Computers & Security 119 (2022) 102758

Contents lists available at ScienceDirect

Computers & Security

journal homepage: www.elsevier.com/locate/cose

On how VoIP attacks foster the malicious call ecosystem

J. Carrillo-Mondéjar[a,*], J.L. Martinez[a], G. Suarez-Tangil[b]

[a] Universidad de Castilla-La Mancha, Albacete, Spain
[b] IMDEA Networks Institute, Madrid, Spain

ARTICLE INFO

Article history:
Received 23 September 2021
Revised 7 May 2022
Accepted 14 May 2022
Available online 16 May 2022

Keywords:
Honeypots
VoIP
Cybercrime
Telephony fraud
Robocalls
Attacks
Underground economy

ABSTRACT

Switched telephone networks are a key and ubiquitous infrastructure. Recent technological advances have integrated modern and inexpensive systems into these networks in order to use the Internet to place calls via Voice over IP (VoIP). The evolution of this technology has also led to an increase in the number and sophistication of the techniques used by criminals to commit fraud. Specifically, with the emergence of VoIP, attackers can now adapt tools commonly used by cybercriminals, such as botnets, to make their attacks more complex and insidious. For example, through bots they can dial multiple numbers automatically, enabling them to target a greater number of victims, and do so more quickly. While recent studies have shed light on how certain parts of this ecosystem work, it is still unclear how attacks on VoIP systems contribute to this type of fraud. This paper presents a novel VoIP honeypot that captures voice interactions, in addition to employing low-level telemetry. With the study of how attackers obtain access to our honeypot and the actions they perform, we present an overview of the most prevalent types of fraud used in this ecosystem, including unique insights into the origin of the attacks and the destination of calls made through our architecture. Finally, we analyze in depth the actions taken to study the different types of telephony fraud.

© 2022 The Author(s). Published by Elsevier Ltd.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

1. Introduction

The invention of the telephone is one of the most important accomplishments of humankind, eliminating the physical barriers of time and location and allowing two people from places miles apart to establish immediate communication. Although it was not designed to be used as an attack vector, unfortunately criminals found a way of using it as such in order to make unwanted or malicious calls to perpetrate extortion, scams and spam campaigns. These calls have been a long-standing and challenging problem, and are responsible for yearly losses of tens of millions of dollars worldwide (Li et al., 2018), besides being very disturbing for telephone users, who feel their privacy is violated.

The recent development of Voice over IP (VoIP) has caused telephone communications to be transferred to a more modern medium, namely the Internet, making it more convenient and cheaper for people to contact each other. These benefits have been well-received not only by ordinary users, but also by cybercriminals, who can now reduce the cost and operational complexity of their telephony campaigns, and this technology allows them to place automated calls with little effort. These calls are typically made in the form of bot-calls, which are carried out by simple software using dialer equipment to generate vast numbers of calls to a given (or randomly chosen) list of phone numbers. Their aim is to make contact with an active recipient and trick them into performing the desired action, ultimately making them fraud victims. Generally, the instructions are transmitted via a pre-recorded message but, in some cases, they can eventually be assigned to a human agent for further interaction, although this becomes a limiting factor, since human agents may not have time to attend to all the connected calls.

In comparison with other attacks such as email spam, voice spam calls are significantly more disturbing because they require immediate attention; when the phone rings, the recipient must decide whether to accept the call and listen to it just judging by the caller ID, which is often “unknown” or even spoofed (Tu et al., 2019). Even if the user ignores or declines the call, spammers can send a prerecorded audio message straight to the user’s voicemail, thus also having the static factor of email spam. In addition, these attacks feel more personal, since, typically, people are less inclined to give their phone number than their email address, so when an attacker is able to contact them, they gain an advantage over the recipient, who normally feels confused. Furthermore, most

* Corresponding author.
E-mail addresses: javier.carrillo@uclm.es (J. Carrillo-Mondéjar), joseluis.martinez@uclm.es (J.L. Martinez), guillermo.suarez-tangil@imdea.org (G. Suarez-Tangil).

https://doi.org/10.1016/j.cose.2022.102758
0167-4048/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

email spam is systematically filtered out by anti-spam algorithms, whereas, at the time of drafting this proposal, this type of countermeasures do not exist in telephony fraud.

Related works attempt to address the problem of malicious calls in two ways: i) statically prior to the call being established; or ii) dynamically during the call. On the one hand, several approaches have performed call request header analysis (Tu et al., 2016). However, this type of measures are easily evaded (e.g., by using a spoofed caller ID (Tu et al., 2019)). On the other hand, recent works have leveraged automated approaches to interactively deal with unsolicited calls. Works such as Li et al. (2018) propose machine learning as a means to prevent malicious calls tailored to the users using features that model the normal behavior expected (e.g., weekday and time). The authors in Sahin et al. (2017) study the effectiveness of an interactive voice response system called Lenny. One of the key lessons in this work is that chatbots can be used for *scambaiting*[1] fraudsters. Finally, other authors have carried out systematic studies to understand technical support scams, uncovering the call centers underpinning this fraud using Web crawlers to find scammers, domains and phones on the Internet (Miramirkhani et al., 2016).

While recent studies have shed light on how certain parts of this ecosystem work, it is still unclear how attacks on VoIP systems contribute to this type of fraud. In this paper, we show how fraudsters leverage VoIP systems to build their infrastructure, and we provide an overview of the current status of VoIP systems and the risks arising when the system itself, or the credentials of its users, is compromised. Some of our findings include:

- Attackers use our system to carry out different types of telephony fraud as well as actions that allow them to obtain greater benefits. In particular, we identify both Toll Evasion and Revenue Share as being two of the most prevalent types of fraud. We also report a large number of cases in which attempts are made to call personal numbers (both landlines and mobiles), which suggests that attacks on VoIP systems are frequently used for fraudulent purposes such as scams or spam such as those described in Sahin et al. (2017). Our work is the first to provide a quantitative and qualitative study of how fraudsters use compromised VoIP systems for *fun and profit*.
- We uncover the tricks fraudsters use to evade restrictions. For instance, there are attempts to call numbers using different prefixes to avoid dial-plan restrictions. By causing an error in the parsing of these numbers, calls to restricted destinations are permitted. Calls to wrong numbers are also used to check fingerprint honeypots, e.g., when an invalid number gives them a ringtone. Learning from these tricks enables us to propose countermeasures against VoIP attacks.
- We find that while a few attacks are random, many others are aware of the context surrounding the target of the attack, namely the *callee*. In particular: (1) fraudsters are aware of the mechanisms used by reputation systems to block connections from TOR nodes or known VPNs, and consequently 50% of IP addresses come from compromised servers that belong to hosting companies; and (2) fraudsters try to maximize the success of their attacks by calling during office hours or hours when there are usually people at home.

The rest of the paper is organized as follows. Section 2 provides an overview of the system, the experiments carried out and the ethical considerations. Section 3 presents the data analysis from our experiments. Section 4 describes the limitations and key findings of our study. Section 5 discusses the proposals from the community regarding malicious calls. Finally, the conclusions are presented in Section 6.

[1] The act of wasting the time and resources of scammers.

2. Methodology

To understand how VoIP services are misused, we deployed a honeypot and monitored the accesses and actions that took place through it. Fig. 1 shows the methodology followed in this work. First, we describe the architecture of our system and the experimental setting. Next, we describe the vulnerable user accounts that were added to our system so that attackers could gain access through them. Finally, we describe the analysis carried out and the ethical principles followed in this work.

2.1. System overview

Our honeypot runs on top of a custom infrastructure with a number of VoIP accounts. The architecture consists of a Virtual Private Server (VPS) hosted in Germany, and we fortified this server to prevent unauthorized access to services that were not relevant to our study. We built an Asterisk-based Private Branch Exchange (PBX) (Sangoma Technologies, 2021), which is a popular open-source communications framework. We decided to use a real server instead of an existing honeypot because honeypots tend to be more limited in terms of interaction and can be detected through simple fingerprinting (Vetterl and Clayton, 2018). We configured our server to allow any type of call from a properly authenticated user and we redirected the calls to an enabled private extension to simulate that the call was being handled correctly. All the accounts were subscribed to a specific dialplan that limited outgoing calls to a predefined set of countries and we restricted calls to premium numbers. The reason to restrict calls to some specific places is to simulate the legitimate behavior of a configured PBX via which it is only possible to call certain destinations.

We populated our VoIP system with a number of user names and passwords, as described in Section 2.2. Each of the accounts deployed belonged to a specific context on the server and was linked to one of the dialplans that were set up. Thus, accounts linked to a dialplan in the US, for instance, could only make calls to the US. As a result, all the calls that matched the established dialplan were redirected to one of our extensions to make fraudsters believe that the calls were successful. Our system then collects logs from Asterisk in order to perform a data analysis as described in Section 2.3.

2.2. Accounts

We designed two experiments whose aim was to differentiate how attackers gained access to our honeypot accounts.

Weak user credentials. We set up a set of 30 user accounts with weak passwords. The user accounts were numerical identifiers designed to reproduce a setting similar to the one used in large companies or universities. Thus, each user (phone) was identified as if it were a phone extension, and this same extension number was used as the password. We only used numerical user IDs and weak passwords in order to promptly capture automated bots running brute force attacks against VoIP services.

Strong user passwords. Our second experiment sets user accounts by using: i) numerical users, similar to how they were created in the previous sections but this time using randomly generated strong passwords; and ii) common users created with the concatenation of a name and surname. For the generation of random names and surnames we used Enron's public email dataset (Klimt and Yang, 2004). This dataset was made public when the investigation into the Enron company ended. We extracted all the names and surnames, and randomly generated

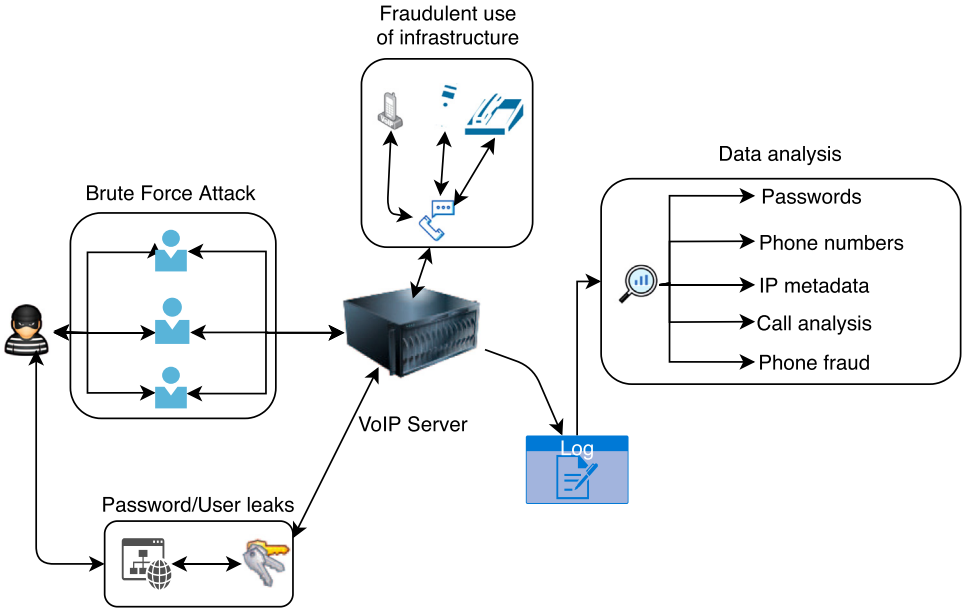


Fig. 1. Methodology.

dummy identities which we used to create users in our VoIP system. For this latter set of users, we generated the passwords by concatenating a long list of different dictionary words to simulate passwords created by the user that can be remembered.

Account leak. The next phase of this experiment was to leak the credentials in a controlled experiment in which we recorded the user name and the source of the leak. We chose a series of paste sites and underground forums to leak credentials since these sites are often used to publish credentials stolen from different Internet services. We leaked 70 accounts to different chosen places, with 30 of them being published on paste sites such as pastebin.com, paste.ee and pastebin.xyz. Another 20 accounts were published in underground forums such as offensivesecurity.net and blackhatworld.com. The remaining 20 accounts were leaked in prominent Russian underground forums. We tried to mimic the way in which information leaks from other services have occurred. As it is not a publicly known service such as gmail, facebook, etc., we also included the IP address and port of our VoIP server.

2.3. Data analysis

In this phase, we performed a data-driven analysis on the actions carried out in our system by the attackers and on the metadata that could be extracted from the attackers that interacted with our honeypot. In particular, our analysis first aimed to characterize the features of the IP addresses we received connections from. We then reversed the password used by the attackers when attacking the authentication mechanism of the VoIP server. Finally, we analyzed both the phone numbers and the metadata inferred from the call attempt. We describe all this in greater detail below.

IP address analysis. We analyzed all the metadata associated with the IP addresses that interacted with our system as well as the open ports of the systems associated with them. We analyzed the location of these addresses to discover the origin of the attack and we also checked whether these addresses appear in lists of known proxies or bots. Finally, we consulted the Autonomous System Name to which the IP addresses belonged as well as the open ports and vulnerabilities presented by the systems that were running behind them. It should be noted that no type of active scanning was carried out on these IP addresses and that

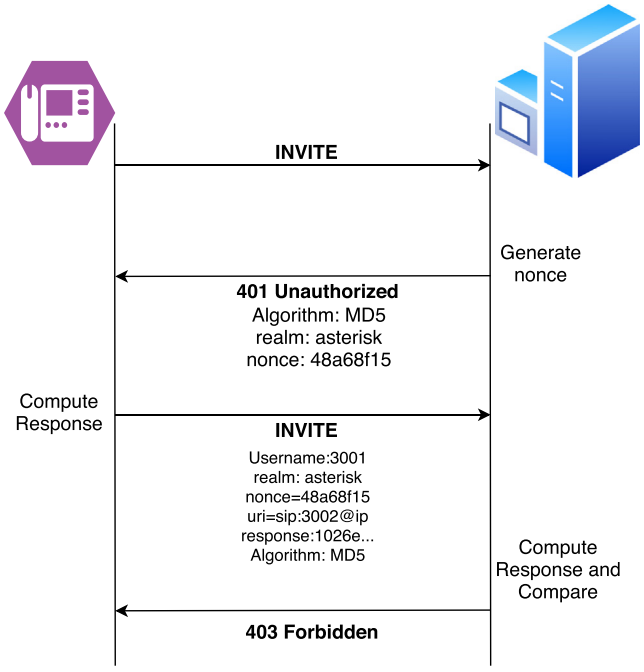


Fig. 2. Authentication message flow in the SIP protocol.

open-source tools available through the Internet were consulted (e.g., Shodan (2022) and IP Intelligence (Getipintel, 2022)).

Passwords. Although in the first version of the SIP protocol the passwords travel in plain text through the network, in version 2 the SIP protocol uses a mechanism based on challenge/response to carry out user authentication (Strand and Leister, 2011). Fig. 2 shows the exchange of messages when a user fails to authenticate through the INVITE method. Unlike the REGISTER method, which registers the device and tells the server where it should direct the calls destined for that user, the INVITE method allows calls to be made without registering the device on the server.

Basically, the client sends an INVITE request to the server, which returns a message saying that it is not authorized, and in

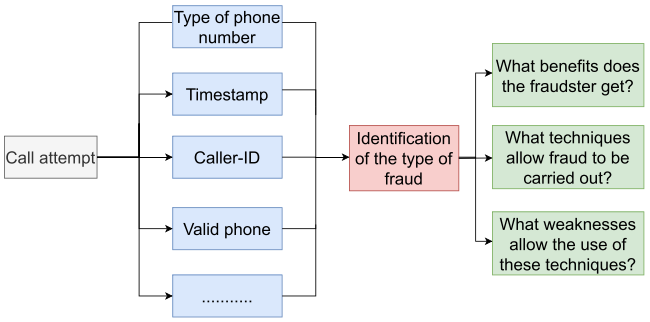


Fig. 3. Methodology followed to identify the type of fraud.

which it transmits a nonce and the realm to authenticate. The client computes the response using Algorithm 1 and sends it back to the server. Finally, the server computes the response and, if it matches the one sent by the user, it will send a correct authentication message, or in this case it will send a message that the authentication has not been carried out successfully.

$$\begin{aligned} A &= MD5(user : realm : password) \\ B &= MD5(method : sip : uri) \\ hash &= MD5(A : nonce : B) \end{aligned} \quad (1)$$

Therefore, the server never sees the plain text password used by the client, which in this case is the password used by the attackers to try to identify themselves to our server. The authentication message flow of the REGISTER method is the same, with only the method parameter changing when computing the response. By using all the data collected from the authentication messages, we obtained all the variables in Algorithm 1 except the password since it is not transmitted in plain text over the network. Therefore, we used a dictionary of the most common passwords to apply brute force and find the most frequently-used passwords on our system.

Phone numbers. We analyzed all the phone numbers that the attackers tried to call through our system. We parsed these numbers to extract information about whether the number was valid based on the patterns that phone numbers follow around the world. We also extracted the location of the phone number based on the country code and the area code, as well as the type of phone number (e.g., land-line or mobile).

*Call analysis and phone fraud.*We analyzed the calls that were made through our system by looking at the timestamp of the calls that were attempted in order to check whether the calls were made while taking into account the local time of the destination or whether they were automated calls that did not use this information. Therefore, on the basis of the type of call, we classified it according to the type of fraud that was being attempted through our system and the benefits that attackers could obtain through this type of fraud. For the identification of the types of fraud we rely on the taxonomy described in Sahin et al. (2017), which systematically explores the different types of fraud in telephone networks as well as the techniques and benefits obtained by fraudsters. Fig. 3 shows how we apply the taxonomy in our scenario to identify the type of fraud that is being carried out in the call attempt. Once the fraud scheme that is being attempted is discovered, we identify the possible benefits as well as the techniques and weaknesses that allow this fraud scheme to be carried out.

2.4. Ethics

We followed the ethical principles for Internet-mediated research in Anabo et al. (2019), which in turn stem from The Belmont Report (Miracle, 2016) and the Menlo Report (Dittrich et al., 2012). We addressed the challenge of obtaining informed consent from unknown Internet users through the *beneficence* principle and by making a risk-benefit assessment. On the one hand, we implemented strict mechanisms to minimize risks, reducing the harm to mere *annoyance*. On the other hand, our honeypot effectively reduces the surface of the attack and prevents criminals from benefiting while depleting the resources of the attackers. We also gained an understanding of the underlying ecosystem that can allow the community to further tackle telephony scams. Finally, by granting attackers access to our VoIP we deterred them from using other compromised VoIP systems. Our risk-benefit assessment was evaluated by our Institutional Review Board (IRB), who considered our decision-making ethical and have approved this study. We next detail the measures we adopted to minimize risks to privacy.

First, we limited the usage of our VoIP system by redirecting calls to a private extension, thus preventing these calls from being forwarded to their actual recipient. Second, we restricted the scope of our experiment by limiting the number of target countries. Although no external calls were made from our system, we simulated that these were being made by calling a private extension created for this purpose, so attackers might believe that these numbers existed and try to call from other systems. Finally, for the analysis of the metadata of all the attackers of the system, no active scan was performed. Instead, third-party services (e.g., Shodan (2022)) were used to obtain information and geolocations.

Table 1
Summary of the actions performed on the VoIP server.

	Calls	
Type	Weak credentials	Strong credentials
Call attempts	306,044	713
Unique phone numbers	211,113	319
Valid phone numbers	10,731	141
Possible phone numbers	34,913	98

3. Data analysis

We ran the experiment for a period of around 4 months for accounts with weak credentials. After those 4 months, we blocked access to those weak accounts and created new accounts with strong passwords, as described in Section 2.2. This experiment lasted around 2 months. During the course of both experiments, we monitored the actions performed on the VoIP server, collecting information for further analysis.

3.1. Overview

We deployed an Asterisk-based PBX server for our experiment. During the first period of the experiment, the server used accounts with weak credentials, while in the second part of the experiment the server used accounts with strong credentials. After parsing the information collected on our system, we detected 2209 unique IP addresses on the server. For the first period, all the accounts that were created in the VoIP system were accessed. On the other hand, out of the 70 accounts we leaked, 20 were accessed, with 12 of these having been leaked in underground forums and 8 on paste sites.

We logged the actions performed with all the accounts, which can be summarized in Table 1. We can see that the total number of call attempts is around 300 K. Unique phone numbers amount to about 200 K and represent the total number of different phones to which calls were attempted. We used the phonenumbers library Drysdale (2022) to validate the phone numbers. This library maps out a taxonomy of phone number patterns of all the countries worldwide and indicates whether the phone number is valid. We considered phone numbers as possibly valid when we saw

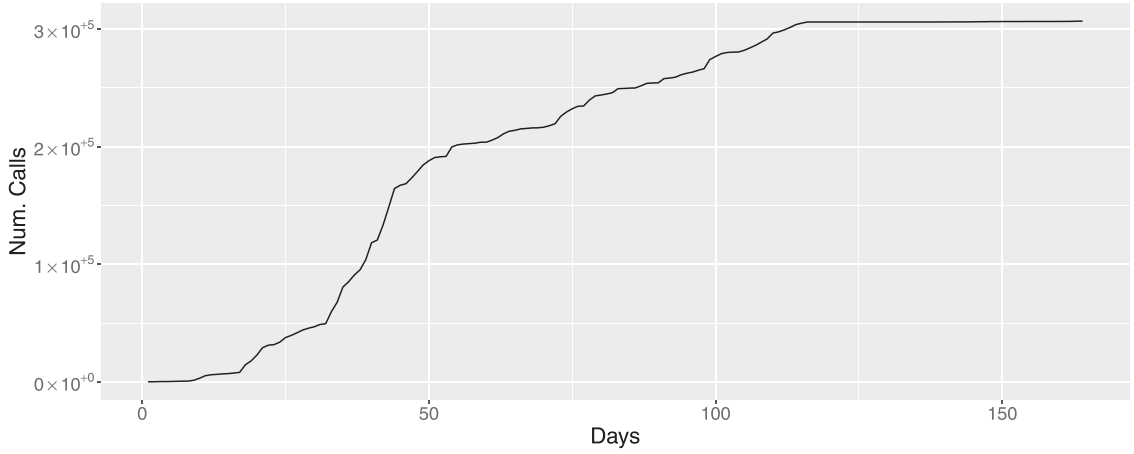


Fig. 4. Cumulative number of calls attempts (weak and strong credentials).

numbers that did not contain the country code but would be valid for a local dialplan in at least one country. For example, a number that does not contain the country code +44 (UK), but is a correct number when calling from a UK dialplan. As these numbers are often correct in different regions, we can not determine the exact destination of the call. The total number of valid and possibly valid phone numbers was around 10 K and 30 K. It is important to note that the library indicates whether the number is valid based on the dialing plan of that country, but it does not mean that this number is currently owned by any user. By observing the results, it can be seen that the number of call attempts is much higher during the time that the weak credentials experiment lasted. These calls are usually generated by bots that are scanning the Internet for active VoIP servers and are looking for valid credentials by brute force and, once they find them, they start generating calls. On the other hand, when using strong credentials, attackers could not successfully access the honeypot using brute force, and therefore we are confident that the access was through the password leaks made through forums or paste sites. As expected, the number of accesses was lower, and there may even have been curious accesses made by attackers who were interested in seeing whether they could call for free.

Regarding the call attempts, we analyzed their timestamps for each day of the experiment. Fig. 4 shows the results of this analysis.

It can easily be seen that, during the experiment with weak credentials, the number of calls increased daily while remaining constant at the end of the experiment, which is when strong passwords were used. In addition, the first calls with weak credentials occurred within a few hours of putting the server online, while with strong credentials, the first calls occurred around 20 days after leaking the credentials of the user accounts.

3.2. IP address analysis

In this section, we analyze the IP addresses that interacted with our system. In total, 1280 and 1188 unique IP addresses interacted with the server with weak and strong credentials, respectively. Between the two sets of IP addresses, 259 match on the two servers. Out of all those IP addresses, only 244 and 19 successfully logged in. The number of IP addresses that interacted with both servers is very similar in terms of the number of different addresses detected. However, it is important to note that the strong credentials experiment started immediately after the weak credentials one finished, and only 259 IP addresses coincide between the two experiments, which seems to indicate that, either the actors that were actively searching for online servers changed their IP address to avoid blacklist bans, or those that might have been compromised servers detected the intrusion and closed their access.

Table 2
Top 20 sources and destinations of calls with weak and strong credentials. The table represents only those calls for which it was possible to extract the phone number's geographical location.

Weak credentials			Strong credentials		
Src	Dst	Calls	Src	Dst	Calls
DE	UK	822	PS	US	78
US	UK	742	PH	CU	33
NL	UK	488	PH	PH	25
NL	FR	407	PH	UK	11
NL	RU	387	PS	ES	8
NL	US	355	PS	IL	6
NL	BO	313	PH	BJ	5
NL	BQ	294	PS	UK	9
US	SG	273	US	US	3
NL	AW	273	US	UK	3
NL	LK	270	US	EG	3
NL	IQ	255	PH	US	3
NL	CW	213	PH	TN	3
NL	AM	207	PH	SI	3
NL	GT	191	PS	JE	3
DE	ZM	176	PS	CH	3
DE	RS	172	PS	FR	3
US	JP	171	PS	AT	3
NL	PE	171	DE	FR	3
NL	JP	171	CH	RU	2

Location. We depict the source of the IP address in Fig. 5. Blue points represent the origin of the IP addresses that interacted in the experiment with weak credentials, green represents the origin of the IP addresses that interacted in the experiment with strong credentials and red represents those IP addresses that matched in the two experiments.

Germany is the most common location with 21.9% of the IP addresses collected, followed by the Netherlands (20.8%), the United States (14.9%), France (12.1%) and the Palestinian Territory (10.2%). It can be seen that there is a large group of IP addresses in the same zone that tried to interact in the strong credentials experiment. This may be due to a side effect of the lack of success of the brute force attack.

We also investigated the origin and destination of the calls that were attempted through our system. We only considered those calls made to valid numbers and from which the geographical location of the telephone number could be extracted. Table 2 shows the top 20 sources and destinations of calls that were attempted

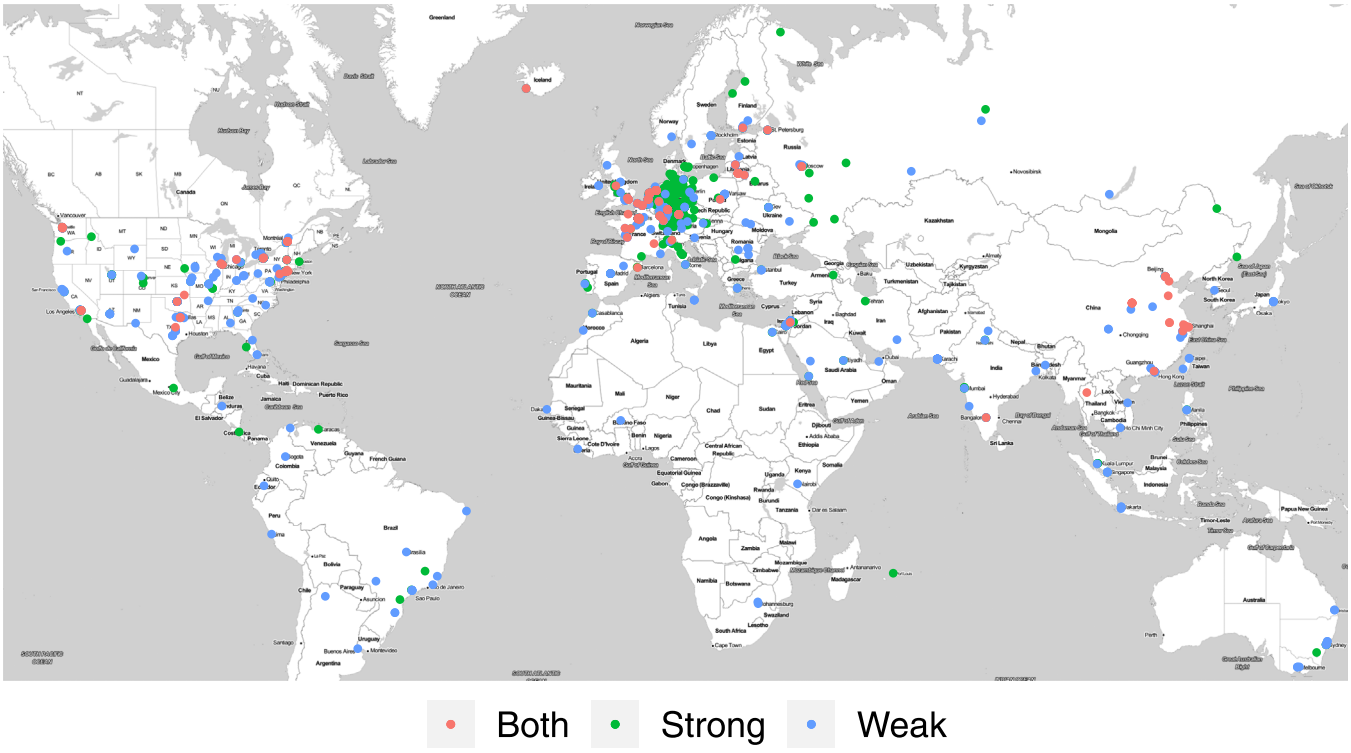


Fig. 5. Geographical map of the location of the IPs that interacted with our system.

with both weak and strong credentials. It can be seen that in the experiment with weak credentials most of the calls were attempted from the Netherlands, while with strong credentials they came from Palestine and the Philippines. It should be noted that this source may not be the actual location of the different actors, but may be hidden behind some proxy, Virtual Private Network (VPN), hostings or similar system with the IP address of those countries.

Hostings. We also analyzed the IP addresses through public services (e.g., VirusTotal or shodan) to search for information and see whether there were relationships between them. Although some IP addresses were previously associated with certain malware samples and domains, there does not seem to be a clear relationship between them. However, on the basis of information about the organization or ISP to which they belong, they do indicate that certain IP addresses were related to each other. By consulting the public databases of IP address information services, we found that around 45% of the IP addresses that appeared in our system belong to hosting services. We plot a graph with the different IP addresses that belong to hosting services together with their autonomous system (AS). Fig. 6 shows the connections between the different IP addresses collected. It can be seen that large groups of IP addresses belong to the same organization (e.g., Online S.A.S or OVH), which seems to indicate that these sites may have been compromised and used by fraudsters to hide their identity or that these hostings are less restrictive when it comes to banning clients that generate fraudulent traffic. We also consulted the IP addresses via an intelligence service (Getipintel, 2022), which uses probabilistic techniques and machine learning models to determine whether an IP address is a proxy, a VPN, a hosting service, tor node, etc. Given an IP, the system returns a value of 1 if the address is explicitly banned from their lists. Otherwise it returns a value between 0 and one, indicating the probability that it is a bad IP. According to the results obtained, 477 are explicitly on their blacklists and 41 have a value greater than 0.98, indicating that these IP addresses have a high probability of being proxies, vpn's, etc.

Table 3
Top 30 most common open ports.

Num. IP	Port	Num. IP	Port	Num. IP	Port
404	80	65	445	46	995
301	443	60	8089	45	465
289	22	58	5985	44	500
163	5060	56	143	44	137
116	3389	55	993	32	8000
92	53	49	1723	32	2000
89	21	49	110	29	5222
85	123	48	111	27	8443
75	25	47	8080	27	2087
73	3306	47	587	27	2082

Open ports. We analyzed the open ports of the IPs that appeared in the logs of the VoIP system. Table 3 shows the most common open ports among the IP addresses collected.

We observe that there are open ports of common services, such as HTTP, SSH, SSL, SMB, FTP or RDP. Also, one of the methods used by attackers to compromise systems is through any of these services, either because of some vulnerability or through the use of weak passwords to access through SSH. In fact, most botnets in the IoT ecosystem infect different devices through the use of weak passwords in services such as SSH or Telnet (Antonakakis et al., 2017; Kambourakis et al., 2017). Another curious circumstance that we found regarding open ports is the appearance of ports 5060 and 3389. The former is a port belonging to UDP and which is the default port used by VoIP systems. This seems to indicate that they are tracking other VoIP systems and searching for (or using) passwords from other VoIP systems to use them on their system and make calls through other users. The latter is the port used by the RDP protocol, which is used for the remote Windows desktop. The RDP protocol contains a recent vulnerability known as BlueKeep (Microsoft, 2019), whose exploit was made public in the summer of 2019 (Proffitt and Wolf, 2019). This vulnerability allows the execution of remote code and could be used by attackers to

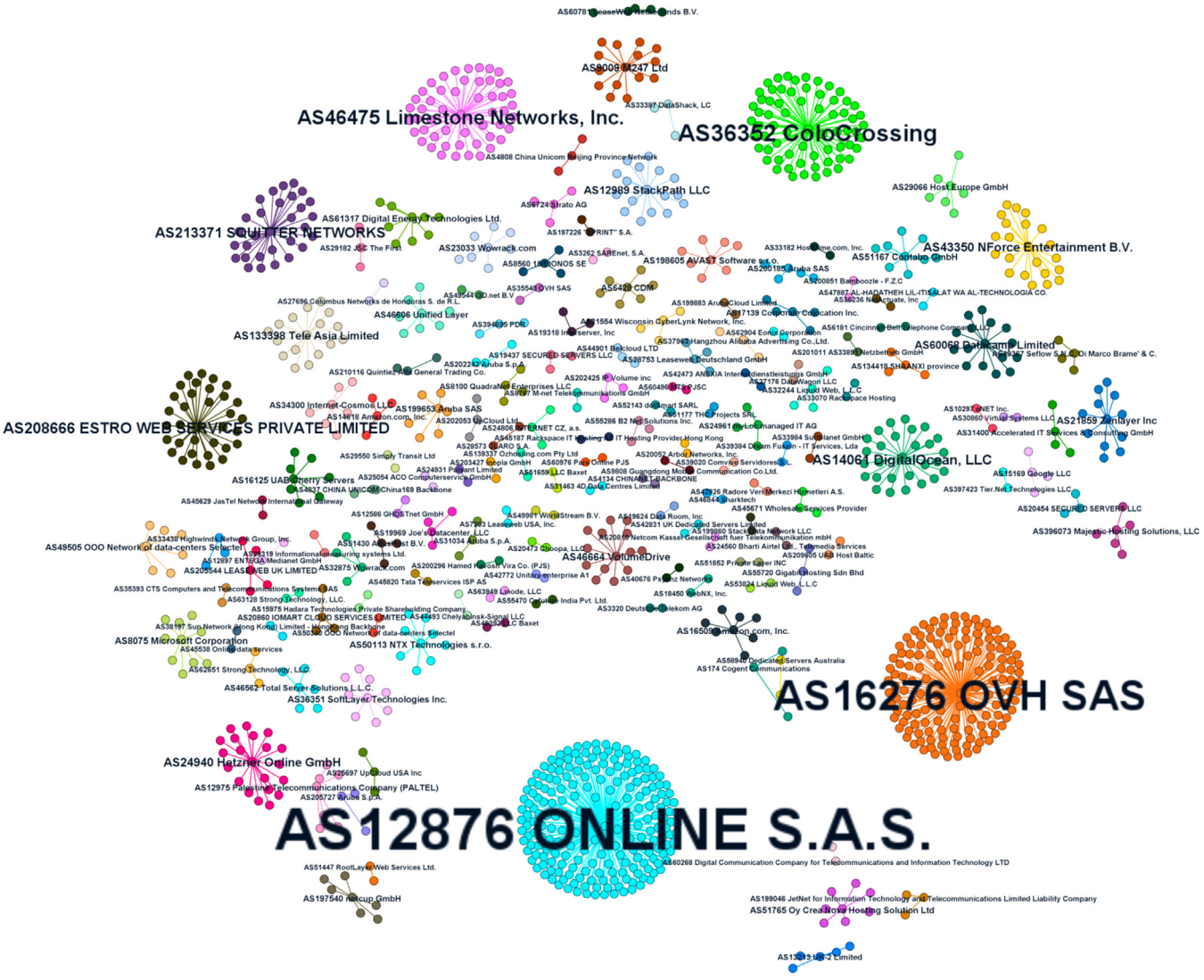


Fig. 6. Relations between the different IPs collected that belong to hosting services.

compromise systems and use them for their interests, such as attacking the user credentials of a VoIP system or as a proxy to carry out their attacks and not reveal their real IP address.

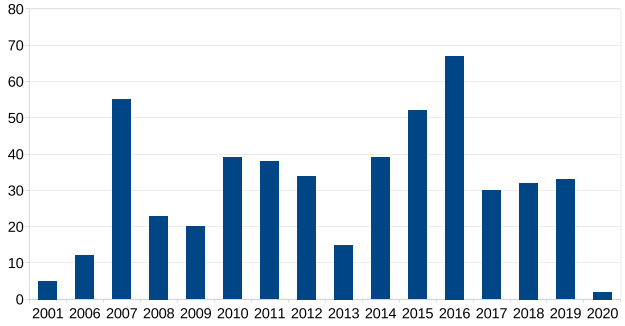
Vulnerabilities. We consulted the IP addresses through open source intelligence services (e.g. Shodan) in order to obtain the vulnerabilities detected in the services that are exposed to the Internet. Of all the information collected, we found 495 different Common Vulnerabilities and Exposures (CVEs). Fig. 7(a) shows the distribution by year of the vulnerabilities that were collected. It can be seen that there are vulnerabilities that were discovered more than 10 years ago, so it seems to indicate that not enough attention has been paid to these servers by their administrators. Also, it can be seen that they are mostly recent vulnerabilities, and that the maximum peak is reached in the vulnerabilities of 2016. It is important to note that these open source intelligence services, shows the current vulnerabilities based on the software version that is running, but do not check whether the vulnerability actually exists. We obtained information on each of the CVEs in the public database of Cvedetails (2022). Each CVE has an associated score that indicates the criticality of the vulnerability and its impact. Fig. 7(b) shows the distribution of CVEs based on the score collected for each one in cvedetails. It can be seen that most of the CVEs have a score between 4 and 8, which corresponds to medium (4–6.9) and high (7–8.99) vulnerabilities. We can also observe that there are several CVEs with a high criticality index (9 and 10). Among the critical vulnerabilities that exist, most of them are related to the Hypertext Pre-Processor (PHP) interpreter. In addition, vulnerabilities appear in Windows systems with the Remote Desktop Protocol (RDP) and the Server Message Block (SMB), which are widely known for the impact they have had throughout the world (s.r.o, 2017; Mitre corporation, 2019).

3.3. Passwords

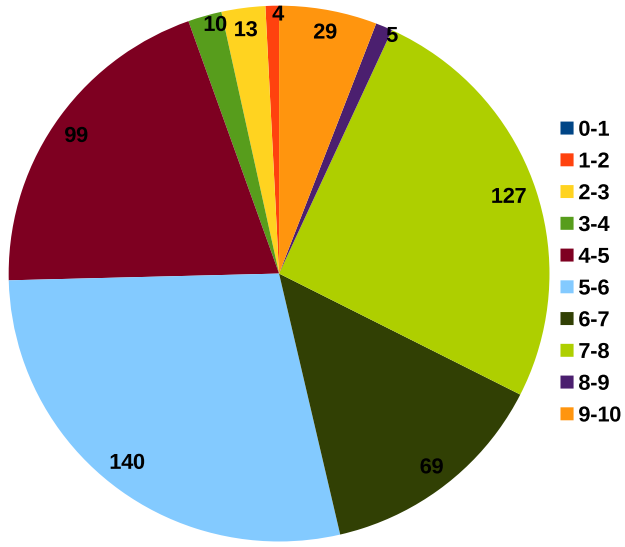
In the Asterisk security event log, you can find the hash computed to carry out the authentication, as well as the nonce sent by the server and the other parameters necessary to calculate the response, except for the password used by the user to authenticate. We parsed the security logs to find each of the fields necessary to compute the response for each of the server's failed attempts. Since we have all the data, except the password and the response hash, brute force can be applied by generating a response for a given password and checking whether the resulting hash is equal to the hash that the user sent in the authentication process. We applied brute force to each of the failed authentication attempts by using a dictionary of the 1000 most common passwords from

Table 4
Top 30 most common passwords used in our system.

Rank	Passwords	Hits	Rank	Passwords	Hits	Rank	Passwords	Hits
1	1234	7944	11	abcd1234	5330	21	asdfg	2330
2	123,456	7006	12	123abc	4911	22	1111	2050
3	1000	6481	13	12,345	3501	23	8888	2023
4	password	6103	14	test123	2921	24	9999	2022
5	pass123	6032	15	qwe123	2540	25	1q2w3e	2018
6	12,345,678	5563	16	asd123	2531	26	1,234,567	2016
7	123,123	5545	17	test	2491	27	111,111	2000
8	abc123	5541	18	password123	2466	28	123,321	1995
9	0000	5442	19	admin123	2423	29	123qwe	1972
10	4321	5431	20	secret	2369	30	123,456,789	1971



(a) CVEs distributed by the year of their appearance.



(b) CVEs distributed by their criticality score.

Fig. 7. CVEs collected from the different IPs that interacted with our honeypot.

Miessler and Haddix (2022). In total, we managed to find 411,497 passwords for all the failed login attempts that occurred on our system (less than 1%). Table 4 shows the top 30 most commonly used passwords of all those found by brute force. The low percentage of cracked passwords could indicate that attackers are using larger dictionaries or even algorithms to generate passwords based on username. In the experiment with weak credentials we used as username and password the SIP extension and all the accounts were compromised, so it is likely that they use an algorithm to create passwords based on the extension they are trying to find.

3.4. Calls and fraud classification

In this section, we present the different destinations of the calls made as well as the types of fraud identified.

Table 5
Types of phone numbers that attackers tried to call.

Type of phone numbers	Valid phones	Possible phones
Mobile	7517	26,649
Landline	1901	1616
Premium rate	514	225
Landline or mobile	357	1393
Universal Access Number	180	138
Personal number	131	21
VoIP	125	317
Shared cost	62	158
Pager	55	55
Toll free	30	1418
Voicemail	0	21

Origin of phone numbers. We analyzed the numbers to which calls were attempted in our system. As we discussed in Section 3, we found a total of 10,872 valid numbers (for both strong and weak credentials) and 35,011 possible ones. We plot the geographical information of the valid phone numbers on a map to depict where the calls that were attempted were directed. Fig. 8 shows a heat map of the different phone numbers found regardless of whether they were called with strong or weak credentials. The numbers can be mapped to the region to which they belong, and not the exact location of the call. We can see from the heat map that most of the unique phone numbers belong mainly to Europe, Africa, Russia and the Havana and Cuba area. However, the most frequently called countries were the United Kingdom, the United States, Russia and France with 3028, 1063, 435 and 428 calls, respectively.

Types of phones. The type of phone numbers which the attackers tried to call was extracted. We used the phonenumbers library to obtain the type of phone number based on the patterns that each country uses for its phone numbers. In some regions, there is no distinction between landline and mobile phone numbers (e.g, the USA), and these cases are included in the landline or mobile numbers category. Table 5 shows the different phone numbers found that were either valid or possible. It can be seen that most of them belong to mobiles or landlines. It is important to note that the possible numbers are local ones that could be valid in some region and, therefore, are not completely accurate since some regions have similar dialing plans and it might be the case, that this number is of another type depending on the region to which it belongs.

The rest of the phone numbers are as follows:

- Premium rate. Telephone numbers through which different services are offered and whose call prices are higher than a conventional call.
- UAN. This is a phone number that allows a company to have several lines associated with it. In this way, when the UAN number is called, one of these lines will be reached depend-

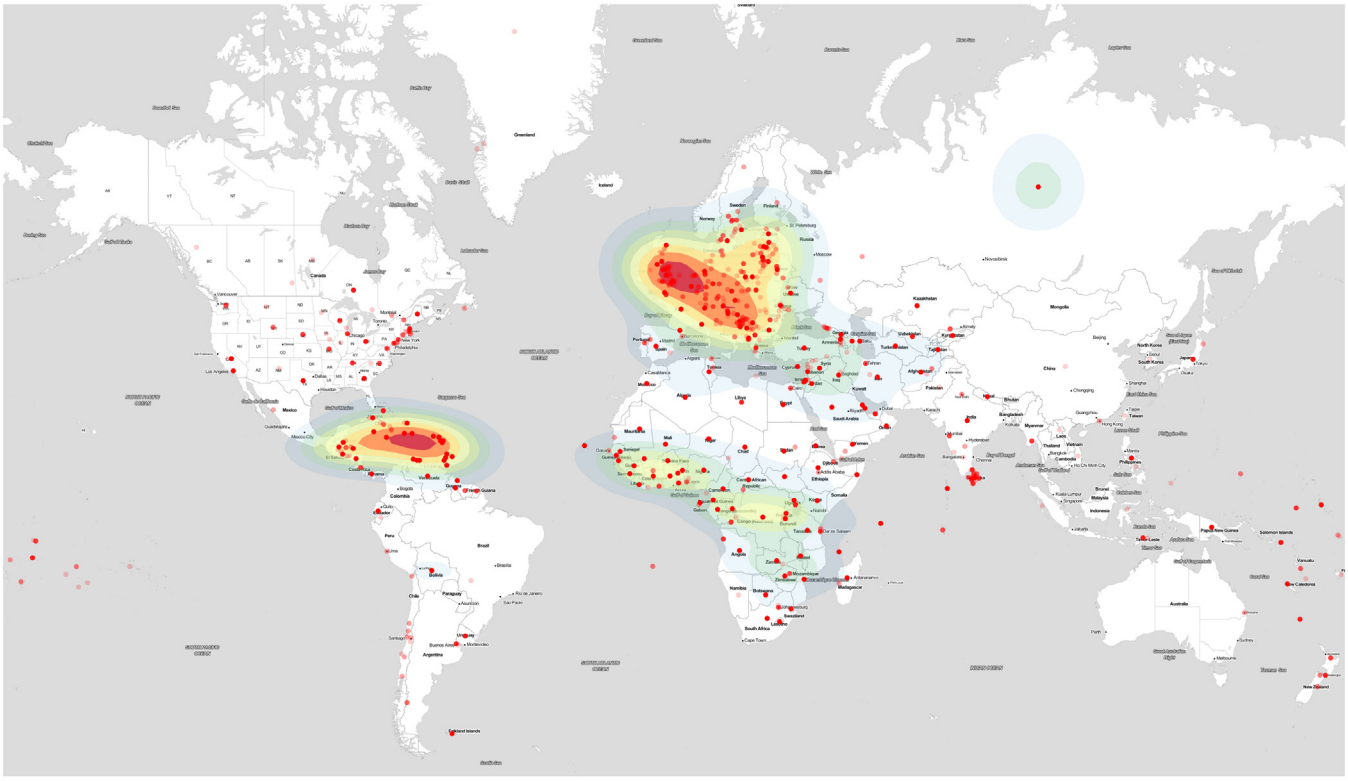


Fig. 8. Geographical map of the location of the phones.

ing on, for example, the geographical location or the time of the call.

- Personal number. This corresponds to virtual phone numbers belonging to the United Kingdom. These phone numbers allow the routing of phone call to another number and have been reported as being used for fraudulent purposes such as fraud or spam Ofcom (2016).
- VoIP. Phone numbers that are assigned to a user instead of a specific phone line. These are contracted through VoIP service providers.
- Shared cost. Telephone numbers for which the cost of the call is divided between the owner of the number and the caller.
- Pager. Phone numbers assigned to devices that are capable of receiving and displaying messages, also known as "beepers" (Wikipedia, 2022).
- Toll free. These are free phone numbers for the person making the call, as the cost is billed to the owner of the phone number.

The rest of the telephone numbers that attackers attempted to call are invalid, that is, the size or dialplan does not match any region. However, most of these numbers do consist of valid numbers that have been concatenated with numbers or symbols at the beginning. For example, the telephone number "*0000000000**4412XXXXXXXX" is not valid, but it can be seen that, if the initial part is omitted, the number that remains is a valid number[2] with country code 44, which is the code for the United Kingdom. This pattern is repeated for all the invalid numbers with different digits, symbols and size. We believe it may be a way to try to evade the established dialplan or to know whether they are in a honeypot. For example, if calls to numbers that are not valid give a ringtone, it may indicate that the system is a honeypot.

[2] We have replaced part of the original number with the letter X so that it cannot be recognized.

Types of fraud. As we mentioned in Section 2.3, we rely on the taxonomy described in Sahin et al. (2017) to identify the different fraud schemes that were attempted through our system. This taxonomy identifies the different weaknesses and techniques that lead to the different types of fraud and the benefits that they can generate. First, we look at the different fraud schemes identified and the possible benefits attackers can obtain through them. Then, we examine the weaknesses and techniques that allow fraudsters to carry out such actions. The fraud schemes detected are the following:

- Toll Evasion Fraud. This type of fraud is the best known and most widely used over recent years in telephone networks. It is based on making calls without having to pay the charges of the call, which are billed to another party. In this specific case, when accessing the user accounts of a PBX to call, it is the owner of the PBX or the user of the compromised account that is charged for the calls. The benefits can range from simply not paying for anonymous calls involving criminal activities to committing other types of fraud. All the calls that were made were with compromised accounts, so the attackers did not pay for the calls.
- Revenue Share Fraud. This type of fraud occurs when an agreement is made between an operator (or a third-party service provider) and a fraudster in which the latter is responsible for making telephone calls to certain numbers owned by the former (e.g., international premium rate service numbers), generating a revenue which is then shared between them. The benefit of this type of fraud is purely economic, and the more call traffic that is generated, the greater the benefit will be for the fraudsters. As can be seen in Table 5, calls were made to premium numbers, which allow attackers to obtain benefits through these special rate numbers. Also, it can be seen in Fig. 8 that a large part of the telephone numbers belong to the area of Havana and Cuba. These numbers maintain a dialplan

similar to that of the United States, although many correspond to special tariffs.

- Voice Spam and Scams. This scheme includes any type of unwanted or abusive calls, such as spam calls or scams. These calls can be made in many ways, from a phishing call posing as a company to telemarketing and robocalling with prerecorded messages. The possible benefits that can be obtained through this type of fraud include collecting information about the users (e.g., whether the phone belongs to a real user), and convincing or deceiving the recipients of the call to carry out some action that could also have some economic benefit for the fraudster. Most calls could be classified as this type of fraud, in which attackers could use compromised VoIP systems to launch scam campaigns or to find new potential victims.
- Wangiri fraud. Also known as callback scam or ping call, this is a specific type of voice spam or scam fraud that consists in making a missed call with the aim of making curious users return the call. A key feature of this type of fraud is the use of caller ID spoofing. We can identify this type of scam by observing whether there was any attempt to modify the caller ID and by checking whether the time from the beginning to the end of the call is just a few seconds. The possible benefit that the attackers obtain from this type of fraud is economic, since the telephone number is owned by the scammers and they obtain a benefit from the calls received. About 200 calls tried to change the caller ID (we discarded those that tried to change it to the anonymous caller ID).

The techniques and weaknesses on which these types of attacks are based within the VoIP system are similar. Basically, part of the problem of fraud over mobile phone lines comes from the interconnection of multiple technologies and the existence of a large number of operators and services. There is a lot of variaton in the regulation and laws concerning telephony between different countries, which has contributed to an increase over recent years in the amount of telephone fraud. In our experiment, attackers gained access through accounts with weak passwords and account leaks on different paste and forum sites. Although this was done on purpose in order to know what happens when fraudsters have access to a system, it often happens that users with little security awareness use weak passwords and unwittingly allow third parties to access their systems. The techniques available for VoIP aim to compromise a PBX server or user accounts to make calls, even simultaneous ones, in order to increase a fraudster's benefits. Fraud schemes can be related to each other since, in this case, all schemes start with avoiding the cost of the calls that are made. It is important to note that no type of fraud was committed through our honeypot since the calls were directed to an internal extension and were not routed to the telephone network.

3.5. Clustering campaigns

In this section, we compute the relationships between the source IP addresses of the calls and the valid telephone numbers receiving the call. We depict the relationships in graph form in Fig. 9. Larger nodes represent source IP addresses, while smaller nodes represent valid phone numbers. The edges connect the IP addresses and phone numbers to which calls were attempted and their thickness represents the number of calls. Black telephone numbers represent those numbers that have a special rate (i.e., international numbers with a premium rate, Montserrat, Barbados, etc.), while black nodes representing the origin of the call indicate that they tried to change the caller id to another phone number (we discard anonymous caller IDs from the representation). The gray clusters represent those IPs that only made one valid call and whose phone number is not related to any other IP addresses.

Now we have grouped the relationships between the source and the destination of the calls, together with the type of fraud, we can analyze the the different clusters. This enables us to draw the following conclusions:

- There are different IPs that are related by the phone numbers they call, as some of them match for most of the phone numbers. This indicates that they belong to similar campaigns. For example, the pink and the green clusters are campaigns that target different objectives. The former cluster focuses mainly on the UK, while the latter cluster has a more international objective, including various countries, such as the US, India or Germany. The former campaign mainly leverages IP addresses belonging to known VPNs or proxies, while the latter uses IP addresses that are more scattered and that include residential IP addresses (i.e., bots connecting directly from the home network of the victims).
- We can see campaigns with a particularly characteristic topology. For instance, there are campaigns with IP addresses that only made calls to a single valid telephone number. This happens when there are campaigns that are targeted. We can also see cases where the phone number was used to check whether the credentials work (i.e., leaked credentials in a forum) or attackers checking whether they could make valid calls (i.e., as a previous step before making a brute-force attack).
- Most of the campaigns were motivated by an economic benefit as their main incentive, mostly through the use of special rate or premium numbers. We can also see campaigns launching unwanted calls to phone numbers of common users, namely SPAM calls.
- Only a few campaigns tried to change the caller ID to another phone number, presumably controlled by the attacker, and if so, it would be the Wangiri fraud. Among these campaigns, only two of them share a destination phone number, implying that the other campaigns can be attributed to different actors.

3.6. Timestamp analysis

In this section, we provide a timestamp analysis focused on calls to the United Kingdom and the United States, which are the two countries that were called the most (based on the numbers obtained as valid and for which it was possible to obtain data on their geographical location). We use local time for both countries on the basis of the time zone to which the phone number belongs. Fig. 10 shows the graph of calls to each country, distributed both by the times at which the call occurred and the day of the week, regardless of whether they were made in the experiment for weak or strong credentials. By looking at the graphs, we can observe that the time of the calls is targeted at the local customs, that is, they are mainly made during office hours and between 8:00 and midnight, when it is common for people to be at home. Also, it can be seen that, for the UK, a percentage of calls were made between 00.00 and 03.59, which could indicate that certain types of bots make calls once they access the system regardless of the time or day.

Regarding the day of the week, we observe that calls were made on any day and that in the UK the number of calls spikes on Thursday, Friday and Saturday. By contrast, in the United States, the number of calls per day is similar, reaching its peak midweek.

4. Discussion

In this section we discuss the limitations of our work, as well as the findings of the study.

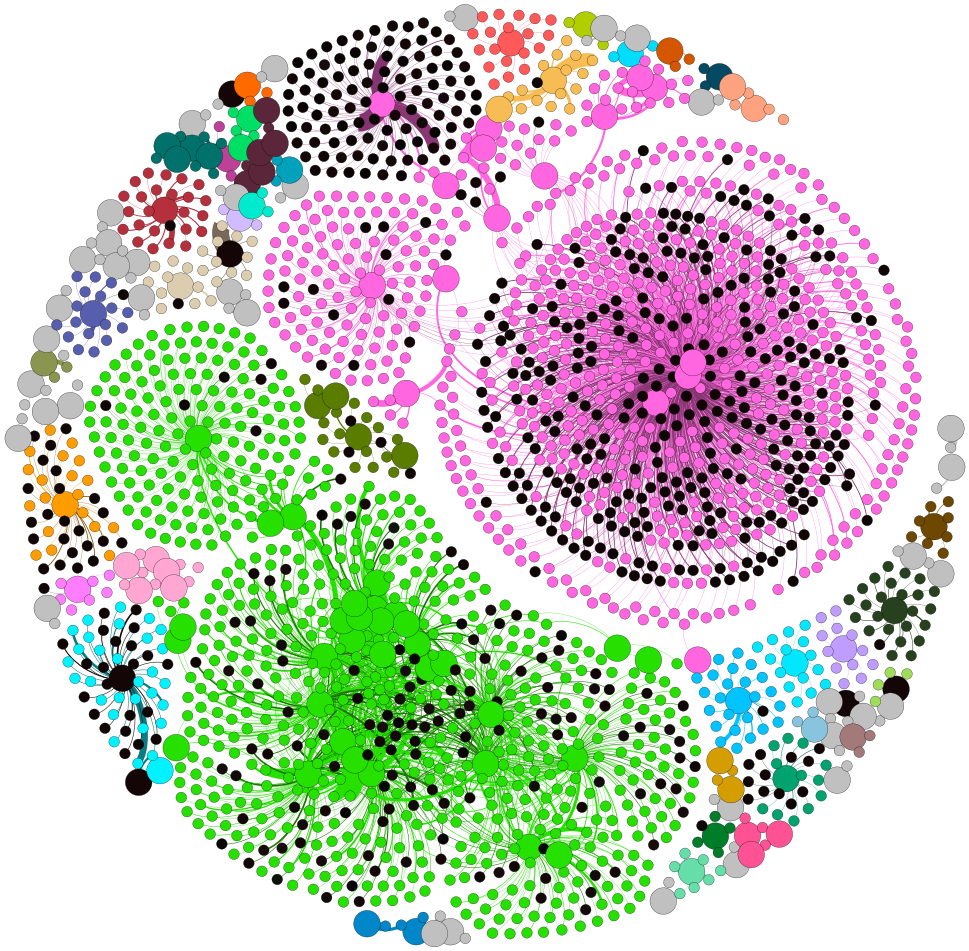


Fig. 9. Clustering of IP addresses and valid phone numbers. The larger nodes represent the IP addresses from which the calls originated and the smaller nodes represent valid phone numbers. The edges connect IPs and phones if there was a call attempt from the former to the latter.

4.1. Limitations

We encountered a number of limitations in the course of the experiments. On the one hand, due to the ethical requirements that this type of experiments have, calls were not routed to the telephone network; instead, those calls that coincided with the dialplan were directed to a private extension to simulate that calls were routed correctly. Some attackers may have been trying to call a phone number they had control over to check whether the calls were being made correctly. This can cause some of the attackers to realize that they are in a system with certain limitations and thus not to carry out any of the actions that they would perform in a real system. On the other hand, when we carried out credential leaks, these were made on public paste sites and in underground public forums that allow the free registration of any user. For example, in private forums, users rely more on the information provided by other users since an invitation is needed to be part of that community.

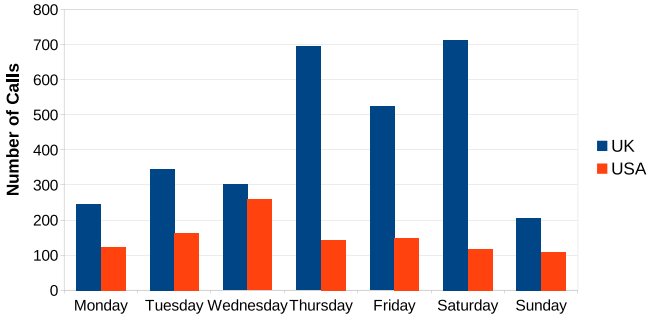
Configuration. We used the Long Term Support (LTS) 16.3.0 version of asterisk as PBX server, enabling the option "alwaysauthreject = no" so that the server response is different when the identification is wrong in the experiment with weak credentials. That is, it shows a different message when the password is incorrect or when the user does not exist. This makes it easier for attackers to identify users through brute force, making it easier for attackers to find usernames and passwords. User accounts are set to a specific dialplan that allows calls to only certain countries (see Sections 2.1 and 2.2), trying to simulate a real system where calls to other countries are not allowed. Using other configurations or dialplans and it may reveal other types of attacks by attackers.

4.2. Key findings

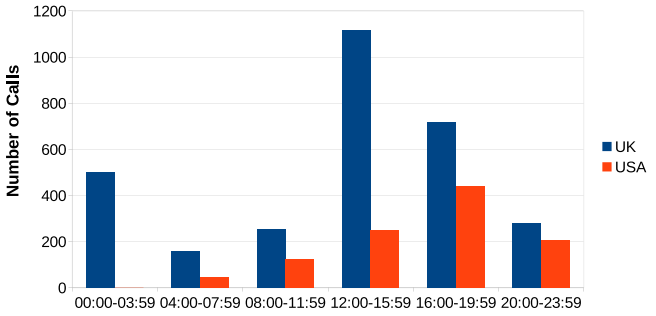
In this work we have shown that VoIP attacks are a threat that underpins part of the malicious call ecosystem. We next summarize the key findings of our work.

Type of attacks. We provided for the first time, a study based on a live measurement of the incoming utterances in the calls that shows how fraudsters use a compromised VoIP system. Among a comprehensive list of different frauds we observed, we found that two types stand out in the wild: Toll Evasion and Revenue Share. First, Toll Evasion enables the commoditization of botnets performing telephony fraud in a similar fashion to *pay-per-install* services sold in underground markets (Thomas et al., 2015). Second, we saw how Revenue Share also adopts successful fraudulent schemes such as "affiliate" marketing (McCoy et al., 2012) or "click fraud" (Stone-Gross et al., 2011) in the realm of telephony.

Depleting. We can see that scammers try to call certain phone numbers several times. This seems to indicate that they obtain phone numbers through leaked or purchased databases (Tu et al., 2016) and attempt to call them once they have access to a system. Related works have shown that depleting is effective against spam (Stringhini et al., 2012). Using a honeypot may help organizations to keep a list of the phone numbers that scammers call and thus



(a) Number of calls made per day of the week.



(b) Number of calls made by time slot.

Fig. 10. Number of calls that were made to the United States and the United Kingdom. The time zone used is the local time of the geographic location of the phone numbers.

minimize the number of spam calls, which can be very annoying (e.g., when there are call attempts to a number whose caller is trying to make other calls at the same time).

Source attacks. Although we know that attackers may use other services to hide their identity, most of the logged IP addresses belonged to Europe. As we saw in Section 3.2, many of them are related and belong to the same organization. Most of these organizations provide hosting services or private and virtual servers for users to set up their own services. Everything seems to indicate that the attackers exploit known vulnerabilities in services hosted on these machines as the software has often not been properly patched. Although nowadays many resources and efforts are invested in raising awareness among users regarding security issues and keeping their systems up to date, there are still many vulnerable devices accessible from the Internet that have not received the necessary patches.

Calls. We observed that most of the calls that were directed to the United States and the United Kingdom were made within the time zone corresponding to each country, thus increasing the chances of their success. However, some calls fell outside the usual time slot, indicating that some bots make calls without assessing the context of the target. In addition, call patterns were found to generate numbers that are invalid. As we saw in Section 3.4, these numbers are made up of valid numbers to which prefixes of different sizes have been added with numbers or symbols. This indicates that the bots are trying to actively bypass the restrictions of the dialplan used by the VoIP server. Introducing prefixes can cause errors in the mechanism used to enforce restrictions, which in turn can let attackers evade these mechanisms and make calls. This behavior, however, is very noisy and can be leveraged to design a system to systematically detect these bots. We can also see that mobile and landline calls are the most targeted type of services as opposed to premium rate services, as was initially expected. This makes it challenging to deploy mitigating actions that rely on the dialplan being used.

Accounts. Fraudsters actively search for user accounts, scanning the Internet for known VoIP server ports to which they try to connect and apply brute-force login credentials. In addition, they also make use of leaked accounts on different sites, and it is more than likely that sales of stolen accounts take place. Although with strong credentials far fewer calls were made compared with accounts with weak credentials, attackers did access the former and similar behaviors were observed.

4.3. Countermeasures

In this section we describe a series of countermeasures to fortify VoIP servers and thus prevent a fraudulent use being made of them.

Accounts. As we have seen, attackers scan VoIP servers and make connection attempts by sweeping users and passwords in order to find weak credentials that allow them to gain access. Depending on the server configuration, it is possible to enumerate valid users based on the response from the server. In the authentication process, the server can respond with the message “Not Found” or “Unauthorized” (Jansky et al., 2017) depending on whether the user is correct or not. A good practice is to configure the server so that the response is the same regardless of whether the user exists but the password is incorrect or the user does not exist. Although the default configuration of Asterisk in the latest versions avoids giving different messages in the authentication process, it can be modified or configured erroneously by the server administrator in the config files. Also, on the basis of the brute force attacks received, we can see that most of them attack numeric users, so using named users would make the task of finding users more difficult for the attackers. Finally, in the event of an attacker obtaining valid credentials, either by brute force or by theft, for example by using malware on mobile devices, a good measure is to continuously monitor the calls made and limit the number of simultaneous calls that can be made through the server. Usually, once they access the server, the attackers try to make numerous simultaneous calls, which can cause great economic losses (Osenbaugh, 2019).

Dialplan. Fraudsters try to make calls to phone numbers in different parts of the world, including attempts to call strange phone numbers in order to avoid the restrictions of the dialplan configured for the user. It is important to implement a correct dialplan configuration and only allow calls to places where you operate or where you usually make calls, including implicit rules for the types of number you want to allow and disable calls to international and premium numbers.

Architecture. To deal with attacks against SIP servers, a good measure is to implement an SIP proxy server (e.g., (Kamailio and Kamailio, 2022)). This allows you to load balance and redirect valid packets to the corresponding SIP server and release the load from the SIP server. Through a proxy, security is increased since it makes it possible to discard and ban brute-force attacks, ban the default user agent of known tools (e.g., Sipvicious (Gauci, 2022) or Sipscan (Endler and Collier, 2022)), provide the same error message in the authentication process[3] regardless of whether the username or password is incorrect and implement other security measures. It is also possible to ban brute-force attempts on the server itself via iptables with solutions such as Fail2ban (Jaquier, (2022)). Another good measure is to change the default port of the SIP server to avoid being found by automatic scanners looking for the default VoIP ports.

[3] See limitation of dictionary attacks with the HTable module in Kamailio at: https://kamailio.org/docs/modules/stable/modules/htable.html##idm49.

5. Related work

In this section we compare this paper with previous related works.

Honeypot systems. Nawrocki et al. (2016) review the state of the art of the different honeypot software systems and the data they collect by making a classification based on the type of interaction that the attackers have as well as the services and applications they emulate. Luo et al. (2017) propose a honeypot system for different Internet of Things (IoT) architectures with the aim of detecting attacks or even zero days at their earliest stage. Similar work is carried out in Pa et al. (2015), in which support is provided for eight IoT architectures and connections based on the Telnet protocol. Unlike previous works, Dowling et al. (2017) propose a method for Wireless Personal Area Networks (WPANs) which enables the detection of different types of attack in the Zigbee protocol, which is part of WPAN. These types of honeypot mainly are mainly focused on detecting attacks that occur on their systems in order to detect new threats or types of attack aimed at compromising a system or violating the privacy of users. Our approach is a real VoIP system, configured to limit attackers and focused on obtaining knowledge through the attacks and actions carried out through it.

Honeypot accounts. Kaur et al. (2018) carry out a review of all the related works on the different approaches designed to combat spam and compromised accounts in social networks by analyzing each of the previous studies and discussing their pros and cons. Bursztein et al. (2014) study the manual hijacking of accounts with a focus on emails and phishing websites intended for Google users. The study focuses on how user credentials are captured by attackers and used once they access the user's account. Onaolapo et al. (2016) carry out a similar investigation with the difference that, instead of focusing on phishing attacks, they use a broader threat model by looking at user credentials in Gmail that have been automatically stolen by malware. They also analyze the behavior of cybercriminals when they gain access through leaked credentials on paste sites and forums. Other studies have focused on the study of user accounts that were under the control of spammers in different social networks. Thomas et al. (2011) investigate the abuse by spammers of the social network Twitter. To do this, they collect a dataset of all tweets from the accounts that were suspended by the tweeter to characterize the behavior of spammers in the social network. Stringhini et al. (2010) analyze the behavior of spammers in social networks using 300 honey accounts in three of the main social networks in order to develop techniques that allow detecting spammers. Unlike our work, they focus on identifying the behavior of user accounts belonging to social networks and email instead of the behavior of attackers in compromised VoIP accounts. In our study, we use "real" accounts that are obtained by attackers either through brute force attacks or from different forums or paste sites, aiming to give the impression that the accounts have been stolen.

Telephone spam and scams. Tu et al. (2016) reviewed the state of the art of the techniques used against telephone scams. They describe the ecosystem of spam calls focusing on the differences with spam in emails. Finally, they analyze the different existing solutions against telephone spam. Miramirkhani et al. (2016) propose the first study of the technical support of the scams and the call centers that are behind them. They build an automatic system to detect phone numbers and domains that are used by scammers. Finally, they make calls to the numbers of the scammers to interact with them and collect statistical details of the techniques and the process used by them. Li et al. (2018) design a set of features to feed machine learning algorithms in order to detect spam calls and scams. As they have no access to mobile telephony infrastructures, they develop a mobile application for users to label calls as malicious and thus enable them to build a dataset to analyze and extract features. Sahin et al. (2017) use a chatbot to connect unwanted calls to the bot, which mimics a person's behavior. Although the chatbot used is not based on any advanced artificial intelligence, it proves very effective, managing to keep calls for 10 min, causing spammers to waste time and resources interacting with a bot. Tu et al. (2019) conduct a study to understand why users fall into traps and end up being victims of scams. To do this, they conduct 10 telephone phishing experiments on 3000 university students. Among all the possible factors that could exist to carry out the deception successfully, they observe that the impersonation of the caller's ID prevailed in most cases. Bullée et al. (2016) conduct a study on the awareness of workers regarding the social engineering attacks they may receive, in this case over the telephone. According to the results obtained, they verify that scam awareness campaigns reduce user exposure only for a short period of time. Prasad et al. (2020) propose a method to classify calls based on the analysis of the audio and metadata of the calls, which allows them to cluster different types of audios and fraud campaigns. In general, previous work has focused on detecting scam calls when they are received and interacting with scammers either by directly calling the phone numbers recorded via the Internet or with a chatbot when receiving calls. On the contrary, our work is focused on how scammers obtain infrastructure or resources to make calls at no cost and what type of actions they perform by monitoring a real VoIP system, thus providing insights into what types of fraud are currently being carried out through compromised VoIP systems.

6. Conclusions

Telephony fraud has been a long-standing and challenging problem since the invention of the telephone. Unfortunately, with the emergence of VoIP, fraudsters now have access to a more powerful technology which allows them to carry out more sophisticated and complex attacks and to target an immense number of victims.

In this paper, we have presented a study of the behavior of criminals when they acquire VoIP user accounts, either through the use of brute force or by using accounts leaked via paste sites or through underground markets. In order to do so, we built a VoIP honeypot and established two experiments: one using accounts with weak credentials, and the other with strong credentials that were previously leaked. Then, we studied how scammers used these compromised accounts and provided a holistic overview of their activity, including provenance and intent. The different types of fraud that were attempted through our systems were identified, as well as the techniques used to increase the possible benefit that the fraudsters can obtain through these attacks.

We identify both Toll Evasion and Revenue Share as being two of the most prevalent types of fraud. This finding can be used to inform law enforcement as well as the public to deploy educational countermeasures. Furthermore, we contextualized fraudulent calls and studied key patterns such as: the time, the location and the hosting used by the fraudsters when connecting to our system. We can see that this context is valuable when deploying technical countermeasures (e.g., depleting fraudsters' resources). The results show that more than 50% of the fraudsters that interacted with the system did so through IP addresses in Europe, with the Netherlands making the highest number of outgoing calls. The UK was the most called country on the basis of the valid phone numbers that were dialed, followed by the USA. Furthermore, it was noted that a large number of the IP addresses detected belonged to hosting services, which seems to indicate that these sites were compromised by fraudsters and used to hide their identity. This assumption is also supported by the fact that, after analyzing these

IP addresses, we found that they had multiple commonly opened ports, as well as others that are less ordinary, such as the one used for the RDP protocol, which is known to have a high criticality vulnerability. In fact, the addresses were scanned for vulnerabilities, obtaining results indicating that most of them were of medium and high criticality.

With this work, we have gained a better understanding of the actions performed by attackers when they have access to a VoIP server. We have also presented a number of countermeasures to tackle this threat. We believe that our findings are key to deploying effective mitigating actions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

J. Carrillo-Mondéjar: Conceptualization, Software, Validation, Data curation, Investigation, Writing – original draft, Visualization, Writing – review & editing. **J.L. Martinez:** Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **G. Suarez-Tangil:** Conceptualization, Methodology, Investigation, Resources, Writing – original draft, Writing – review & editing, Supervision.

Acknowledgments

This work has been supported by the Ministry of Economic Affairs and Digital Transformation, Spain under project RTI2018-098156-B-C52, by the Regional Government of Castilla-La Mancha under the project SBPLY/17/180501/000353 and SBPLY/21/180501/000195, by the Spanish Education, Culture and Sports Ministry under grant FPU 17/03105, and by the “Ramon y Cajal” Fellowship RYC-2020-029401-I.

References

Anabo, I.F., Elexpuru-Albizuri, I., Villardón-Gallego, L., 2019. Revisiting the belmont report’s ethical principles in internet-mediated research: perspectives from disciplinary associations in the social sciences. Ethics Inf. Technol. 21 (2), 137–149. doi:10.1007/s10676-018-9495-z.

Antonakakis, M., April, T., Bailey, M., Bernhard, M., Bursztein, E., Cochran, J., Durumeric, Z., Halderman, J.A., Invernizzi, L., Kallitsis, M., Kumar, D., Lever, C., Ma, Z., Mason, J., Menscher, D., Seaman, C., Sullivan, N., Thomas, K., Zhou, Y., 2017. Understanding the mirai botnet. In: 26th USENIX Security Symposium (USENIX Security 17). USENIX Association, Vancouver, BC, pp. 1093–1110. https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/antonakakis

Bullée, J.-W., Montoya, L., Junger, M., Hartel, P.H., 2016. Telephone-based social engineering attacks: an experiment testing the success and time decay of an intervention. In: SG-CRC, pp. 107–114.

Bursztein, E., Benko, B., Margolis, D., Pietraszek, T., Archer, A., Aquino, A., Pitsillidis, A., Savage, S., 2014. Handcrafted fraud and extortion: manual account hijacking in the wild. In: Proceedings of the 2014 Conference on Internet Measurement Conference - IMC ’14. ACM Press, Vancouver, BC, Canada, pp. 347–358. doi:10.1145/2663716.2663749.

Chronicle. Virustotal. https://www.virustotal.com/.

Cvedetails. CVE security vulnerability database. https://www.cvedetails.com/.

Dittrich, D., Kenneally, E., et al., 2012. The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research. Technical Report. US Department of Homeland Security.

Dowling, S., Schukat, M., Melvin, H., 2017. A ZigBee honeypot to assess IoT cyberattack behaviour. In: 2017 28th Irish Signals and Systems Conference (ISSC), pp. 1–6. doi:10.1109/ISSC.2017.7983603.

Drysdale, D. Phonenumbers: Python version of Google’s common library for parsing, formatting, storing and validating international phone numbers.https://github.com/daviddrysdale/python-phonenumbers.

Endler, D., Collier, M. Hacking tools: sipscan. https://en.kali.tools/all/?tool=1253.

Getipintel. Free Proxy / VPN / TOR / Bad IP Detection Service via API and Web Interface | IP Intelligence. https://getipintel.net/.

Gauci, S. Kali tools:Sipvicious package description.

Jansky, T., Čejka, T., Bartoš, V., 2017. Hunting sip authentication attacks efficiently. In: Tuncer, D., Koch, R., Badonnel, R., Stiller, B. (Eds.), Security of Networks and Services in an All-Connected World. Springer International Publishing, Cham, pp. 125–130.

Jaquier, C. Fail2ban. https://www.fail2ban.org/.

Kambourakis, G., Kolias, C., Stavrou, A., 2017. The mirai botnet and the IoT zombie armies. In: MILCOM 2017 - 2017 IEEE Military Communications Conference (MILCOM), pp. 267–272. doi:10.1109/MILCOM.2017.8170867. ISSN: 2155-7586

Kaur, R., Singh, S., Kumar, H., 2018. Rise of spam and compromised accounts in online social networks: a state-of-the-art review of different combating approaches. J. Netw. Comput. Appl. 112, 53–88. doi:10.1016/j.jnca.2018.03.015.

Klimt, B., Yang, Y., 2004. Introducing the enron corpus. CEAS.

Li, H., Xu, X., Liu, C., Ren, T., Wu, K., Cao, X., Zhang, W., Yu, Y., Song, D., 2018. A machine learning approach to prevent malicious calls over telephony networks. In: 2018 IEEE Symposium on Security and Privacy (SP), pp. 53–69. doi:10.1109/SP.2018.00034.

McCoy, D., Pitsillidis, A., Grant, J., Weaver, N., Kreibich, C., Krebs, B., Voelker, G., Savage, S., Levchenko, K., 2012. Pharmaleaks: understanding the business of online pharmaceutical affiliate programs. In: 21st {USENIX} Security Symposium ({USENIX} Security 12), pp. 1–16.

Luo, T., Xu, Z., Jin, X., Jia, Y., Ouyang, X. IoTCandyJar: towards an intelligent-interaction honeypot for IoT Devices, Black Hat, 2017.

Microsoft, 2019. CVE-2019-0708 | Remote Desktop Services Remote Code Execution Vulnerability. https://portal.msrc.microsoft.com/en-US/security-guidance/advisory/CVE-2019-0708.

Mitre corporation, T., 2019. CVE common vulnerabilities and exposures. https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2019-0708.

Miessler, D., Haddix, J. g0tmi1k., danielmiessler/SecLists. https://github.com/danielmiessler/SecLists.

Miracle, V.A., 2016. The belmont report: the triple crown of research ethics. Dimens. Crit. Care Nurs. 35 (4), 223–228.

Miramirkhani, N., Starov, O., Nikiforakis, N., 2016. Dial one for scam: analyzing and detecting technical support scams. CoRR abs/1607.06891607.06891. http://arxiv.org/abs/1607.06891.

Nawrocki, M., Wählisch, M., Schmidt, T. C., Keil, C., Schönfelder, J., 2016. A survey on honeypot software and data analysis. CoRR abs/1608.06249http://arxiv.org/abs/1608.06249.

Ofcom. Consumers warned about 070 missed call scam.

Onaolapo, J., Mariconti, E., Stringhini, G., 2016. What happens after you are pwnd: understanding the use of leaked webmail credentials in the wild. In: Proceedings of the 2016 Internet Measurement Conference. ACM, New York, NY, USA, pp. 65–79. doi:10.1145/2987443.2987475. Event-place: Santa Monica, California, USA

Osenbaugh, J. Telecom fraud on the rise: what enterprises need to know.

Pa, Y.M.P., Suzuki, S., Yoshioka, K., Matsumoto, T., Kasama, T., Rossow, C., 2015. Iotpot: analysing the rise of IoT compromises. 9th USENIX Workshop on Offensive Technologies (WOOT 15). USENIX Association, Washington, D.C.. https://www.usenix.org/conference/woot15/workshop-program/presentation/pa

Prasad, S., Bouma-Sims, E., Mylappan, A.K., Reaves, B., 2020. Who’s calling? Characterizing robocalls through audio and metadata analysis. In: 29th USENIX Security Symposium (USENIX Security 20). USENIX Association, pp. 397–414. https://www.usenix.org/conference/usenixsecurity20/presentation/prasad

Proffitt, C, Wolf, D. Bluekeep:Havoc on the horizon.

Project, T. K. S. S. Kamailio sip server. https://www.kamailio.org/w/.

Sahin, M., Francillon, A., Gupta, P., Ahamad, M., 2017. Sok: fraud in telephony networks. In: 2017 IEEE European Symposium on Security and Privacy (EuroS P), pp. 235–250.

Sahin, M., Relieu, M., Francillon, A., 2017. Using chatbots against voice spam: analyzing lenny’s effectiveness. In: Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017). USENIX Association, Santa Clara, CA, pp. 319–337. https://www.usenix.org/conference/soups2017/technical-sessions/presentation/sahin

Sangoma Technologies, 2021. Open Source Communications Software | Asterisk Official Site. https://www.asterisk.org/home.

Shodan. Shodan is the world’s first search engine for internet-connected devices. https://www.shodan.io/.

s.r.o, C. H. Multiple vulnerabilities in microsoft windows SMB server. https://www.cybersecurity-help.cz/vdb/SB2017031416.

Stone-Gross, B., Stevens, R., Zarras, A., Kemmerer, R., Kruegel, C., Vigna, G., 2011. Understanding fraudulent activities in online ad exchanges. In: Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference, pp. 279–294.

Strand, L., Leister, W., 2011. Improving SIP authentication.

Stringhini, G., Egele, M., Zarras, A., Holz, T., Kruegel, C., Vigna, G., 2012. B@bel: leveraging email delivery for spam mitigation. In: Presented as Part of the 21st {USENIX} Security Symposium ({USENIX} Security 12), pp. 16–32.

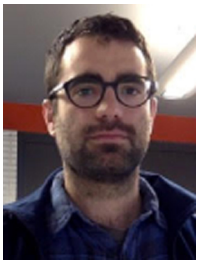
Stringhini, G., Kruegel, C., Vigna, G., 2010. Detecting spammers on Social Networks. In: Proceedings of the 26th Annual Computer Security Applications Conference. ACM, New York, NY, USA, pp. 1–9. doi:10.1145/1920261.1920263. Event-place: Austin, Texas, USA

Thomas, K., Grier, C., Song, D., Paxson, V., 2011. Suspended accounts in retrospect: an analysis of twitter spam. In: Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference. ACM, New York, NY, USA, pp. 243–258. doi:10.1145/2068816.2068840. Event-place: Berlin, Germany

Thomas, K., Huang, D., Wang, D., Bursztein, E., Grier, C., Holt, T.J., Kruegel, C., McCoy, D., Savage, S., Vigna, G., 2015. Framing dependencies introduced by underground commoditization. Workshop on the Economics of Information Security.
Tu, H., Doupé, A., Zhao, Z., Ahn, G., 2016. Sok: everyone hates robocalls: a survey of techniques against telephone spam. In: 2016 IEEE Symposium on Security and Privacy (SP), pp. 320–338. doi:10.1109/SP.2016.27.
Tu, H., Doupé, A., Zhao, Z., Ahn, G.-J., 2019. Users really do answer telephone scams. In: 28th USENIX Security Symposium (USENIX Security 19). USENIX Association, Santa Clara, CA, pp. 1327–1340. https://www.usenix.org/conference/usenixsecurity19/presentation/tu
Vetterl, A., Clayton, R., 2018. Bitter harvest: systematically fingerprinting low- and medium-interaction honeypots at internet scale. 12th USENIX Workshop on Offensive Technologies (WOOT 18). USENIX Association, Baltimore, MD. https://www.usenix.org/conference/woot18/presentation/vetterl
Wikipedia. Pager. https://en.wikipedia.org/wiki/Pager.

Javier Carrillo-Mondéjar received a B.Sc. degree in Computer Science and a Master's degree in Advanced Computer Science from the University of Castilla-La Mancha, Spain, in 2016 and 2017, respectively. Currently, he is enrolled full-time on the Ph.D. Program in Advanced Information Technology at this university. In 2016, he joined the Computer Architecture and Technology Group of the Informatics Research Institute of Albacete (I3A) as a researcher. His research interests are related to malware detection and classification techniques, as well as the methods used in malware to spread and remain hidden in computer systems. He has also been a visiting researcher at King's College London and the University of Jyväskylä.

Jose Luis Martínez received his M.Sc. and Ph.D. degrees in Computer Science and Engineering from the University of Castilla-La Mancha (Spain) in 2007 and 2009, respectively. In 2005, he joined the Department of Computer Engineering at the University of Castilla-La Mancha, where he was a researcher in the Computer Architecture and Technology group at the Albacete Research Institute of Informatics (I3A). In 2010, he joined the department of Computer Architecture at the Complutense University in Madrid, where he was an assistant lecturer. In 2011, he rejoined the Department of Computer Engineering of the University of Castilla-La Mancha, where he is currently a full professor. His research interests include video coding and transcoding, and topics related to security. He has also been a visiting researcher at the Florida Atlantic University, Boca Raton (USA), and the Centre for Communication System Research (CCSR), at the University of Surrey, Guildford (UK). He has over 100 publications in these areas in international refereed journals and conference proceedings.



Guillermo Suarez-Tangil is with IMDEA Networks Institute. His research focuses on systems security and malware analysis and detection. In particular, his area of expertise lies in the study of smart malware, ranging from the detection of advanced obfuscated malware to the automated analysis of targeted malware. Previously, he was Lecturer (Assistant Professor) at King's College London (KCL). Before joining KCL, he was senior research associate at University College London (UCL), where he was also actively involved in other research areas involved with detecting and preventing Mass-Marketing Fraud.