of the structural properties of an AS with respect to tolerance on IoT botnet activity.

By contrast to previous pieces of work, we herein conduct a novel study derived by the profiling of diverse Internet measurements gathered in a 15-month period that highlights the influence of AS inter-domain routing properties on malicious activity as stemmed by Mirai and Mirai-like IoT botnets. Moreover we provide insights on the adequacy of widely used IP blacklists for tracking IP addresses that participate in adversarial events. Therefore, the contributions of this work are:

1) A novel macroscopic view on the influence of AS-level relationships with respect to Mirai-like IoT botnet propagation.
2) Compilation of the most important AS attributes that frequently embrace Mirai-like botnet activity.
3) Assessment of IP blacklisting efficiency as used by Regional Internet Registries and ASes in the context of tracking IoT botnet activity.

The remainder of this paper is structured as follows: Sec-

tion II provides background information on the taxonomy of Autonomous Systems. Section III provides an overview of related work. Section IV describes the datasets and methodology used in this work. Section V is dedicated on presenting our findings. Finally, Section VI summarises and concludes this work.

II. Background

A. Taxonomy of Autonomous Systems (ASes)

An Autonomous System (AS) is a collection of IP prefixes managed by network administrators and operating under a single and well-defined routing strategy. Each AS is assigned an ASN (Autonomous System Number) by its respective Regional Internet Registry (RIR) and serves as a unique identifier for its network. An AS can be recognized into three broad categories; (i) single-homed stub, (ii) multi-homed, and (iii) transit. A stub AS is connected to only one AS whereas a multi-homed AS is connected to several other ASes. A transit AS is a conduit between other ASes that are connected to it.

As visualized in Fig.1, ASes form business relationships that can be divided into two main categories: (i) peer-to-peer (p2p) and (ii) customer-to-provider (c2p). In the c2p scenario, an AS needs to purchase transit services for any traffic headed to the rest of the Internet that the AS does not own or cannot access through its customers. Under the p2p relationship, two peer ASes obtain access to each others' customers, typically on a quid pro quo basis. Furthermore, inter-AS traffic on the Internet is often routed based on the commercial relationships that exist between the ASes. Fig. 1 represents a simple Internet topology showing the relationship between ASes. The aforementioned business relationships are critical aspect in defining routing polices between administrative ASes.

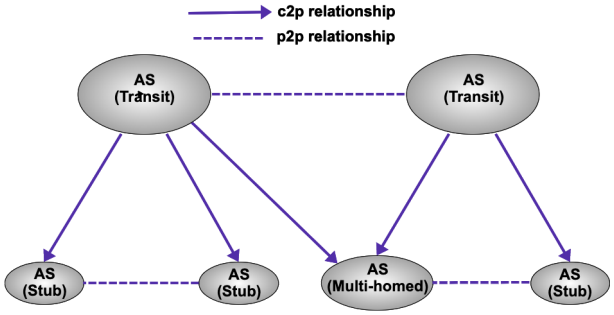

Fig. 1: Business relationships amongst ASes.

B. Border Gateway Protocol (BGP)

For the purpose of exchanging routing information between ASes, interdomain routing protocols such as BGP are deployed. BGP was developed to regulate route selection and packet forwarding across ASes. The BGP router keeps a table with the path (AS path) in order to reach a certain IP prefix. One of the primary reasons ASes employ BGP for interdomain routing is to allow their own policies to be transmitted to their neighbours and, ultimately, across the entire Internet. One of the most distinguishing characteristics of the interdomain routing protocol is that it enables each AS to define its own administrative policy for determining the optimum route, as well as for broadcasting and accepting route announcements.

III. Related Work

Several studies investigated the distribution of illicit activities over countries as well as aggregated units of resources for botnets, such as ASes and IP address spaces (e.g., [9]–[11]). However, no work to our knowledge has investigated the correlation between the structural properties of an AS based on its inter-domain routing policies against diverse CTI feeds as we do in this work.

Work described in [12] focuses on developing security schemes for rating AS reputation. Hence, they attempt to identify malicious or poorly managed networks. However, there are no insights on network attributes that frequently embrace malicious activities. In parallel, the study in [7] proposed different reputation metrics that are entirely focused on the concentration of abuse while taking into consideration some features of hosting providers. By contrast with the aforementioned pieces of work, we examine the structural characteristics of ASes in order to determine the influence of these characteristics explicitly on IoT botnet activity.

The work in [2]–[5] focused on identifying the general properties of botnets and revealed that they had a particularly heavy concentration in a small number of countries, showing the most ASes harbor malicious activities. Nonetheless, both studies do not provide insight into the prevalence of botnets among ASes and do not consider the individual prefixes advertised by ASes, as we show in this work.

IV. Dataset Description & Methodology

A. Dataset Description

Our study is built on measurements from a network of 40 globally distributed attack honeypots, IP address reputation data, inter-domain BGP routing and topology data.

Attack Honeypots: We collected cyber threat intelligence (CTI) data generated by attack honeypots. The data was generated by Bad Packets' globally-distributed network of honeypots placed on 16 unique autonomous systems in 16 countries. These honeypots detect active botnets by emulating hundreds of vulnerable IoT devices, including IP cameras, smart home devices and consumer-grade routers frequently targeted by botnets that scan the internet and engage in malicious activity. In order to further monitor Mirai-like IoT botnet activity, we employ sinkhole domains previously used by DDoS botnet threat actors. The honeypots are deployed on a diverse set of network providers (ASNs) spread across over 16 countries and unique autonomous systems.

Incoming traffic from malicious actors targeting the honeypots is captured and further indexed using Splunk. Mirai-like activity is determined via a fingerprinting method comparing the TCP SYN sequence values with the IP address value. Any Mirai-like activity initiated by a given infected host has the

CTI data			BGP data			
Observation Period			Shadowserver	CAIDA's ASRank		
01/07/2020 - 22/09/2021			Collected prefixes	Customers	Providers	Peers
IP addresses	ASes	Malicious events	7,024,624	96,464	23,390	218,094
661,921	9,521	3.2M				

TABLE I. Summary of CTI and BGP data, CTI feeds collected from 40 globally distributed attack honeypots run by Bad Packets. BGP data obtained from Shadowserver and CAIDA, representing the total number of prefixes and AS links in our dataset.

sequence value of the first TCP SYN packet to be equal with the senders IP address [3].

IP address reputation: We leveraged some of the most commonly-used blacklists implemented by Internet registries and ISPs for botnet activities, phishing and spam. We combined this data with 661,921 IP addresses identified by the honeypots in order to identify ASes that showed an abnormally high level of harmful activities. Namely we used; (i) Spamhaus [1],(ii) Barracuda [2], (iii) Spam Open Relay Blocking System (SORBS) [3], and (iv) Composite Blocking List (CBL) [4].

BGP routing data: BGP data was gathered from CAIDA's ASRank [5] project. We retrieved the AS's neighbours and AS rank for each AS in our dataset. We used this data to retrieve the degree for each AS and pinpoint the types of their neighbours. As shown in [6], the AS degree metric is an effective heuristic for estimating the magnitude of an AS and its routing capability. It also shows the importance of the AS regarding global traffic routing in the Internet.

Shadowserver data: We utilised data provided by the Shadowserver Foundation to reveal the advertised IP prefixes for each AS in our dataset. Shadowserver provides an ASN report containing all the routed Classless Inter-Domain Routing (CIDR) for an AS. As per Table I the total number of prefixes advertised by ASes in our dataset is 7,024,624.

As summarised in Table I, our observations were taken from 661,921 distinct IP addresses which caused approximately 3.2M malicious events ranging from scanning and infection located across 9,521 Autonomous Systems (ASes) collected over 15 months, from July 2020 to September 2021. In addition, Table I represents the number of IP prefixes collected from Shadowserver advertised by all ASes in our dataset and the number of links, including customers, providers and peers connected to the ASes.

B. Methodology

IP address mapping: achieved by mapping a set of IP addresses defined as B involved in botnet activity to their originating ASes (ASx). Hence, $occ(B \in ASx)$ represents the total number of B announced by ASx.

IP address space: in order to illustrate the prevalence of IoT-based botnet over AS's prefixes, we retrieve the advertised prefixes (Pm) for each B to determine the number of prefixes for ASx involved in botnet activity $(Pm \in ASx)$. We subsequently identify the total number of advertised prefixes for each ASx by using Shadowserver data, resulting in $(Pt \in ASx)$. We obtain the abuse rate of ASx as follows:

$$M(ASx) = \frac{\#(Pm \in ASx)}{\#(Pt \in ASx)} \tag{1}$$

IP Blacklist effectiveness: we measure the effectiveness of a blacklist by computing the coefficient variation (CV) of $occ(B \in ASx)$ and L, where L indicates the presence of B in the blacklist. CV represents a statistical indicator for the dispersion of data points around the mean. In our case, a smaller score implies that a high proportion of IPs is blacklisted, whereas a high score indicates a low ratio of IPs is observed by blacklists. Hence the CV for a given ASx is denoted as:

$$CV(ASx) = (SD/\mu) \tag{2}$$

where SD is the sample standard deviation and μ is the sample mean.

AS degree: based on our BGP inter-domain measurements and defined by aggregating the number of connected neighbours to ASx, including customers, peers and providers. Thus, the total connected links to ASx are defined as:

$$\mathbb{AS}(d) = \mathbb{R}(p) + \mathbb{R}(c) + \mathbb{R}(r) \tag{3}$$

where R(p) are the total peers connected to AS, R(c) are the total customers connected to an AS and R(r) are the total providers connected to a given AS.

Malware Payload profiling: in this step, we analyse and cluster the malicious Mirai-like payloads instrumenting botnet activities. We treat each observed CTI log as a textual representation in order to construct a group of documents and apply Natural Language Processing (NLP) to adequately profile malware binary strings. Initially payloads are segmented using whitespaces to return the segments as tokens. In order to convert chunks of text into meaningful numerical representations, we apply the TF-IDF (Term Frequency - Inverse Document Frequency) vectoriser. The TF-IDF approach enables us to determine the weight for each document, as well as determine the importance of a token in a set of documents. Hence, the TF-IDF weight is applied to numerically represent payloads by building a document term matrix comprised of all the segmented tokens in all documents. The TF-IDF weight is computed by:

[1]Spamhaus:https://www.spamhaus.org/
[2]Barracuda: https://www.barracudacentral.org/
[3]SORBS: http://www.sorbs.net/
[4]CBL: https://www.abuseat.org/
[5]CAIDA:https://www.caida.org/projects/ark/

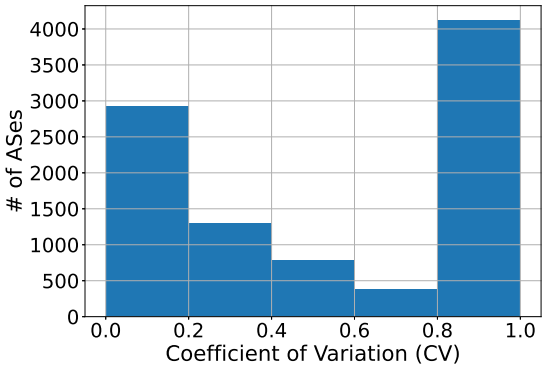

Fig. 2: CV distribution of correlating malicious IP addresses over all observed ASes with our IP blacklist data feeds.

$$TF - IDF = TF * IDF \tag{4}$$

$$tf_{i,j} = \frac{tf_{i,j}}{\sum_{t \epsilon d} f_{t,d}} \tag{5}$$

$$idf_i = log(\frac{N}{df_i}) \tag{6}$$

where TF is used to calculate the frequency with which the term occurs in each payload in our dataset. IDF is used to calculate the occurrence of unusual terms across all payloads. Terms that occur infrequently in our dataset get a high IDF score. Finally, N represents the total number of logs and d the entire number of logs in our CTI feeds.

In order to further cluster payload distributions we utilise the k-means algorithm to relate malware variant groups based on the AS degree of membership and the distance between logs. The sum of squared distance between each point and the centroid in a cluster is calculated by:

$$WCSS(K) = \sum_{j=1}^{k} \sum_{x_i \varepsilon cluster j} \|x_i - \bar{x_j}\|^2 \tag{7}$$

where $\bar{x_j}$ is the sample mean in cluster j. The optimal number of clusters was determined through the use of the elbow method by examining the WCSS distribution over different trials of the k-means clustering process as we also discuss later on in Section V.

V. Results

A. AS degree and botnet presence

The degree of an AS indicates the number of ASes directly connected to a given AS and considered its neighbours. We found that ASes with a high number of Mirai-like botnet-related malicious IP addresses are more likely to have a lower degree of ASes. When we consider the number of providers connected to ASes, our analysis shows that ASes connected to 8 providers or less tend to host a high portion of malicious IPs. It was revealed that ASes of such characteristics hosted 70% of IPs observed within our attack honeypots. In addition, a high number of these ASes have not appeared in any of the IP blacklists used within this work as shown in Fig. 2. In

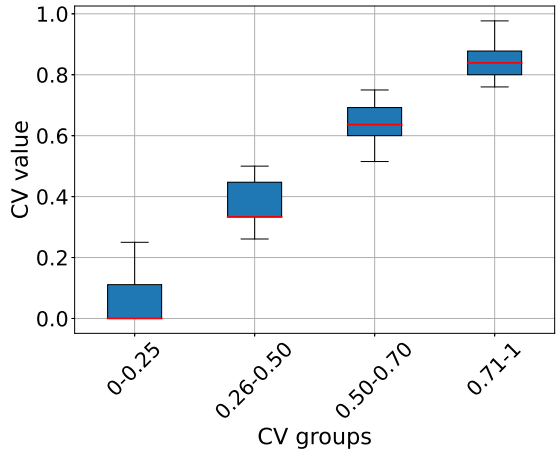

Fig. 3: Four distinct CV groups with respect to the effectiveness of commercial IP blacklists on tracking botnet-related IP addresses across the examined ASes; the two groups with CV values greater than 0.6 represent 70% of the examined ASes indicating that more than 90% of botnet addresses were not captured by IP blacklists.

particular, Fig. 2 indicates more than 70% (i.e., $0.0 < CV < 0.5$) of the malicious IP addresses residing over various ASes to be partially or fully detected by some IP blacklists whereas more than 70% to not be reported at all (i.e., $0.5 < CV < 1$). Through Fig. 3 we identify four distinct AS groups with respect to their CV scores and verify that the first two groups with $CV < 0.5$ have a median value at the bottom of the first quartile. Hence it indicates that botnet-related IP addresses in ASes belonging into these first two groups were matched with IP blacklist data over a higher certainty. On the other hand, the latter two groups have a median value corresponding at the bottom of the second quartile indicating that high CV values were more frequently obtained and thus botnet-related IP addresses were not matched with our blacklist data feeds.

Through the conducted analysis focusing on IP reputation revolving around addresses that originated from lower and high AS degree, it was revealed that IP blacklist databases observed 70% of IPs from ASes with low degree. Our cross-correlation also highlights that 90% of IP addresses listed were from ASes with high degree. Evidently, attackers prefer to target ASes that have a lower AS degree and avoid ASes with a high degree. We argue that attackers tend to adopt such behaviour in order to evade ISP monitors such as blackhole routes since such routes suppress bidirectional communication between an attacker and a victim.

IoT botnets are controlled by malicious actors that instrument various entities, including bots and loader servers such as to launch malicious activities. Such entities typically reside in particular ASes and comprise a set of contiguous IPv4 addresses seen as a single network prefix. We stratify ASes into two groups with respect to their abuse ratio and further measure the structural properties for each group as shown in Table II.

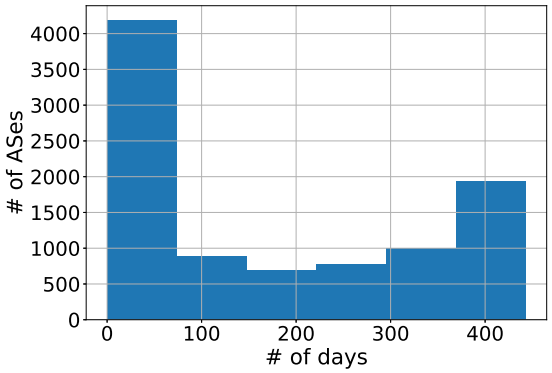

Fig. 4: Duration for ASes participating in botnet activity, where ASes with long duration have a low CV score.

Abuse ratio	ASes degree	customer	Peer	provider	Hosted malware downloaders
0-50	41	12	26	2.6	1.7
51-100	16	3	11	1.9	5

TABLE II. Abuse ratio of AS prefixes with respect to the average AS degree, number of customers, peers and provider ASes for each ratio group.

B. Botnet payload distribution over ASes

Following manual mining of our correlated data feeds and the grouping depicted in Table II, we identify that ASes with a low AS degree host a high proportion of malware downloaders. Furthermore, our assessment on the AS temporal duration with respect to active botnet activity in Fig. 4 indicates that 50% of ASes were active for less than 100 days. The identified ASes also obtained an average CV score of 0.5 demonstrating that a low proportion of botnet-related IP addresses residing in these networks were captured by IP blacklists. ASes active for more than 100 days obtained an average score of 0.40 suggesting that malicious IP addresses participating in botnet activity were more likely to be captured by IP blacklists. In general, we observe that the majority of botnet activity is instrumented under the objective to evade particular AS security policies and they ensure to transfer critical entities such as malware uploaders in around a 3-month period. Through further thorough analysis in our datasets, we have also determined the type of malicious activities performed by infected IPs and mapped them to their origin AS and their corresponding BGP advertised IP prefixes. It was revealed that 26% of the observed ASes had more than 50% of their IP prefixes participating in botnet-related activity. Thus, more than half of these networks were actively involved over various cyber-attacks as triggered by Mirai-like IoT botnets. Moreover, the attributes of the aforementioned ASes depict that malicious actors prefer to target and abuse ASes with a low number of providers with respect to their BGP routing policies. The insight distilled from this observation empowers the opinion that a large portion of ASes embraces malicious activity due to minimal security practices.

Via filtering botnet payloads from our CTI feeds we compose a dataset resulted by the application of TF-IDF as

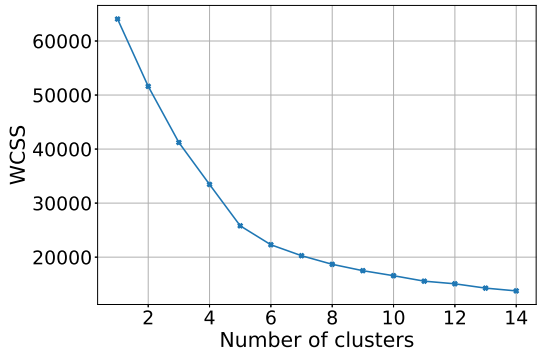

Fig. 5: Assessment of the sum of squared errors (WCSS) distribution for selecting the most optimal number of clusters.

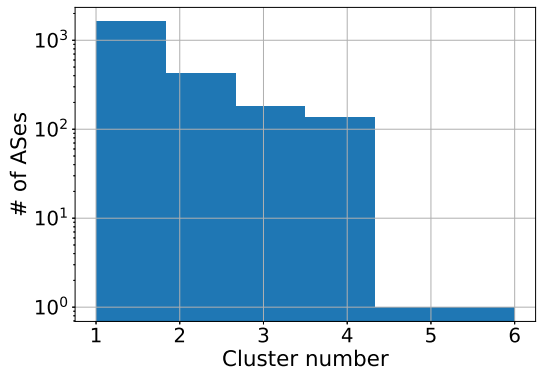

Fig. 6: 70% of ASes are involved in botnet activities by concentrating on targeting one cluster whereas the remaining are observed to target multiple clusters.

described in Section IV. We further employ k-means clustering in order to gain insight into the distribution of Mirai-like botnet payloads amongst ASes. Through the use of the elbow method we identify the optimal number of clusters after an iterative assessment of cluster number values k with respect to the sum of squared errors (WCSS) for each k. The value k was selected at the "elbow", i.e., the point of inflection on the curve which provides a good indicator of the optimal point. In our case, the optimal number of clusters for the data was six as represented in Fig. 5. The k-means outputs are mapped to ASes in order to analyse the dispersity of malicious activities with respect to payload distribution. As shown in Fig. 6 a high proportion of ASes are typically abused by malicious actors to send malicious payloads targeting only one cluster. Evidently, such ASes have certain structural properties in terms of the number of connected providers. Specifically, the mean of connected providers for cluster one is two, which is where the majority of ASes in our dataset reside.

As discussed in Section IV, we compose a feature set resulted by the application of TF-IDF in our honeypot datasets and aim to profile the distribution of Mirai-like malware payload across the observed ASes. In order to reduce the high-dimensionality of our feature space we have also employed the Principal component analysis (PCA) method and further clustered the Principal Components (PC) with the highest

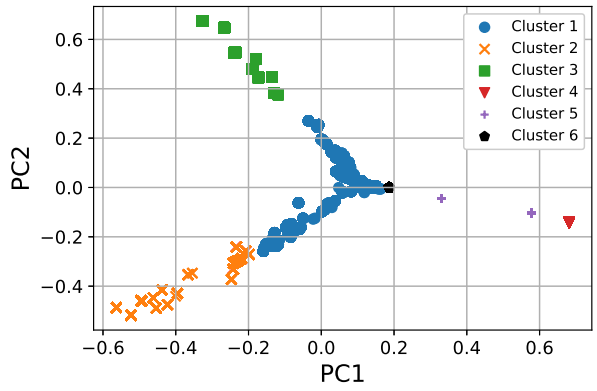

Fig. 7: Clusters of the main malware variants underpinning botnet propagation in our analysis as resulted by applying PCA over our TF-IDF payload feature set.

variance. As shown in Fig. 7, samples comprising cluster 1 relate to Mirai-like malware that exploits a common Android Debug Bridge (ADB) vulnerability over the TCP port 5555 such as to compromise Android devices. We identify that such compromised devices were in the majority used to launch DDoS attacks or perform cryptocurrency mining. Similarly, the samples comprising cluster 6 target ADB and through an alternative vulnerability aim particularly for shell access through injecting Powershell scripts. It is worth mentioning that exploited devices through this technique could spread the malware much more aggressively to other systems that were previously connected to them via SSH. Payload samples in cluster 2 are droppers aiming to detect the operating environment such as to profile vulnerabilities. Samples in cluster 3 target devises manufactured by AVTECH via a set of 164 TCP ports which have not been used by other malware variants. We can arguably claim that bootmasters adapt this behaviour in order to conduct zero-day exploits. The samples comprising cluster 4 include multiple malware variants targeting vulnerabilities on DVR cameras (e.g., MVPower, D-Link, Vacron). These samples particularly target to kill userspace Linux processes and cannibalize on core kernel-space system calls. Finally, in cluster 5 there is a variety of malware built to target particular CPU architectures destined for IoT and embedded systems (e.g., ARM7).

VI. Conclusion

The work described in this paper measured the influence of AS structural properties on Mirai-like IoT botnet activity. Through a novel macroscopic measurements analysis that aggregated inter-domain routing data with CTI feeds and diverse IP blacklist data for a 15-month period we showcase AS attributes that embrace botnet activity. We demonstrate that commonly and widely used IP blacklist databases were incapable at tracking concentrated botnet activity as hosted over more than 60% of the ASes analysed in our datasets. In addition, our analysis reveals that ASes with a low AS degree host a high proportion of malware downloaders. Thus, we witness a technique that is commonly applied by botmasters such as to evade visibility by high degree ASes that usually have better security policies in place. We argue that the findings in this work may contribute significantly towards the design of next generation data-driven attack detection and situational-awareness solutions.

Acknowledgement

The authors would like to thank Bad Packets LLC, CAIDA and Shadowserver Foundation for providing their datasets. This work has been supported in part by the PETRAS National Centre of Excellence for IoT Systems Cybersecurity, which has been funded by the UK EPSRC under grant number EP/S035362/1.

References

[1] Arman Noroozian, Jan Koenders, Eelco Van Veldhuizen, Carlos H Ganan, Sumayah Alrwais, Damon McCoy, and Michel Van Eeten. Platforms in everything: Analyzing {Ground-Truth} data on the anatomy and economics of {Bullet-Proof} hosting. In *28th USENIX Security Symposium (USENIX Security 19)*, pages 1341–1356, 2019.

[2] Hatem A Almazarqi, Angelos K Marnerides, Troy Mursch, Mathew Woodyard, and Dimitrios Pezaros. Profiling iot botnet activity in the wild. In *2021 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE, 2021.

[3] Owen P Dwyer, Angelos K Marnerides, Vasileios Giotsas, and Troy Mursch. Profiling iot-based botnet traffic using dns. In *2019 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE, 2019.

[4] Angelos K. Marnerides, Vasileios Giotsas, and Troy Mursch. Identifying infected energy systems in the wild. In *Proceedings of the Tenth ACM International Conference on Future Energy Systems*, e-Energy '19, page 263–267, New York, NY, USA, 2019. Association for Computing Machinery.

[5] Manos Antonakakis, Tim April, Michael Bailey, Matt Bernhard, Elie Bursztein, Jaime Cochran, Zakir Durumeric, J Alex Halderman, Luca Invernizzi, Michalis Kallitsis, et al. Understanding the mirai botnet. In *26th {USENIX} security symposium ({USENIX} Security 17)*, pages 1093–1110, 2017.

[6] Matthew Luckie, Bradley Huffaker, Amogh Dhamdhere, Vasileios Giotsas, and KC Claffy. As relationships, customer cones, and validation. In *Proceedings of the 2013 conference on Internet measurement conference*, pages 243–256, 2013.

[7] A. Noroozian, M. Korczynski, S. TajalizadehKhoob, and M. Van Eeten. Developing security reputation metrics for hosting providers. In *Proceedings of the 8th USENIX Conference on Cyber Security Experimentation and Test*, CSET'15, page 5, USA, 2015. USENIX Association.

[8] Arman Noroozian, Elsa Turcios Rodriguez, Elmer Lastdrager, Takahiro Kasama, Michel Van Eeten, and Carlos H Gañán. Can isps help mitigate iot malware? a longitudinal study of broadband isp security efforts. In *2021 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 337–352. IEEE, 2021.

[9] Eugenio Nerio Nemmi, Francesco Sassi, Massimo La Morgia, Cecilia Testart, Alessandro Mei, and Alberto Dainotti. The parallel lives of autonomous systems: Asn allocations vs. bgp. In *Proceedings of the 21st ACM Internet Measurement Conference*, pages 593–611, 2021.

[10] Samaneh Tajalizadehkhoob, Rainer Böhme, Carlos Ganán, Maciej Korczyński, and Michel Van Eeten. Rotten apples or bad harvest? what we are measuring when we are measuring abuse. *ACM Transactions on Internet Technology (TOIT)*, 18(4):1–25, 2018.

[11] Matej Zuzčák and Petr Bujok. Causal analysis of attacks against honeypots based on properties of countries. *IET Information Security*, 13(5):435–447, 2019.

[12] AU Prem Sankar, Prabaharan Poornachandran, Aravind Ashok, RK Manu, and P Hrudya. B-secure: a dynamic reputation system for identifying anomalous bgp paths. In *Proceedings of the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications*, pages 767–775. Springer, 2017.