VB2021
localhost

7 - 8 October, 2021 / vblocalhost.com

RANSOMWARE: A CORRELATION BETWEEN INFECTION VECTORS AND VICTIMS

Doina Cosovan, Cătălin Liță, Jue Mo & Ryan Sherstobitoff
SecurityScorecard, Romania & USA

dcosovan@securityscorecard.io
cvaleriu@securityscorecard.io
jmo@securityscorecard.io
rsherstobitoff@securityscorecard.io

ABSTRACT

Ransomware attacks have increased exponentially recently. Some companies have even started to take out insurance against them.

Nowadays it is not as easy as it was in the past to hide the fact that you've been breached, especially if the breach is a result of a ransomware infection or leads to a ransomware infection.

This is because the attackers behind more than 20 different ransomware families have started to threaten to publicly expose the data belonging to companies unwilling to pay the ransom. Most of the attackers use Tor domains to disclose the identity of the companies they've infected, as well as to upload files they've stolen before starting the encryption process.

In this paper we analyse the techniques that lead to companies being infected with ransomware, in an attempt to find a way to figure out if an entity is a potential future ransomware victim and what it can do to minimize the chances of being hit.

Ransomware infects systems through other malware families or exploit kits, vulnerable services, spam campaigns, and so on. By correlating the victims of these malware families and exploit kits, the entities running vulnerable services, as well as the entities that have poor email hygiene, with the victims of ransomware attacks, we can estimate the risk those exposures added to the probability of ransomware infection.

There are two ways of finding victims of ransomware attacks. Non-paying victims can be found by crawling the Tor websites maintained by the attackers, while both paying and non-paying victims can be found by sinkholing ransomware families that use multiple command-and-control domains and don't register all of them.

For the ransomware families for which we can find both paying and non-paying victims as a result of having information from both the attacker's website and our sinkholes, we can determine the percentage of paying victims.

Even if a company doesn't get infected with ransomware itself, a third-party entity, such as a supplier, can get infected with ransomware, thus allowing the attackers access to the same data the company shared with that third party. The initial company can later be blackmailed, using the threat of making that data public to competitors – as happened with *Apple* recently. Therefore, in order to protect themselves, it is also important for companies to monitor their third-party entities for how vulnerable they are.

1. INTRODUCTION

1.1 Ransomware leaks websites

In 2019, the SunCrypt ransomware started to exfiltrate the data of infected companies before encrypting their files so that they could blackmail the companies into paying the ransom even if they had a backup of their data. Since then, the majority of ransomware families have started to follow the same practice.

Figure 1 illustrates the evolution of the number of companies per month for which data was leaked publicly during the period of January 2020 to May 2021 across 37 ransomware families (AKO, Astro Team, Avaddon, BABUK LOCKER, CL0P, Conti, Cuba, DarkSide, DoppelPaymer, Egregor, Everest, Grief, LV, LockBit, Lorenz, MAZE, Marketo, Mount Locker, N3tw0rm, NEMTY, Nefilim, NetWalker, Noname, Pay2Key, Payload.bin, Prometheus, Pysa, Ragnar_Locker, Ragnarok, RansomEXX, Ranzy Locker, Sekhmet, Sodinokibi / REvil, SunCrypt, SynACK, Team Snatch and XING LOCKER).

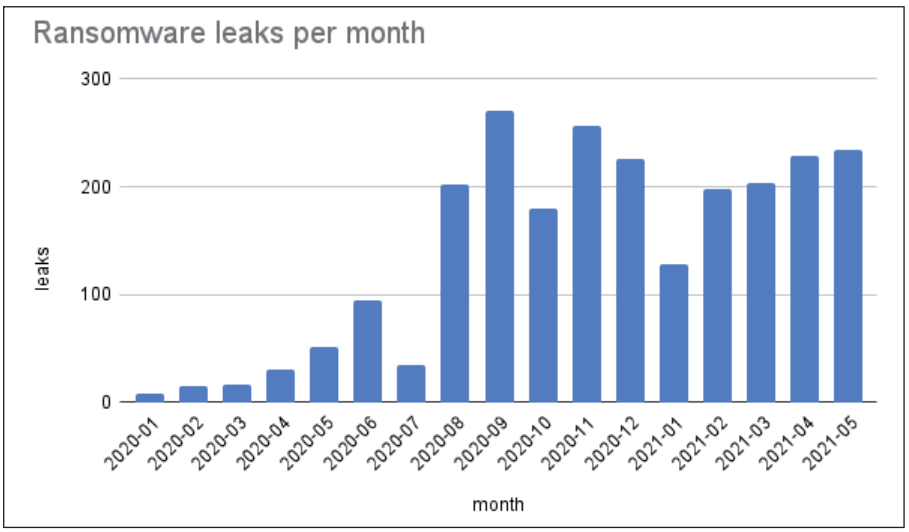

Figure 1: Number of companies per month whose data was leaked during January 2020 to May 2021.

The peaks are due to ransomware families appearing or starting to employ this practice, while the lows are due to some malware families retiring or rebranding. However, from February 2021 the number of leaks stabilized and has slowly but steadily been increasing. The number of companies whose data was leaked is only a part of the total number of companies infected, as it doesn't include the companies that paid the ransom and thus avoided being listed on the leaks websites.

Since we will be focusing on the ransomware families that have been active in the past three months, Figure 2 shows the number of leaks posted during March, April and May 2021 for the top 10 ransomware families ordered by the number of leaked companies during the three-month period. From this image it can be seen that some ransomware families, like Conti, Avaddon and Prometheus, seem to show an increase in activity, while others, like Cl0p, show a decrease in activity over this period.

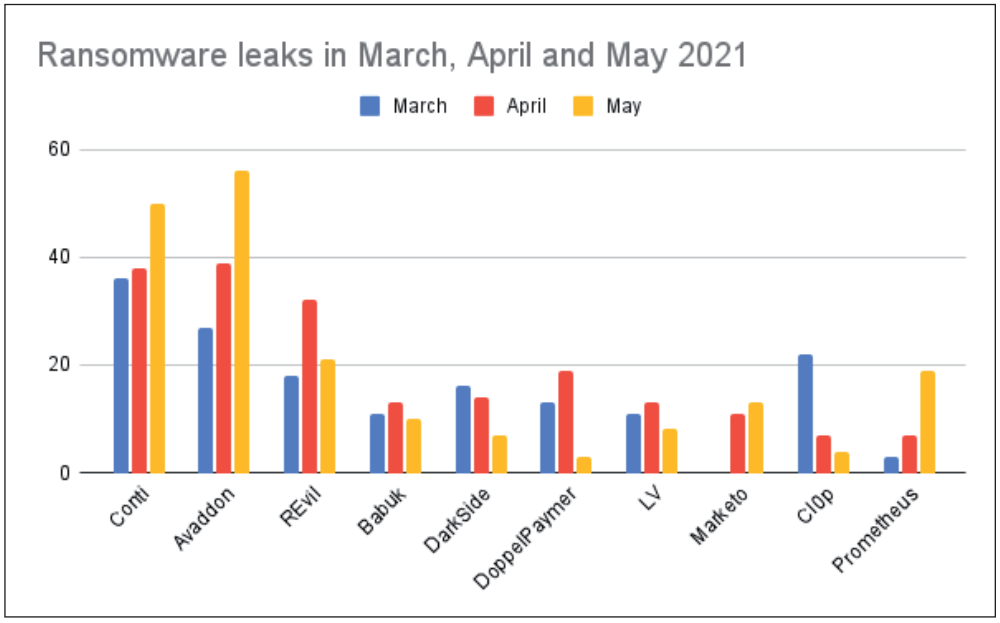

Figure 2: Number of companies whose data was leaked by top 10 most prolific ransomware families during March, April and May, 2021.

1.2 Ransomware infection process

A ransomware infection consists of multiple phases:

1. Initial access.
2. Local operations such as privilege escalation, potentially credential harvesting, and lateral movement.
3. Data exfiltration to a location under the attackers' control.
4. Data encryption, which can be followed by making the exfiltrated data publicly available by uploading the exfiltrated data on a leaks website if the victim is not willing to pay the ransom.

In this paper, we want to shed light on the most common ways in which these activities unfold.

2. DATA GATHERING

We collected 748 companies for which data was leaked on ransomware websites across 25 ransomware families (Astro Team, Avaddon, BABUK LOCKER, CL0P, Conti, DarkSide, DoppelPaymer, Everest, Grief, LV, Lorenz, Marketo, Mount Locker, N3tw0rm, Nefilim, Noname, Payload.bin, Prometheus, Pysa, Ragnar_Locker, Ragnarok, RansomEXX, Sodinokibi (REvil), SynACK, XING LOCKER) since 1 March 2021.

For the companies for which a domain was not specified, we added it manually. Then, we obtained the digital footprint of those domains. By 'digital footprint' we mean all the IP addresses we managed to collect that belong to the corresponding company. The size of the affected companies in terms of their digital footprint was rather diverse. Figure 3 illustrates the number of companies for each range of the number of IP addresses associated with them.

Throughout this paper, we analyse the data for all 748 companies for the period 1 March 2021 to 31 May 2021. As an exception, we decided to look into NetFlow data only for the 600 companies with the smallest digital footprint (fewer than 1,000 IPs).

As data sources we used NetFlow, sinkholes, spam traps, scanning data, and various leak sources.

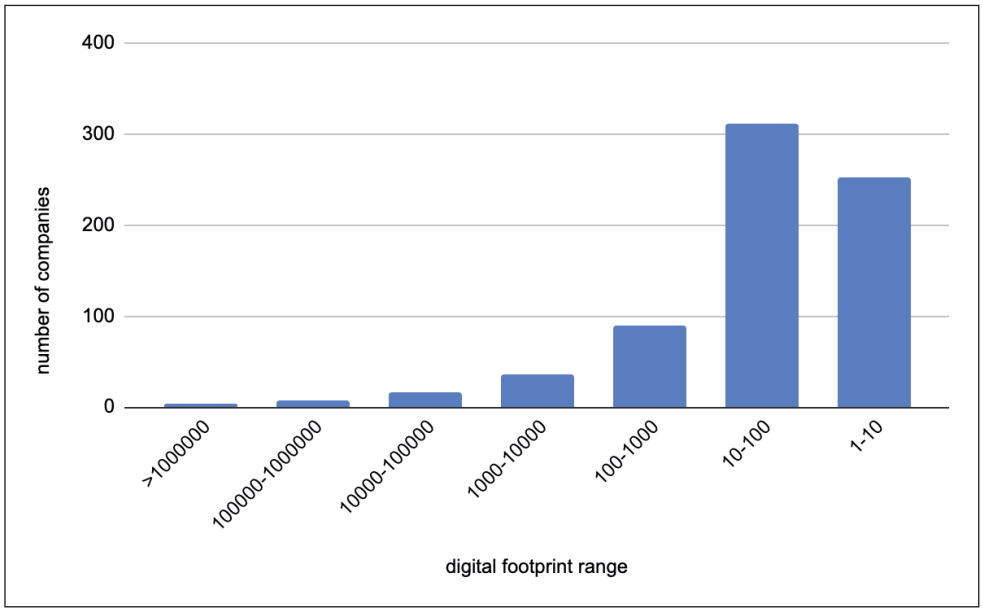

Figure 3: Number of companies per range of number of IP addresses.

3. INITIAL ACCESS

3.1 Exposed services

Looking at our scanning results, we observed exposed services in 258 companies. Figure 4 shows the top 10 exposed services sorted by the number of leaked companies exposing them. The Remote Desktop Protocol (RDP), Virtual Network Computing (VNC) and Server Message Block (SMB) services, known to be abused by ransomware groups, are among the top 10. We can also see in the top 10 the presence of SQL databases (*MySQL*, *Postgresql*, *Microsoft SQL*), email services (POP3, IMAP) and file services (FTP).

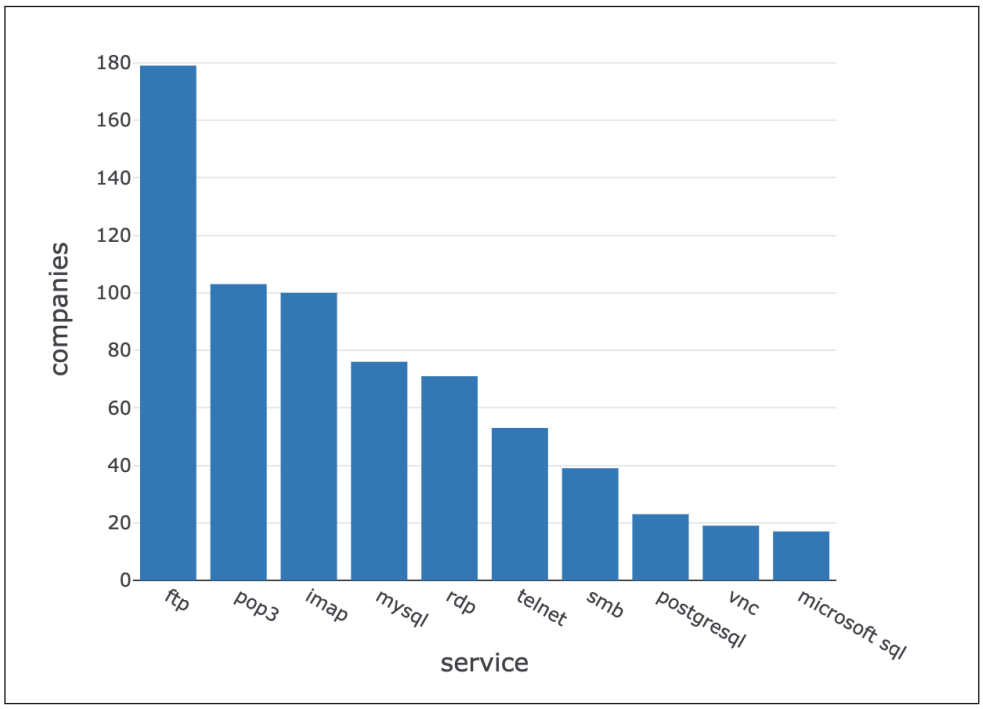

Figure 4: Top 10 exposed services by number of leaked companies.

However, being exposed doesn't necessarily mean being abused. Therefore, we decided to look into the NetFlow data for the IP addresses corresponding to these companies in order to spot which of the services were actually abused.

In order to discover the abused services, we first computed the number of flows for each pair (server IP address, server port) during each hour. Then we computed the 99th percentile and the maximum value for the number of hourly rows. We considered a service to be abused if the maximum value was five times the value of the 99th percentile.

Figure 5 shows the top 15 abused ports. Some services such as HTTPS (443), HTTP (80), SSH (22) and FTP (21) might appear among the top abused ports because they are frequently used for data exfiltration. The services HTTPS (443) and HTTP (80) could also be abused by attackers in order to discover and take advantage of service or application vulnerabilities.

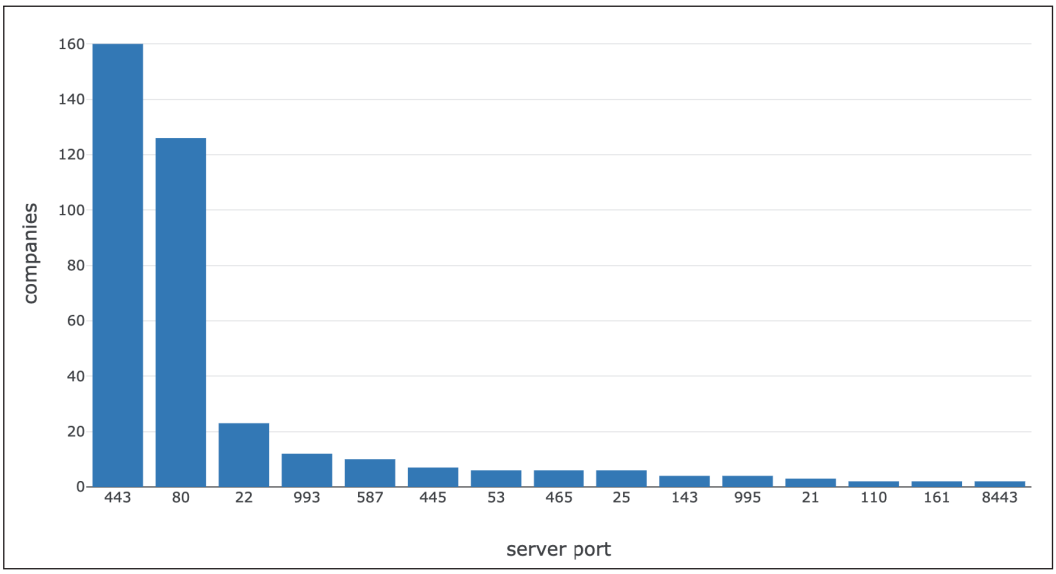

Figure 5: Top 15 abused services per number of leaked companies.

When looking at some particular cases for port 443, we observed multiple cases in which the spike in the number of flows for port 443 of IP addresses belonging to leaked companies happened shortly after the leak was made public and was accompanied by a spike in the amount of transferred data. We have a few theories about this. If the server's files are infected, the website can be traversed as an open directory. This means either that the company downloaded the encrypted files from the server in this way or that a third party, learning about the leak, downloaded the encrypted data afterwards in the hope that they would be able to decrypt the data at a later date. We would like to mention that the IP addresses downloading the data were not associated with the given company but with the *Rackspace Cloud*. Since in this case we are talking about victims of Avaddon, which closed down recently and made its decryption keys publicly available, the potential third party got lucky.

One such case is illustrated by Figure 6, showing the hourly number of flows, and Figure 7, showing the hourly amount of exchanged data in GB. Another theory is that the attackers wanted to exfiltrate as much data as possible after they realized that they were not going to get paid.

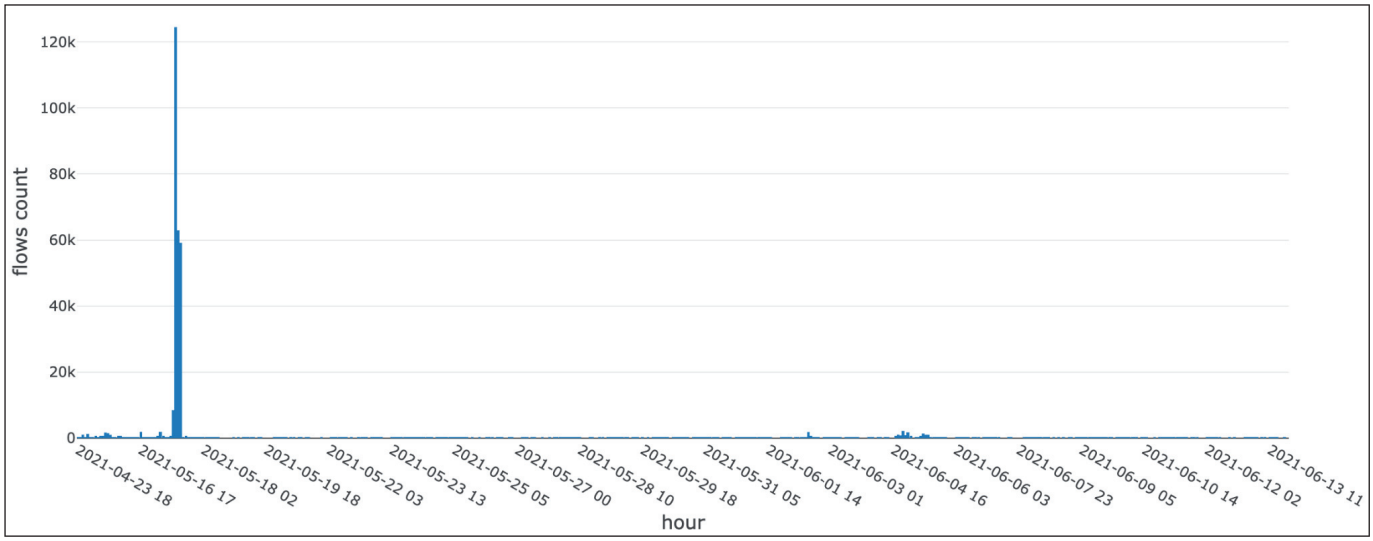

Figure 6: Hourly number of flows on port 443 for a post-leak data exfiltration.

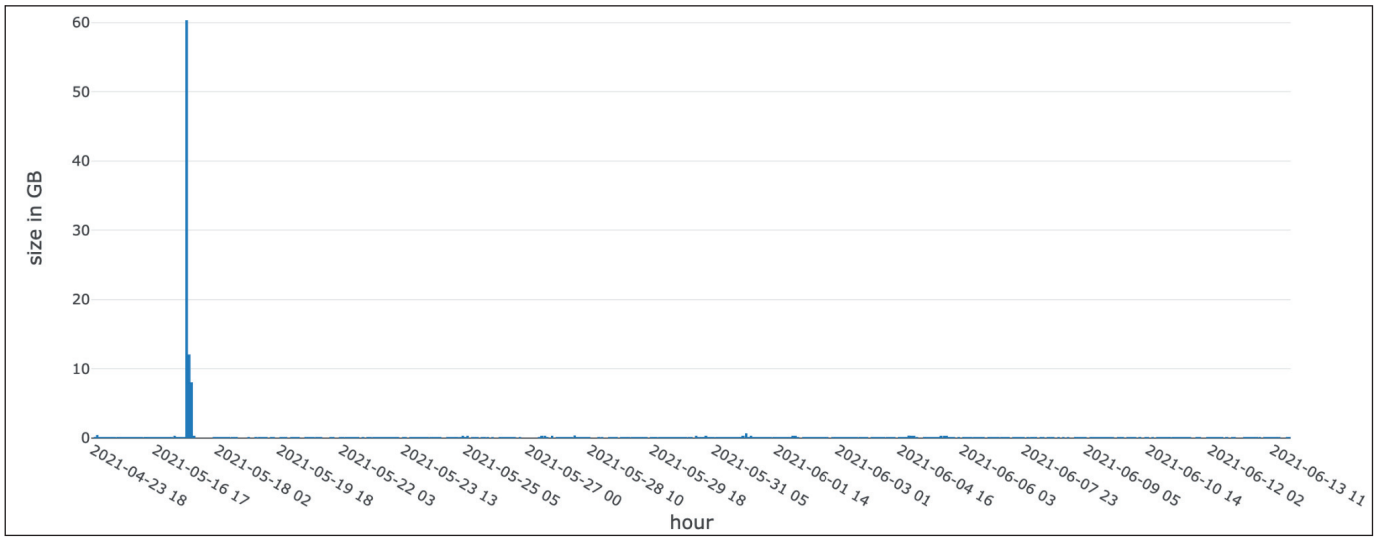

Figure 7: Hourly size of flows in GB on port 443 for a post-leak data exfiltration.

We observed some cases with brute forcing and exfiltration over RDP. One such case is a company infected with the LV ransomware. The point of entry was RDP. Figure 8 shows the number of RDP connections per hour, and we can see thousands of connections per hour. The brute forcing happened between 22 and 30 March, with a short break on 29 March. After the brute force, the attacker started the exfiltration, also over RDP, on 6 April. There were at least nine RDP connections with 4GB of exchanged data each until 31 May, as can be seen in Figure 9. We observed a similar behaviour on two other IP addresses belonging to the same company. The leak was posted on the website in June.

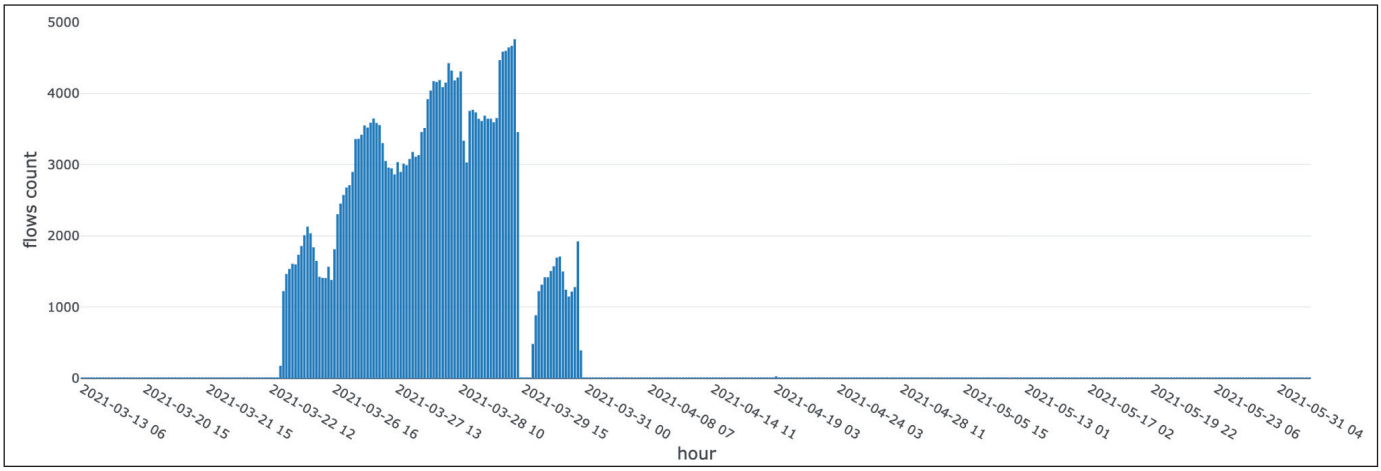

Figure 8: RDP scanning of an LV victim.

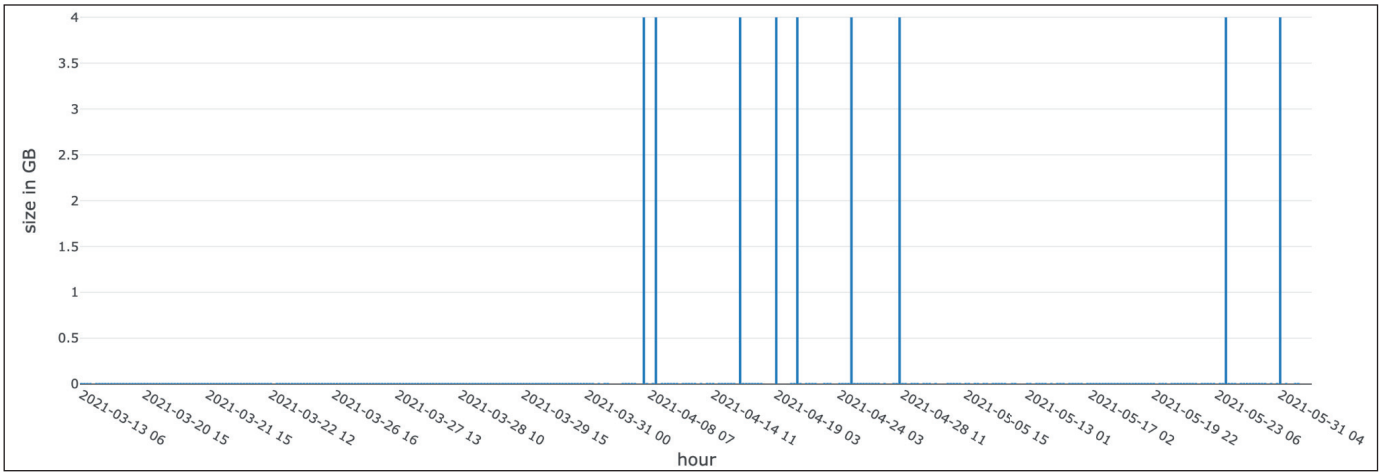

Figure 9: Data exfiltration via RDP of an LV victim.

We found a very similar situation for a Grief victim. Interestingly, the times of the RDP scanning and the times of the data exfiltration were approximately the same and the date when the company data was leaked matched in those two cases. The RDP brute forcing and data exfiltration for the Grief victim is illustrated in Figures 10 and 11, respectively.

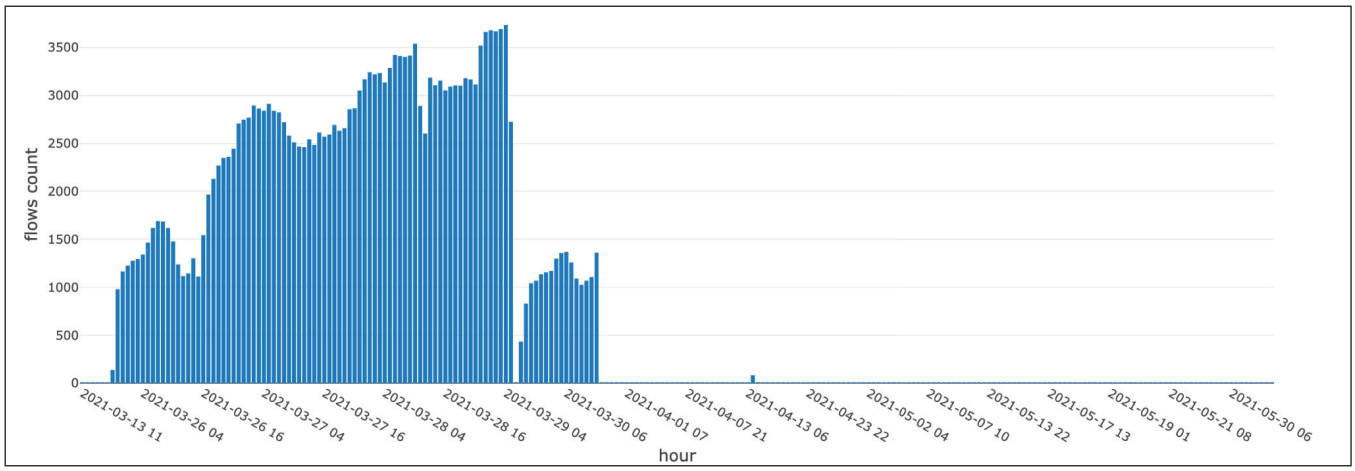

Figure 10: RDP scanning of a Grief victim.

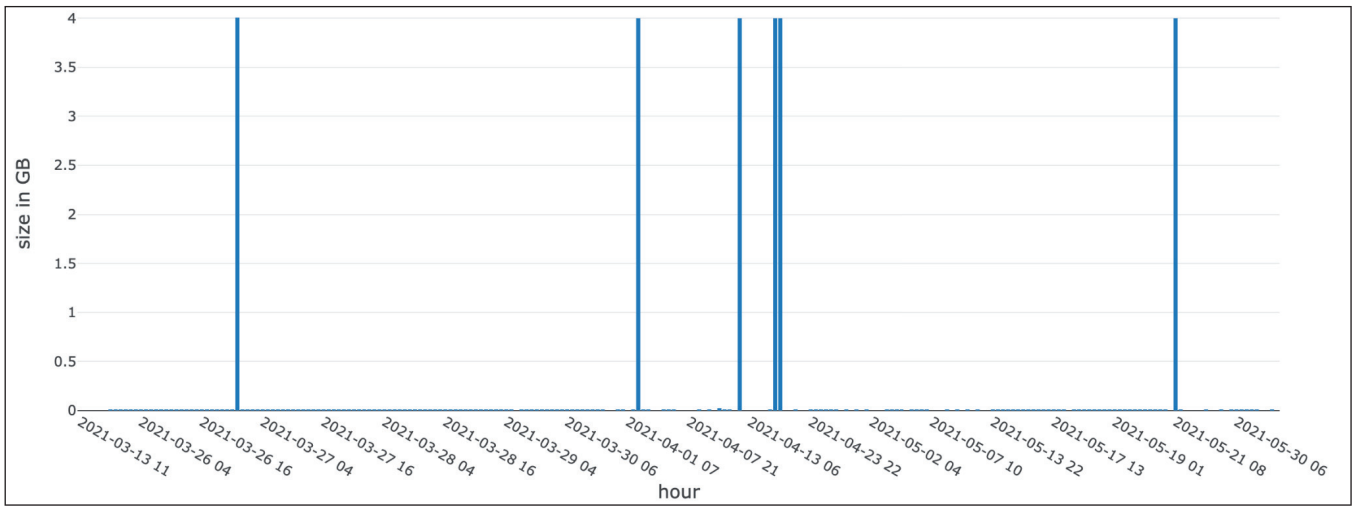

Figure 11: Data exfiltration via RDP of a Grief victim.

3.2 Vulnerabilities

1,947 different Common Vulnerabilities and Exposures (CVEs) were discovered across the leaked companies and 248 companies had at least one CVE.

Furthermore, 555 of the CVEs have publicly or commercially available proof-of-concepts (POCs). Figure 12 shows the top 15 CVEs with POCs, ordered by the number of companies vulnerable to them.

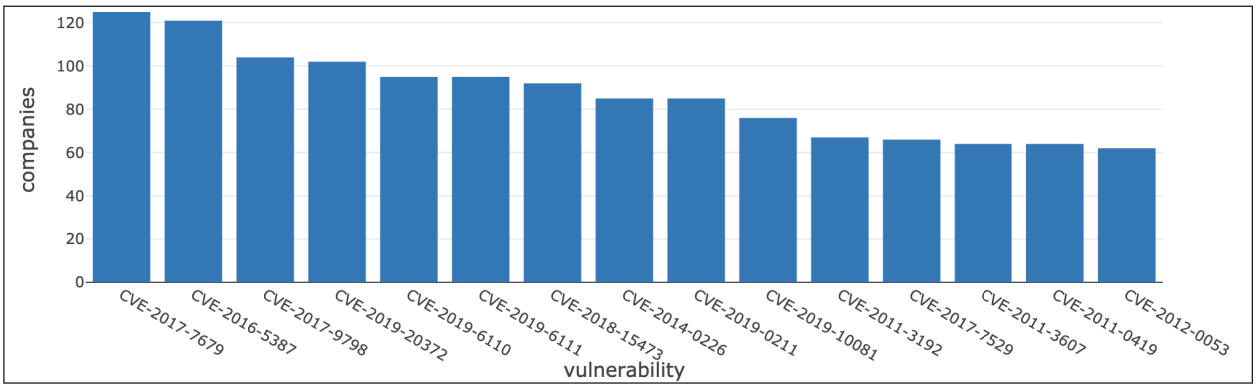

Figure 12: Top 15 CVEs with POCs by number of leaked companies.

These are mostly vulnerabilities for *Apache*, *NGINX* and *OpenSSH*. They allow the trackers to read unauthorized content (CVE-2017-9798, CVE-2019-20372), overwrite files such as SSH's credentials file (CVE-2019-6111), perform SSH user enumeration (CVE-2018-15473), execute arbitrary code with root privilege (CVE-2019-0211, CVE-2011-3607), obtain sensitive credential information (CVE-2017-7529, CVE-2014-0226), or execute arbitrary code (CVE-2014-0226).

29 of the CVEs are remote code execution (RCE) vulnerabilities. Figure 13 shows the top 15 RCE vulnerabilities sorted by the number of leaked companies having them.

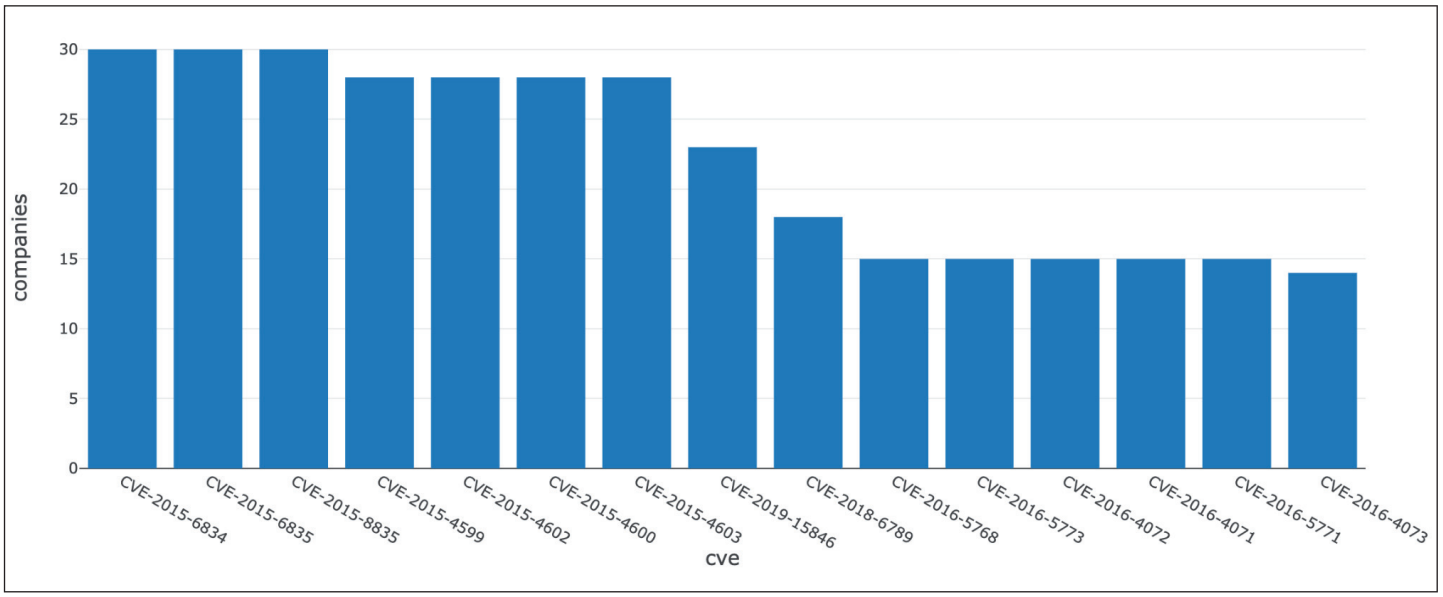

Figure 13: Top 15 RCE CVEs by number of leaked companies.

However, we note that this is the current state and that prior to the ransomware infection and leak, the companies could have had even more vulnerabilities.

3.3 Spam

Like other types of malware, ransomware can send spam in order to harvest credentials or install malicious software. Poor email hygiene leads to an increased risk of getting infected in this way. We decided to investigate how susceptible the ransomware-infected companies were to this infection vector by looking at two things:

- Company mail flow best practices
- Availability of company email addresses to spammers

The path the email takes from the Internet to a mailbox and vice versa is called mail flow. With regard to the company mail flow best practices, we decided to look at the Sender Policy Framework (SPF) records of the companies. SPF is a mechanism used to prevent spoofing of sender email addresses. We argue that companies that do not have SPF records or that have SPF records that are too permissive are more susceptible to phishing campaigns because their employees can receive emails that appear to be coming from within their company.

We discovered that 377 companies (50%) had SPF issues. Figure 14 shows the SPF issues sorted by the number of companies having them, with the issues being as follows:

- SPF record softfail – the company has an SPF record, but it soft fails, meaning that the host will mark the email as an SPF failure, but will accept it.
- SPF record missing – the company doesn't have an SPF record.
- SPF record malformed – the company has an SPF record, but it is malformed.
- SPF record wildcard – the company has an SPF record, but it contains wildcards.

In order to check the availability of the companies' email addresses to spammers, we looked into our spam traps to see if email addresses for the company domains appear in spam. We observed 13 companies in this situation in the past month alone: email addresses for 10 of the companies appear as senders, while email addresses for three of the companies appear as recipients. The emails contain phishing, scams or malware. The emails received from company addresses are

impersonations, taking advantage of SPF-related issues. For example, one of the companies was impersonated to send mails with malware attachments such as Agent Tesla and Snake Keylogger.

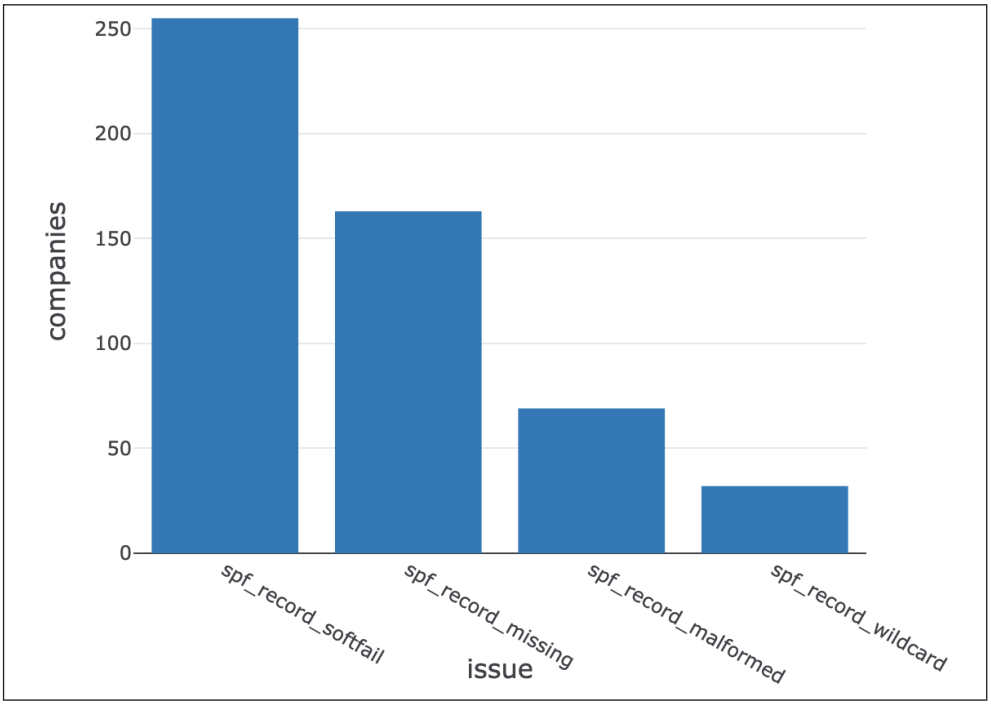

Figure 14: Number of leaked companies per SPF issue type.

3.4 Leaked credentials

Figure 15 shows the top 10 leak sources sorted by the number of companies that were victims of ransomware. We found leaked records for 528 (70%) of the companies.

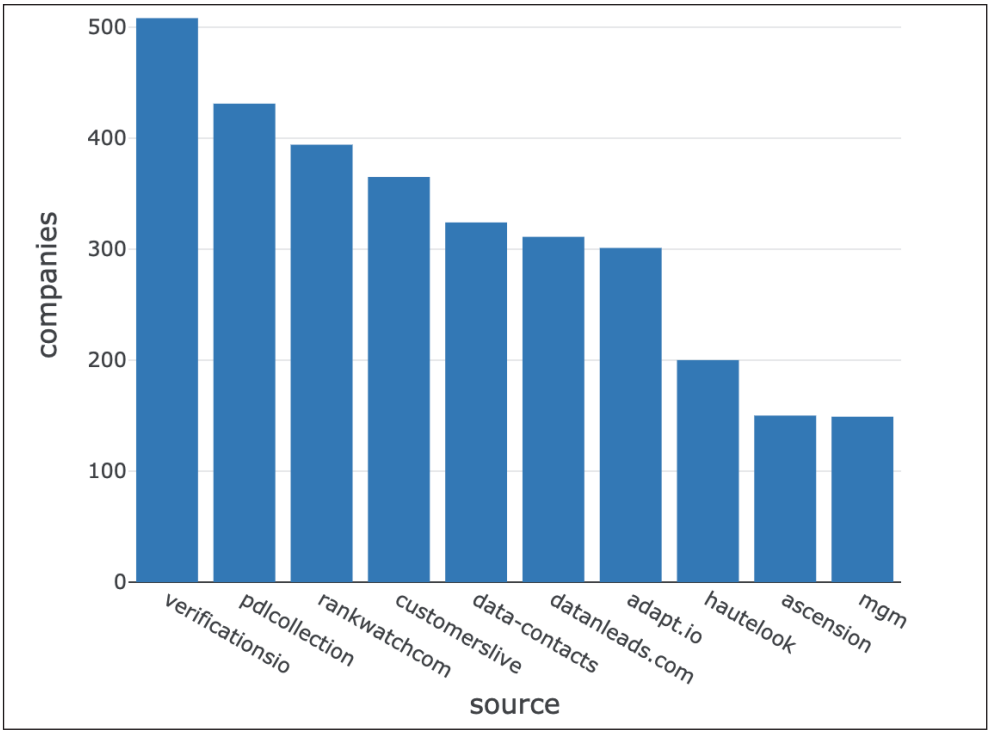

Figure 15: Number of leaked companies per leak source.

As many as 508 companies were found in the leaks from *Verifications.io* [1]. *Verifications.io* is a service used to check the validity of email addresses. A user can submit a list of email addresses they want to check and *Verifications.io* will

send emails to those addresses and check if it receives back an error. In February 2019, a *MongoDB* database of 150GB of data was leaked, containing approximately 800 million email addresses, which were correlated with other personally identifiable information such as ZIP code, phone, gender, date of birth, and user IP address. Since 508 of the 748 (68%) ransomware-infected companies had email addresses in that database, the attackers could be checking them using this service before sending phishing emails to targeted companies. However, given the number of email addresses the database contains, it might be expected to find such a big percentage of the leaked companies in there regardless.

A large number of companies (431 out of 748, amounting to 58%) were also found in the *pdlcollection* leak from *People Data Labs* (*PDL*), a data broker which mostly has data scraped from *LinkedIn*. This seems like a good source for attackers to find email addresses of the employees of targeted companies.

RankWatch is an SEO marketing platform whose *MongoDB* database was leaked, exposing approximately 40 million email addresses and other information. 394 companies (53%) were found there.

The other top leak sources are:

- *Customerslive* – a marketing-related database.
- *Data-contacts* – an exposed *ElasticSearch* instance, possibly related to the medical field.
- *Datanleads.com* [2] – a website that sells data on company and people.
- *Adapt.io* [3] – a website that provides a business contact database.
- *Hautelook* [4] – the fashion shopping site *HauteLook* was breached in 2018, exposing over 28 million unique email addresses alongside names, genders, dates of birth and passwords stored as bcrypt hashes.
- *Ascension* – an *ElasticSearch* database exposing millions of financial and banking documents.
- *MGM* – a breach containing email and physical addresses, names, phone numbers and dates of birth with 3.1 million unique email addresses.

3.5 Application security

577 companies were found to have issues related to web application security. The top 10 most prevalent issues among the leaked companies are listed in Figure 16, where the names indicate:

- csp no policy – there is no content security policy (CSP), making the websites susceptible to cross site scripting (XSS) and data injection attacks.
- hsts incorrect – incorrect HTTP Strict Transport Security (HSTS), which tells the browser to only contact a website through HTTPS, not HTTP.
- xss incorrect – incorrect value for X-XSS-Protection response header, which protects against reflected XSS attacks.
- content type incorrect – incorrect value for X-Content-Type-Options response header, which indicates that MIME types should not be changed and thus prevents MIME type sniffing.
- frame opts incorrect – incorrect values for X-Frame-Options response header, which indicates whether or not a browser should be allowed to render a page in a <frame>, <iframe>, <embed> or <object>, thus useful in avoiding clickjacking attacks by ensuring that their content is not embedded into other sites.
- unsafe sri – SubResource Integrity (SRI) enables browsers to check that resources are delivered without unexpected manipulation.
- domain missing https – the domain doesn't have HTTPS.
- insecure https redirect – site redirects to a domain in a way that limits the security provided by HTTPS and HTTP Strict Transport Security (HSTS) headers, which leaves users vulnerable to being redirected to a spoofed / malicious version of the site.
- http in redirect chain – an HTTP URL is present in the redirect chain, which leaves users vulnerable to being redirected to a spoofed / malicious version of the intended destination site.
- csp too broad – has a CSP policy but it is too permissive.

These issues allow attackers to perform XSS, clickjacking, MIME-type sniffing, resource manipulation, and man-in-the-middle attacks.

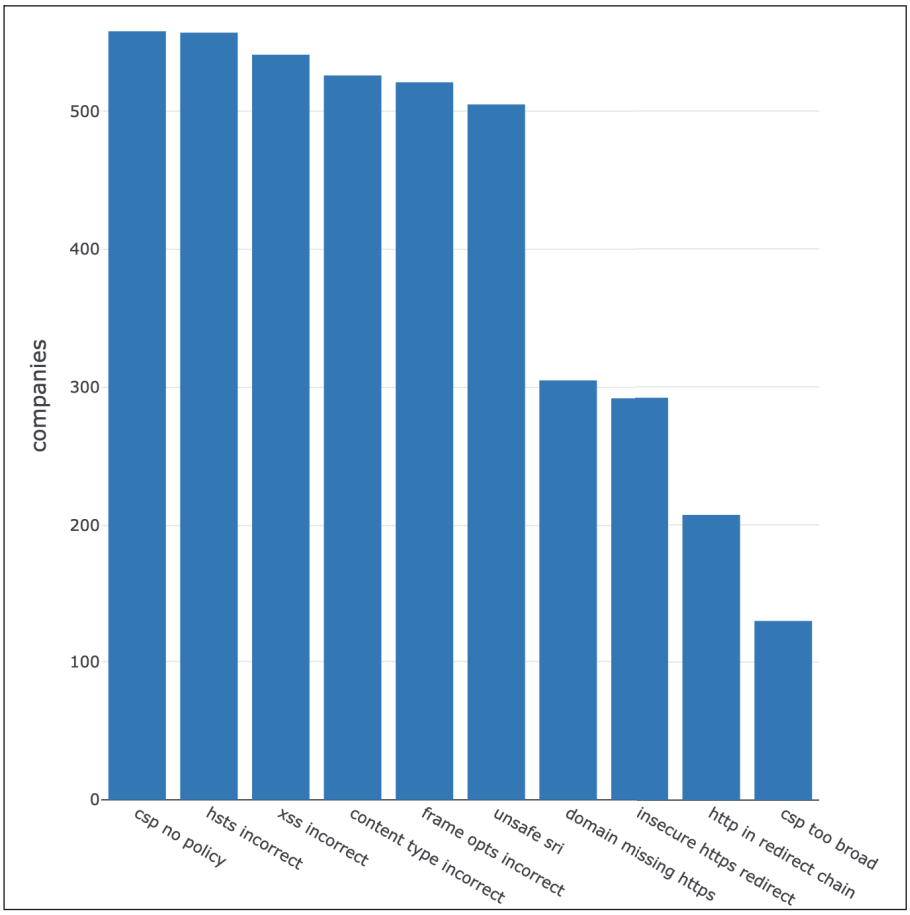

Figure 16: Number of leaked companies per application security issue.

3.6 Malware

We first looked in our sinkholes for requests coming from infected systems belonging to the analysed companies. Figure 17 shows the top 10 malware families ordered by the number of infected leaked companies. It illustrates that some companies still have infections with old malware families, such as Conficker, which is more than 10 years old.

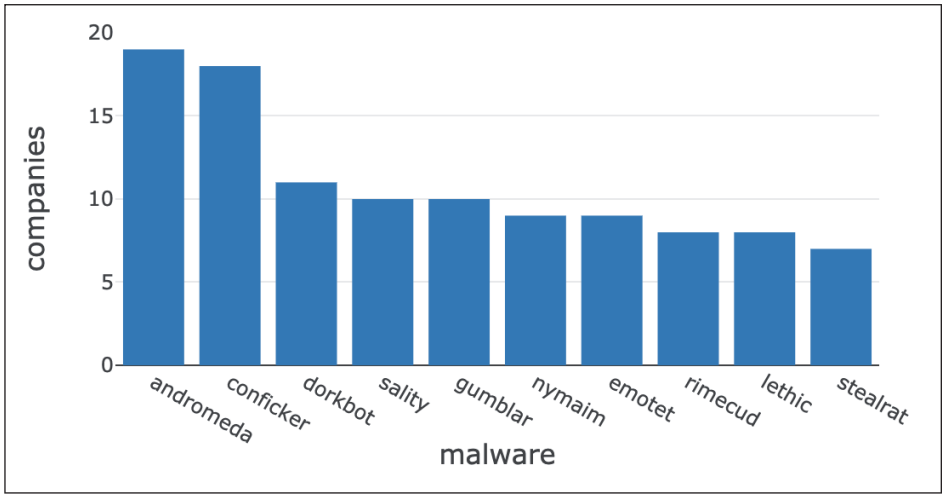

Figure 17: Top 10 malware infections from sinkhole sorted by number of leaked companies.

However, we decided to go a step further. We knew that in OSINT there are reported cases where particular malware families are installing ransomware. The following associations have been reported:

- Trickbot – Conti [5]
- Emotet – Ryuk [6]

- Dridex – DoppelPaymer [7]
- Qakbot – Egregor [8]
- Icedid – Egregor [9], [10]
- Ursnif – Egregor [10]
- Phorpiex – Avaddon [11]
- Bazarloader – Conti [12]
- SDBot – Cl0p [13]
- Buer – Ryuk [14]
- Pyxie – RansomExx [15]
- Vatet – RansomExx [15]

We decided to check the prevalence of these cases among the leaked companies. In order to do so, we collected command-and-control servers for these malware families.

We collected the command-and-control servers from extracted malware configurations, from Internet scanning, and from behavioural reports. Some malware families use IP addresses (Trickbot, Emotet, Dridex, Qakbot), while others use domains (Phorpiex, Buer). In the case of domains, we resolved them to IP addresses. Then we searched for indications of communication between these command-and-control servers and the IP addresses belonging to the leaked companies in NetFlow data. Figure 18 shows our findings.

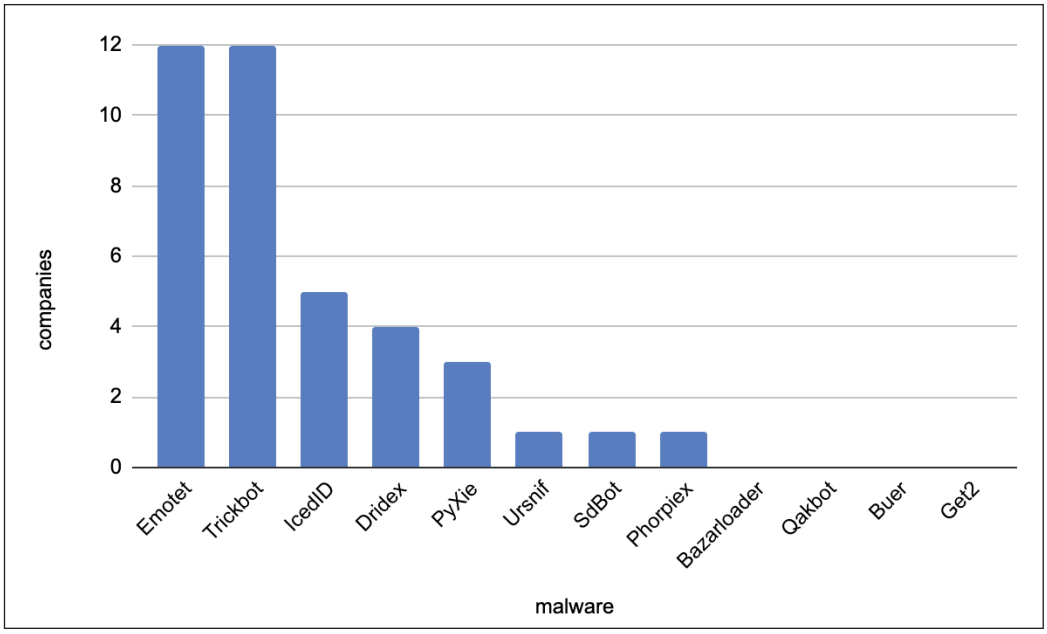

Figure 18: Number of leaked companies per malware associated with ransomware.

4. LOCAL OPERATIONS

4.1 Lateral movement – CobaltStrike

Since CobaltStrike is known to be particularly popular at the moment for lateral movement, we decided to apply the same technique for it.

In this particular case we stumbled upon IP addresses hosted on shared infrastructure. The problem is that, even if the server might be a command-and-control server for CobaltStrike, since other, perhaps legitimate domains resolve to it too, we can't differentiate in the NetFlow records which domain the client IP address is contacting. Therefore, we decided to filter out the IP addresses to which more than 20 domain names resolve. We collected 4,914 CobaltStrike C2 servers, out of which 89 were contacted by the infected companies. However, after filtering out the ones hosted on shared infrastructure, there remained only 30. With these restrictions, we still found 25 companies infected with CobaltStrike. Figure 19 shows the distribution of these 30 infections across the ransomware families.

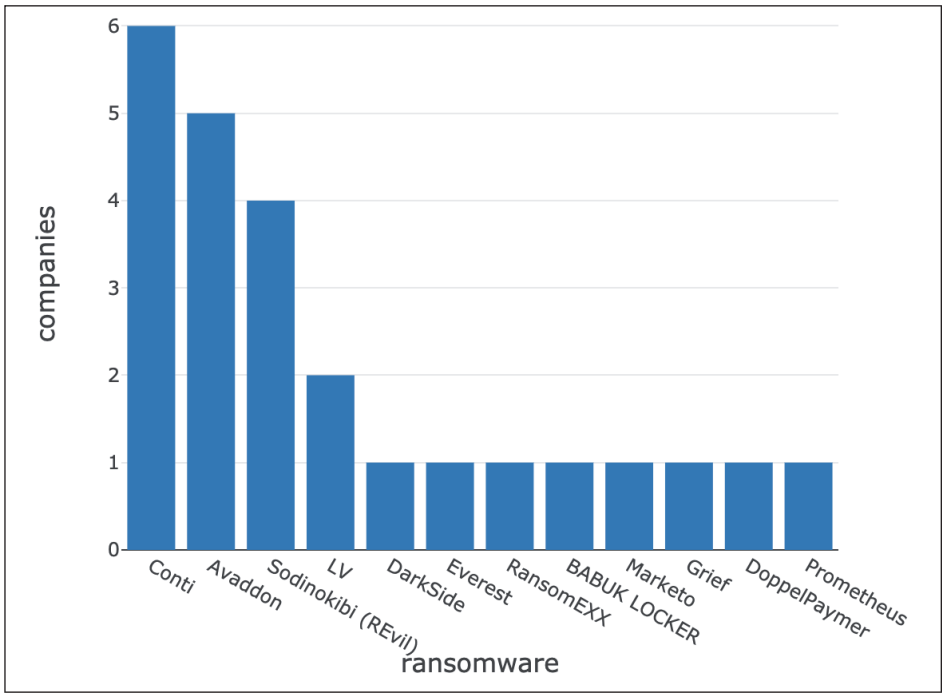

Figure 19: Number of CobaltStrike infections per ransomware.

5. DATA EXFILTRATION

When infecting a company, exfiltrating information prior to encrypting it has become a standard practice for ransomware operations. We have observed various ways of performing the data exfiltration. They differ in the infrastructure and services used.

In terms of infrastructure, we observed attackers using services provided by cloud solutions such as *MEGA* and *Rackspace*, but also setting up individual servers on *DigitalOcean*. When it comes to network protocols, services such as RDP, SSH, FTP, HTTPS and SMB were used.

5.1 MEGA Cloud and HTTPS

Many Sodinokibi victims have their data exfiltrated to the *MEGA Cloud* servers. In one example of such a victim, the data was exfiltrated to three *MEGA* IP addresses over HTTPS between 3:00 a.m. and 3:00 p.m. on 21 March. Figure 20 shows all the HTTPS traffic from the company's IP address while Figure 21 shows only the HTTPS traffic to the three *MEGA* IP address ranges over the same period. The leak was published in April.

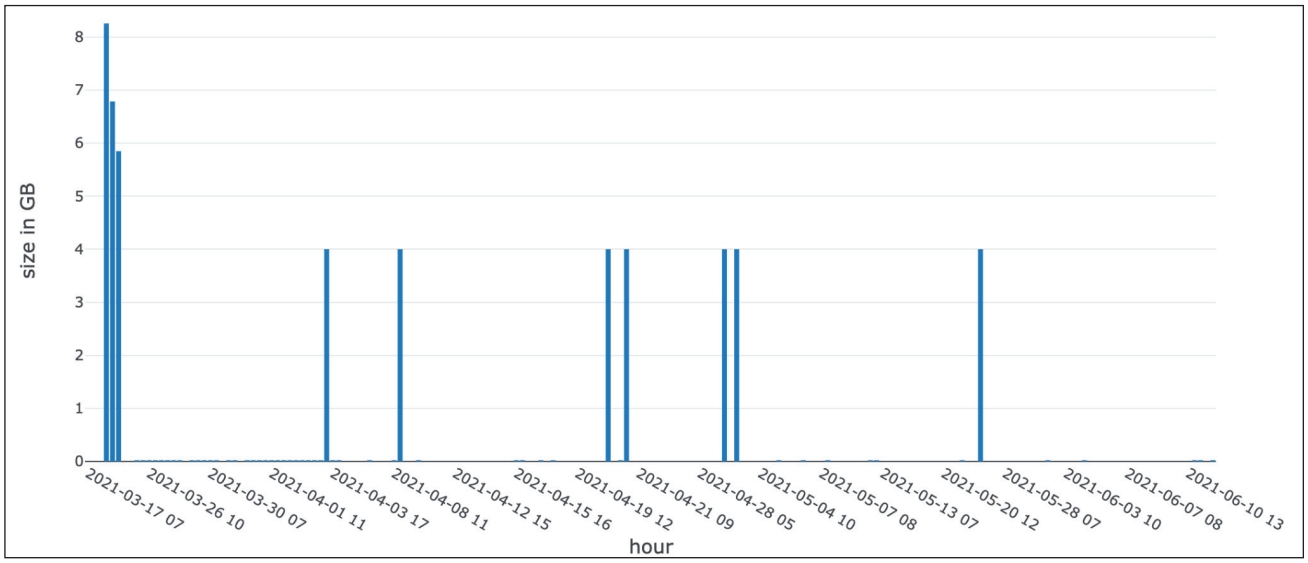

Figure 20: All HTTPS traffic from the exfiltrated IP address.

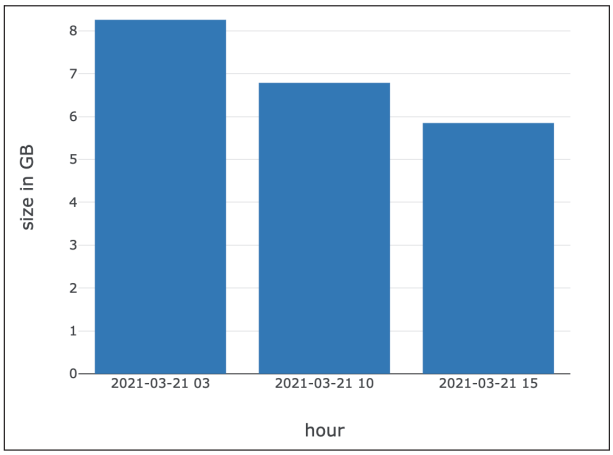

Figure 21: HTTPS traffic to three MEGA IP addresses from the exfiltrated IP address.

We observed at least 52 different companies having IP addresses communicating with *MEGA* IP addresses across 20 different ransomware families, but mostly from Conti and Avaddon.

5.2 Rackspace and SMB

For a victim of the LV ransomware, we observed data exfiltration to *Rackspace* over SMB. Figure 22 shows all the SMB traffic from the company's IP address while Figure 23 shows only the SMB traffic to three *Rackspace* hosting IP address ranges over the same period. The exfiltrated data amounted to a few TB in under a week. The exfiltration happened between 8 and 13 April, while the leak was announced on the website a few days later in April.

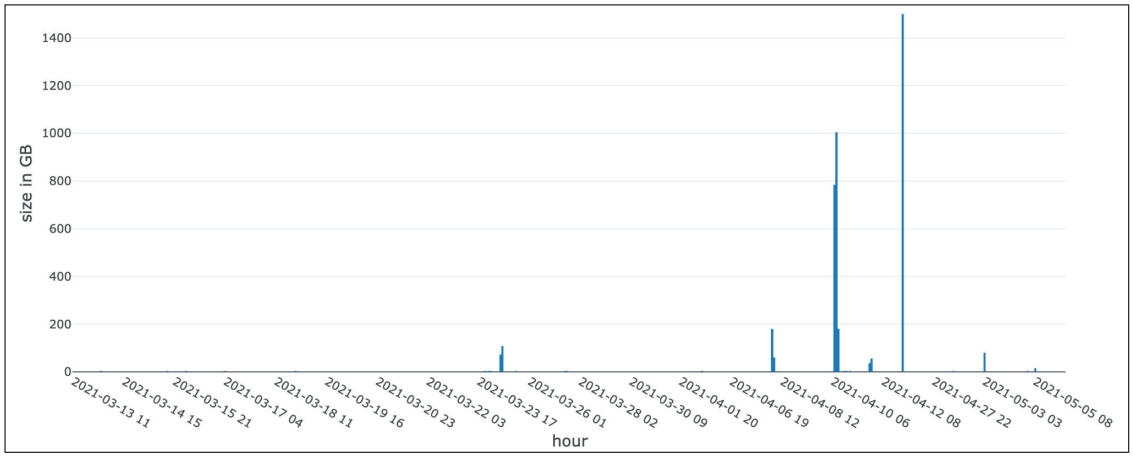

Figure 22: All SMB traffic from the exfiltrated IP address.

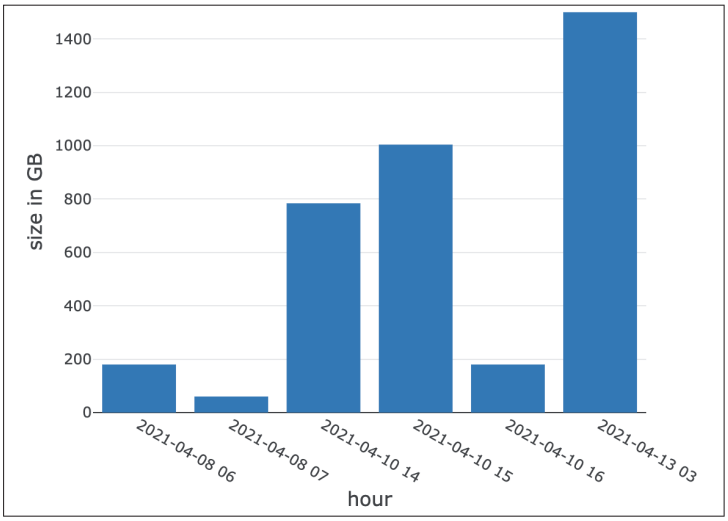

Figure 23: SMB traffic to three Rackspace IP addresses from the exfiltrated IP address.

5.3 DigitalOcean and SSH

Exfiltration over SSH can be observed in victims of DarkSide. Figure 24 shows the hourly exchange of GB during the exfiltration process to a server located at *DigitalOcean*. The infected system contacts the *DigitalOcean* server over SSH.

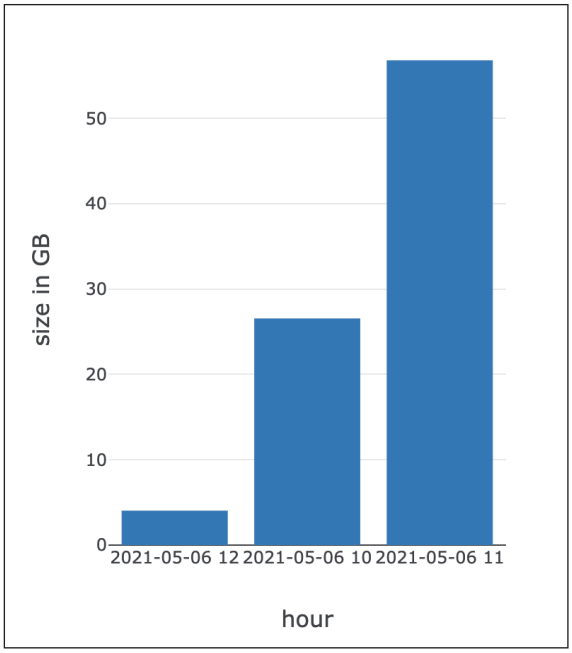

Figure 24: Exfiltration over SSH to DigitalOcean.

5.4 RDP

For the LV ransomware we observed data exfiltration happening over RDP soon after a brute forcing attack happened on RDP a few days before, indicating that the brute force was successful.

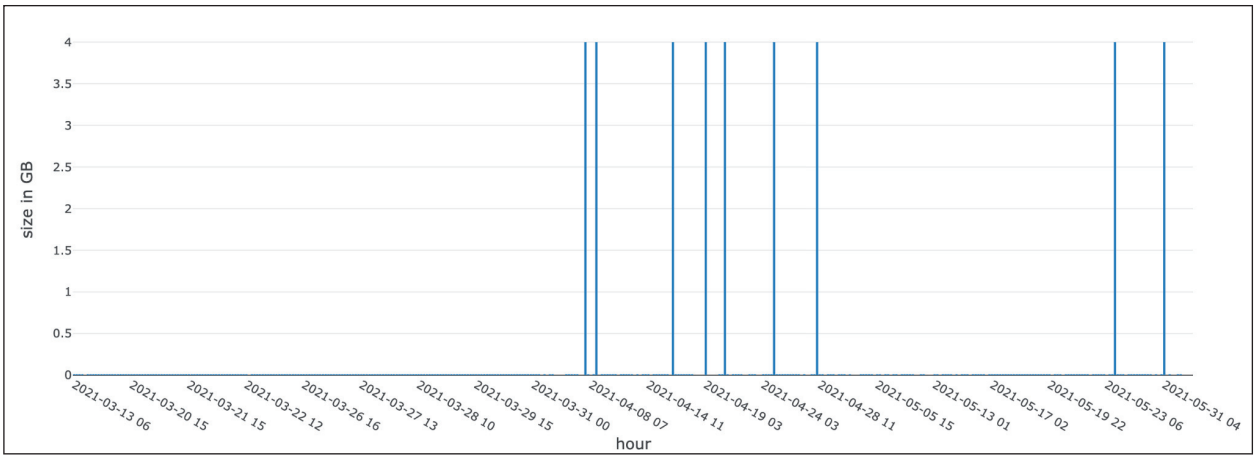

Figure 25: Exfiltration over RDP.

6. PAYING VERSUS NON-PAYING VICTIMS

Sodinokibi is an interesting case in which we have both the company name of the non-paying victims from the leaks website as well as paying victim IP addresses from our sinkhole. This allows us to figure out:

- How long it takes for the leak to appear on the website after the infection occurred.
- How many of the victims pay the ransom and how many don't pay, ending up with leaked information.

We first collected all the Sodinokibi IP addresses contacting our sinkholes during February – May 2021 and kept only the ones that have the URL form known to be used by Sodinokibi:

https://<c2_domain>/<uri_part_1>/<uri_part_2>/<random_resource_name>.<ext>

where:

- <part_1> is one of ‘wp-content’, ‘static’, ‘content’, ‘include’, ‘uploads’, ‘news’, ‘data’, ‘admin’
- <part_2> is one of ‘images’, ‘pictures’, ‘image’, ‘temp’, ‘tmp’, ‘graphic’, ‘assets’, ‘pics’, ‘game’
- <random_resource_name> has an even length between two and 18 characters and consists of only lowercase letters ranging from a to z
- <ext> is one of ‘jpg’, ‘png’, ‘gif’

Out of 20,014 IP addresses contacting the Sodinokibi sinkholed domains, 14,375 had the expected URL format. After finding a few of the IP addresses belonging to the leaked companies in the sinkhole data, we observed that the infected systems only made a handful of requests and only during a few hours. Therefore, we also filtered out the IP addresses that contacted the Sodinokibi domains for a period longer than 24 hours. We suspect these are sandboxes or malware researchers frequently running Sodinokibi samples.

For example, the following requests seem to be made by malware researchers because they use exactly the same URL to make queries to two different sinkholed domains from two different IP addresses during a period of multiple days:

```
x.x.x.101 2021-03-27T13:38:58 *k.com POST /uploads/images/scpgdiufwh.png HTTP/1.1
x.x.x.213 2021-04-23T10:15:35 *k.com POST /uploads/images/scpgdiufwh.png HTTP/1.1
x.x.x.213 2021-04-24T08:20:54 *k.com POST /uploads/images/scpgdiufwh.png HTTP/1.1
x.x.x.213 2021-04-24T07:17:38 *k.com POST /uploads/images/scpgdiufwh.png HTTP/1.1
x.x.x.213 2021-04-23T09:38:14 *k.com POST /uploads/images/scpgdiufwh.png HTTP/1.1
x.x.x.101 2021-03-27T13:01:58 *k.com POST /uploads/images/scpgdiufwh.png HTTP/1.1
x.x.x.213 2021-04-24T14:08:17 *k.com POST /uploads/images/scpgdiufwh.png HTTP/1.1
x.x.x.101 2021-03-27T13:03:56 *v.com POST /include/images/ahpw.gif HTTP/1.1
x.x.x.213 2021-04-24T07:19:48 *v.com POST /include/images/ahpw.gif HTTP/1.1
x.x.x.213 2021-04-23T10:17:36 *v.com POST /include/images/ahpw.gif HTTP/1.1
x.x.x.101 2021-03-27T13:41:00 *v.com POST /include/images/ahpw.gif HTTP/1.1
x.x.x.213 2021-04-24T08:23:30 *v.com POST /include/images/ahpw.gif HTTP/1.1
x.x.x.213 2021-04-24T14:10:22 *v.com POST /include/images/ahpw.gif HTTP/1.1
x.x.x.213 2021-04-23T09:40:19 *v.com POST /include/images/ahpw.gif HTTP/1.1
```

We ended up with 13,629 unique IP addresses.

Then, we collected all the Sodinokibi leaks that were published during March – May 2021 and computed the digital footprint of each leaked company.

The reason we went one month further back in the sinkhole records compared to the leaks records is because the infections happen earlier than the leaks as the attackers and the victims take some time to negotiate the terms. Therefore, in order to find the victims of leaks from March to June 2021 in the sinkhole, we needed to search them before March as well. We decided to include February infections, too.

We searched for those IP addresses that appear in both datasets in order to see what is usually the time interval between the infection and the leak. Out of the 82 companies for which Sodinokibi published leaks we found 12 in our sinkholes. There are some possible explanations for this:

- The infected system might have contacted one of the other 1,000 domains hard coded in the Sodinokibi samples.
- The reporting feature might have been disabled in the configuration.
- The infected IP addresses might not be part of the digital footprints we collected.

By correlating the dates of the infections and the dates of the leaks for those 12 companies we observed that the interval can range between two days and one month.

Also, we attributed the IP addresses contacting the sinkholes in order to obtain companies that were infected, but not leaked. After obtaining the initial list of companies that made requests to our sinkholes with valid Sodinokibi format, we only kept the companies with a small digital footprint (with fewer than 1,000 IP addresses), because most of the leaked infections had a small digital footprint and we also wanted to remove big telecommunication companies from our analysis. In the end we had a list of 130 companies that had valid requests to our sinkholes.

According to an interview provided by a REvil representative to a Russian blogger, approximately one third of the companies they breach pay the ransom in order to avoid being listed on the leaks website [16]. From our results, we could say that our numbers are close to the reported ones because there are 130 potentially infected companies and 87 companies present in leaked data in the last three months.

7. RECOMMENDATIONS

To prevent or at least minimize the chances of getting hit by a ransomware attack, we recommend the following:

- Credentials: don't reuse credentials, use strong passwords, use a password manager, use two-factor authentication, don't make your phone number and email address easily accessible.
- Exposed services: minimize the exposed digital footprint, use strong authentication where applicable, put services behind a VPN where applicable.
- Vulnerabilities: patch them immediately, especially if they provide remote code execution or privilege escalation and if there is a publicly or commercially available proof of concept.
- Malware: use an up-to-date anti-malware product.
- Exfiltration: use an intrusion detection system (IDS) / intrusion prevention system (IPS) / firewall, and monitor for and block large amounts of outgoing data.

CONCLUSIONS

Although ransomware has been the centre of media attention for the past couple of months, we shouldn't ignore the other types of malware. The fact that the last step of a ransomware attack involves notifying the victim that the attack took place and demanding a ransom means that ransomware attacks are very visible. However, there are other types of malware whose actions are not as visible but this doesn't mean they are less dangerous. Unlike in ransomware attacks, which have visible and immediate consequences, other attacks can have delayed consequences. In ransomware attacks, the victim knows when the attack happened and can investigate the traces in order to understand the tactics, techniques and procedures used by the attackers and document them in order to help other potential victims. In other types of attacks, you might not even know what hit you and when.

In the same way that the initial ransomware attacks, that consisted only of data encryption, evolved into double extortions (consisting of both exfiltration and encryption of data), the double extortion may evolve into exfiltration alone. From the current landscape, it seems as if exfiltration is a bigger motivator to pay the ransom than encryption. Encryption is starting to become a secondary activity and exfiltration the main one. The attackers might move in the direction of only doing the exfiltration.

With regard to the initial access, ransomware is not necessarily all that different from other malware, it makes use of phishing emails, exposed vulnerable services, application vulnerabilities, and other malware families. While there may be some preferred services and malware families, the attackers can and will easily adapt to use others once the currently used ones are no longer feasible.

More importantly, since most attackers are running a ransomware-as-a-service, the same malware is distributed and used by different threat actors according to their own preferences. This means that even across the victims of the same ransomware, there will be a variety of infection and exfiltration techniques.

What seems rather constant is the usage of data exfiltration and the presence/usage of CobaltStrike. A combination of these two seems to be the best indicator of a ransomware infection about to happen.

REFERENCES

[1] Diachenko, B. 800+ Million Emails Leaked Online by Email Verification Service. Security Discovery. March 2019. https://securitydiscovery.com/800-million-emails-leaked-online-by-email-verification-service/.

[2] Hacken Proof. New Data Breach exposes 57 million records. November 2018. https://blog.hackenproof.com/industry-news/new-data-breach-exposes-57-million-records.

[3] Hacken. How Sensitive is Your Non-Sensitive Data. October 2018. https://hacken.io/industry-news-and-insights/how-sensitive-is-your-non-sensitive-data/.

[4] Williams, C. 620 million accounts stolen from 16 hacked websites now for sale on dark web, seller boasts. The Register. February 2019. https://www.theregister.co.uk/2019/02/11/620_million_hacked_accounts_dark_web.

[5] Abrams, L. Ryuk successor Conti Ransomware releases data leak site. BleepingComputer. August 2020. https://www.bleepingcomputer.com/news/security/ryuk-successor-conti-ransomware-releases-data-leak-site/.

[6] Cybereason. A One-two Punch of Emotet, TrickBot, & Ryuk Stealing & Ransoming Data. April 2019. https://www.cybereason.com/blog/one-two-punch-emotet-trickbot-and-ryuk-steal-then-ransom-data.

[7] Stone-Gross, B.; Frankoff, S.; Hartley, B. BitPaymer Source Code Fork: Meet DoppelPaymer Ransomware and Dridex 2.0. CrowdStrike Blog. July2019. https://www.crowdstrike.com/blog/doppelpaymer-ransomware-and-dridex-2/.

[8] Abrams, L. QBot partners with Egregor ransomware in bot-fueled attacks. BleepingComputer. November 2020. https://www.bleepingcomputer.com/news/security/qbot-partners-with-egregor-ransomware-in-bot-fueled-attacks/.

[9] Abdo, B.; McKeague, B.; Ta, V. So Unchill: Melting UNC2198 ICEDID to Ransomware Operations. FireEye. February 2021. https://www.fireeye.com/blog/threat-research/2021/02/melting-unc2198-icedid-to-ransomware-operations.html.

[10] Santos, D.; Barbehenn, B.; Falcone, R. Threat Assessment: Egregor Ransomware. Palo Alto Networks. December 2020. https://unit42.paloaltonetworks.com/egregor-ransomware-courses-of-action/.

[11] Cybereason. Cybereason vs. Avaddon Ransomware. April 2021. https://www.cybereason.com/blog/cybereason-vs.-avaddon-ransomware.

[12] Available Solution for Conti Ransomware. Trend Micro. https://success.trendmicro.com/solution/000286405.

[13] Santos, D. Threat Assessment: Clop Ransomware. Palo Alto Networks. April 2021. https://unit42.paloaltonetworks.com/clop-ransomware/.

[14] Gallagher, S. Hacks for sale: inside the Buer Loader malware-as-a-service. Sophos News. October 2020. https://news.sophos.com/en-us/2020/10/28/hacks-for-sale-inside-the-buer-loader-malware-as-a-service/.

[15] Tracy, R.; Schmitt, D. When Threat Actors Fly Under the Radar Vatet, PyXie and Defray777. Palo Alto Networks. November 2020. https://unit42.paloaltonetworks.com/vatet-pyxie-defray777/3/.

[16] ELITE RUSSIAN HACKERS REVIL: HOW TO EARN 100 MILLION USD ON RANSOMWARE? | Russian OSINT. https://youtu.be/ZyQCQ1VZp8s.