Delft University of Technology

Cash for the Register? Capturing Rationales of Early COVID-19 Domain Registrations at Internet-scale

Pletinckx, S.R.G.; Habben Jansen, G.J.; Brussen, A.; van Wegberg, R.S.

DOI
10.1109/ICICS52457.2021.9464572

Publication date
2021

Published in
2021 12th International Conference on Information and Communication Systems, ICICS 2021

Citation (APA)
Pletinckx, S. R. G., Habben Jansen, G. J., Brussen, A., & van Wegberg, R. S. (2021). Cash for the Register? Capturing Rationales of Early COVID-19 Domain Registrations at Internet-scale. In M. Alsmirat, A. Almaaitah, Y. Jararweh, & J. L. Mauri (Eds.), *2021 12th International Conference on Information and Communication Systems, ICICS 2021* (pp. 41-48). Article 9464572 (2021 12th International Conference on Information and Communication Systems, ICICS 2021). https://doi.org/10.1109/ICICS52457.2021.9464572

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

2021 12th International Conference on Information and Communication Systems (ICICS) | 978-1-6654-3351-8/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICICS52457.2021.9464572

Cash for the Register? Capturing Rationales of Early COVID-19 Domain Registrations at Internet-scale

Stijn Pletinckx, Geert Habben Jansen, Arjen Brussen, Rolf van Wegberg
Cyber Security Group
Delft University of Technology
Delft, the Netherlands
{s.r.g.pletinckx, g.j.habbenjansen, a.brussen}@student.tudelft.nl, r.s.vanwegberg@tudelft.nl

***Abstract*—The COVID-19 pandemic introduced novel incentives for adversaries to exploit the state of turmoil. As we have witnessed with the increase in for instance phishing attacks and domain name registrations piggybacking the COVID-19 brand name. In this paper, we perform an analysis at Internet-scale of COVID-19 domain name registrations during the early stages of the virus' spread, and investigate the rationales behind them. We leverage the DomainTools COVID-19 Threat List and additional measurements to analyze over 150,000 domains registered between January 1st 2020 and May 1st 2020. We identify two key rationales for covid-related domain registrations. Online marketing, by either redirecting traffic or hosting a commercial service on the domain, and domain parking, by registering domains containing popular COVID-19 keywords, presumably anticipating a profit when reselling the domain later on. We also highlight three public policy take-aways that can counteract this domain registration behavior.**

***Index Terms*—Domain registrations, Rationales, COVID-19**

I. Introduction

The COVID-19 pandemic introduced novel incentives for criminal business opportunities. In chaos adversaries thrive and find ways to monetize the circumstances – COVID-19 was no exception. For example, the John Hopkins hospital provided a popular online interactive map of how COVID-19 spreads globally. Quickly, cyber criminals targeted this map and built a malicious Java application, which required to be run locally to get access to the 'map' [1]. As with many attack vectors and tools, the malware was also offered as a service. Cyber criminals were selling the tool starting at $200 [2].

Likewise, new domain name registrations 'related' to the COVID-19 pandemic surged [3]. Instinctively, this increase in domain registration could be attributed to malicious attempts, like phishing campaigns, piggybacking the COVID-19 'brand name'. However, in the past we have seen globally-known viruses, like Ebola, being exploited in domain name registration for non-malicious practices, such as domain reselling [4], a practice wherein domains, like stocks, are expected to increase in value and be sold off later.

This paper aims to investigate the trends in covid-related domains. To be precise, we intend to uncover the rationales behind domain registrations, and identify their malicious intent. In particular, we will take a look at some of the 'early adapters' when it comes to registering covid-related domain names. Such early phases are crucial to recognize and anticipate on to minimize damage and handle the situation properly. As such, this work can be generalized to any exploitation of a salient event for domain name registration.

For our analysis we build on the DomainTools COVID-19 Threat List [5]. We make the following contributions:

- We present an analysis at scale of the registrations of covid-related domains before and during the early stages of the COVID-19 pandemic. We analyze over 150,000 domains registered between January 1st 2020 and May 1st 2020.
- We show a clear influence by media coverage of the pandemic on the registration of covid-related domains.
- We uncover that most covid-related domains registered during the pandemic have two key rationales:
  - Online marketing, by either redirecting a batch of registered covid-domains to one specific domain, or by directly registering a commercial service behind a covid-related domain.
  - Domain parking, by buying domains containing popular COVID-19 keywords without attaching any service to the domain, possibly with the intent of reselling them for a profit.
- We suggest three public policy takeaways that can be deduced out of the trends reported in this paper. Concretely, from the background of the COVID-19 pandemic, we propose general measures aimed at preventing large scale abuse of domain registration based on the popularity of a specific event.

The rest of this paper is structured as follows. Section II describes the lens through which we will study the domains, and the rationales we want to uncover. In Section III we elaborate on the data set used for this study. Section IV presents our results, which will be reflected upon in Section V along with a number of public policy takeaways. We cover some of the related work in Section VI, and conclude in Section VII.

Authorized licensed use limited to: TU Delft Library. Downloaded on June 30,2021 at 12:21:59 UTC from IEEE Xplore. Restrictions apply.

II. Domain Registration

Much like a memorable company name or a catchy slogan, a well-chosen domain name can increase traction. As such, rather than registering a domain name to alleviate the need of remembering an IP address, it can also be exploited to boost traffic. This concept highly alters the potential value a domain can have. Just as with an asset on the stock market, domain names can produce fluctuating values depending on their performance, coverage by the media, and desirability (supply and demand). It is through this lens that we aim to analyze the registration of domain names regarding the COVID-19 pandemic.

Concretely, we look at covid-related domains by treating them as assets. Are they being used for boosting traffic? Do actors invest in them? Is the registration trend susceptible to media coverage and key-word popularity? Viewing the registration of domain names through this lens, and within the context of a pandemic, allows us to get a much broader view on the influence of topical events on the domain registration process, along with its general trends.

A. Rationales

We aim to uncover three rationales behind the registration of covid-related domains: redirecting traffic, domain parking, and selling goods.

Redirecting Traffic For many people, hosting a website functions as a direct source of income. For instance by selling services through one's webpage, or through advertisements. For others, hosting a website can serve as an indirect source of income. Here, the domain functions as a marketing resource to redirect potential customers to your direct point of sale –e.g., your physical store. In both cases profits are often influenced by the number of visitors –more page visits equals more profit. We hypothesize that a first rationale for covid-related domain registrations is to leverage (brand)name popularity to redirect traffic towards a particular domain.

Domain Parking Domain parking is the act of registering a domain at a registrar without actually using that domain for any service, such as hosting a website. Motivations depend on the individual behind the domain, but general purposes of this practice are to either 'claim' a domain for later development, or to sell it in the near future with the goal of obtaining a profit – i.e., 'cybersquatting'. As a second rationale, we hypothesize that covid-related domains are being used for cybersquatting.

Selling Goods As a final rationale, we hypothesize that covid-related domains are registered simply for selling goods. In particular, during the early days of the lockdowns, many shortages in stores were caused due to obsessive hoarding by people [6]. Next to the general trend in online shopping, we therefore speculate that Shopify webstores and Amazon affiliation links on these covid-related domains try to tap in to the need for buying supplies online.

III. Data

Our measurement methodology consists of two steps. First, we take the public COVID-19 Domain Threat List by DomainTools [5] as a starting point for our analysis. Second, we collect additional data on domains (n=153,515) in the threat list. This dataset is freely available and contains a list of malicious domains as determined by DomainTools. In what follows we describe the attributes of the DomainTools threat list, alongside it's limitations, and report on our additional data collection approach.

A. DomainTools Threat List

The Threat List constructed by DomainTools is a list of domains whose domain name is related, in one way or the other, to the COVID-19 pandemic. This relationship is purely semantical, and hence provides no guarantee that the domain is associated to (illegitimate) initiatives concerning COVID-19. For our analysis we use the version of May 1st, containing 153,515 entries spanning from January 1st 2020 to May 1st 2020. The list is updated every day by DomainTools, both by adding as well as removing domain names. As a result, some divergence may exist compared to the version of May 1st used in this research. In the DomainTools Threat List, each entry consists of three fields:

- *Domain Name* – Shortened URL string containing domain name and top level domain (TLD).
- *Registration/Scan date* – The date the domain was registered, or if not available, the date DomainTools first detected the domain.
- *Domain Risk Score* – A score computed by DomainTools. A minimum score of 70 is required for a domain to appear in the dataset.

The domain risk score for a given domain depends on multiple factors [7], including the domain's proximity to other known malicious domains, and the similarity to malicious site archetypes based on machine learning methods. However, the exact method used to compute the domain risk score is not made public by the DomainTools team.

While the Threat List seems to be the more complete list regarding covid-related domains, it does have some inherent limitations. Primarily, it lacks in-depth information concerning the registration of the domain such as owner, registrar, expiration date, etcetera.

Second, the list provides no further information about the reason a domain is considered a threat, other than the DomainTools Risk Score. This Risk Score is calculated by a custom algorithm from DomainTools whose inner workings are not public. What is known, however, is that the algorithm consists of two parts: one that determines the proximity of a domain to known bad domains, and another that uses machine learning to determine how similar certain features of the domain are to domains used for spam, phishing, or malware [7].

We resepect the choice of DomainTools to keep this algorithm private, given that perhaps adversaries could adapt their modus operandi to lower their risk score. However, knowing that similarities with other malicious domains is a factor for the score, some information could be disclosed. For example, a similarity score with other domains or basic clustering could help for analysis and research on these domains.

Authorized licensed use limited to: TU Delft Library. Downloaded on June 30,2021 at 12:21:59 UTC from IEEE Xplore. Restrictions apply.

Third, the list lacks data about the content of the domains, as well as potential redirects of the domain names. Given the motivation for the Threat List is to stop malicious exploitation of the COVID-19 crisis, one could expect similar content behind the domains as part of larger campaigns leveraging the 'brand name' of the virus.

To mitigate some of the shortcomings of the data set, we perform additional look-ups which we will describe next.

B. Additional Data Collection

In order to gather more information about the domains, we performed additional data gathering to detect page re-directs and page keywords. Additional data collection was performed between May 6th and May 17th using a custom Python script.

Page re-directs are used to point domains to different services. We detected these redirects by checking the HTTP status code for a 3xx response, and saving the domain name that is redirected to.

Page keywords were used to naively filter parked domains and shops by inspecting the HTML contents of the page. For basic detection of parked domains, the landing page content was tested for the strings *"parked-content.godaddy.com"* and *"domain was recently registered at Namecheap."*, common strings for parked domains registered at GoDaddy and NameCheap respectively. Basic detection for shops was done by checking the source for Amazon affiliate links or a *"Powered by Shopify"* mark.

While this list of keywords is by all means non-exhaustive, it does allow us to capture a lower bound on the prevalence of parked domains and shops that leverage the COVID-19 pandemic.

Gathering the additional data sometimes resulted in failing requests, often caused by the domain being unreachable at the moment of the request. In order to perform further analysis on meaningful data, we decide to ignore all domains where both the lookup for the IP and the page content failed, which can give a strong indication that the domain is offline. Filtering on this criteria leaves 132,469 domains out of the original 153,514. In total 21,046 domains, or 7.8%, are dropped due to failing requests.

Our additional lookups provide only a select set of extra data, leaving out many other possible sources to explore. Furthermore, our lookups consist of only one snapshot in time, meaning that changes in a domain's content and setup are not captured in this study. We elaborate more on this at the end of this paper in the Limitations section.

IV. Results

In this section we present the results of our analysis. Specifically, we show how the three rationales presented in Section II can be captured from the attributes of domain registrations in our data. Before doing so, we first report on the general registration trend and its characteristics to build up the background against which these rationales are observed.

TABLE I
Top 10 redirected-to web pages

Redirected-to URL	# Redirects
injurysolutions.com	452
www.domainchop.com	172
wildcard.hostgator.com	155
www.oyzta.com	90
Dead link to image	81
howtopreventcoronavirusinfection.com	74
nahahealth.com	56
www.google.com	55
www.rts.com/covid-19-resources	52
5e92fae5935f0.site123.me	38

A. Domain Registration Trend

Figure 1 illustrates the registrations of covid-related domains in the period spanning January 1st 2020 up until May 1st 2020. In the figure, several COVID-19 media events are depicted. In general we observe that registrations start off moderately around January and February, but then sharply increase in March. After this peak, the trend slowly seems to decrease over the subsequent weeks.

1) Official Name Release Affects Domain Semantics: On the 11^{th} of February, the World Health Organization (WHO) officially declared the virus to be named 'COVID-19' (**Co**rona **Vi**rus **D**isease 20**19**). In Figure 1, this day is clearly noticeable by the sudden spike in domain registration around mid February. 86% of the domains in our data registered on that day contain 'covid' in their domain name. Before this day, however, on average 3% of the daily registered domains contained the name covid, while after the official name release this number is 46%. This shows a clear increase, and demonstrates how media coverage can influence the trends of domain registration.

B. Capturing Rationales

1) Redirecting Traffic: In total, we found that 17,487 (13.2%) domains served as redirects towards other domains. Table I shows the top 10 redirected-to domains. As a case study, we take a deeper look at the top redirected-to domain: injurysolutions.com.

Visiting its website, Injury Solutions seems to be a Las Vegas based chiropractic physician, specializing in car accidents. If we inspect the domains that point to injurysolutions.com, we find a pattern among the hostnames. These hostnames most often either directly, or indirectly, make a reference to Las Vegas, by listing the city's name or one of its sports teams, such as for example the Las Vegas Knight, along with a reference to COVID-19. The TLDs seem to be chosen descriptively more than conforming to industry standards, with among the most popular TLDs being `.vegas`, `.games`, `.flights`, and `.voyage`. Especially the 109 occurrences of the `.vegas` TLD shows the marketing strategy of the registrant.

Authorized licensed use limited to: TU Delft Library. Downloaded on June 30,2021 at 12:21:59 UTC from IEEE Xplore. Restrictions apply.

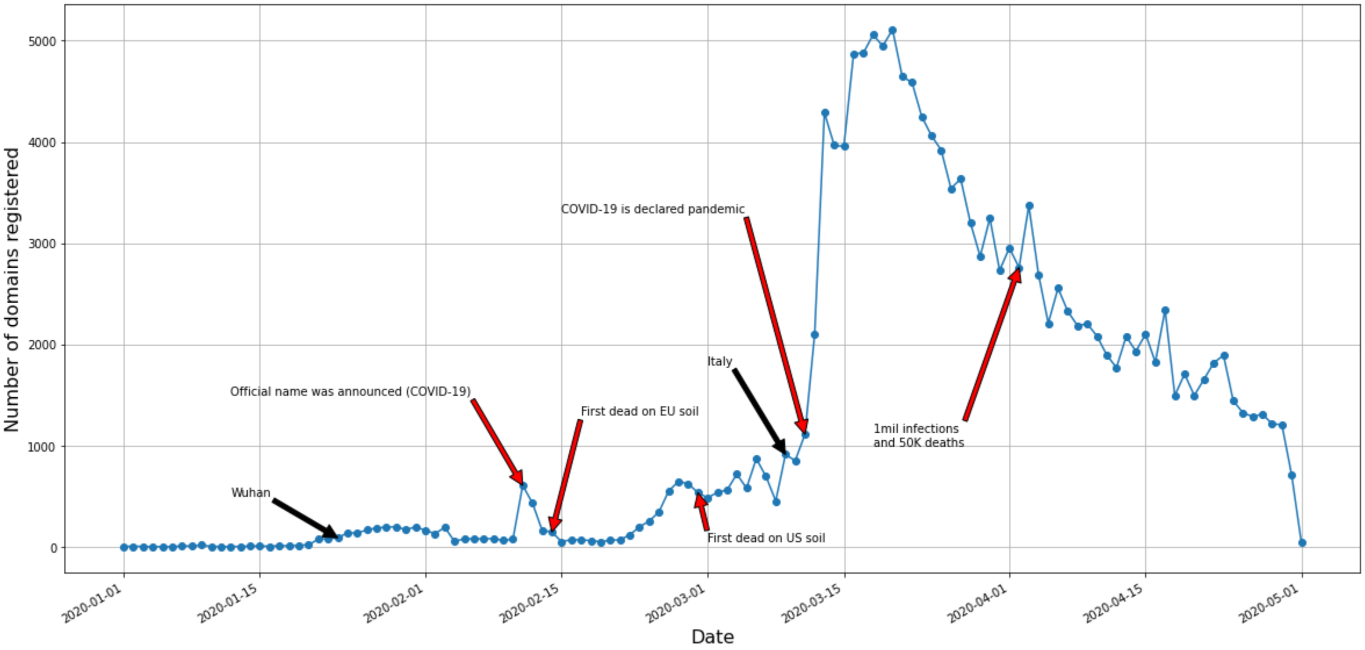


Fig. 1. COVID-19 related domain registrations between January 1^{st} 2020 until May 1^{st} 2020. Red arrows indicate key media-events, black arrows specific lockdowns.

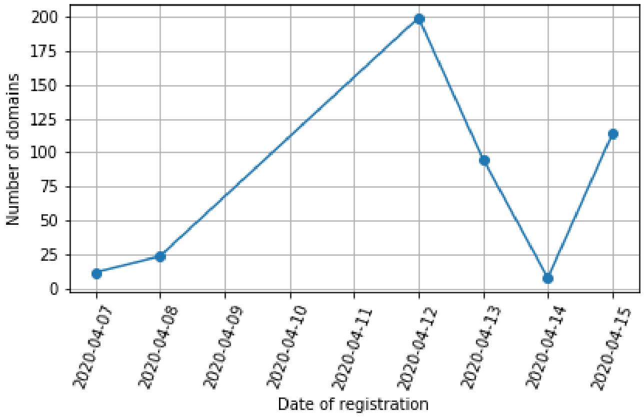


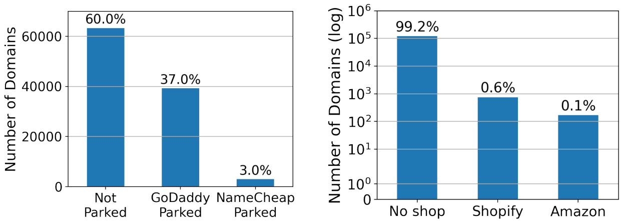
Fig. 2. Amount of registered/detected domains that redirect towards injurysolutions.com per day.

Additionally, we discover that all domains were registered at the same registrar –NameSilo, LLC– which not only provides domain name registration, but also domain hosting services.

Figure 2 illustrates the registration/detection date of each domain in the dataset that points towards injurysolutions.com. The amount of new domains seems to be low first, but then rises sharply a few days later, after which it decreases again for one day. This might indicate that the registrant of these domains was testing the waters before investing heavily in new covid-related domains.

The findings above suggest that most domains redirecting to Injury Solutions are deliberately used as part of a marketing campaign that exploits the COVID-19 crisis.

2) Domain Parking: As discussed in Section II, parked domains are a potential indicator for cybersquatters trying to make profit of reselling domain names in high demand. Figure 3a shows the distribution of parked domains. What stands out is that 40% of the domains appears to be parked and has, hence, no service residing behind its name. This is only checked on the index page of the domain. Hence, other parts of the domain could still be used for malicious purposes, as indicated by the DomainTools Risk Score.



(a) Parked Domains (b) Webshops

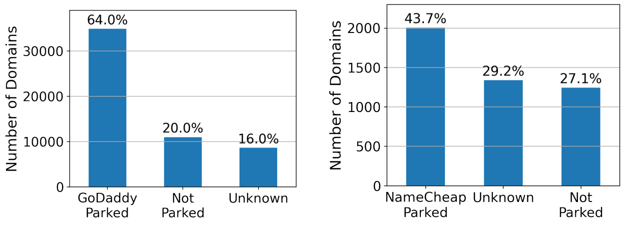
Fig. 3. (a) Percentage of domains parked on GoDaddy or on NameCheap. Domains that are either not-parked, or parked at a different registrar construct the ‘Not Parked’ bar. (b) Number of domains in dataset actively hosting Shopify webshops or Amazon affiliation links.

Please note that the percentages in Figure 3a are calculated by merely looking at the standard pages of two registrars, allowing us to obtain a lower bound. This means that in the 60% currently labeled as ‘Not Parked’, potential parked domains could still be found.

To emphasize the ubiquity of parked domains in the DomainTools data set, we plotted the distribution for these two registrars. Figure 4a shows the distribution of parked domains

Authorized licensed use limited to: TU Delft Library. Downloaded on June 30,2021 at 12:21:59 UTC from IEEE Xplore. Restrictions apply.

for all GoDaddy-registered domains. Respectively, Figure 4b shows the distribution of parked domains for all NameCheap-registered domains. Both figures show that parked domains take up the majority of all the registered covid-related domains at both registrars. We suspect different registrars to contain a similar trend. Note that the 'Unknown' bar represents the domains for which our request did not get a valid web page back. Ergo, these entries could be either parked or non-parked.



(a) GoDaddy

(b) NameCheap

Fig. 4. Distribution of domains registered at GoDaddy and NameCheap that are either parked or not parked. The 'Unknown' bar represents the domains that did not give back a valid web page to our request, and could hence either be parked or non-parked.

One counterargument for the hypothesis that these parked domains are used for cybersquatting could be that the owner of the domain simply did not put any service on his/her domain yet due to time restrictions. However, analyzing the dates of registering the parked domains, we see an almost identical trend as depicted in Figure 1. Ergo, the amount of domains that are parked decreases as we go further back in time, but it is highly likely that the reason for this is that these domains simply follow the global trend of all domains in the DomainTools Threat List.

3) Selling Goods: Figure 3b shows the number of domains hosting Shopify webstores or containing Amazon affiliation links. The keyword filter marked 746 domains as a Shopify store and 168 domains containing Amazon affiliation links. An additional 758 domains used an x509 certificate for *myshopify.com*, and therefore are also classified as Shopify stores. Noteworthy is the big difference between Shopify and Amazon domains, indicating that hosting a web shop is more popular than publishing a link to an external marketplace or web shop.

V. Discussion

The results presented in the previous chapter show a clear exploitation of the omnipresence of the COVID-19 'brand'. In this section, we discuss the controversial rationales for domain registrations, as well as some public policy takeaways and 'lessons learned'. To end, we discuss a bit further the role of threat lists, and provide some limitations of the results from this study.

A. Questionable Rationales

Diverting traffic through HTTP redirects, parking domains, and selling goods on Amazon, can all be non-malicious activities and have no direct legal consequences. Only when used in particular for malicious purposes, these actions become punishable. While this study shows no direct proof for such malicious intent, the rationale behind domain registrations described in this paper is highly questionable.

We have shown that around 13% of the domains studied purely function as a redirect towards other domains. When analyzing the top redirecting domains, it became clear that most actors use this mechanism as means of boosting traffic to one or more services. While this is again no violation of any law, nor a flaw of system, it does portray an act of impropriety by leveraging the topic of a global crisis to promote services completely unrelated to the COVID-19 pandemic.

B. Public Policy Take-aways

Leveraging current events to increase the success rate of malicious campaigns is no novel phenomenon. Recent work on the effectiveness of phishing campaigns has presented numerous examples of online scams exploiting the ubiquity of a particular matter [8]. The paper by Williams & Polage notes that prominent events may influence the emergence of such scams, of which phishing is a notorious example.

What we can extrapolate from this is that salient events fuel the modus operandi of attacks based on social engineering – i.e., familiarity increasing credibility. Thus, encountering such topics in the content of a phishing email, ransom demand, or domain name, will likely keep occurring, to which the COVID-19 pandemic is no exception. As such, we identify three preventative measures that can counteract domain registration rationales described in this paper. While we present these measures in the context of a pandemic, we would like to stress the parallels with any global event wherein large-scale exploitation of 'brand awareness' is used.

Registration Policy By the end of March, domain registrar NameCheap stated that it would no longer allow registrations of domain names related to COVID-19 [9]. This resulted in a significant decrease of new registrations per day. In the future, these restrictions on domain name registrations can be applied pro-actively and across the sector. Of course, (malicious) registrations will remain a 'cat and mouse game' between registrants and the registrar's blacklist.

Domain Reselling Trying to predict popular domain names is not only a challenge for the registrants, but also for the registrars. Even with the previous policy in place, many will still succeed into acquiring a domain. We speculate that a lot of these registrations are exploited for domain reselling, treating the domain name like an asset on a fictitious stock market. This incentive can easily be rendered useless if domain reselling platforms prohibit transactions concerning current popular key words – e.g., COVID. A major advantage of this policy would be that it can be put into place after the fact, at a time when the community has established which domain names are 'lucrative' at the moment.

Cross-sector cooperation To battle questionable rationales in domain registrations, and to unlock the potential of the aforementioned policies, global and cross-sector cooperation is necessary. For example, when official news regarding names,

Authorized licensed use limited to: TU Delft Library. Downloaded on June 30,2021 at 12:21:59 UTC from IEEE Xplore. Restrictions apply.

vaccines, measures, etcetera is released, registrars should take the necessary actions in order to avoid abuse in domain name registrations. Here, registrars should work together, and not compete for selling popular domain names piggybacking the 'brand name' of a virus. Finally, this cooperation should include domain reselling platforms in order to prevent domain owners who did manage to register a domain to make any profit from reselling.

C. Transparency of Threat Lists

While there clearly exists value in flagging domains to prevent any potential harm, the results of this paper have shown that such categorizations should be adhered to carefully. Especially during times of disarray, such as the COVID-19 pandemic, it is of crucial importance that the right domains remain available. Simply flagging all domains containing the word 'covid' would obviously lead to a lack of information being distributed, presumably causing an increase in chaos.

The DomainTools Threat List tries to mediate in this trade-off by providing a Risk Score to each domain, calculated by their proprietary algorithm. One of the limitations of this study is that we only looked at the index page of a domain. While we did not find any direct proofs of malicious behavior, it does not mean malicious activities are not pursued originating from paths other than the index page. Nonetheless, merely a Risk Score will not differentiate between the two. It would therefore help, especially in times of turbulence, to have more information available on the reason why a domain is granted a specific Risk Score. More transparency would be a tremendous benefit in assuring the spread of legitimate information.

D. Limitations

We briefly touched upon the limitations of the data set in Section III. We now discuss the impact of these limitations on our results, as well as other limitations stemming from our analysis.

The data set used for this study has a bias. The DomainTools Threat List – our starting point – only contains domains with a Risk Score of 70 or higher, and therefore does not provide an entire overview of the full landscape.

Also our analysis has limitations. Apart from analyzing web page content for parked domains and Amazon/Shopify affiliations, we did not perform an in-depth content analysis at scale. This limits us to draw conclusions on, for example, what people are selling on their Shopify page, and whether it is related to COVID-19. Possibly, such analysis could also reveal some of the intent behind the domains that were not categorized by one of our rationales. Additionally, because we limited the additional lookups to a demarcated set of keywords, the numbers reported in this paper function as a lower bound.

Another limitation of the analysis is the fact that we only looked at the index page of a domain. Although our results suggest that nothing malicious seem to happen on these pages, the DomainTools Risk Score indicates that the domains are not benign. Though we do not exclude the possibility of false-positives in their list, it is very likely that malicious activity is happening on other paths of the domain.

We also did not look into specific registrars known for domain parking. As we were interested in the general trend, and expected a lot of actors to just jump on the bandwagon for registering covid-related domains. Another yet to be answered question is what happens with resold domains and if they are actually bought by others; is domain parking in this context a good use of resources, or is it a waste of time and money? Our analysis did not follow-up further on the parked domains to see how they do in the domain-reselling market. This could be a topic for future work.

VI. Related Work

Some work on the impact of covid-19 in the cyber world has already been done [10]–[12], of which the study of Kawaoka et al. [13] is most closely related to ours. In their work, the authors also examined covid-related domains (using a different data set), specifically for their correlation with covid-19 outbreaks, as well as through content analysis. Similar to our study, the authors could show a clear correlations with domain registration trends and public events. Furthermore, the authors could also show that very little domains were actually used for malicious purposes, which is again in line with our findings.

Another similar study on the authenticity of covid-related domains examined 303 websites and demonstrated that indeed a big number of domains were 'squatting', a result we have proven in this paper on a much larger scale [14]. Furthermore, the authors argue for the need of a restricted TLD for covid-related domains, which we see as a great addition to our policy suggestions from Section V.

Media influence on cyber criminals and cyber attacks has been reported before by Ghiëtte and Doerr [15]. In their work, the authors show how media coverage correlates with an increase in port scan traffic from unseen actors. This increase in probes showed how adversaries are quick to jump on a novel attack vector, reported by the public news, to take profit before systems all around can patch their software.

Also HTTP redirects have been studied before, specifically for its use in malicious drive-by-download attacks. Mekky et al. developed a methodology that detects the use of HTTP redirects towards malicious websites [16]. This phenomenon has also been covered by other studies [17], [18], and [19].

Another popular use of HTTP redirects is in combination with "typosquatting", which is the act of registering misspelled versions of popular domain names. An analysis was done on such domains [20]. The authors show that 20% is being used for static redirects, and that typosquatting often goes hand-in-hand with domain parking. Similar conclusions regarding typosquatting were drawn from a study by Wang et al. [21].

The concept of domain parking has been studied in earlier work with the aim of both understanding the ecosystem, as well as developing detection mechanisms. Alrwais et al. have shown that within the ecosystem of domain parking services, lot's of fraud exist aimed at exploiting both domain owners, as well as advertisers [22]. Their analysis consisted of infiltrating within the ecosystem, and tracking the monetary flow to

Authorized licensed use limited to: TU Delft Library. Downloaded on June 30,2021 at 12:21:59 UTC from IEEE Xplore. Restrictions apply.

estimate gains and losses. Furthermore, Vissers et al. [23] performed an extensive analysis of parked domains, concluding that these domains offer a lot of exposure to malware and elaborate scams. Additionally, the authors provide a highly accurate classifier that detects such parked domains on a client site level.

A more extensive study on domain registration behavior has been done by Coull et al. [24]. In their work, the authors show how parking popular or in-demand domain names can make large amounts of financial profit. Apart from advertisement, the authors also mention domain reselling as a highly lucrative business. As such, we suspect similar motives for the parked domains found in this study.

VII. Conclusion

In this paper we investigated the early trends in covid-related domain registration, leveraging the DomainTools Threat List and additional measurements.

We were able to link the trend of registering covid-related domains to defining events reported in the media during the first months of the pandemic. Additionally, we observed two key rationales for domain registration. First, we have shown how covid-related domain names are used for marketing purposes by redirecting a set of domains to one website. Likewise, we detected various Shopify instances and Amazon affiliation links, possibly serving the increasing demands of supplies during the COVID-19 crisis. Second, we demonstrated that the majority of the domains appear parked, rendering them without any direct use or service. We suspect that these parked domains are used for 'cybersquatting', meaning that the owner aims to make a profit on reselling the domain name later on.

Next, we believe that these rationales can be counteracted by 1) having a strict policy of covid-related domain registrations, 2) regulating the reselling of covid-related domains, and 3) having a global cross-sector cooperation. Although these three policies are COVID-19 specific, we can draw parallels to other exploitations of salient events.

References

[1] "John Hopkins University statement on malware disguised as covid-19 map." https://releases.jhu.edu/2020/03/17/johns-hopkins-university-statement-on-malware-disguised-as-covid-19-map/. Accessed: 2020-05-18.

[2] "Popular johns hopkins coronavirus site targeted by hackers: report." https://www.marketwatch.com/story/popular-johns-hopkins-coronavirus-site-targeted-by-hackers-report-2020-03-12. Accessed: 2020-05-19.

[3] "Sipping from the coronavirus domain firehose." https://krebsonsecurity.com/2020/04/sipping-from-the-coronavirus-domain-firehose/. Accessed: 2020-06-07.

[4] "This man just sold ebola.com for $200,000." https://www.theverge.com/2014/10/24/7058543/this-man-just-sold-ebola-for-200000. Accessed: 2020-05-15.

[5] "The domaintools threat list." https://www.domaintools.com/resources/blog/free-covid-19-threat-list-domain-risk-assessments-for-coronavirus-threats. Accessed: 2020-05-01.

[6] "Covid-19: Why hoarding supplies is human nature, according to a psychologist." https://www.weforum.org/agenda/2020/04/evolution-coronavirus-covid19-panic-buying-supplies-food-essentials/. Accessed: 2020-05-17.

[7] "Domain risk score overview." https://www.domaintools.com/resources/videos/domain-risk-score-overview. Accessed: 2020-05-01.

[8] E. J. Williams and D. Polage, "How persuasive is phishing email? the role of authentic design, influence and current events in email judgements," *Behav. Inf. Technol.*, vol. 38, no. 2, pp. 184–197, 2019.

[9] "Namecheap blocks registration of domains with 'coronavirus' and 'vaccine' in the name." https://www.theverge.com/2020/3/25/21194417/namecheap-coronavirus-covid-19-domain-name-ban-registrar-abuse. Accessed: 2020-07-30.

[10] A. Feldmann, O. Gasser, F. Lichtblau, E. Pujol, I. Poese, C. Dietzel, D. Wagner, M. Wichtlhuber, J. Tapiador, N. Vallina-Rodriguez, O. Hohlfeld, and G. Smaragdakis, "The lockdown effect: Implications of the covid-19 pandemic on internet traffic," in *Proceedings of the ACM Internet Measurement Conference*, IMC '20, (New York, NY, USA), p. 1–18, Association for Computing Machinery, 2020.

[11] A. Lutu, D. Perino, M. Bagnulo, E. Frias-Martinez, and J. Khangosstar, "A characterization of the covid-19 pandemic impact on a mobile network operator traffic," in *Proceedings of the ACM Internet Measurement Conference*, IMC '20, (New York, NY, USA), p. 19–33, Association for Computing Machinery, 2020.

[12] M. Candela, V. Luconi, and A. Vecchio, "Impact of the covid-19 pandemic on the internet latency: A large-scale study," *Computer Networks*, vol. 182, p. 107495, 2020.

[13] R. Kawaoka, D. Chiba, T. Watanabe, M. Akiyama, and T. Mori, "A first look at COVID-19 domain names: Origin and implications," in *Passive and Active Measurement - 22nd International Conference, PAM 2021, Virtual Event, March 29 - April 1, 2021, Proceedings* (O. Hohlfeld, A. Lutu, and D. Levin, eds.), vol. 12671 of *Lecture Notes in Computer Science*, pp. 39–53, Springer, 2021.

[14] N. Tombs and E. Fournier-Tombs, "Ambiguity in authenticity of top-level coronavirus-related domains," *Special Issue on COVID-19 and Misinformation 1 the Harvard Kennedy School (HKS) Misinformation Review*, 2020.

[15] V. Ghiëtte and C. Doerr, "How media reports trigger copycats: An analysis of the brewing of the largest packet storm to date," in *Proceedings of the 2018 Workshop on Traffic Measurements for Cybersecurity*, WTMC '18, (New York, NY, USA), p. 8–13, Association for Computing Machinery, 2018.

[16] H. Mekky, R. Torres, Z. Zhang, S. Saha, and A. Nucci, "Detecting malicious http redirections using trees of user browsing activity," in *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, pp. 1159–1167, 2014.

[17] Z. Li, K. Zhang, Y. Xie, F. Yu, and X. Wang, "Knowing your enemy: understanding and detecting malicious web advertising," in *the ACM Conference on Computer and Communications Security, CCS'12, Raleigh, NC, USA, October 16-18, 2012* (T. Yu, G. Danezis, and V. D. Gligor, eds.), pp. 674–686, ACM, 2012.

[18] J. Zhang, C. Seifert, J. W. Stokes, and W. Lee, "ARROW: generating signatures to detect drive-by downloads," in *Proceedings of the 20th International Conference on World Wide Web, WWW 2011, Hyderabad, India, March 28 - April 1, 2011* (S. Srinivasan, K. Ramamritham, A. Kumar, M. P. Ravindra, E. Bertino, and R. Kumar, eds.), pp. 187–196, ACM, 2011.

[19] M. Cova, C. Krügel, and G. Vigna, "Detection and analysis of drive-by-download attacks and malicious javascript code," in *Proceedings of the 19th International Conference on World Wide Web, WWW 2010, Raleigh, North Carolina, USA, April 26-30, 2010* (M. Rappa, P. Jones, J. Freire, and S. Chakrabarti, eds.), pp. 281–290, ACM, 2010.

[20] T. Moore and B. Edelman, "Measuring the perpetrators and funders of typosquatting," in *Financial Cryptography and Data Security, 14th International Conference, FC 2010, Tenerife, Canary Islands, Spain, January 25-28, 2010, Revised Selected Papers* (R. Sion, ed.), vol. 6052 of *Lecture Notes in Computer Science*, pp. 175–191, Springer, 2010.

[21] Y. Wang, D. Beck, J. Wang, C. Verbowski, and B. Daniels, "Strider typo-patrol: Discovery and analysis of systematic typo-squatting," in *2nd Workshop on Steps to Reducing Unwanted Traffic on the Internet, SRUTI'06, San Jose, CA, USA, July 7, 2006* (S. M. Bellovin, ed.), USENIX Association, 2006.

[22] S. Alrwais, K. Yuan, E. Alowaisheq, Z. Li, and X. Wang, "Understanding the dark side of domain parking," in *23rd USENIX Security Symposium (USENIX Security 14)*, (San Diego, CA), pp. 207–222, USENIX Association, Aug. 2014.

[23] T. Vissers, W. Joosen, and N. Nikiforakis, "Parking sensors: Analyzing and detecting parked domains," in *22nd Annual Network and Distributed System Security Symposium, NDSS 2015, San Diego, California, USA, February 8-11, 2015*, The Internet Society, 2015.

Authorized licensed use limited to: TU Delft Library. Downloaded on June 30,2021 at 12:21:59 UTC from IEEE Xplore. Restrictions apply.

[24] S. E. Coull, A. M. White, T. Yen, F. Monrose, and M. K. Reiter, “Understanding domain registration abuses,” *Comput. Secur.*, vol. 31, no. 7, pp. 806–815, 2012.

Authorized licensed use limited to: TU Delft Library. Downloaded on June 30,2021 at 12:21:59 UTC from IEEE Xplore. Restrictions apply.