Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
The Covid-19 threat landscape

Aditya K Sood, Srilekha Talluri, Amit Nagal, Ruthvik Reddy SL, Bharathasimha Reddy D and Rohit Chaturvedi, F5

Since the advent of Covid-19, the increase in working from home, online education and entertainment via online platforms has ramped up the number of Internet users worldwide. According to Forbes, Internet hits surged by 50-70%.

This phase shift and entertainment via online platforms has ramped up the number of Internet users. Since the advent of Covid-19, the increase in working from home, online education, and social distancing has led to an increase in the number of Internet users. This has led to an increase in the number of malicious actors exploiting this opportunity.

There have been many phishing and malware-related attacks that have harvested the Covid-19 theme in recent times. A few examples will help highlight the seriousness of the problem. The University of British Columbia (UBC) received a fake Covid-19 survey with the title ‘Covid-19 Mandatory HR Survey’ and convinced a user to download document files from public file-sharing platforms like Dropbox and Box. And we saw phishing emails that impersonated the Ministry of Health with the subject line ‘Corona Virus Latest Updates’.

Phishing emails don’t always carry attachments. Sometimes, to evade detection from mail-filtering security devices, the threat actor puts URLs in the email body and lures the user to visit a malicious site. Recently, such an attack was seen where Urology Cancer Research & Education (UCARE) Oxford received mail that purported to be Covid-19 vaccine invitations. The mail message had the subject line ‘NHS Vaccination’ and the malicious link was placed in the body.

Overall, attackers have opted for well-designed social engineering techniques to exploit the Covid-19 pandemic for monetary gain. In the following sections, we discuss the chosen methodology covering data collection and applied analytics techniques followed by analysis.

Methodology

Our research is primarily focused on an empirical analysis of the domain names registered by adversaries engaged in Covid-19 cybercrime. In this section, we highlight the methodology for data collection, including a basic overview of techniques chosen for empirical analysis.

We used the Covid-19 specific domains that were created in the timeframe of January 2020 to May 2021. We collected data for analysis in multiple ways:

- Data shared by third parties and partners working collaboratively to build intelligence.
- In-house data collection mechanisms utilising techniques such as malware analysis, network traffic analysis and reverse engineering to gather data related to malicious domains.
- Utilising open-source data sets to enhance the coverage for analysis. For example, we used open-source data provided by DomainTools.

We created data sets and optimised them for empirical analysis. We chose to go to the Moon.

Tools and techniques

To conduct the analysis of suspicious Covid-19 scam domains, we opted for several techniques.
The team created automated scripts and tools to collect the domain meta-data for obtaining better context about the characteristics of the Covid-19 domains. We conducted data de-duplication tasks to remove the duplicated entries, followed by data normalisation to create database tables storing well-structured data for active querying. We focused primarily on conducting exploratory and predictive analysis based on the collected data to answer what happened during Covid-19 cyber attacks, including what to expect in the future while dealing with Covid-19 scams.

We also used other core data techniques such as:
- Frequency Analysis (FA) to create a tag cloud representing frequently used words.
- FP-growth (frequent pattern-growth) – an advanced version of the Apriori algorithm, to correlate metadata and map associations between frequent patterns in data sets. It uses a compact data structure and eliminates repeated database scans to improve execution speeds.

Overall, all the techniques listed above resulted in obtaining granular insights into the abuse of Covid-19 registered domains used for cybercrime.

**Threat model**

We have followed a coherent approach to focus our analysis on the Covid-19 domains registered for a myriad of attacks and fraudulent activities conducted by adversaries. If the domain name was registered and used for Covid-19 cybercrime, we treated that domain as malicious. The Covid-19 domain may have been used for the following set of cyberthreats:

- Attackers conducting large-scale distributed or targeted phishing attacks and using Covid-19-themed domains to host phishing pages.
- Fraudsters involved in online cybercrime, who registered domains with reference to Covid-19 to spread fear and extort money from end users.
- Adversaries conducting drive-by-download attacks and forcing end users to visit Covid-19 themed domains to spread infections by distributing malicious code.
- Botnet operators registering Covid-19-related domains and hosting command and control (C&C) panels to operate and manage compromised machines running bots.
- Cyber criminals using Covid-19-themed registered domains for hosting software to collect stolen data from end-user systems, mainly for data exfiltration operations.

In all of the above scenarios, the attackers require domains to be registered to trigger nefarious operations on the Internet. Considering the above threat models, our research focused entirely on Covid-19 domains to unearth the trends in the cyberthreat landscape.

Next, we present a detailed analysis of Covid-19-themed cyberthreats that occurred recently.

**Geographical landscape**

In this section, we’ll investigate how Covid-themed domains are hosted across the world. The hosting countries represent where the IP address is geographically located. The domains’ IP addresses and locations are obtained using a Geo DNS query. We have portrayed the data in a granular manner, highlighting the countries that are hosting Covid-specific domain names. In Figure 1, countries (locations) hosting a higher number of malicious Covid-19 domains are indicated with circular points. The sizing of points is relative to the number of malicious Covid-19 domains hosted in the associated countries. Figure 2 shows the percentages of domains hosted in the top 10 countries.

If you look at the above insights, you can observe that North America and Europe act as a locus for hosting Covid-19 malicious domains. The top countries that are being used for hosting the domains are countries with high incomes. Clearly, the US has the maximum share, with around 74% of the total domains hosted. The top 10 countries contribute to nearly 94% of the total share.

**Live and dormant domains**

Next, we’ll look at the present state of active and inactive Covid-19 malicious domains used for cybercrime. The ‘active’ domains here represent domain names...
that are still valid and used by adversaries, whereas ‘inactive’ domains represent those that are no longer used, have been de-registered or are not used to host suspicious content related to Covid-19. Figure 3 gives an overview.

From Figure 3, you can easily determine that thousands of Covid-19 domains used for cybercrime are still active in the wild, which means that attackers are continuing to abuse the Covid-19 theme for conducting cybercrime. At the same time, you can also map several domains that are found to be inactive. This points to the fact that attackers are not using the same Covid-19-themed domains every time, but instead are registering new domains. Overall, more domains are still registered and active on the Internet, showing how effectively the attackers are exploiting the Covid-19 theme.

Time analysis

In this section, we’ll study the trend in the creation of Covid-19 malicious domains.

With the pandemic, the need for information related to coronavirus has increased tremendously. Worldwide, people started to search for the subject online to gather more information. Malicious actors were aware of this fact and exploited it accordingly. Figure 4 depicts the trend in the creation of Covid-19-themed suspicious domains.

If you check the graph, you can observe a spike in March and April 2020. At that point in time, the Covid-19 pandemic was in its initial stages and infection rates were significantly high. Although the number of Covid-19 domain registrations declined after the spike period, there is still some consistency in the number of new domains registered per day to date. However, you can easily see that, for the last nine months, adversaries were still registering Covid-19-themed malicious domains in order to conduct online scams and potentially advanced attacks. This trend still holds true.

TLDs and IP ranges

Now, let us investigate the subnets and top level domain (TLD) trends in the active Covid-19-themed domains. These IP address spaces are extracted using domain name resolution obtained by performing DNS queries to the Covid-19-themed domains.

If you look at Figure 5, the top IP address space used for hosting fraudulent domains is 34.64.0.0/10, which belongs to Google, and the second space (184.168.0.0/16) belongs to Go Daddy. This tells us that the scammers are taking advantage of public clouds and hosting services to quickly deploy a web portal to be used for malicious operations and move on to other hosting providers when needed, which includes registering new domains.

Also, Figure 6 shows the distribution of the top TLDs that are being used by Covid scam domains. Analysing the TLDs used by adversaries while registering malicious Covid-19-themed domains helps you
to understand the DNS root zone associated with the namespace. Clearly, ‘.com’ is the most exploited TLD, followed by ‘.org’ and ‘.net’. The fact that adversaries are targeting these TLDs further resonates with the fact that commercial, organisational and network DNS namespaces are being utilised in large numbers for creating Covid-themed domains.

**Frequency analysis**

In this section, we present the lexical analysis of the various domain names used for Covid-19 fraudulent activities. The domains we obtained from the data are parsed using a custom dictionary and the frequency of each word is depicted in the form of a word cloud where the largest size represents the highest frequency. The word cloud is depicted in Figure 7 and Table 1 shows the frequencies of the top 10 words used in the registration of Covid-19 domains.

You can observe that most of the domains used for Covid-19 scams use words like Covid19, Covid, corona, virus, vaccine, and so on to dupe people in the guise of government agencies or humanitarian organisations. Additionally, this indicates how efficient attackers are using a Covid-19 theme to launch social engineering attacks and trick users into providing sensitive information that can be further used to launch more cyber attacks and conduct financial frauds for monetary gain.

**Association analysis**

Now we will take a deep dive into association analysis to understand the choice of words chosen by adversaries who registered Covid-19 fraudulent domains. Association analysis helps you to understand the relation and affinity between various Covid-related words and how similar domains are registered. Generally, you can create small subsets with a similar set of words.

In association analysis, each domain is considered as a transaction and the words in it are like items. The data was sent to an FP growth tree or association analysis module. We confined the analysis to pair relations and made a connectivity graph using the list of generated pairs (satisfying a threshold). The size and nodes in the graph are based on the frequency of the words in Covid-19 domains. The relations between the words are portrayed using a network graph, as shown in Figure 8.

Figure 8 shows a strong association with keywords corona, Covid-19 with other keywords. The detailed statistical analysis is shown in Table 2. Table 2 gives a sample view of the association analysis of Covid-19 domain-related words. Here we have Confidence, which is the measure of the occurrence of Consequents given the Antecedents. The RHS support is the frequency of instances of the Consequents in the transaction, where the transaction is the set of words used in the domain. For example, the domain ‘novelcoronavirus-cure.com’ has the transaction as (‘corona’, ‘cure’, ‘novel’, ‘virus’).

The Lift is a useful measure to determine the strength of the association between Antecedents and Consequents. It determines mutual dependence. Lift is 1 when both are independent and it is defined as follows:

\[
\text{Lift (A } \rightarrow \text{ C)} = \frac{\text{Confidence(A } \rightarrow \text{ C)}}{\text{Support(C)}} = \frac{\text{Support(A, C)}}{\text{Support(A) * Support(C)}}
\]

The Conviction determines which part of the association rule has the upper hand – whether the Antecedent drives the Consequent or vice versa. In Table 2, you can see that all the Conviction values are negative, which implies the Consequents here are having the upper hand. As an analogy, when it comes to mobile phones and screen protectors, typically mobile sales drive the sales of screen protectors. In that example, mobile phones have the upper hand in the association. Positive
Conviction implies that Antecedents drive the Consequents or vice versa. The Conviction is computed as follows:

\[
\text{Conviction}(A \rightarrow C) = \frac{1 - \text{Confidence}(A \rightarrow C)}{1 - \text{RHS}\_\text{support}}
\]

The association rules help build the connectivity graph to understand the relationship between various subsets containing items that occur together. The visualisation of association rules provides an overview of words used in Covid-19 domain names and how these are related, thereby helping us to understand the bigger picture concerning the selection and registration of Covid-19 domains.

### Inference

The Covid-19 pandemic demonstrated how cybercrime at its core remains largely the same, but criminals change the narrative. Since people are curious to know about Covid-related information from the Internet, attackers use this curiosity to carry out nefarious operations and unethical activities. A significantly large set of new domains have been registered with terms related to Covid-19 and the response to it, such as ‘Covid19’, ‘corona’, ‘virus’, ‘testing’, ‘vaccine’ etc. These suspicious domains reference web content that promises vaccines and other aids, often in the guise of legitimate government agencies or humanitarian organisations, and in this way attempt to steal credentials and infect systems with malware.

### Conclusion

In this research, of the nearly 100,000 domains that were analysed, most of the domains used words such as ‘Covid19’, ‘corona’, ‘Covid’, ‘virus’ (Table 1). In 2020, Google search trends showed that coronavirus and election results were the top searches. Similarly, Covid-19 domain registrations are in sync with the search trends for the coronavirus, which show that attackers are exploiting human behaviour and people’s response to the crisis.

Most of the domains are hosted in North America and Europe, which are the countries with sound and scalable economies. Attackers took advantage of cloud-hosting services such as Google Cloud and Go Daddy to host Covid-19 fraudulent domains because cloud hosting services provide the benefit of shared hosting. The many-to-many mapping that often occurs in a cloud environment can make IP-based firewalls ineffective.

While ‘.com’ is the best domain suffix for search engine optimisation and is the largest, oldest and most trusted TLD, it is being exploited the most by attackers. That holds true with Covid-specific cyber attacks as well. Overall, malicious actors exploited and abused end users’ desire to understand the Covid-19 pandemic to conduct cybercrime.

### About the authors

Dr Aditya K Sood is a senior director of threat research and security strategy in the Office of the CTO at F5. He has research interests in cloud security, IoT security, malware automation and analysis, application security and secure software design. Sood obtained his PhD from Michigan State University in computer sciences. He is also the author of the books Targeted Cyber Attacks and Empirical Cloud Security.

Srilekha Talluri is a software engineer at F5 Networks. She received her BE in computer science from BITs Pilani Hyderabad in 2020. She is a part of the Innovate F5 Project.
Prevent hybrid working from fuelling fraud

James Richardson, Bottomline

Online and remote ways of working have created a warm and welcoming environment for fraudsters. And yet many businesses say they should be doing more to beat fraud. Now it’s time for them to take action.

Most people would probably agree that being a victim of fraud should never, ever be an accepted aspect of managing a business. But when doing the research for the most recent ‘Payments Barometer’, a staggering 55% of respondents commented that they see financial loss as ‘part and parcel’ of running a business.¹

This would be easier to accept if it was straightforward for organisations to recover fraud losses — or even if businesses really believed they were already doing everything possible to defeat the fraudsters. But in a world where organisations successfully recovered only 20% of their losses to fraud (albeit rising from 13% in 2019 and 17% in 2020), a full 70% of respondents admitted they could be doing more to mitigate fraud risk. And these aren’t small businesses. Around three-quarters of the large (1,000 to 9,999 employees) and enterprise (10,000+) organisations that were surveyed also made this admission.

References
1. Beech, Mark. ‘Covid-19 pushes up Internet use 70% and streaming more than 12%, first figures reveal’. Forbes, 25 Mar 2020. Accessed Aug 2021. www.forbes.com/sites/markbeech/2020/03/25/Covid-19-pushes-up-Internet-use-70-streaming-more-than-12-first-figures-reveal/?sh=4671d3253104.
2. Kumaran, Neil; Lugani, Sam. ‘Protecting businesses against cyberthreats during Covid-19 and beyond’. Google Cloud, 16 Apr 2020. Accessed Aug 2021. https://cloud.google.com/blog/products/identity-security/protecting-against-cyberthreats-during-Covid-19-and-beyond.
3. Segura, Jerome. ‘Fake Covid-19 survey hides ransomware in Canadian university’. Malwarebytes, 28 Oct 2020. Accessed Aug 2021. https://blog.malwarebytes.com/cybercrime/2020/10/fake-Covid-19-survey-hides-ransomware-in-canadian-university-attack/.
4. Bisson, Dave. ‘NHS Warns of New Covid-19 Vaccine-Related Phishing Campaigns’. Cybereason, 1 Feb 2021. Accessed Aug 2021. www.cybereason.com/blog/nhs-warns-of-new-Covid-19-vaccine-related-phishing-campaigns.
5. McNee, Sean. ‘Covid-19 Threat list retired – domain risk assessments for Coronavirus threats’. DomainTools, 30 Jun 2021. Accessed Aug 2021. https://www.domaintools.com/resources/blog/free-Covid-19-threat-list-domain-risk-assessments-for-coronavirus-threats.
6. Sood, Aditya; Zeadally, Sherali. ‘Drive-by download attacks: a comparative study’. IT Professional, Sep/Oct 2016, vol.18, pp.18-25. Accessed Aug 2021. www.computer.org/csdl/magazine/it/2016/05/mit2016050018/13rRuye-Ckea.
7. ‘Coronavirus’. Google Trends. Accessed Aug 2021. https://trends.google.com/trends/explore?date=2019-12-01%202021-05-25&q=%2Fm%2F01cpuy.