

# Mitigating the Impact of Spams by Internet Content Pollution

Sabyasachi Roy\* Abhinav Pathak\* Y. Charlie Hu  
Purdue University, West Lafayette, IN 47907  
{roy0, pathaka, ychu}@purdue.edu

## 1 Introduction

In recent years, there has been a steep rise in the amount of unsolicited-emails (spams) [11]. Such mails overwhelm users’ mailboxes, consume server resources and cause delays to mail delivery. Many techniques [2, 10, 12, 5, 13] have been used for mitigating spams. Despite the plethora of schemes proposed, all of them have the cardinal problem of **false positives** which compromises the reliability of emails. Furthermore, many of such schemes are plagued with security, privacy, deployability and transparency issues.

In this project, we propose a new spam-mitigation approach that is orthogonal and complementary to the previous schemes: *it reduces the spams reaching the mailboxes of real users by misleading spammers into spamming non-existing mailboxes*. To send spams, spammers need a set of victim addresses ( $V$ ). From the perspective of spammers’ resource utilization, it is imperative that the set  $V$  consists largely of valid addresses. It has been observed that to create the set  $V$ , spammers primarily use two techniques: (1) crawling the Internet (homepages, newsgroups) [9], and (2) guessing email addresses with the hope of hitting on valid ones [11]. In this paper, we (1) hypothesize that majority of spams reaching a user’s mailbox is because the user’s address is harvested; (2) perform experiments to confirm hypothesis; and (3) propose a simple, *false positives free* scheme that mitigates the impact of spam on individual mailboxes by poisoning the address harvesting of spammers.

## 2 Methodology

We first study the relative fraction of spams due to the two most common techniques: harvesting and random guessing of users’ mailboxes. Our study is based on collected data for the number of spams arriving at individual mailboxes for a set of mailboxes ( $M$ ) at Purdue. An email is considered to be a spam if it is flagged as a spam by SpamAssassin [7]. The set  $M$  contains two subsets: long existing addresses ( $L$ ), recently created (created on the day of starting the experiment) addresses ( $N$ ). Thus,

monitoring the addresses in  $L$  gives us the average number of spams ( $l$ ) that arrive at a mailbox due to harvesting as well as random guessing, whereas monitoring the ones in  $N$  gives us the average number of spams ( $n$ ) that arrive at a mailbox due to random guessing only as such addresses do not have any trace in the Internet, and as a result the only way to reach such addresses is random guessing. Thus, we can estimate the fraction of spams due to harvesting as  $f_h = \frac{l-n}{l}$ .

Another subtlety that needs accounting is that spammers can use SMTP [6] responses to spams to randomly guessed addresses to discover newly created addresses. For example, an “user unknown” error message [6] means that the corresponding address does not exist. In other words, random guessing itself can be used as a harvesting mechanism. To discount the above effect from the calculation of  $f_h$ , we monitor the spams reaching the set  $N$  only until it becomes harvested. The event of becoming harvested is recognized by a heuristic that relates it to a sudden surge in the number of spams. The intuition is confirmed in Section 3.

## 3 Preliminary Results

Figure 1 shows the number of spams received by an individual address (mailbox) per day. We show the data worth 30 days corresponding to two mailboxes, one belonging to a faculty member ( $M1 \in L$ ) and the other belonging to a student ( $M2 \in L$ ). The reason for choosing these two types is that the extent of harvesting of a particular address depends on its availability throughout the Internet. It is likely that different people will have different availability of their addresses. The curve “RG” shows the number of spams estimated due to random guessing (from the set  $N$ ). We can observe that the num-

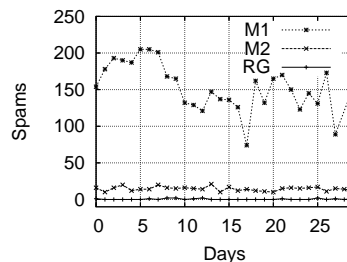


Figure 1: Spams/day at 3 types of mailboxes.

\*Students

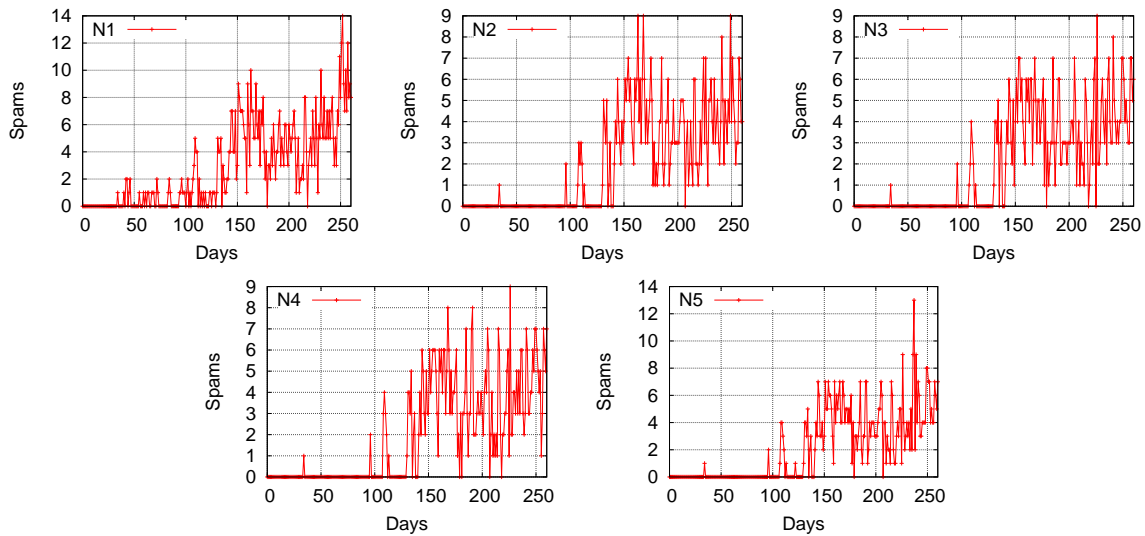


Figure 2: A surge in spams marks harvesting.

ber of spams received at M1 (average 150 per day) is much larger than M2 (average 15 per day). Moreover, the proportion of the spams due to harvesting is much more than spams due to random guessing (average 0.5 per day). This proportion is 99.7% for M1 and 96.7% for M2 calculated using the formulae for  $f_h$ . This shows that the *majority of the spams that common users receive at their mailboxes is due to harvesting*.

Figure 2 shows that the five mailboxes  $N1 - N5$  in the set  $N$  mostly remained idle (receiving 0 – 2 spams per day) for a period of 70 days and then suddenly experienced a rise in the spams. This confirms the intuition behind the heuristic mentioned in Section 2. Following this observation, we considered data for addresses in  $N$  only until 70 days into the experiments in calculating  $f_h$ .

## 4 Proposed Solution and Discussions

We propose a spam-mitigation scheme that leverages the spammers’ resource limit, a fact that is exploited by several previous work [1, 4, 3, 8]<sup>1</sup>. It works by polluting the Internet by advertising (through web pages and newsgroups) a large number of dummy addresses (addresses that do not correspond to any real user but are not bounced). Such dummies poison the address feed (the set  $V$ ) of spammers, and thus mislead the spammers into wasting resources on dummies. This scheme will *mitigate the spams due to harvesting*, which is the *dominant spam contributor* (Section 3). This scheme is transparent, compatible with other anti-spam techniques, and more importantly, does not adversely affect legitimate emails.

We envisage and discuss several major issues in a real

<sup>1</sup>Direct spammers are primarily bounded by their upload bandwidth, while botnets tend to limit their spamming activities to evade detection.

deployment of such a system below.

**Optimized Spamming** Evidence from a spam sink-hole, that we have set up, suggests that spammers often optimize their resources by sending a single spam to a large number of recipients by using cc/bcc facilities provided by SMTP. Our scheme can counter such attempts by aggressively dropping emails whose destinations consist of a fraction of dummies.

**Server Load** Our scheme does not lead to increased server load because spammers are resource bound [1, 4, 3, 8]. All our scheme does is to divert spammers’ resources to unfruitful work. Also, once the scheme is widely deployed, there will not be any uneven load distribution among mail servers of different organizations.

**Advertisement** The circulation of email addresses in the Internet and the consequent harvesting of them by spammers are a consequence of the activities of the users associated with those addresses. Thus, to effectively mislead spammers, the advertisement of dummies needs to be extensive and mimicking real addresses.

## References

- [1] A. Back. Hashcash. <http://www.cypherspace.org/hashcash>.
- [2] G. Cormack and A. Bratko. Batch and online spam filter comparison. In *Proc. of CEAS*, 2006.
- [3] C. Dwork, A. Goldberg, and M. Naor. On memory-bound functions for fighting spam. *CRYPTO, Lecture Notes in Computer Science*, 2729, 2003.
- [4] C. Dwork and M. Naor. Pricing via processing or combatting junk mail. *CRYPTO, Lecture Notes in Computer Science*, 740, 1992.
- [5] S. Garriss, M. Kaminsky, M. J. Freedman, B. Karp, D. Mazieres, and H. Yu. Re: Reliable email. In *Proc. of NSDI*, 2006.
- [6] Rfc 821: Simple mail transfer protocol. <http://www.faqs.org/rfcs/rfc821.html>.
- [7] J. Mason. Reverse-engineering spammer tactics. <http://spamassassin.apache.org/presentations/2004-09-Toorcon/html>.
- [8] P. Nelson, K. Dallmeyer, L. Szybalski, T. Palazr, and M. Wieher. Spamlot: A toolkit for consuming spammers resources. In *Proc. of CEAS*, 2006.
- [9] M. Prince, B. Dahl, L. Holloway, A. Keller, and E. Langheinrich. Understanding how spammers steal your email addresses: An analysis of the first six months of data from project honeypot. In *Proc. of CEAS*, 2005.
- [10] A. Ramachandran, D. Dagon, and N. Feamster. Can dns-based blacklists keep up with bots? In *Proc. of CEAS*, 2006.
- [11] A. Ramachandran and N. Feamster. Understanding the network-level behavior of spammers. In *Proc. of SIGCOMM*, 2006.
- [12] M. Waldfish, J. D. Zamfirescu, H. Balakrishnan, D. Karger, and S. Shenker. Distributed quota enforcement for spam control. In *Proc. of NSDI*, 2006.
- [13] J. Yeh and A. Harnly. Email thread reassembly using similarity matching. In *Proc. of CEAS*, 2006.