THE DOMAINTOOLS REPORT

SPRING 2017

INTRODUCTION

In the DomainTools Reports, we explore various "hotspots" of malicious or abusive activity across the Internet. To date, we have analyzed such varied markers as top level domain (TLD), Whois privacy provider, domain age, patterns of registrant behavior, and more. In each case, we found patterns across our database of over 300 million (315M+ as of this writing) active domains worldwide; these patterns helped us pinpoint nefarious activity, at a large scale, in ways that are similar to methodologies used by security analysts and threat hunters at smaller scales to expose threat actor infrastructure.

It is a truism that threat actors (e.g. malware, phishing) and "nuisance actors" (e.g. spam) often follow observable patterns in their habits. Network defenders use this to their advantage to map infrastructure being used against them; a simple example is the nefarious domain registrant who uses the same registration information for multiple domains that are used to mount attacks. Similar patterns play out at Internet scale, and some of these patterns have already proven useful in predicting whether a given set of domains, sharing a particular pattern, are likely to be risky.

For this report we re-ran some of our original analyses to see how the world has changed—or stayed the same—since publication of our first report in the spring of 2015. By identifying trends, we may be able in turn to extrapolate from them to develop high-level predictions about how threat actors may use Internet infrastructure in the future.

KEY FINDINGS

TOP-LEVEL DOMAINS

The top level domains (TLDs) with the highest concentrations of malicious activity are **a brand-new slate this year,** some with truly alarming concentrations of bad domains (for example, **over 60% of the domains in .science have been blacklisted)**. Many of this year's top-ten were not yet open for registration in 2015. Because new TLDs have been coming online at a great rate, we expect to continue to see a lot of churn in the rankings of the TLDs in future reports.

WHOIS PRIVACY

Among Whois privacy services, there was also considerable change; **only three out of the 2017 top-ten were in the tops in 2015**. One provider, onamae.com, leads in concentration of malicious activity and also has a relatively high volume, with **over 111,000 malicious domains** as of this writing. It is worth bearing in mind that the use of Whois privacy, by itself, is not a strong predictor of maliciousness in domains.

FREE EMAIL PROVIDERS

Among free email providers (which are represented in the registrant email fields in domain Whois records), many of this year's top-ten for concentration of malicious domains were also in the 2015 top rankings. One, however, stands out: **mynet.com was unranked in 2015 and rose to 1st place in 2017, with over 60% of the associated domains on blacklists**.

IP GEOGRAPHY

The geographical concentrations of badness also changed, with **four of this year's top-ten not having been ranked at all in 2015**. The top country for concentration of blacklisted domains was, in 2015 and 2017**, Cambodia, where approximately 80% of domains hosted in the country have been blacklisted**.

METHODOLOGY

As in our original report, we examined four domain characteristics, to see what patterns emerged in the amounts and rates of nefarious activity tied to those characteristics: **TLD, Whois privacy provider** (for those domains registered with privacy), **free email provider** (for registration contact email addresses), and **IP geolocation** of the IP addresses associated with the domains. Using well-known blacklist providers, we analyzed the counts of blacklisted domains versus neutral domains, for each of the four characteristics. This gave us both absolute numbers of bad domains and ratios of good to bad. Because all of the domain Whois records in our database are parsed and the database is cross-indexed, we can perform analysis on these characteristics at the scale of nearly the entire Internet. One of the data points for this report, IP address, does not come from Whois records; the DomainTools databases include a variety of data types reflecting registration, hosting infrastructure, web server metadata, and more.

We looked at four particular types of nefarious activity: **spam, phishing, botnet, and malware**. To be sure, some domains may not fall neatly into one of the categories—for example, a phishing domain might host malware, and might even receive botnet callbacks for command and control. Regardless, this approach allowed us to identify concentrations as well as high absolute populations of nefarious domains along each of the dimensions we studied.

We created some minimum thresholds in order to look at meaningful hotspots. For example, if a given attribute has two domains and both are blacklisted, the 100% concentration level is not necessarily interesting, since one's odds of sending traffic to one of those domains are quite low. The thresholds we used were as follows:

>> TLD: at least 1,000 malicious domains
>> Privacy providers: at least 50,000 total domains
>> Free email providers: at least 500 malicious domains
>> Geography: at least 100 domains

Why not use the same thresholds for all? Each criterion had its own intrinsic properties that informed our thresholds. For geography, for example, we chose a low threshold because high concentrations of malicious domains in a given country could be an interesting signal for a researcher or network defender.

Each analysis will contain a **DomainTools Report VCP Chart** (Volume, Concentration, Proportion), which show the following data for each attribute:

>> Ranking of overall numbers of malicious domains
>> Ranking of concentration of malicious domains
>> How each item stacks up against the overall averages
>> Proportions of each type of badness within each item

In the VCP Charts, items in the upper right quadrant are of particular interest, as they are both above average in volume and concentration. These are the best indicators of malicious domains when looking at that attribute. Each chart plots the total number of malicious domains on the X-axis vs the concentration of malicious domains on the Y-axis, using a logarithmic scale. Each mark is a pie chart showing the relative proportion of types of malicious activity found in each of those items. The total size of the pie charts represents the relative volume of malicious domains. The crossing grey lines show 95% confidence intervals around the averages for each axis.

DISCUSSION

TOP LEVEL DOMAIN: BRAND NEW HOTSPOTS OF BADNESS

There has been in recent years a profusion of new generic top level domains (gTLDs) such as .bike, .guru, and over a thousand others. These gTLDs allow considerable customization and specialization of domain names, and have represented the greatest change to the namespace since the inception of the Internet. But since new gTLDs are continuing to open for registration, and receive populations of names, we have observed a great deal of churn in the statistics.

NOW: 2017

2017	TLD
1	science
2	study
3	racing
4	stream
5	men
6	review
7	click
8	gdn
9	download
10	cricket

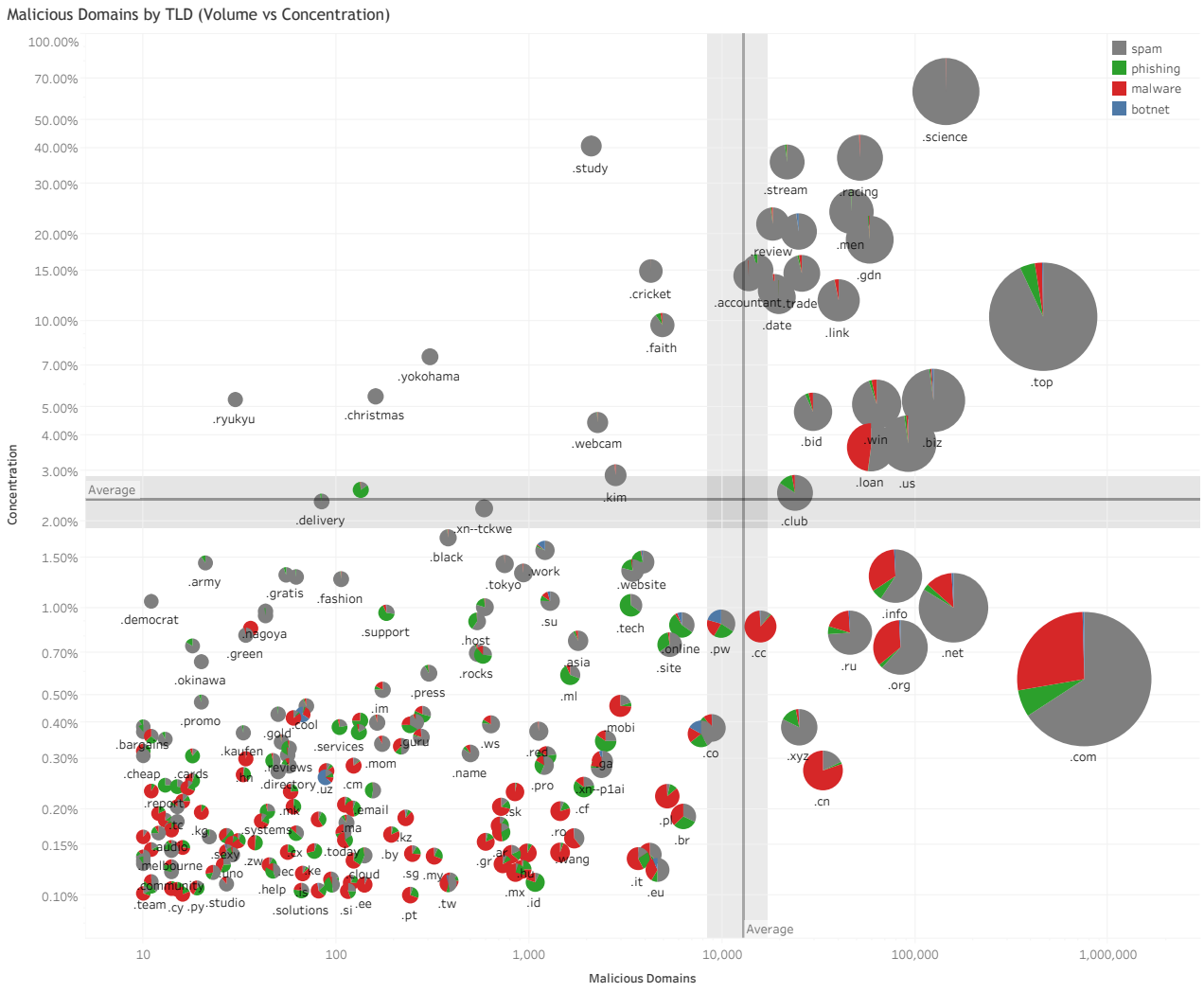
THEN: 2015

2015	TLD	2017	Position Change
1	link	15	-14
2	cf	140	-138
3	us	25	-22
4	biz	21	-17
5	rocks	63	-58
6	asia	59	-53
7	club	29	-22
8	ru	57	-49
9	ga	99	-90
10	co	89	-79

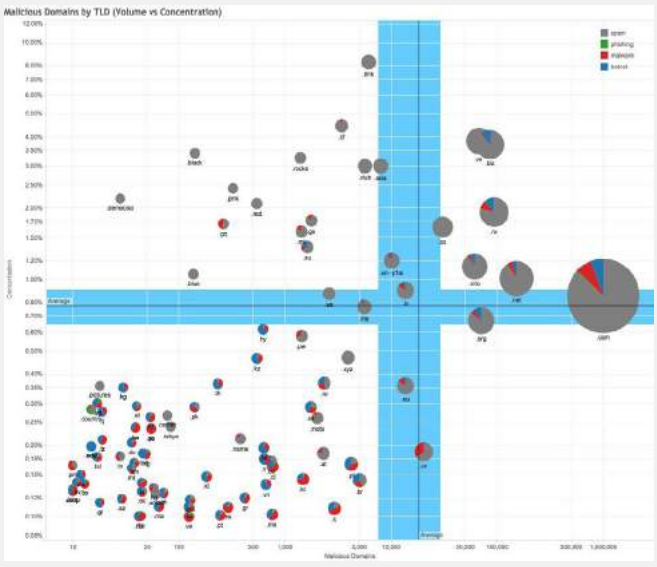
All of the top 10 are new gTLDs, whereas in 2015, some of the top TLDs had existed for years (.biz, .us, .ru, .ga, and .co).

None of the 2017 top-ten TLDs was in meaningful operation in 2015. We expect a lot of continued churn among the TLDs for the foreseeable future as the gTLDs continue to fill out, but that should not stop investigators from paying attention to the top-ten for this year. While this ranking should by no means be construed as an indictment of any gTLD, it helps illuminate patterns of activity by malicious actors. If a given domain is in .science or .study, this in and of itself does not mean the domain is especially likely to be malicious or abusive. Having said this, it is worth noting that in .science, of the ~230,000 domains in the TLD, over 144,000 have been blacklisted, and even more noteworthy, perhaps, is that the blacklisted domains in .science are dominated by a single registrant. Similarly, the blacklisted domains in the .racing TLD are also largely the work of a single registrant entity.

Combined with other evidence, being in one of these TLDs raises the overall statistical risk/nuisance profile of the domain compared to the general population of domains, or to other TLDs such as .com, .net, etc.



2015 Chart for Comparison



What's Changed?

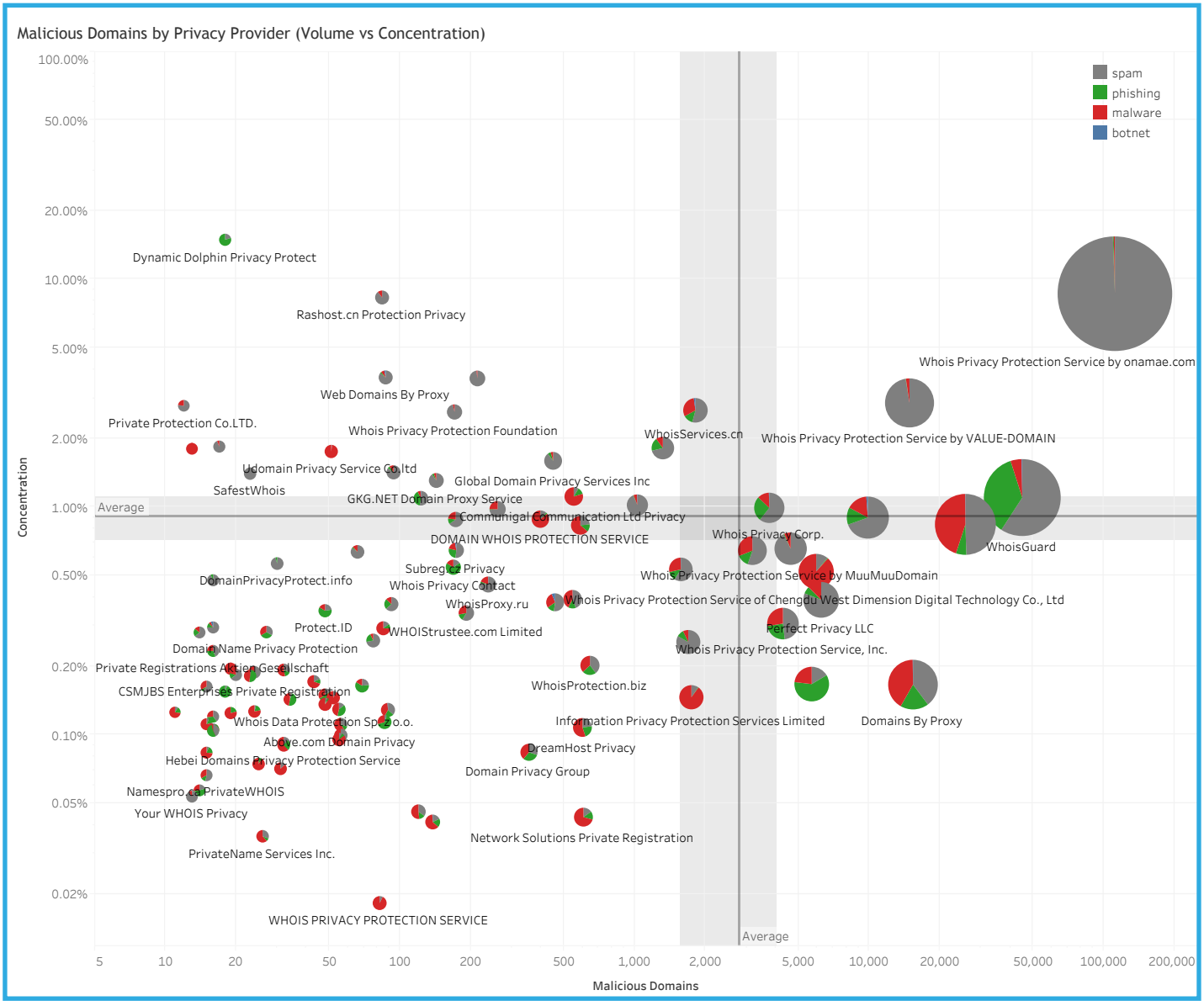
The 2017 VCP chart shows that .com is a bit lower on the concentration axis, and that malware is a larger proportion of the total badness in .com than in 2015. It's also interesting to note that the average for concentration has changed as well, so .com has moved down both in terms of absolute concentration and relative to average concentration for 2017. Notable, too, is that while .top didn't make the top 10 for concentration, it has a large absolute volume of bad domains—and its ~10% concentration is not trivial.

PRIVACY PROVIDERS: PERNICIOUS PROXIES?

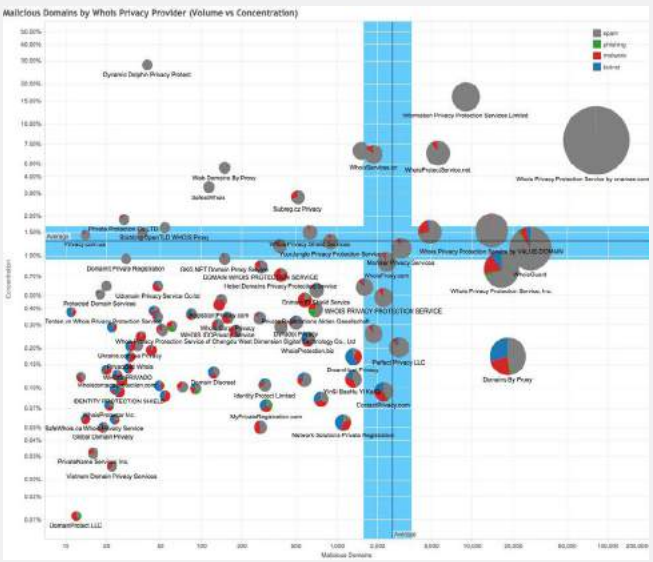
Our research has found that the simple use of Whois privacy is not a reliable signal of badness. While there is a very slightly higher incidence of private registration among domains blacklisted for malicious or abusive activity, it is not statistically significant. However, when it comes to which specific privacy provider is used, we have observed much stronger signals. These are the Whois privacy providers with the highest rates of malicious activity associated with them:

2017	Privacy	2015	Position Change
1	Whois Privacy Protection Service by onamae.com	2	1
2	Whois Privacy Protection Service by VALUE-DOMAIN	4	2
3	WhoisServices.cn	NR	n/a
4	WhoisProtectService.net	3	-1
5	WhoisGuard	8	3
6	Moniker Privacy Services	7	1
7	Whois Privacy Corp.	5	-2
8	PrivacyProtect.org	11	3
9	YinSi BaoHu Yi KaiQi	28	19
10	Domain ID Shield Service	14	4

Because all but one of these providers also existed at a meaningful level in 2015, we can also show their changes in position. This year's inauspicious champion, the service provided by onamae.com, was #2 last time around. Clearly, there is some convergence between the preferences of miscreants and the services offered by this provider, though what that specific convergence may be is beyond the scope of our research. There have also been some significant rises in the ranks, with YinSi BaoHu Yi KaiQi rising 19 places, and WhoisServices.cn coming literally out of nowhere to claim the #3 spot in our rankings. Researchers and network defenders performing forensics on known-bad or suspicious domains that are privacy-protected would do well to take the specific provider into account when assessing the risk profile of those domains.



2015 Chart for Comparison



What's Changed?

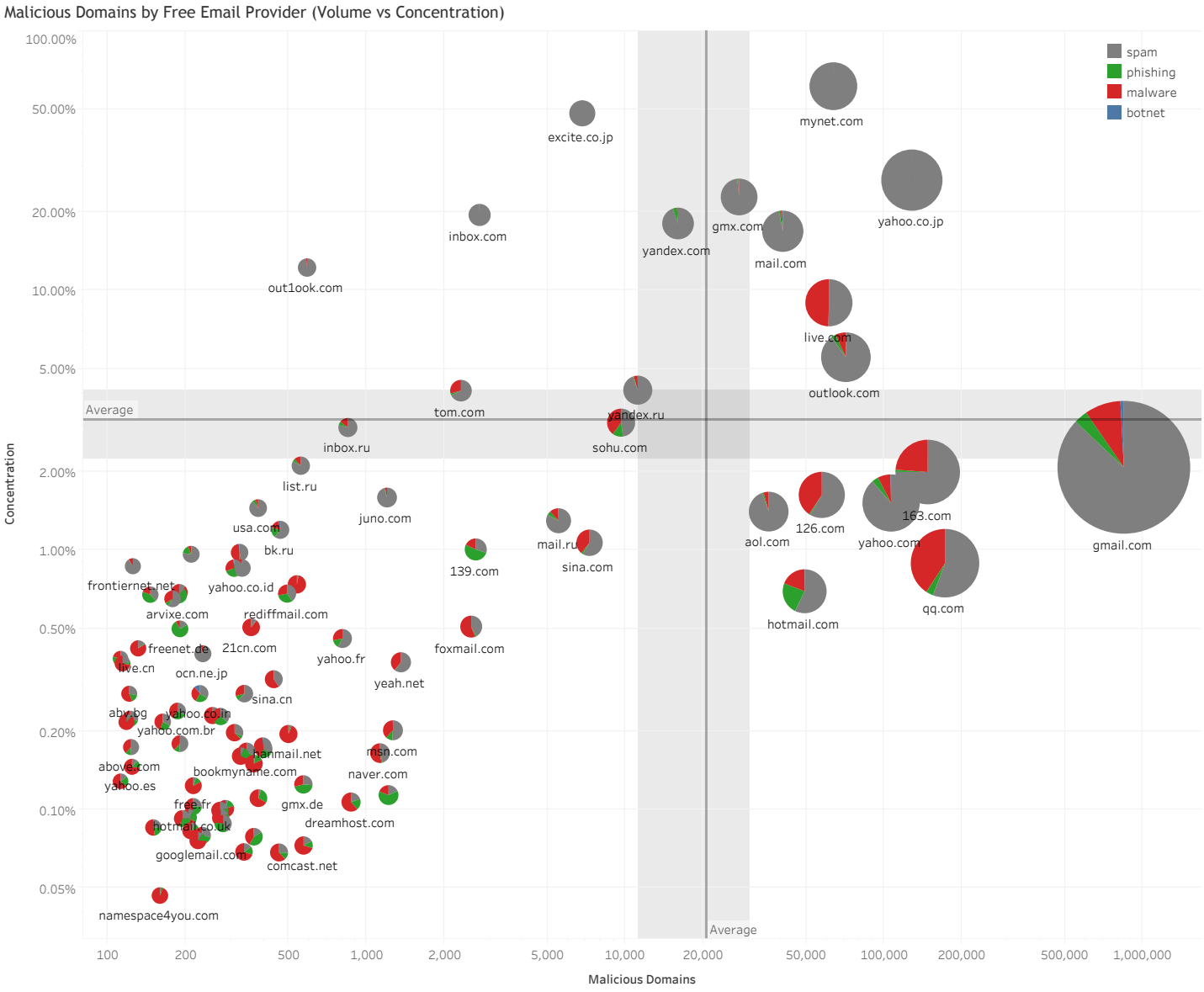
The upper-right quadrant was vacated by Information Privacy Protection Services and WhoisProtectService.net—perhaps they have lost favor with some nefarious actors. What hasn't changed may stand out more: the relative positions and sizes of onamae and VALUE-DOMAIN are very similar in both charts. VALUE-DOMAIN does appear to have diversified a bit, however, with phishing being a larger contributor to its total in 2017 than in 2015.

FREE EMAIL PROVIDERS

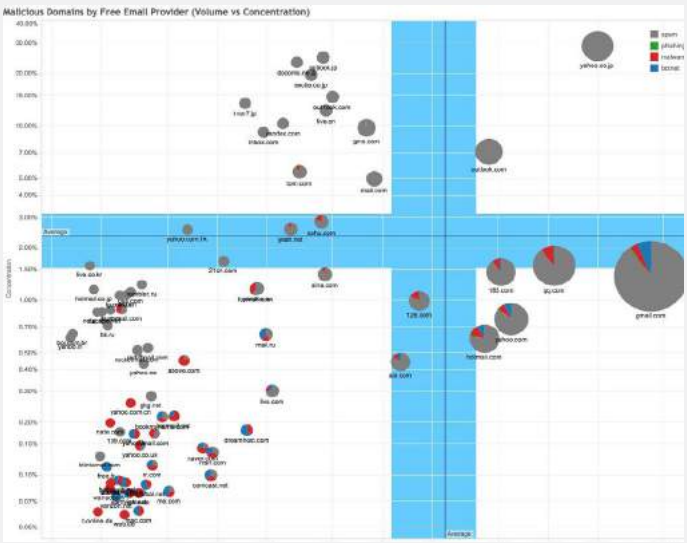
Another datapoint that we analyzed in the Whois records of the populations of neutral and nefarious domains was the email provider favored by the domain registrants for the contact information given at registration. As with privacy providers, the mere use of a free email provider is not in and of itself a predictor of badness. Countless neutral (and downright beneficial) domains are registered with email addresses from gmail, yahoo, and other major free email providers. But, also as in the case of the Whois privacy providers, some free email providers were much more strongly represented in the population of malicious domains than others. Of all 104,000+ domains registered with mynet.com email addresses, for example, almost 64,000 have been "convicted" of illicit activity. Similarly to what we observed in the TLDs, a small number of registrants (which in fact could be aliases of the same entity) dominates the blacklisted domain entries using mynet.com email addresses.

2017	Email	2015	position change
1	mynet.com	NR	n/a
2	excite.co.jp	4	2
3	yahoo.co.jp	1	-2
4	gmx.com	8	4
5	inbox.com	9	4
6	yandex.com	7	1
7	mail.com	12	5
8	out1ook.com	5	-3
9	live.com	27	18
10	outlook.com	10	0

Yet another parallel between email and Whois privacy providers is that there was not as much churn in the top ten as we saw with TLDs. Seven of the top ten were also in 2015's list. But there were a couple of significant rises: mynet.com came from not being represented in 2015 all the way to the dubious distinction of top slot this year, and live.com showed a significant increase in the rates of unsavory domains linked to it. While it bears repeating that the use of any of these providers, including mynet.com, is not proof that a domain is dangerous, many of the actual concentrations are extremely high. Only one of the top ten had a lower than 10% incidence of observed bad activity among the domains connected to it. Our research, especially in the development of the DomainTools Reputation Engine, shows that many more domains with close associations to blacklisted domains, are nefarious. This means that it is a near-certainty that the actual incidence of malicious activity is higher than the numbers show. This can be described in simple form as "guilt by association:" if a given registrant, using (for example) a mynet.com email address, registers 100 domains and 90 of them have been convicted, it is all but certain that the other 10 are also bad. Registrants of bad domains very rarely have a mix of legitimate and illegitimate holdings.



2015 Chart for Comparison



What's Changed?

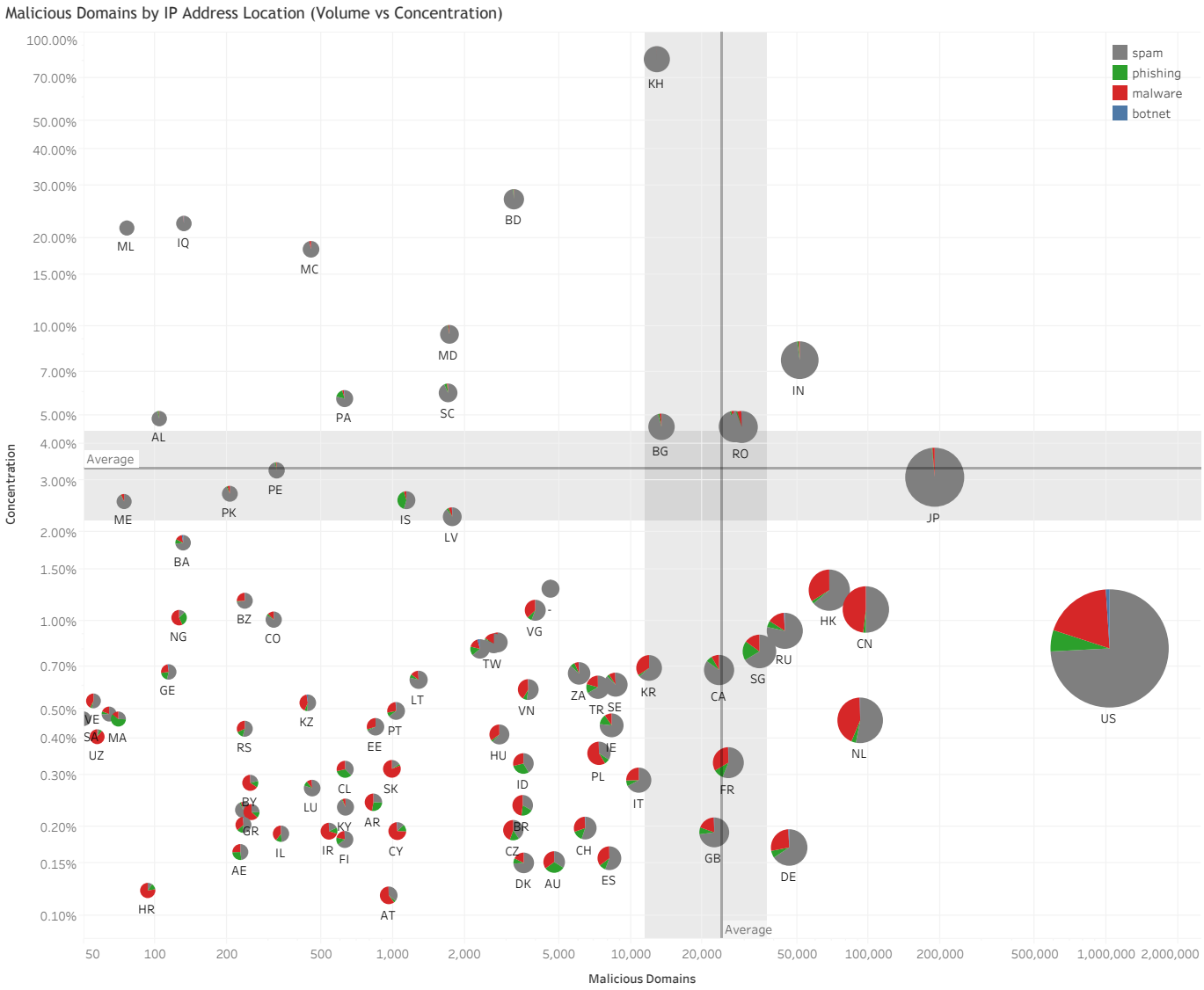
The "quadrant of shame," where the highest combinations of concentration and volume are found, picked up several entrants in 2017 compared to the previous report. While gmail and yahoo.co.jp retained relatively similar volumes, concentrations, and proportions, the entrance of mynet.com and the rise of mail.com, live.com, gmx.com, and outlook.com is immediately apparent.

IP GEOGRAPHY

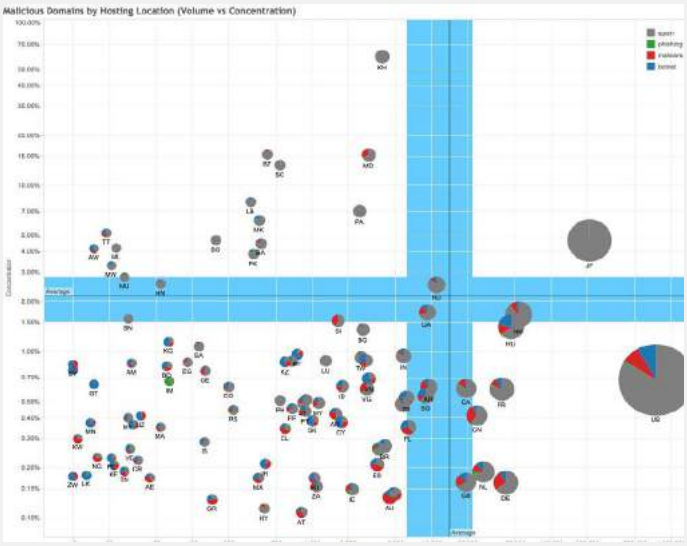
The country in which a domain is hosted can be of obvious forensic or defensive interest to security teams. When network or host logs show communication from a trusted network to a country that is a geopolitical hotspot, or at any rate one in which the organization has no business interests, SOC teams often go on alert. For this reason, we chose a low threshold, just 100 domains, in our analysis of malicious activity rates. If a given country had a low number of overall domains but a high concentration of illicit ones, we reasoned, then by definition any traffic to those domains would be of interest. These are the country codes with the highest concentrations of illicit domains:

2017	Code	Country	2015	position change
1	KH	Cambodia	1	0
2	BD	Bangladesh	NR	n/a
3	IQ	Iraq	NR	n/a
4	MC	Monaco	NR	n/a
5	MD	Moldova	3	-2
6	IN	India	20	14
7	SC	Seychelles	4	-3
8	PA	Panama	6	-2
9	AL	Albania	NR	n/a
10	UA	Ukraine	13	3

As with TLDs, this top-ten list saw a great deal of churn over the last two years. Bangladesh, Iraq, Monaco, and Albania all went unrepresented in our 2015 report because they did not meet the threshold (and thus they did not receive any ranking at all for that year). Cambodia, repeating in the #1 position, has an alarming 80% concentration: of the approximately 16,000 domains hosted there, almost 13,000 have already been observed to be illicit. In the case of Cambodia, the vast majority of these "convictions" are for spam, so it is perhaps not as alarming from the security point of view as Moldova, where a significant proportion of the blacklisted domains are associated with malware.



2015 Chart for Comparison



What's Changed?

The rise of India in volume and concentration stands out, while the other significant geographies of the US and Japan stayed relatively constant between the two reports, in all three dimensions (volume, concentration, proportion). Though less dramatically than India, Romania also rose in terms of absolute numbers, and the proportion of spam compared to the other categories also rose in Romania-hosted domains.

CONCLUSIONS AND NEXT STEPS

As in each of the previous editions of the DomainTools Report, we have found patterns in Internet-wide analysis of Whois and hosting records that both locate (logically and physically) and characterize concentrations of unsavory activity. The TLD space is in a very clear state of rapid change, but even if our next TLD analysis shows a large change in the top ten, the extremely high blacklist rates in this year's group (none had a concentration of less than around 15%) makes those TLDs worth watching. Likewise, geography and free email provider showed some very high concentrations of malicious activity, making them valuable forensic or defensive criteria in the examination of domains seen in traffic logs. Only the Whois privacy providers showed relatively low concentrations, with all but the #1 provider showing sub-10% concentrations.

We will continue to monitor both the absolute numbers and the trends of these four domain attributes, as well as others that we have examined in other editions of the DomainTools report (such as domain age, name server domain age, registrant behavior, and more), to help paint an ever more-detailed picture of the logical and physical hotspots of dangerous or nuisance activity on the Internet. In the meantime, we hope that this analysis proves helpful to researchers, network defenders, and any other reader with an interest in the ever-changing Internet security landscape.

ABOUT DOMAINTOOLS

DomainTools helps security analysts turn threat data into threat intelligence. We take indicators from your network, including domains and IPs, and connect them with nearly every active domain on the Internet. Those connections inform risk assessments, help profile attackers, guide online fraud investigations, and map cyber activity to attacker infrastructure. Fortune 1000 companies, global government agencies, and leading security solution vendors use the DomainTools platform as a critical ingredient in their threat investigation and mitigation work.

For more information about DomainTools' data and products, please visit our website at **www.domaintools.com** or give us a call at **206-838-9020**.

WORLD LEADER IN DOMAIN AND DNS INTELLIGENCE

- >> 310 Million+ Current Domain Names
- >> 13 Billion+ Domain and IP Observations
- >> 10 Billion+ Historical Domain Profiles
- >> 200K Domain Observations Per Second
- >> 5 Million+ New Domain Profiles Daily