ISSN:0975-9646

Effect of User-Unknown Email addresses in spammers' lists

Rajiv Mahajan[1], Dr. Surjit Singh[2], Pavitar Singh[3]

[1]Department of Information Technology, Amritsar College of Engineering & Technology, Amritsar
[2]Department of Information Technology, Amritsar College of Engineering & Technology, Amritsar
[3]Department of Computer Sc & Engineering, Amritsar College of Engineering & Technology, Amritsar

Abstract:

Spam emails are an extra burden over the network and also another Issue for the security of the internet. Today lot of spam mail are coming from the different resources. Many approaches and method are there to filter the spam mail from the inbox like IP blacklisting, Content filtering, greylisting .This Paper describe solutions and analyse the bandwidth of the network by removing the unknown user from the email spammers lists. It is based on an ongoing real world experiment.

1. INTRODUCTION

Spam presents a significant challenge to users, Internet service providers, states, and legal systems worldwide. The costs of spam are significant and growing, and message volume threatens to destroy the utility of electronic mail communication. There is an urgent need to stop this because as per the survey report of December 2006, the Email Sender and Provider Coalition (ESPC) conducted a survey in conjunction with marketing research firm Ipsos to provide insight into the email behaviors of today's consumers. The ESPC surveyed a random sample of 2,252 Internet users from top U.S. ISPs like AOL, MSN/Hotmail, Yahoo!, Lycos, Excite , Gmail, Netscape, Compuserve in order to gauge consumers' behaviors and views toward spam, unsubscribe features and emerging anti-spam technologies.
The results showed that the average American is extremely email-savvy, and most have very specific opinions on email and spam and how to manage both. 73 percent of respondents have used email for six or more years and over 80 percent check their email at least once per day. Those surveyed also showed a familiarity and affinity for using "Report Spam" and "Unsubscribe" features, with over 80 percent of respondents using each of them to manage their inboxes.

Additionally, the results indicate a clear desire by consumers for greater support from ISPs, email providers, and marketers so that they can more easily control their mail experience. Most would like to see tools like "Unsubscribe" and "Report Fraud" buttons (90 percent and 80 percent respectively) added to their email programs. 53 percent of respondents claimed they would be more likely to open and read email if the sending company was certified with an icon displayed in the email inbox.

The message to senders and ISPs/mail program providers is clear. For senders, building trust and confidence are a priority, and these best practices should be followed:

- Give careful attention to the "FROM" address and "SUBJECT" line of emails.
- Make it easier to "unsubscribe" than to "report as spam."
- Use the information provided by recipients who report spam to understand WHY they are dissatisfied with your email program.
- Examine third-party options for certifying your practices.

For ISPs and mail program providers, providing more tools for consumers to control their inboxes is essential. These providers should consider:

- Adding "report fraud" and "unsubscribe" functions to the email interface. Further, consumers would support the sharing of fraud and spam data regionally and globally.
- Giving consumers the opportunity to provide more feedback on why they are reporting email as spam.

Further, consider sharing that information with senders so they can reevaluate their mailing programs.

- Working with senders to provide options for notifying consumers that a sender's practices or reputation has been certified by a 3rd party in the inbox.

The survey results indicate a high awareness and knowledge of the "Report spam" function and its purpose.

- Approximately 83 percent of respondents indicate that they have used a "Report Spam" button.
- 80 percent decide whether to click on the "Report Spam" or "Junk" button without opening the actual message;
- 73% base the decision on "FROM"
- 69% base the decision on "SUBJECT"

- 79 percent of panelists indicate they use the "Report Spam" button when they don't know who the sender is.
- Just 20 percent admit to using the "Report Spam" button as a quick way to unsubscribe.
- 66 percent were willing to provide additional information on why they were reporting something as spam.

Unsubscribing:
Similarly, consumer responses indicate a familiarity and understanding of the unsubscribe process.

- 82 percent of panelists use the unsubscribe features provided when they want to stop receiving email from a company from which they had previously requested to receive email.
- Trust in unsubscribe is high with 71 percent of panelists indicating that they believe unsubscribe links work, and 48 percent of respondents reporting that they use unsubscribe links even when they don't recognize the sender.

Consumer views about their email programs:

Consumers clearly want more tools with which to fight spam and phishing threats.

- 90 percent of panelists indicate that they would appreciate having an "Unsubscribe" button built directly into their email program and indicated they would use such a feature if it were added to their email program
- 80 percent of panelists believe there should be a "Report Fraud" button in their email program.
- Nearly 70 percent believe that information gained from a "Report Fraud" button should be shared across North America; and further nearly 70 percent believe such information should be shared worldwide.

Consumer views about their Junk Folder:

Overall, panelists report that the mail they request to receive is not getting lost in their junk folders.

- 64 percent of panelists report that they rarely or never see messages that they've requested in their bulk boxes.
- 80 percent of panelists report that 5 percent or less of their messages that they requested or wanted to receive land in the bulk folder.

Consumer views about certification of email:

Overall, consumers are looking for help in determining which senders they can trust.

- Respondents would support senders having their practices and policies certified by 3rd parties.
- Respondents are considerably more likely to open and read email from senders whose practices are certified by a 3rd party and identified in the inbox with an icon. While 53 percent would be more likely to open and read such identified email, just 18 percent would not be more likely to open and read the message.

1.1. Methods for Removal

Blacklisting is Perhaps the simplest method in which each incoming request's IP-address on a SMTP-server is tested against a list of known spamming hosts. Almost all big email-providers have already been blacklisted on at least some of the widely available blacklists [3][4].

Content-filters applied to the header and / or the body of a mail message. Filtering is based on a "bad-word-list and scoring-mechanisms to weight words for fine-tuning and maintenance. Spammers are reported to register mail accounts with online services known to have spammed filtering and to test their spam against those filters. This leads to a permanent "one-step-behind"-situation for filters, no matter how advanced content-filtering becomes [5].

Greylisting is another methods which is used in these day which forcing the sending MTA of a message to resend it after a short time. Mostly spam is sent through zombies, usually Windows-PCs infected with some worms like own SMTP-engine, which is usually quite simple. It is difficult to handle through this method and therefore consider this condition as an error and stop delivery.

1.2. Modifying SMTP

The disadvantages of reactive anti-spam-methods discussed above brought the discussion on fixing one of the real causes for spam: SMTP lacks authentication. This offers spammers the chance to remain hidden and to evade lawsuits. So the key approach is to implement some kind of authentication and authorisation. Beside some side-effects seen on current methods, like breaking intended mail-forwarders, the real problem is to enforce the modified standards world-wide by Preventing Harvester, Obfuscation Tar pit.

2. REMOVE AN EMAIL-ADDRESS

There are some efficient, compatible, standard-conform and barrier free ways to obfuscate email addresses. Email obfuscation has to be considered effective, but it has only been tested to work as long as the address to protect has not been published before [7]. [13] Suggested that later obfuscation of an address might also reduce the amount of spam received by Frequency of email-address changes and removing e-mail address from the WebPages. This Section describes and analyse the effect of User Known removal from the e-mail spammers lists.

2.1. User unknown

This approach is to return "User unknown" within the SMTP-dialogue. This approach is only useful if the

error message is generated in the SMTP dialogue, i.e. as soon as the sending party sent the "RCPT TO:"-line, the server should respond with the SMTP status code 550 "User unknown" [21]. This is different to generating the error message later and sending an email containing an error message to the email address claimed to be the sender address, as those addresses are often invalid or forged. The author of this paper for example receives some hundred bounce mails due to an email address of his being abused as sender address. Due to this, rejecting spam with an error message email is not good practice, although often practised.

This so called bounce spam is mostly avoided by not accepting a spam mail during the SMTP-dialogue, as now most spammers use bulk mail software or so called bots they control directly. If they would send their spam through open relays, the relaying server would generate the error message to the presumed sender, thereby generating bounce spam. But, as open relays are blocked at most sites and spammers are interested in getting their spam through spam filters, they avoid using open relays.
Bots, as computers infected with some backdoor software allowing to remote control them and thereby offering the possibility to abuse them as bulk mailers, are called, have usually their own minimalistic SMTP engine implemented. This engine would understand an error-message and stop trying to deliver that mail. This is quite the same mechanism, grey listing relies on. Looking at the offers of producers of bulk mailers, one often discovers some kind of "add-in" or "plug-in" for that bulk mailer. Those plug-ins either offer "subscription management", which is basically some solution to take care of "unsubscribe"-clicks or mails and bounces, if delivered, or they offer an "email validate", first of all validating email addresses syntactically and then by connecting to the remote server and issuing either a "VRFY" or "RCPT TO:" command within the SMTP dialogue1.

Considering this, the inventor of SPONTS, an anti spam appliance [22], [20] decided to reject spam with an "User unknown" message. He claims this would reduce spam on the long term.

To verify this idea, a test server has been set up. On this machine, a mini SMTP server written in Perl has been installed. This server is started from the internet super server xinetd2, thus reducing network interaction to basic input-output-handling, but with a strong impact on performance.

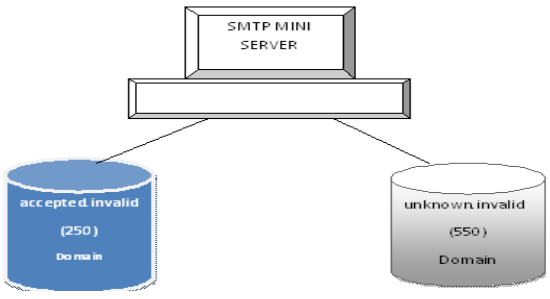

Fig-1: Layout of Setup

The server has been configured to accept email for two domains. For one of them (called accepted.invalid in this document), it will always return SMTP status 250 "message accepted" and for the other domain (called unknown.invalid) SMTP status 550 "user unknown". Both domains have been heavily spammed before; they received an average of 30898 spam mails daily. Then, the mail software has been changed to stop accepting messages for the domain unknown,invalid. Both, before and after changing the behaviour of unknown. invalid, a lot of probing of email addresses has been logged. Almost 20% of all connection attempts were stopped after trying some RCPT TO: within the SMTP dialogue. This supported the presumption that spammers will test their email addresses for validity and will probably take action on unavailable addresses.

During a four month test period, the amount of spam received on each domain kept growing (Table 1).

Durtaion (in Days)	Accepted.invalid	Unkown.invalid
Start	22959	22825
After 15	28430	26590
After 30	31896	33876
After 45	34738	33882
After 60	37856	36132

Table 1 Average daily spam count received per domain

There is no significant difference between the amounts of spam received on either domain. It therefore seems as if "User unknown" is currently ignored by most bulk mailers. However, testing will continue and more up to date results will be communicated and discussed in the final paper and at the conference. Those results are in contrast to results from an informal long-term test with a heavily spammed email address under a domain of the author's: The mail server has been configured to stop accepting mail for this address, after this address received 2350 spam mails daily.

3. CONCLUSION

The method, to send "user unknown", has no reproducible results as far as its effectiveness is concerned. It is also simpler, to stop publishing an email address rather than to implement sending "user unknown" specifically on spam mails: Implementing a spam filter returning "user unknown" during the SMTP dialogue upon delivery of a spam message is far from trivial and requires some advanced techniques. In a real world environment, the existing "spots"-appliance's [20] basic concept is to first finish the SMTP dialogue and, if the message is considered to be spam, to temporarily reject it. This real-world solution has its disadvantages: Most email validates used by spammers would interrupt the first connection after the RCTP TO: has been accepted. The spam-filtering mechanism would not come into effect, as there is no message to be identified as spam. The validator would also reject

another email identified by the same three tuple, even if this new message is not spam.

4. REFERENCES

[1] Cormack, G. and T. Lynam. TREC 2005 Spam Track Overview. in The Fourteenth Text Retrieval Conference (TREC 2005). 2005. Gaithersburg, MD, USA.

[2] Carreras, X. and L. Marquez. Boosting Trees for Anti-Spam Email Filtering. in 4th International Conference on Recent Advances in Natural Language Processing (RANLP-2001). 2001

[3]McWilliams, Brian, AOL lands on spam blacklist, Sebastopol, http://spamkings.oreilly.com/archives/2005/04/aol_lands_on_sp.html, 2005

[4]McWilliams, Brian, SpamCop blocking some Gmail servers, Sebastopol, http://spamkings.oreilly.com/archives/2006/01/, 2006

[5] Zhou, Y., M.S. Mulekar, and P. Nerellapalli. Adaptive Spam Filtering Using Dynamic Feature Space. in 17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'05). 2005.

[6] Graham-Cumming, J. The Spammers' Compendium. 2006 15 May 2006 [cited 2006 May]; Available from: http://www.jgc.org/tsc/.

[7]Eggendorfer, Tobias, Methoden der präventiven Spambekämpfung im Internet (in German: Methods of preventive spam abatement in the internet), Master thesis, Fernuniversität in Hagen, München, Hagen, 2005

[8]Wu, C.-T., K.-T. Cheng, et al. Using visual features for anti-spam filtering. in IEEE International Conference on Image Processing, 2005 (ICIP 2005). 2005.

[9] Damiani, E., S.D.C.d. Vimercati, et al. P2P-based collaborative spam detection and filtering. in 4th IEEE International Conference on Peer-to-Peer Computing (P2P'04). 2004. Zurich, Switzerland.

[10] Albrecht, K., N. Burri, and R. Wattenhofer. Spamato - An Extendable Spam Filter System. in 2nd Conference on Email and Anti-Spam (CEAS'05). 2005. Stanford University, Palo Alto, California, USA.

[11] Yerazunis, W.S., S. Chhabra, et al., A Unified Model of Spam Filtration 2005, Mitsubishi Electric Research Laboratories, Inc: 201 Broadway,Cambridge, Massachusetts 02139, USA.

[12] Ma, W., D. Tran, et al. Detecting Spam Email by Extracting Keywords from Image Attachments. in Asia-Pacific Workshop On Visual Information Processing (VIP2006). 2006. Beijing, China.

[13] Center for Democracy and Technology, Why am I getting all this spam?, Washington, D.C., 2003, http://www.cdt.org/speech/spam/030319spamreport.pdf

[14] Tran, D., W. Ma, and D. Sharma. Fuzzy Normalization for Spam Email Detection in Proceedings of SCIS & ISIS. 2006.

[15] Tran, D., W. Ma, and D. Sharma. A Noise Tolerant Spam Email Detection Engine. in the 5th Workshop on the Internet, Telecommunications and Signal Processing (WITSP'06). 2006. Hobart, Australia.

[16] Ma, W., D. Tran, and D. Sharma. Detecting Image Based Spam Email by Using OCR and Trigram Method. in International Workshop on Security Engineering and Information Technology on High Performance Network (SIT2006). 2006. Cheju Island, Korea.

[17] Tran, D., W. Ma, et al. A Proposed Statistical Model for Spam Email Detection. in Proceedings of the First International Conference on Theories and Applications of Computer Science (ICTAC 2006). 2006.

[18] Sakkis, G., I. Androutsopoulos, et al., A Memory-Based Approach to Anti-Spam Filtering for Mailing Lists. INFORMATION RETRIEVAL, 2003. 6(1): p. 49-73.

[19] Chuan, Z., L. Xianliang, et al., A LVQ-based neural network anti-spam email approach. ACM SIGOPS Operating Systems Review, 2005. 39(1):p. 34 - 39.

[20] IKU AG, Sponts / UCE: Nachhaltige Spam-Abwehr (in German: Lasting spam defence), Saarbrücken, http://www.sponts.de/uce.jsp, 2006

[21]Klensin, John (Editor), RFC2821: Simple Mail Transfer Protocol, o. A., 2001, http://www.ietf.org/rfc/rfc2821.txt

[22]Eggendorfer, Tobias, Spezialfilter. Antispam-Appliance mit Langzeitwirkung (in German: Special filter: Anti spam appliance with long term effects) in: Linux Magazin 09/2004, Linux New Media, München, 2004