Evolution of IoT Linux Malware: A MITRE ATT&CK TTP Based Approach

1st No Given Author

2nd No Given Author

***Abstract*—In recent years, attacks against Internet of Things devices have increased by 59% says [1]. In this work, we investigate the evolution of malware that emerged in the last two years by taking advantage of the MITRE ATT&CK framework to deliver an analysis methodology based on this structure.**

We analyzed 14 distinct malware families that were discovered in the period by major security vendors and our threat intelligence investigations.

In this paper, we propose a methodology to keep track of threats capability evolution using the MITRE ATT&CK framework. Our research aims to extend the current knowledge of Linux malware in the IoT domain and deliver a different analysis point of view.

The findings presented in this paper about what changed, for example, what techniques are removed from the malware implementation, support the benefit of this analysis and tracking methodology to study the evolution of malware.

***Index Terms*—IOT Linux Malware, TTP, MITRE ATT&CK, evolution**

I. Introduction

Internet of Things (IoT) devices are more and more part of everyday life, the demand and the of such appliances is increasing fast together with the development of new devices, and so the possible security concerns expand with them. Internet of Things threats are growing fast: according to Zscaler [2] the attacks against IoT devices increased by 700% in December 2020 with respect to the pre-Covid-Sars-19 pandemic period.

Nowadays IoT devices are used in a lot of different environments, from industry to healthcare [3], thus the security concerns not only affecting the data or the network but, also regarding the safety of human operators of these appliances.

Despite all the energy cybersecurity researchers invest in studying new threats and keeping up with state-of-the-art malware and attacks techniques, the majority of work, that studies the evolution of malware, are limited to a narrow aspect of the overall problem. Some researches focus on one specific threat and its updates. Others often only consider the implementation details and its changes by putting particular attention on code aspects and implementation choices to determine the malware progression.

This research aims to study the evolution of techniques and functionalities used by IoT Linux malware by exploiting the power of the MITRE ATT&CK matrix to understand and characterize current threats and provide insight into upcoming ones. Indeed, we propose a different analysis methodology leveraging the ATT&CK framework and covering also the pre- and post- exploitation aspects of an attack.

We set up an analysis environment to retrieve the results from 14 different malware families. The samples were collected among attacks that were first seen in the last two years. The families have been selected to achieve an high variety of distinct malware implementation and history. Since we wanted to test this analysis methodology we needed a lot of unique IoT Linux malware families to be able to simulate correctly the real scenario.

Section II gives a brief introduction to IoT Linux malware and the MITRE ATT&CK framework. In Section III, we report the data gathering methodology, the sources of the collected data, and the composition of our data set. The analysis approach is described in Section IV, while results are reported in Section V. Section VI presents the related works. Finally, Section VII will present some limitations and issues we identified in our study and conclude the paper. Conclusions are drawn in Section VIII.

II. Background and Terminology

In this section, some background knowledge on IoT Linux malware and the MITRE ATT&CK framework is provided.

A. Internet of Things Security

IoT devices potential can go from simple and basics MCU-based devices such as sensors to more powerful one like routers and Network-Attached Storage (NAS). In this research, we focused on the security of more complex machines like video recorder devices, NAS, and router, that typically can be a more valuable target for attacks. The less powerful devices (e.g. sensors, actuators, etc) generally are low-performance and cheap and are outside the scope of this paper.

We emphasis on malware families that target IoT devices running a Linux Kernel based operating system (OS). Most of the known families aim at building large networks of infected machines they can control, also called botnets. Moreover, some actors take advantage of these devices, in particular NAS, which may have unsecured services exposed to the internet: as stated by [4] the 20% of the total devices they tested have the FTP service active and the anonymous user enabled, thus they are exposed to the Internet. These machines are a profitable target for ransomware gangs because they often

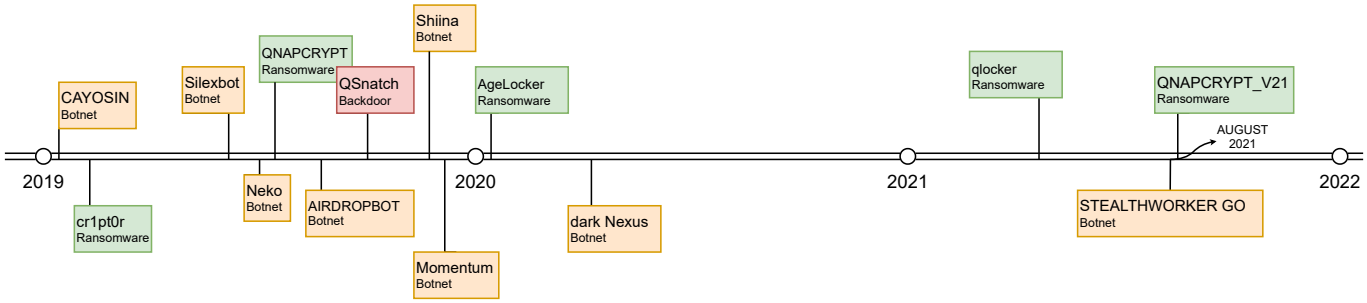

Fig. 1. Data set composition first seen date timeline.

contain sensitive and valuable data. Since these devices are exposed to the network and usually always turned on they can be an interesting target of IoT Linux malware. Indeed, it is not uncommon that these devices are not properly secured: 83% of home routers are vulnerable due to insufficient firmware or security updates according to [5]. Moreover, as stated by Deloitte [6] 70% of the devices use the vulnerable factory-set default usernames and passwords.

In addition, as said by Edelson [7], NAS vendors assume that these kinds of devices are installed in a protected network, for instance secured by firewall or security monitoring systems, but of course, this may not always be the case. The author debates over some protection deployed by NAS vendors claiming that these countermeasures are more oriented to the data protection side rather than the network security.

B. MITRE ATT&CK Framework

The MITRE ATT&CK is a public, knowledge-based repository of adversary tactics and techniques created by the MITRE Corporation [8] to create a standardized taxonomy for the sharing of malware information. The repository structure reflects each step of an adversary's attack lifecycle [9], and the Tactics and Techniques model provides a common terminology for both the defensive and offensive side of cybersecurity. Furthermore, the MITRE ATT&CK standard allows researchers and professionals to understand and analyze how a specific adversarial behavior is used to achieve a goal and better respond to it.

From the latest update, ATT&CK for Enterprise contains 14 Tactics, 185 Techniques, and 367 Sub-techniques [10]. A *tactic* (e.g. Initial Access) describes the reason for performing an action (e.g. use of default credentials), a *technique* (e.g. Brute Force) describes how the tactical action gets achieved, and *sub-techniques* (e.g. Password Guessing) are a more specific description of a particular technique. We call a TTP (Tactics, Techniques, Procedures) chain a group of different techniques from various tactics used in an attack. A TTP gets identified by an id (e.g. T1190 identifies Exploit Public-Facing Application) that allows interactions in an automated fashion.

For example, QSnatch is a first stage backdoor that implements, among others, some defense evasion techniques. This characteristic can be described using the ATT&CK matrix as shown in the Table I.

TABLE I
EXAMPLE OF A MAPPED CHARACTERISTIC OF QSNATCH MALWARE FAMILY

		ID
Tactic	Defence Evasion	TA0005
Technique	Impair Defenses	T1562
Sub-technique	Disable or Modify Tools	T1562.001
Techinique	Hide Artifacts	T1564
Sub-techinique	Hidden Files and Directories	T1564.001

III. ANALYZED DATA

We used a variety of methodologies and resources to build the data set with the samples we have analyzed. Our data set is composed of 14 distinct malware families. The chosen time frame for collecting the samples is from January 2019 to August 2021. Moreover, we decided to avoid analyzing all malware variants in a family as they share lots of details, because we wanted an high number of unique and different families that do not share code among them in our experiment.

A. Malware samples Collection

In order to collect the samples, we started from a database of IoT malware updated over the years based on reports and technical blogs from security researchers. Then, we selected those meeting our criteria: targeting IoT Linux devices, first seen in the two year window, and be a different family with respect to the ones already selected. About the earliest samples, what we already had was enough: information about older samples is widely available. The main challenge was to find more recent threats. The first approach was to check all the major cyber-security blogs to find fresh information, but this was not very successful. As a result, we resorted to some threat intelligence knowledge and some retroactive hunting techniques that were more proficient.

Another challenge we faced was to avoid selecting malware that was just a variation of an already chosen one. Since late 2016, when the Mirai Botnet [11] source code was leaked, an increased number of its variants has been seen in the wild [12]. Mainly because of the high availability of the original code, even non-expert have been able to write down a botnet [13]. Due to this high Mirai code reuse, we decided to exclude from our analysis Mirai variants, besides one family,

named Shiina, that exploits a large number of vulnerabilities [14]. This decision was based on the fact that an unbalanced data set, meaning an unfair amount of equally implemented botnets, would bring noisy results when compared with less numerous families.

Table II displays malware that exploits at least a security vulnerability. It is interesting to notice that older samples take advantage of more vulnerabilities than more recent ones.

TABLE II
NUMBER OF VULNERABILITIES USED BY EACH SAMPLE WE COLLECTED

YEAR	MALWARE	# of VULNERABILITIES
2019	momentum	14
	Shiina	13
	Neko	8
	QNAPCRYPT	3
	cr1pt0r	1
	Qsnatch	
	AIRDROPBOT	0
	CAYOSIN	
	Silexbot	
2020	dark Nexus	1
	Agelocker	0
2021	QNAPCRYPT_V21	4
	qlocker	1
	STEALTHWORKER GO	0

B. Selected Malware Families

In this paragraph, we are going to briefly delineate the families that compose our data set. A data set summary is displayed in Fig. 1 with respect to the first seen date of each family.

CAYOSIN is a botnet that came out in early 2019. This botnet has been sold online for 20$ per month. It is a Qbot based code with some functionalities similar to Mirai, such as the watchdog service, tables, and some random strings [15] [16].

Cr1pt0r was first seen in early 2019. It is a ransomware targeting D-link NAS devices, but a recent update shows that the malware operator also offers decryption keys for the Synology brand [17]. The analysis for this ransomware was challenging. The main problem was that the program is a 32-bit ELF compiled for the arm architecture and the executable is stripped. Then, when we finally achieve to emulate it we need to fully understand its behavior: without the right file in the host, it would not run. The ransomware uses the Sodium crypto library [18] and, in particular, the `curve25519xsalsa20poly1305` algorithm for asymmetric encryption. A particular aspect of this malware is that it behaves differently whether the private key file is found on the host or not. If the file is there, the sample will automatically try to decrypt, otherwise, it will encrypt with the public key.

Discovered in mid-2019, **Silexbot** is a botnet that did not end up carrying out massive Distributed Denial of Service (DDoS), but the impact it had on infected devices was to make them unusable by wiping the memory and destroying the firmware. The author of this malware claimed that their motivation was to prevent other attackers from using infected devices as part of a botnet. The alleged author behind Silexbot is a 14-year-old from Europe [19]. This was an atypical malware as its lifespan was only some days and it targeted multiple architectures.

The malware, to reach its purpose, first flashes the memory with random data, then it makes the device inaccessible over the network, and finally, it wipes the permanent storage. The functionalities recall those of BrickerBot [20] because of some code similarities found.

First seen in 2019, **Neko** is a botnet targeting routers of several brands [21]. Differently from previous cases, this family was UPX-packed [22] with its magic number tampered, to prevent the botnet from being unpacked and analyzed. Moreover, Neko is capable of executing backdoor commands, kill processes, and scan for devices vulnerable to a set of exploits.

QNAPCRYPT is a ransomware also known as *eCh0raix* that came out in mid-2019. At first, it was targeting Synology and QNAP NAS devices in different campaigns. Then, in September 2020, it evolved to combine exploits for both brands in a single executable [23]. It is written in the Go Language, and its code is very basic. After being executed, it tries to connect to the command and control server via *Tor*. The encryption function retrieves 32 random characters from a hard-coded string to create an AES-256 key. Moreover, the malware checks if there is another instance of itself already running. It is also capable of killing some processes such as the web-servers `nginx`, `apache2`, and many more. The program encrypts every file contained in a list of absolute paths and selected file extensions.

AirDropBot is a botnet that targets different IoT devices discovered in mid-2019. This malware, according to VirusTotal, gets detected mostly as a Mirai variant but, as said in [24]. Even if the malware's author did take some ideas from a pre-existing botnet code, it is not considered a variant. We found a ELF file of this malware that is packed with a custom UPX-like packer, and it also implements an additional code obfuscation defense.

Qsnatch is a backdoor targeting QNAP NAS first seen in October 2019 [25]. It is the only backdoor malware family we found. It is a bash script compiled with the bash compiler SHC [26]. To extract the script we executed the sample inside the GNU debugger `gdb`, and then we statically analyzed it. We found out that, to connect to a C&C (Command and

Control), the malware implements a Domain Generation Algorithm (DGA) based on a date. When the program finds the right domain, it will download the second stage. Thus, to find the real impact of this malware, first the DGA needs to be reversed to download and analyze the second stage. Apart from this, Qsnatch exfiltrates some configuration and system data. Moreover, it steals the OS password by implementing a fake login interface that memorizes the password before passing them to the real login engine [27].

Shiina is a botnet first seen in late 2019. It is the only Mirai variant we decided to include in the analysis because it uses a wide set of vulnerabilities according to [14]. It targets a variety of devices: NVR [1], DVR [2] and Routers by different brands. It downloads and executes many exploits from `hxxp://ililililililililil[.]hopto[.]org`. We found a Pastebin [28] with all the exploits this variant implements.

Besides the number of vulnerabilities used, this is a very simple and basic malware. Like the original Mirai, it uses BusyBox [29] to execute commands on the infected machine. Furthermore, it implements some simple Defence Evasion techniques by deleting files on the host.

Momentum targets multiple CPU architectures such as Arm, Intel, MIPS, Motorola 68020, and more. This botnet was discovered at the end of 2019 by [30]. It is capable of connecting to an Internet Relay Chat (IRC), registering itself, and accepting commands from the C&C. Then, the botnet operators can control the system by sending messages in the IRC channel. Besides its DDoS capability, the malware can also open a proxy on a specified IP, change the client's name, and disable or enable packets from the host.

First seen in January 2020, **Agelocker** is a ransomware targeting QNAP NAS devices. The AGE (Actually Good Encryption) algorithm is used to encrypt the files [31]. The defense mechanism implemented by this program kills all the instances of some security monitoring systems running, such as the `wazuh-agent` [32]. The C&C for this ransomware gets retrieved with a request to a Pastebin link. Indeed, we suppose that this malware can be activated at any time just by changing the IP address in the note.

Dark Nexus is a botnet that appeared in April 2020 [33]. This malware can carry out many different kinds of DDoS attacks. Moreover, it periodically performs a `GET` request to its C&C to download the latest version to ensure execution of always up-to-date code. Guided by its C&C, it uses the Telnet port 23 and random IP address to propagate. If the telnet service is running, it tries brute-forcing the credentials.

[1]Network Video Recorder
[2]Digital Video Recorder

Qlocker is a ransomware discovered in early 2021 [34]. It targets NAS devices, in particular the QNAP ones. Written in Python, its main script uses an RSA public key to encrypt a randomly generated password used to encrypt files with the 7zip tool. To retrieve the data the victim has to send the base64 encoded result via an `HTTP` form in a `.onion` domain. At this point, the ransomware operator checks the base64 encoded data with its RSA private key to retrieve the Bitcoin wallet and, after the payment, the password to open the archive and restore the files is sent to the victim.

STEALTHWORKER GO, also called *GoBrut*, is a botnet written in Go, it was first reported in 2019. When it emerged, it was a malware that conducted attacks against e-commerce websites, then in 2020 the malware evolved, and it targeted also Windows and Linux servers that were running popular services such as phpMyAdmin, WordPress, and more [35]. In August 2021, it changed again. The latest version now also targets Synology NAS devices. As Synology states [36], this may be the first stage to deploy more malicious code in the victim machine.

IV. Analysis Design and Implementation

In this section, the analysis design and process are explained. At first every IoT malware family was mapped in MITRE ATT&CK Enterprise Linux matrix [37], then the collected TTP data were analyzed.

The analysis involves malicious programs first seen in the past two years, so we considered the period from January 2019 to August 2021. This decision was made to deliver contemporary results and to have future-looking research.

A. Mapping Techniques Tactics and Procedures

Once we had our list of candidates, we started analyzing them with both static and dynamic analysis. The goal was to describe the malicious programs as detailed as possible with the ATT&CK matrix. We spent a lot of time on each family to completely understand its functionality and features.

Older malware is well documented, but most recent families were more difficult to find because many of them have not been yet disclosed to the general public.

Regarding the malware analysis, every time we encountered a capability, we searched in the ATT&CK matrix the right technique that represents it and the most explicit TTP. This was challenging because not all techniques used by these programs are easily linkable with the ones represented in the MITRE framework. It was hard to describe some malware abilities because the implementation possibilities are much more than those representable with the ATT&CK matrix.

As we proceeded with the analysis every piece of knowledge we gained was also stored in a custom instance of OpenCTI. OpenCTI is an open-source platform that allows threat intelligence researchers to handle and share knowledge [38]. Moreover, this platform easily visualizes data and provides a powerful tool to understand the "inner" relations between advanced threats, malware, attacks, and more. It is a tool that

could put together structured information, such as TTPs, as well as less organized knowledge, such as notes and inferences from researchers.

B. TTP analysis

After we mapped out all the TTPs from the families they have been stored in a Comma Separated Values (CSV). At this point, the analysis is performed by a custom Python script that takes as input the data file and proceeds with the information extraction. This automation allowed us to replicate the analysis multiple times with different requirements and retrieve different results each time.

V. Results

This section describes the results and the findings we discovered during this research. We tried to analyze every detail about each family targeting IoT devices, so the considerations in this section are based also on the actual scenario and IoT security state-of-the-art.

We present an interesting analysis point of view that exploits the power of ATT&CK framework and taxonomy. So, researchers can understand the trend of IoT malware evolution based on the addition and removal of capabilities. In the first part, we show a general overview of all types of malware we analyzed, then we focus more on specific malware categories.

A. Most prevalent ATT&CK TTPs used by IoT malware

The most prevalent techniques used by IoT malware are displayed in the Table III below. The table shows off the ten most used TTPs. As we can see the top one is about *File and Directory Discovery* (T1083) and belongs to the *Discovery* tactic. Actors use this ability to enumerate files and directories in the victim machine. This can be achieved with shell commands or by interacting with the native API.

TABLE III
Top 10 used TTPs

ATT&CK Phase	TTP	Count
discovery	T1083	10
command-and-control	T1071.001	9
initial-access	T1133	8
execution	T1059.004	7
impact	T1498.001	
credential-access	T1110.001	6
discovery	T1057	
execution	T1106	5
impact	T1486	
defense-evasion	T1070.004	4
lateral-movement	T1210	
persistence	T1053.003	

Another interesting TTP is *External Remote Services* (T1133) in the *Initial Access* tactic. This technique describes the method to enter the system, in particular by exploiting remote services such as SSH or VPN access. The reason behind this result is probably the low-security of IoT devices [39], and the fact that often the default password is not changed [40] [41], so with brute force technique: the device can be easily remotely accessed. In fact, one more heavily used TTP is *Brute Force: Password Guessing* (T1110.001) for Credential Access, confirming our belief that the login is unsecured.

In Table IV, instead, the most used TTP for each phase is shown. We want to focus on the last entries in his table, so the less instrumented capabilities. Since our data set is composed of only one backdoor family, this is the only program that collects, exfiltrates, and escalates privileges. In fact, if we compare this table with Tables VI and VII, we notice that neither of the two other malware types use these capabilities.

As seen before, the most implemented abilities are those allowing the attacker to control and spread the malware and gain information about the targets. The C&C is often implemented over a web protocol such as `HTTP` or `HTTPS`: the T1071.001 TTP (*Application Layer Protocol: Web Protocols*) is one of the most present.

TABLE IV
This table shows the most prevalent TTP for each category

ATT&CK Phase	Most prevalent TTP	Count
discovery	T1083	10
command-and-control	T1071.001	9
initial-access	T1133	8
execution	T1059.004	7
impact	T1498.001	7
credential-access	T1110.001	6
lateral-movement	T1210	4
defense-evasion	T1070.004	4
persistence	T1053.003	4
exfiltration	T1020 T1041	1
privilege-escalation	T1078.003 T1543	1
collection	T1005, T1074.001, T1119, T1560.001, T1560.003, T1602.002	1

B. Most prevalent TTPs in Botnet and Ransomware families

In this part we emphasize the most exploited technique for every phase with respect to the malware typology. From Table VI and Table VII we observe that there are some similarities and some clearly different implemented strategies.

The impact phase has the same techniques for every family, since they are all the same kind of malware. Ransomware has the goal to encrypt the system and obtain money from the victim in exchange for decrypted files, so five out of five use the *Data Encrypted* for Impact (T1486) technique. While, botnet's objective is to create a large network of infected machines that the botnet owner can control. In fact, all the malware have as impact: *Network Denial of Service* (T1498). Besides one family named Silexbot, that had as intent to make the victim devices unusable by wiping and destroying the disk and firmware, the techniques linked to this are displayed in the Table V.

Meanwhile, the phases that are not even implemented are the ones that involve gaining data as a part of the attack. For instance, collection and exfiltration tactics.

In addition by comparing the command and control phase, we can notice that they have the same TTP as more prevalent. The technique is T1071.001 (*Application Layer Protocol: Web*

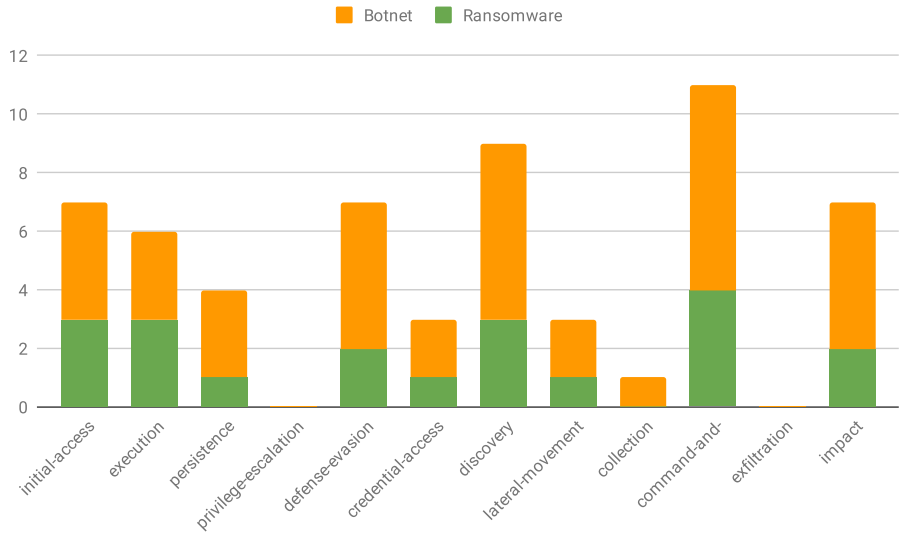

Fig. 2. Number of unique and different TTP used by Botnet and Ransomware by MITRE ATT&CK Tactics.

TABLE V
SILEXBOT IMPACT TACTIC MAPPED

Technique ID	Techinique name
T1561.001	Disk Wipe - Disk Content Wipe
T1561.002	Disk Wipe - Disk Structure Wipe
T1495	Firmware Corruption
T1529	System Shutdown/Reboot

TABLE VI
THIS TABLE SHOWS THE MOST PREVALENT TTP FOR EACH CATEGORY FOR THE BOTNET SAMPLES.

ATT&CK Phase	Most prevalent TTP	Count
impact	T1498.001	7
command-and-control	T1071.001	5
discovery	T1057	5
initial-access	T1133	5
execution	T1106	4
lateral-movement	T1210	3
defense-evasion	T1027.002	2
persistence	T1053.003	2
collection	T1602.002	1
credential-access	T1110.004	1
exfiltration privilege-escalation	x	8

Protocols). In our opinion, this is the most implemented kind of C&C, because it is the easier way to control the program by sending commands over the network.

C. TTPs added to IoT malware families over the past 2 years

In the latest years, we have noticed an increasing number of TTPs in the persistence phase. This change in implementation is probably due to the growing attention given to IoT security. While, in the past, it was sufficient to install the malware in the system, in recent times it requires more work to bypass

TABLE VII
THIS TABLE SHOWS THE MOST PREVALENT TTP FOR EACH CATEGORY FOR THE RANSOMWARE SAMPLES.

ATT&CK Phase	Most prevalent TTP	Count
impact	T1486	5
discovery	T1083	5
defense-evasion	T1070.004	3
command-and-control	T1071.001	3
execution	T1059.004	3
initial-access	T1133	3
credential-access	T1110.001	2
lateral-movement	T1210	1
persistence	T1053.003	1
collection exfiltration privilege-escalation	x	5

new security features. In order to be able to stay as long as possible, the programs need to implement some persistence techniques: the most used is *Scheduled Task/Job* especially with the *Cron* process (T1053.003). The attacker abuses this software utility to perform time-based execution of malicious code, and possibly to execute processes at system startup. Moreover, this application allows running programs under the context of a specific user as part of lateral movement or privileges escalation.

D. TTPs removed to IoT malware families over the past 2 years

In recent years, lateral movement capabilities are not implemented as much as before. Our analysis highlighted the drop of two techniques linked with the *Lateral Movement* tactic: *Remote Services* (T1021) and *Exploitation of Remote Services* (T1210). This is possibly caused by the fact that latest malware, such as Dark Nexus, have the command and control

server that takes care of the botnet propagation. Related to this, also the discovery of network information, System Network Configuration Discovery (T1016), is no longer enforced.

We observed also that privilege escalation was not an interest of IoT malware authors. We think that the benefits gained from executing these kinds of malware, with higher privileges, in these circumstances, are not worth further implementations.

E. Comparison between two Malware categories

As illustrated in Fig. 2 we can see the different approaches in implementation between Botnet and Ransomware. The figure represents the number of different techniques used in each attack lifecycle phase, so the higher is the number the greater is the number of non-identical TTPs in that tactics.

The botnet samples use more different TTPs, this probably is due to the widespread use of such malware, which brings more innovation through the years. While ransomware probably requires fewer changes to be as effective as before.

An interesting result is the difference between the number of distinct techniques used in the *Discovery* tactic; we suppose that this happens because botnets also have capabilities to spread themselves and rely on this ability. Conversely, ransomware needs less information about the system, it needs only information about file and directories.

These results highlight also the difference in the C&C tactic. The botnet needs a more developed command and control server to be able to control the bot network as the aim of this malware is to build a network of controlled machine, rather than ransomware that does not often need a C&C to be fully working.

VI. Related Work

Most existing work on the malware evolution focuses on a specific kind of malware [42] or presents code-based similarity analysis between samples. Some research uses machine learning to understand the correlation and evolution of programs, but very few focus on Linux IoT malware. To the best of our knowledge, we are the first to focus on the pre- and post-exploitation aspects of IoT Linux malware and make the most of the MITRE ATT&CK framework to deliver a different analysis methodology. Moreover, we build a comparison between different malware families based on the extrapolated TTPs that characterize them.

An interesting work [43] that exploits the power of the ATT&CK framework describes a new approach that uses hierarchical clustering to deduce technique associations that point out technique inter-dependencies in a TTP chain.

Cozzi et al. [44] describe in detail and develop a strategy to understand Linux malware on a large-scale data set. The analysis is a high-level study on the shared practices between different malware families. Another work from Cozzi et al. [45] better focuses on tearing down all the aspects of IoT malware. They apply function-level and code-base similarity between samples to show the correlations between malware. Although they state important points about code reuse among different families and AV detection failures, they rely only on the malware implementation detail. The main difference with our work is that we focus more on the overall aspect of an attack, including the malware's initial access and its final impact.

Another similar work is proposed by Torabi et al. [46] that analyzes deeply the relationship between IoT malware using a string-based similarity approach, showing also the correlation between the malware development and the Covid-Sars-19 pandemic, where a high evolution of malware has been observed. They leverage Natural Language Processing (NLP) techniques (e.g. word tokenization) to process and extract meaningful strings and apply a combination of Jaccard and overlap similarity coefficients to provide their results which highlighted high code reuse between different samples. Another work that researches IoT Linux malware is [47]. Dang et al. focus on fileless attacks on Linux-based IoT devices but with emphasis on the infection stage only.

A work related to ours is from Alrawi et al. [48]. The authors propose a novel framework to understand the lifecycle of IoT malware and compare the findings with malware targeting other kinds of devices such as desktop and mobile. Moreover, the researchers extended the investigation to other aspects rather than code-based comparison only. The main difference with our study is the data set. The one used by the authors is larger, and as they state: the majority of samples are Mirai variants, which we avoided since we wanted to have a set of families that do not share pieces of code, to provide a complete experiment of the proposed methodology. Moreover, they do not fully map malware capabilities to all the ATT&CK tactics.

To study the evolution of malware there are researches on using machine learning to detect and learn similar malware behavior. Wadkar et al. [49] use a support vector machine to detect the evolution of malware. Also, Tupadha et al. [50] implement some machine learning models to evaluate the features distances between different malicious programs. Even if these are prominent studies, their focus is constrained on Windows samples only. Indeed, it is hard to find some machine learning applications in studying the evolution of IoT malware largely because this is a new and fast-growing phenomenon. Although, some applications exist, like [51], but are focused in the detection of Linux-based IoT malware.

VII. Limitation and Future work

We are aware that our research may have two limitations. The first is the number of malicious programs analyzed, and the second is the unbalanced data set, in the number of families per category we have.

These limitations underline the difficulty of gathering unique IoT Linux malware that does not take code from other malware. Older malware is more prone to be found in multiple variants, while recent malware programs are less shared among the community, so it's harder to find any information or code to analyze.

Despite these limitations, our goal was not to produce a large-scale study on IoT malware but rather to propose a

methodology that exploits the power of the MITRE ATT&Ck taxonomy that can be applied to a larger set of data. Still, further data collection is required to determine more precisely changes in IoT malware behavior with this analysis methodology.

The unavailability of a large number of unique malware families brings to have a data set that is not well balanced. In our research, we had a large number of botnet samples and a relatively small number of ransomware ones. It is obvious that not having the same quantity for both will result in a biased comparison between the two categories, but we tried to not rely on numbers only.

The authors are aware that these families are only a part of the overall malware targeting these devices, which include both known but under NDA and unknown malicious programs. Even so, the data we have represents the reality, more botnet samples are available because of the growth of this phenomenon [52], the simplicity of writing one with a limited skill-set, and the huge accessibility of already crafted code. Moreover, the ransomware attacks increased dramatically over the last year [53] [54], since many victims are more prone to pay [55]. In addition, there is a lack of different kinds of malware and the reason could be that IoT devices are not worth creating sophisticated software.

Our results are encouraging and we believe this work puts the basis to better understand and analyze upcoming malware. Future work will focus on finding new malware categories to explore and newer samples, this will help us to better finalize our belief in evolution that will occur in IoT malware. We would extend the research to keep up with present-day threats to make the analysis even more precise.

VIII. Conclusions

This paper has investigated the evolution of the Internet of Things Linux Malware based on changes in TTPs. This work aims to propose a different analysis point of view for this growing phenomenon. The results achieved encourage us to follow this research to find a more detailed evolution path of IoT Linux Malware.

This research provides results on the evolution of IoT threats based on the ATT&CK framework. As reported in the section above, some changes in the techniques implemented are probably the reaction to the advancements in security. The findings highlight the introduction of some defense mechanisms that, in our opinion, support the fact that awareness on IoT security is growing. Furthermore, the results suggest that botnet malware is more prone to instrument different kinds of techniques and change rapidly. While ransomware is less predisposed to changes and even the ones that seem different in the end exploit the same techniques.

The analysis, though, has some limitations: the number of malware analyzed is little, but we focused on delivering a different investigation methodology rather than a large-scale study. In addition, another challenge that needs to be addressed in future works is the unbalanced sample data set: a well-balanced data set will result in more trustworthy results.

References

[1] SonicWall Inc, "Mid-Year Update: 2021 SonicWall Cyber Threat Report," 2021. [Online]. Available: https://www.sonicwall.com/resources/white-papers/mid-year-2021-sonicwall-cyber-threat-report/

[2] Help Net Security. (2021) IoT malware attacks rose 700% during the pandemic. [accessed August 18, 2021]. [Online]. Available: https://www.helpnetsecurity.com/2021/07/20/iot-malware-attacks-rose/

[3] A. Chacko and T. Hayajneh, "Security and privacy issues with iot in healthcare," *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 4, no. 14, 07 2018, copyright - © 2018. This work is licensed under http://creativecommons.org/licenses/by/3.0/ (the "License"). Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License; Last updated - 2020-04-06. [Online]. Available: https://www.proquest.com/scholarly-journals/security-privacy-issues-with-iot-healthcare/docview/2306345400/se-2?accountid=14390

[4] C. Burgos. (2017) Is your NAS exposed to the Internet? [accessed August 19, 2021]. [Online]. Available: https://www.securityartwork.es/2017/05/05/is-your-nas-exposed-to-the-internet/

[5] N. Daube. (2019) One of the Greatest Threats Facing the Iot: Router Security. [accessed August 19, 2021]. [Online]. Available: https://www.cyberdefensemagazine.com/router-security/

[6] Deloitte. (2020) Internet of Things (IoT) The rise of the connected world. [accessed August 19, 2021]. [Online]. Available: https://www2.deloitte.com/content/dam/Deloitte/in/Documents/technology-media-telecommunications/in-tmt-IoT_Theriseoftheconnectedworld-28aug-noexp.pdf

[7] E. Edelson, "Security in network attached storage (nas) for workgroups," *Network Security*, vol. 2004, no. 4, pp. 8–12, 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1353485804000650

[8] MITRE. [accessed August 15, 2021]. [Online]. Available: https://www.mitre.org/

[9] R. Brewer. (2017) The six stages of a cyber attack lifecycle. [accessed August 18, 2021]. [Online]. Available: https://www.helpnetsecurity.com/2017/03/06/cyber-attack-lifecycle/

[10] MITRE ATT&CK. (2021) Updates - April 2021. [accessed August 13, 2021]. [Online]. Available: https://attack.mitre.org/resources/updates/updates-april-2021/index.html

[11] M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. A. 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 *26th USENIX Security Symposium (USENIX Security 17)*. Vancouver, BC: USENIX Association, Aug. 2017, pp. 1093–1110. [Online]. Available: https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/antonakakis

[12] T. Seals. (2019) Mirai Botnet Sees Big 2019 Growth, Shifts Focus to Enterprises. [accessed August 18, 2021]. [Online]. Available: https://threatpost.com/mirai-botnet-sees-big-2019-growth-shifts-focus-to-enterprises/146547/

[13] Threat Intelligence Team. (2018) Seven new Mirai variants and the aspiring cybercriminal behind them. [accessed August 18, 2021]. [Online]. Available: https://blog.avast.com/hacker-creates-seven-new-variants-of-the-mirai-botnet

[14] A. Remillano and J. Urbanec, "New mirai variant uses multiple exploits," May 2019. [Online]. Available: https://www.trendmicro.com/en_us/research/19/e/new-mirai-variant-uses-multiple-exploits-to-target-routers-and-other-devices.html

[15] Malware Must Die. (2019) CAYOSIN DDoS Botnet - A Qbot base upgraded with Mirai codes. [accessed August 16, 2021]. [Online]. Available: https://imgur.com/a/4YxuSfV

[16] P. Paganini. (2019) Cayosin Botnet: a deeper look at this threat supported by the psychological profile of the "youngsters-wannabe-hackers" Rolex boasters. [accessed August 16, 2021]. [Online]. Available: https://securityaffairs.co/wordpress/80858/cyber-crime/cayosin-botnet-mmd.html

[17] I. Ilascu. (2019) Cr1ptT0r Ransomware Infects D-Link NAS Devices, Targets Embedded Systems. [accessed August 17, 2021]. [Online]. Available: https://www.bleepingcomputer.com/news/security/cr1ptt0r-ransomware-infects-d-link-nas-devices-targets-embedded-systems/

[18] Libsodium documentation. [accessed August 17, 2021]. [Online]. Available: https://libsodium.gitbook.io/doc/

[19] C. J. Dietrich. Silexbot bricks IoT devices - A detailed look at its distribution. [accessed August 16, 2021]. [Online]. Available: https://chrisdietri.ch/post/silexbot-brick-iot-devices/

[20] Wikipedia contributors. (2021) BrickerBot — Wikipedia, The Free Encyclopedia. [accessed August 16, 2021]. [Online]. Available: https://en.wikipedia.org/w/index.php?title=BrickerBot&oldid=1002459487

[21] A. Remillano II and J. Urbanec. (2019) Neko, Mirai and Bashlite Target Routers, Devices. [accessed August 16, 2021]. [Online]. Available: https://www.trendmicro.com/en_us/research/19/h/back-to-back-campaigns-neko-mirai-and-bashlite-malware-variants-use-various-exploits-to-target-several-routers-devices.html

[22] The UPX Team. UPX the Ultimate Packer for eXecutables. [accessed August 16, 2021]. [Online]. Available: https://upx.github.io/

[23] R. Nigam, H. Zhang, and Z. Zhang. (2021) New eCh0raix Ransomware Variant Targets QNAP and Synology Network-Attached Storage Devices. [accessed August 17, 2021]. [Online]. Available: https://unit42.paloaltonetworks.com/ech0raix-ransomware-soho/

[24] unixfreaxjp - Malware Must Die. (2019) MMD-0064-2019 - Linux/AirDropBot. [accessed August 16, 2021]. [Online]. Available: https://blog.malwaremustdie.org/2019/09/mmd-0064-2019-linuxairdropbot.html

[25] C. Cimpanu. (2020) CISA says 62,000 QNAP NAS devices have been infected with the QSnatch malware. [accessed August 16, 2021]. [Online]. Available: https://www.zdnet.com/article/cisa-says-62000-qnap-nas-devices-have-been-infected-with-the-qsnatch-malware/

[26] K. Günther. SHC Shell Compiler. [accessed August 16, 2021]. [Online]. Available: https://www.linux-magazine.com/Online/Features/SHC-Shell-Compiler

[27] SecurityScorecard, "QSnatch Technical Report," SecurityScorecard, Tech. Rep., 2020. [Online]. Available: https://securityscorecard.com/resources/qsnatch-technical-report

[28] . [accessed August 16, 2021]. [Online]. Available: https://pastebin.com/UGe1vNEJ

[29] Denys Vlasenko. [accessed August 16, 2021]. [Online]. Available: https://www.busybox.net/

[30] A. Zahravi. (2019) Momentum Botnet's Newest DDoS Attacks and IoT Exploits. [accessed August 16, 2021]. [Online]. Available: https://www.trendmicro.com/en_us/research/19/l/ddos-attacks-and-iot-exploits-new-activity-from-momentum-botnet.html

[31] F. Valsorda and B. Cartwright-Cox. (2019) age. [accessed August 16, 2021]. [Online]. Available: https://github.com/FiloSottile/age

[32] Wazuh Inc. WAZUH. [accessed August 16, 2021]. [Online]. Available: https://wazuh.com/

[33] R. Lakshmanan. (2020) Dark Nexus: A New Emerging IoT Botnet Malware Spotted in the Wild. [accessed August 16, 2021]. [Online]. Available: https://thehackernews.com/2020/04/darknexus-iot-ddos-botnet.html

[34] L. Abrams. (2021) Qlocker ransomware shuts down after extorting hundreds of QNAP users. [accessed August 16, 2021]. [Online]. Available: https://www.bleepingcomputer.com/news/security/qlocker-ransomware-shuts-down-after-extorting-hundreds-of-qnap-users/

[35] P. Arntz. (2021) Check your passwords! Synology NAS devices under attack from StealthWorker. [accessed August 17, 2021]. [Online]. Available: https://blog.malwarebytes.com/botnets/2021/08/check-your-passwords-synology-nas-devices-under-attack-from-stealthworker/

[36] Synology Inc. (2021) Synology® Investigates Ongoing Brute-Force Attacks From Botnet. [accessed August 17, 2021]. [Online]. Available: https://www.synology.com/en-global/company/news/article/BruteForce

[37] MITRE ATT&CK. MITRE ATT&CK Enterprise Linux Matrix. [accessed August 18, 2021]. [Online]. Available: https://attack.mitre.org/matrices/enterprise/linux/

[38] (2021) OpenCTI - website. [accessed August 18, 2021]. [Online]. Available: https://www.opencti.io/en/

[39] J. Wurm, K. Hoang, O. Arias, A.-R. Sadeghi, and Y. Jin, "Security analysis on consumer and industrial IoT devices," in *2016 21st Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2016, pp. 519–524.

[40] Ophtek LLC. (2018) Default Passwords: The Biggest Weakness in IoT Security. [accessed August 19, 2021]. [Online]. Available: https://ophtek.com/default-passwords-biggest-weakness-iot-security/

[41] C. Cimpanu. (2017) 15% of All IoT Device Owners Don't Change Default Passwords. [accessed August 19, 2021]. [Online]. Available: https://www.bleepingcomputer.com/news/security/15-percent-of-all-iot-device-owners-dont-change-default-passwords/

[42] M. Humayun, N. Jhanjhi, A. Alsayat, and V. Ponnusamy, "Internet of things and ransomware: Evolution, mitigation and prevention," *Egyptian Informatics Journal*, vol. 22, no. 1, pp. 105–117, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1110866520301304

[43] R. Al-Shaer, J. M. Spring, and E. Christou, "Learning the Associations of MITRE ATT amp; CK Adversarial Techniques," in *2020 IEEE Conference on Communications and Network Security (CNS)*, 2020, pp. 1–9.

[44] E. Cozzi, M. Graziano, Y. Fratantonio, and D. Balzarotti, "Understanding linux malware," in *2018 IEEE Symposium on Security and Privacy (SP)*, 2018, pp. 161–175.

[45] E. Cozzi, P.-A. Vervier, M. Dell'Amico, Y. Shen, L. Bilge, and D. Balzarotti, "The tangled genealogy of iot malware," in *Annual Computer Security Applications Conference*, ser. ACSAC '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–16. [Online]. Available: https://doi.org/10.1145/3427228.3427256

[46] S. Torabi, M. Dib, E. Bou-Harb, C. Assi, and M. Debbabi, "A Strings-Based Similarity Analysis Approach for Characterizing IoT Malware and Inferring Their Underlying Relationships," *IEEE Networking Letters*, pp. 1–1, 2021.

[47] F. Dang, Z. Li, Y. Liu, E. Zhai, Q. A. Chen, T. Xu, Y. Chen, and J. Yang, "Understanding Fileless Attacks on Linux-Based IoT Devices with HoneyCloud," in *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 482–493. [Online]. Available: https://doi.org/10.1145/3307334.3326083

[48] O. Alrawi, C. Lever, K. Valakuzhy, R. Court, K. Snow, F. Monrose, and M. Antonakakis, "The circle of life: A large-scale study of the iot malware lifecycle," in *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, Aug. 2021, pp. 3505–3522. [Online]. Available: https://www.usenix.org/conference/usenixsecurity21/presentation/alrawi-circle

[49] M. Wadkar, F. Di Troia, and M. Stamp, "Detecting malware evolution using Support Vector Machines," *Expert Systems with Applications*, vol. 143, p. 113022, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417419307390

[50] L. S. Tupadha and M. Stamp, "Machine Learning for Malware Evolution Detection," 2021.

[51] H.-T. Nguyen, Q.-D. Ngo, and V.-H. Le, "IoT Botnet Detection Approach Based on PSI graph and DGCNN classifier," in *2018 IEEE International Conference on Information Communication and Signal Processing (ICICSP)*, 2018, pp. 118–122.

[52] S. N. Thanh Vu, M. Stege, P. I. El-Habr, J. Bang, and N. Dragoni, "A Survey on Botnets: Incentives, Evolution, Detection and Current Trends," *Future Internet*, vol. 13, no. 8, 2021. [Online]. Available: https://www.mdpi.com/1999-5903/13/8/198

[53] J. Coker. (2021) Ransomware Attacks Grew by 485% in 2020. [accessed August 20, 2021]. [Online]. Available: https://www.infosecurity-magazine.com/news/ransomware-attacks-grow-2020/

[54] S. K. Skelton. (2021) Ransomware attacks increase dramatically during 2021. [accessed August 20, 2021]. [Online]. Available: https://www.computerweekly.com/news/252504676/Ransomware-attacks-increase-dramatically-during-2021

[55] K. Lab. (2021) Over half of ransomware victims pay the ransom, but only a quarter see their full data returned. [accessed August 19, 2021]. [Online]. Available: https://www.kaspersky.com/about/press-releases/2021_over-half-of-ransomware-victims-pay-the-ransom-but-only-a-quarter-see-their-full-data-returned