14 September 2017

This is a public comment about the study "Statistical Analysis of DNS Abuse in gTLDs," commissioned by ICANN's Competition, Consumer Trust, and Consumer Choice Review Team ((https://www.icann.org/news/announcement-2017-08-09-en). We make these comments in our capacities as cybersecurity professionals and researchers.

In general, the paper is a fine piece of work. It appears to be careful in its planning and execution, and it marshaled an impressively large set of data. The authors deserve credit for making an important contribution to the ICANN community's understanding of abuse issues, and the paper should be considered seriously.

One of the study's most important observations is that a great deal of abuse tends to be concentrated at a relatively small number of registrars and registries. We have documented this phenomenon for phishing domains in our Global Phishing Surveys for the Anti-Phishing Working Group over the past nine years. The new study confirms the phenomenon across several types of abuse.

The question is *why* abuse tends to be concentrated at a few registrars and in a few registries. Unfortunately the paper's analysis of "how different structural and security-related properties of new gTLD operators influence abuse counts" is the weakest part of the study. Below we comment on this topic.

The authors concluded that "inferential analysis reveals that abuse counts primarily correlate with stricter registration policies." It is true (and intuitive) that TLDs with open registration policies have more abuse in them, while TLDs that have restricted availability (or are completely closed) have less abuse in them. However, that is probably not the *primary* and most important correlation. Instead, the critical determining factor is probably *price*. What we see is that cheap, unrestricted TLDs are the ones that tend to attract abuse.

gTLDs that have registration restrictions also tend to sell at higher retail prices. These domains offer exclusivity, which is charged for accordingly. In restricted gTLDs, the combination of higher price *and* restricted access discourages registrations by malefactors.

Historically we have seen documented examples of how very low domain prices have attracted swarms of abuse. These include:

- The abuse problems enabled by free and low-priced .INFO domains in 2003-2006.
- In March 2007, CNNIC significantly reduced the cost of .CN domains to one yuan (US$0.13). The low price helped .CN grow explosively, but there was collateral damage: by August 2007 phishers were registering large numbers of .CN domains for their own use. A price increase and the imposition of a more restrictive registration policy ended the spate of abuse.
- Free domain registrations in .TK attracted a great deal of phishing, spam, and other types of abuse from 2006 to 2012.
- In 2015 through 2017, a great deal of abuse appeared in new gTLDs that sold at the low end of the market. At various points these included .XYZ, .SCIENCE, and more recently .TOP.

The new gTLD program created intense new competition, and some registries and registrars decided to compete based on low price. It is not surprising that abuse then appeared at those registries and registrars. As we have observed in the past: cybercriminals are rational actors driven by a profit motive.

Some require large numbers of domains. They often use funds in their possession rather than stolen credit card numbers to register domains, and are therefore price-sensitive since they're spending money they have in-hand. And they naturally choose to register where the registrars and/or the TLD operators are inattentive or ineffective at fighting abuse. Low price can be a differentiator in the market, but low price can enable a great deal of abuse and harm, especially when it's not coupled with well-executed abuse monitoring and mitigation.

The paper's authors noted: "In our models, we also considered the name of the registry operators to capture systematic differences in the policies of registries across new gTLDs such as pricing, bulk registration options, etc. In other words, we tested the correlation between registry operators and domain abuse counts. However, we did not find any statistically significant effects on the abuse counts."

This conclusion is contradicted by other data that suggests that abuse tends to cluster amongst the gTLDs of certain registry operators (not back-end providers) and registrars. Some TLD operators are "portfolio players" that run many TLDs each. Each portfolio operator has its own abuse profile. A portfolio operator may apply similar pricing and abuse-handling strategies across multiple gTLDs in its portfolio, thereby raising or lowering the abuse risk in that set of gTLDs.

Regarding registrars, there has been quite a bit of documented evidence that particular resellers have sold large numbers of abusive domains. Thus *some* registrars with reseller models have had issues with abuse out-of-line with their peers, while others have not.

The paper's authors noted that "In future work, we plan to collect detailed data on registration policies across all new gTLDs and perform a more fine-grained analysis on factors that may also influence abuse counts." We agree that the relationship between abuse and pricing, and the relationship between abuse and registry and registrar business practices (including abuse monitoring and mitigation), deserve further scientific study. These are important but under-studied areas that need detailed research.

As the authors noted, the study's methodology under-counts the number of malicious domain registrations. The authors counted a registration as "malicious" only if it was blocklisted within three months of creation. We recently documented how phishers sometimes "age" their domains in order to get better domain reputation scores. Similarly, we often see spammers wait four or more months before using their domains, and this activity sometimes involves many domains. We believe that virtually all domains used for spamming (such as those listed by SURBL and Spamhaus) were "maliciously registered," by spammers. Such blocklists tend to list very few compromised domains. Similarly, the vast majority of "command and control" domains used by malware are maliciously registered, since the miscreants behind those malware attacks need to retain full control of their infrastructure, particularly their domains.

Sincerely,

--Rod Rasmussen
  Principal, R2 Cyber
--Greg Aaron
  Vice-President, iThreat Cyber Group