Anticipating New Spam Domains Through Machine Learning

Originally published at https://www.unite.ai/anticipating-new-spam-domains-through-machine-learning/

Researchers from France have devised a method for identifying newly-registered domains that are likely to be used in a 'hit and run' fashion by high-volume email spammers – sometimes, even before the spammers have sent out one unwanted email.

The technique is based on analysis of the way that that the Sender Policy Framework (SPF), a method of verifying email provenance, has been set up on newly-registered domains.

Thanks to the use of passive DNS (Domain Name System) sensors, the researchers were able to obtain near real-time DNS data from Seattle-based company Farsight, yielding SPF activity for TXT records for a range of domains.

Using a class weight algorithm originally designed for processing imbalanced medical data, and implemented in the scikit-learn machine learning Python library, the researchers were able to detect three quarters of the pending spam domains within moments, or even in advance of their operation.

The paper states:

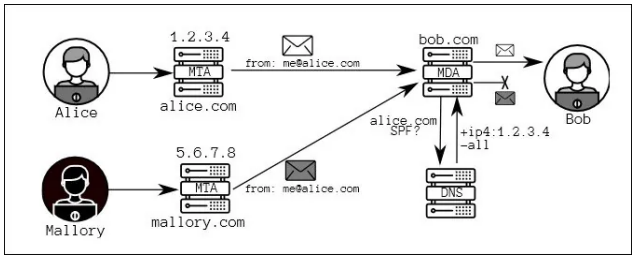
> *'With a single request to the TXT record, we detect 75% of the spam domains, possibly before the start of the spam campaign. Thus, our scheme brings important speed of reaction: we can detect spammers with good performance even before any mail is sent and before a spike in the DNS traffic.'*

The researchers claim that the features used in their technique could be added to existing spam detection systems to increase performance, and without adding significant computation overhead, since the system relies on SPF data passively inferred from near real-time DNS feeds that are already in use for different approaches to the problem.

The paper is titled *Early Detection of Spam Domains with Passive DNS and SPF*, and comes from three researchers at the University of Grenoble.

SPF Activity

SPF is designed to avoid the spoofing of email addresses, by verifying that a registered and authorized IP address has been used to send an email.



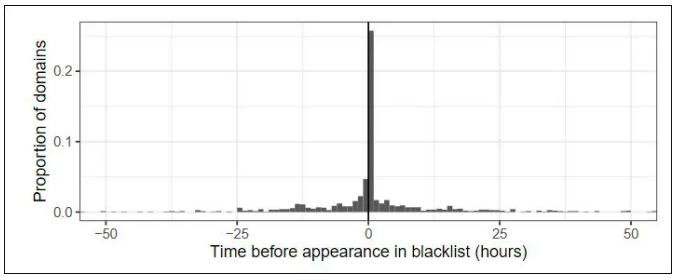
In this example of SPF, 'Alice' sends a benign email to 'Bob', while the attacker 'Mallory' tries to impersonate Alice. Both are sending mail from their own domains, but only Alice's server is registered to send Alice's mail, so Mallory's spoof is thwarted when his fake mail fails SPF verification. Source: https://arxiv.org/pdf/2205.01932.pdf

Other methods of email verification include DomainKeys Identified Mail (DKIM) Signatures, and Domain-based Message Authentication, Reporting, and Conformance (DMARC).

All three methods must be registered as TXT records (configuration settings) at the domain registrar for the authentic sending domain.

Spam and Burn

Spammers exhibit 'signature behavior' in this regard. Their intention (or, at least, the collateral effect of their activities) is to 'burn' the reputation of the domain and its IP addresses by blasting out bulk mail until either action is taken by the network providers selling these services; or the associated IP addresses are registered with popular spam-filter lists, making them useless for the current sender (and problematic for the future owners of the IP addresses).



A narrow window of opportunity: the time, in hours, before a new spam domain is banned and made useless by SpamHaus and various other monitoring services.

When the domain location is no longer practicable, the spammers move on to other domains and services as necessary, repeating the procedure with new IP addresses and configurations.

Data and Methods

The domains studied for the research cover the time period between May and August of 2021, as provided by Farsight. Only freshly registered domains were considered, since this accords with the *modus operandi* of the persistent spammer.

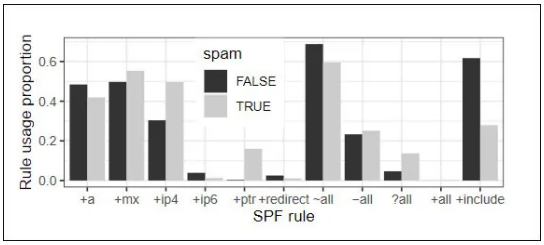
The domain list was built over data from the ICANN Central Zone Data Service (CZDS). Blacklist information from the SURBL and SpamHaus projects was used to effect near real-time identification of potentially problematic new domain registrations – though the authors concede that the imperfect nature of spam lists can lead to benign

domains accidentally being categorized as potential sources of bulk mail.

After capturing DNS TXT queries to the newly registered domains found in the passive DNS feed, only queries with valid SPF data were retained, providing the ground truth for the algorithms.

Stage	Queries	Unique domains	Spam domains
1. Traffic to new domains	399M	14M	0.8%
2. SPF traffic	36M	1.4M	1.5%
3. Ground truth	26M	40,224	5.9%

SPF has a number of usable features; the new paper has found that while 'benign' domain owners most commonly use the *+include* mechanism, spammers have the highest usage of the (now deprecated) *+ptr* feature.

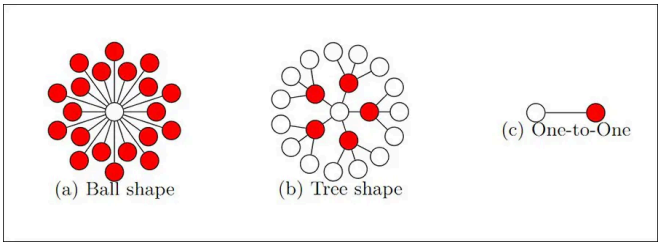


SPF rule usage of spammers, compared to standard usage.

A +ptr lookup compares the IP address of the sending mail to whatever records exist for an association between that IP and the hostname (i.e. GoDaddy GDDY 0.74%). If the hostname is discovered, its domain is compared to the one that was first used to reference the SPF record.

Spammers can exploit the apparent rigor of +ptr to present themselves in a more credible light, when in fact the resources needed to conduct at-scale +ptr lookups cause many providers to skip the check entirely.

In short, the way that spammers use SPF in order to secure a window of opportunity before the 'blast and burn' operation begins, represents a characteristic signature that can be inferred by machine analysis.



Characteristic SPF relationships for spam domains.

Since spammers often move to very nearby IP ranges and resources, the researchers developed a relationship graph to explore the correlation between IP ranges and domains. The graph can be updated almost in real time in response to new data from SpamHaus and other sources, becoming more useful and complete over the course of time.

The researchers state:

> *'The study of these structures can highlight potential spam domains. In our dataset, we found [structures] in which dozens of domains used the same [SPF] rule and the majority of them appeared on spam blacklists. As such, it is reasonable to assume that the remaining domains are likely to have not yet been detected or are not yet active spam domains.'*

Results

The researchers compared the spam domain detection latency of their approach to SpamHaus and SURBL over a 50-hour period. They report that for 70% of the spam domains identified, their own system was faster, though conceding that 26% of the identified spam domains did appear in the commercial blacklists in the following hour. 30% of the domains were already in a blacklist when they appeared in the passive DNS feed.

The authors claim an F1 score of 79% against ground truth based on a single DNS query, while competing methods such as Exposure can require a week of preliminary analysis.

They observe:

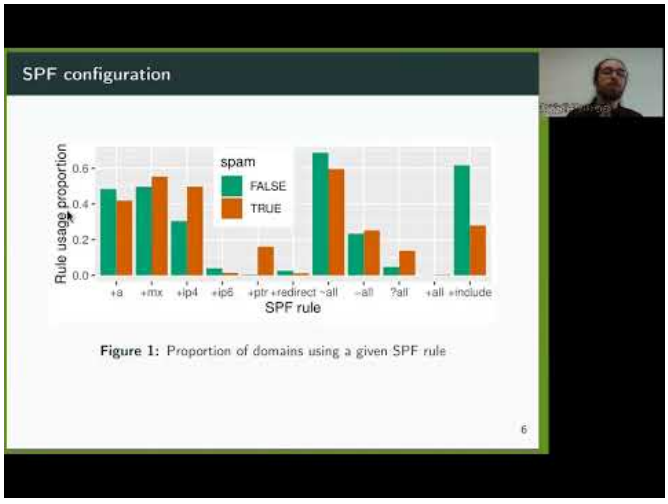
> *'Our scheme can be applied at early stages of a domain life cycle: using passive (or active) DNS, we can obtain SPF rules for newly registered domains and classify them immediately, or wait until we detect TXT queries to that domain and refine the classification using hard-to-evade temporal features.'*

And continue:

> *'[Our] best classifier detects 85% of spam domains while keeping a False Positive Rate under 1%. The detection results are remarkable given that the classification only uses the content of the domain SPF rules and their relationships, and hard to evade features based on DNS traffic.*
>
> *'The performance of the classifiers stays high, even if they are only given the static features that can be gathered from a single TXT query (observed passively or actively queried).'*

To see a presentation on the new method, check out the embedded video below:

Paper: Early Detection of Spam Domains with Passive DNS and SPF



First published 5th May 2022.

Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.
Portfolio site: martinanderson.ai
Contact: martin@martinanderson.ai

Discover More