Spam 2.0: The Problem Ahead

Vidyasagar Potdar, Farida Ridzuan, Pedram Hayati, Alex Talevski,
Elham Afsari Yeganeh, Nazanin Firuzeh, and Saeed Sarencheh

Anti Spam Research Lab, Digital Ecosystems and Business Intelligence Institute
Curtin University of Technology, Australia
http://asrl.debii.curtin.edu.au
{v.potdar,a.talevski}@curtin.edu.au,
{farida.mohdridzuan,pedram.hayati}@postgrad.curtin.edu.au,
{yeganeh,sarenche,n_firoozeh}@iasbs.ac.ir

Abstract. Webspam is one of the most challenging problems faced by major search engines in the social computing arena. Spammers exploit weaknesses of major search engine algorithms to get their website in the top 10 search results, which results in higher traffic and increased revenue. The development of web applications where users can contribute content has also increased spam, since many web applications like blogging tools, CMS etc are vulnerable to spam. Spammers have developed targeted bots that can create accounts on such applications, add content and even leave comments automatically. In this paper we introduce the field of webspam, what it refers to, how spambots are designed and propagated, why webspam is becoming a big problem. We then experiment to show how spambots can be identified without using CAPTCHA. We aim to increase the general understanding of the webspam problem which will assist web developers, software engineers and web engineers.

Keywords: Webspam, CAPTCHA, Spambot, anti-spam, spambot navigation, Spam 2.0, Pligg spam.

1 Introduction

Spam in the context of email is defined as "*unsolicited, anonymous, commercial and mass email messages*". Spam originated via email and one of the first spam email dates back to early eighties, when a lawyer sent out an advertizing email on a newsgroup. Since then spam has evolved into what we know as spam today. A spammer is defined as "*an entity that is involved in spamming*". Spammers use many different mediums to spam web users, drifting from the traditional email approach to new approaches that are termed as *webspam or as we call it Spam 2.0 [31, 32]*.

Webspam refers to the techniques employed by spammers to spread spam via websites in contrast to using emails [33]. Spammers now use blogs, forums, wikis [30, 34] or even develop their own websites to post advertizing material. Overall the motivation is still the same i.e. to generate revenue, increase page rank, promote product or services and steal user information [1].

D. Taniar et al. (Eds.): ICCSA 2010, Part II, LNCS 6017, pp. 400–411, 2010.

Spammers use a number of techniques to drive traffic to their websites and one of those is to fine tune their websites to deceive search engines in ranking their websites higher [29]. It is quite often seen that when you search for a particular keyword, you are taken to a website which does not relate to what you are looking for, but instead it is an advertizing page designed by spammer [36]. Such websites are carefully crafted to make the search engines believe that it is providing genuine content by implementing keyword stuffing, incorporating fresh content and several other strategies.

With sophisticated anti-spam techniques, it is now possible to get rid of the majority of spam; however a small percentage of spam can still escape these filters. Spammers rely on this small percentage of spam to attract their targets as they expect a portion of people to respond to their spam content. It is this response rate that keeps the spammers active [2]. Hence spammer aims to keep broadcasting very many spam messages on a regular basis as it increases their chances to find new targets. Spam is growing at a rapid pace and has become a big industry, mainly because it costs very little to send out millions of spam messages electronically [3]. According to [4], the cost to post an advertizing comment on a blog is very marginal, making spam campaigns extremely profitable particularly with favourable conversion rates.

This paper focuses on understanding how webspam operates and looks into the implications of webspam on productivity at work, consumption of network resources etc. We then study why spam is such a difficult problem to solve, where we look at technical, social and economic factors that affect spam. Later we outline the current solutions used to tackle spam and finally discuss our experimental results where we analyze spambot behavior.

1.1 How Does Webspam Operate?

Webspam operates by using a number of spambots. A *spambot* is a piece of code that is designed to post advertising comments on web applications like forums [35] & blogs. A collection of spambots forms a network that is referred as a *Botnet*, which is a network of infected machines that operate under the command of a *Botmaster*. Botnets normally use infected hosts to transmit spam [5]. Once the machine is infected it becomes a spam agent. Spammers have developed sophisticated web spamming tools that can automatically identify websites that host a particular type of web applications like blogs from Wordpress, forums from SMF, PHPBB etc.

Spambots can be categorized into *applications specific bots* that target web applications like Pligg, SMF, PHPBB, Wordpress etc and *websites specific bots* targeting websites with high traffic e.g. Amazon, CNET etc. Bots that target specific applications are customized so that can only be used to spam a subset of websites that are developed using a particular web application. These bots do their job perfectly and at times leave no trace of their activity or their origin. We will discuss later in our paper how the honeypot that we developed was able to capture these targeted spambots.

The harvesting activity i.e. finding out targets to spam is also done intelligently and to some extent the web applications are to be blamed. The simplest techniques to find whether a particular website is developed using a specific application is to look for unique text e.g. *"# Published News # Upcoming News # Submit a New Story"* is used

in Pligg. This makes it easy to find targets and start spamming. The Botmaster usually does this job and sends individual bots to the selected sites for spamming.

On the other hand, website specific bots are extremely customized, usually developed for one particular website. The spammers study the structure of the website and develop an attack strategy to craft a bot that can bypass all the anti-spam mechanisms used. Even though these kinds of bots take a longer time to be develop they provide extremely good returns and are most difficult to detect. Hence understanding bot behaviour is becoming a key challenge for the research community and developing a system that can filter out bots from their behaviour is where the technology should be heading.

1.2 What Are the Implication of Webspam?

There have been several reports outlining the loss in labor productivity and network resources [6-10]. Although majority of these studies focused on email spam, this equally applies to webspam also. According to Nucleus research [8], in 2003 an average employee received 13.3 spam messages per day, which equated 6.5 minutes of their time to read and delete. This research also mentioned that for every 72 employees a company lost one employee to spam. In general, spam decreases individual's productivity directly/indirectly. The cost implications of spam show a very gloomy picture too.

- Nucleus research reported the average cost of spam per employee per year is $874
- [7] Estimated that spam is costing organizations $75 billion globally.
- [9] Reported that the labor loss caused by spam mail amounted to 21.6 billion dollars per year.
- [6] Reported that companies lost 20 billion dollars to buy additional servers in 2003 to manage spam.

Spam has implication on the global climate change too. According to a report by McAfee, the global energy consumption to process spam e-mail in 2008 was 33 billion kilowatt-hours (kWh), which can otherwise be used to support at least 2.4 million homes in a year [10]. Having understood the problem and its implications, we now look at why spam has become such a difficult problem to solve.

2 Why Is Spam a Difficult Problem to Solve?

The problem to address spam is not just technical in nature but it has economic and social dimensions too. Therefore, a spam management solution should incorporate technical, economic and social aspects too. In this section we highlight these issues in detail to understand why spamming is a difficult problem to solve.

2.1 Technical Dimension

Number of Bots

Internet is flooded by sheer number of bots that originate from numerous locations, with different technologies and strategies. Even though some of them can be easily

defeated, intelligent bots can recognize dead links, fake email addresses, identify spam traps etc making this a difficult task. Botnets are evolving rapidly because spammers continuously develop new techniques to hide bots [11], hence detecting bots is becoming harder. Moreover, the number of bots is increasing as well; about 30,000 new machines are infected and become bots every day [12]. According to Microsoft research the total number of bot-accounts signed up in hotmail in Jan 2008 is more than three times the number in Jun 2007 [11]. Here are some blacklists from different references showing the number of bots that have been detected as spammers [13-16].

Openness of the Internet

Openness of the Internet is a one key factor that allows spam to proliferate. Spammers can easily take this opportunity to manipulate this freedom since spammers have the same opportunity as other users in using any web application. So, if other users are allowed to write hundreds of comments everyday, spammers could do the same as long as their behavior matches that of a real human. Since no prior authorization is required to post a comment on a blog or forum, spammers can write advertizing comments on blogs without a problem. This inherent freedom allows spammers to take the risk as they have nothing to lose.

Gray Area between Spam and Non-spam

There is no distinct definition to differentiate real content from spam content, since there is no clear boundary between the two. One piece of information that is spam for one user may not be spam for the another, hence defining this boundary is a challenging task that has two major problems:

1. *False-positive*: it occurs when a legitimate content is flagged as spam.
2. *False-negative*: it occurs when spam content is flagged as genuine content.

The damage that can result from false-positives can be very serious and anti-spam solutions should be wary of this problem [18].

Vulnerable Web Applications

Many web applications are vulnerable to spam because these applications did not consider any anti-spam measures when they were built. Numerous web developers blindly use open source software to fast track development cycles; however they do not consider the impacts of adopting a web application without introducing any anti-spam measures. As a result, the number of vulnerable commercial or noncommercial websites has increased exponentially. The majority of these web applications provide features for users to contribute content, which is intelligently exploited by spammers. In addition, personal websites designed by novice developers are extremely vulnerable to spamming as well since they are not familiar with spam attack algorithms. Other than that there is no dedicated phase during the design and development stages of websites that caters for spam control.

Defensive Approach to Spam Management

Current spam management techniques take a defensive approach to address this problem. Many algorithms detect spam once it has entered the system. This approach

is very dangerous in the case that spammers combine their spam content with a virus. Thus there needs to be a change in thinking on how spam is to be managed.

Intelligent Bots
Spambots have become extremely intelligent lately, they are not only capable of finding vulnerable websites and launch targeted attacks on but also devise real time strategies to get the highest return on their spam campaigns. Botnet's are widely spread across the web and largely operated by unsuspecting users' infected machines. Spam works in the favor of bots in three ways, *firstly* these machines are not blacklisted, so traffic originating from these sources is not blocked, *secondly* spammers do not need to pay for the resources and *thirdly* prosecuting the owners of infected machines is very difficult in the current legal system. Other than that, bot detection itself is a problem because of two reasons [19]: *firstly* bot's attacks are transient and *secondly* one specific bot does not send all the spam traffic.

Dynamically Adapting Spam Protection Schemes
Despite a lot of research focusing on developing better spam protection schemes, the nature of spam itself is changing rapidly to adapt to these new schemes. Spammers and the anti-spam companies are evolving simultaneously to adapt to the changing environment in the following ways;

⇒ spammers use keyword stuffing to fill their pages with specific keywords and anti-spam algorithms provide a list of the most used keywords which are being used by spammers.

⇒ spammers try to use unknown IP addresses for spamming and anti-spam software's provide blacklisted IP addresses from where spam originates. As a counter measure, spammers keep changing IP addresses or use proxy servers or infected machines to send spam.

⇒ spammers switch the nature of spam from text to graphic and anti-spam software's now use OCR to interpret textual content from images to detect spam.

Spammers are creative enough to create better content to prevent it from being detected easily by the anti-spam software's. Hence, while spam is changing dynamically, spam still remains an unsolved problem.

Trial and Error
Spambot can be easily created, in fact just a few lines of code is required to create a simple spambot. While the cost of creating spam is relatively cheap, not all spambots are good enough to breakthrough the anti-spam filters. Nevertheless, spamming is mostly based on trial and error. Given nearly infinite resources from infected machines, it is relatively easy for spammers to spam until the desired number of websites are infected within a given spam campaign.

2.2 Economic Dimension

Part Time Spamming Provides Good Income
Some people get hired to spam others and it is a very popular practice in developing countries like India [20], [22-24]. Spam jobs are advertised as email processing job

that looks legal, however the actual idea is to get the spam comments or emails sent out in bulk. Most spamming jobs are advertized as home based work where an individual can work anytime that s/he wants and only an email account is required to do this job. We found some advertisement paying as low as 1 INR i.e. 2 cents per comment added to a website [20]. Since this job is performance driven, one gets paid more if one can spam more. This monetary benefit attracts a large number of people to the spamming business.

Cost to Spam Is Low
Another reason why spamming is difficult to control is that the cost to spam is very low. The majority of the cost is incurred by the infected machines. Unlike sending paper mail, the cost of sending spam is very minimal and the spammer ultimately profit even if a negligible rate of receivers respond to their advertisement [21].

2.3 Social Dimension

Human Spammers
As mentioned previously all bots may not be automated. Humans may be employed in developing countries to break strong CAPTCHA's which are used to limit the access of bots to sites. CAPTCHA generates textual or audio tests that are easy for human to solve, but difficult for bots [25].

Low Level of Awareness
It is also necessary to enhance people's awareness towards fighting spam. Spammer use general topics that interest wide range of the community such as cheap medicines, low priced air fares, low rate mortgages. In other words spammers abuse low level awareness of web users to spam e.g. spammer may suggest buying special product like medicines cheaper from a particular site and people believe it to be true and follow the links to a scam website operated by a spammer. Hence a general lack of awareness amongst the majority of web community is a key factor why spam is flourishing.

3 How Is Spam Managed Currently?

Currently spam is managed by three main approaches; these are Detection Approach, Prevention Approach & Attack Approach. All these approaches have had limited success so far and a reliable spam management strategy should include all the three approaches to achieve maximum spam protection. We now explain each of these approaches in detail.

3.1 Detection Approach: Spam Content Is Identified and Filtered Out

This approach is one of the first approaches developed and implemented to manage spam. The basic idea here is to identify and filter out spam content from genuine content. To achieve this, several techniques have been proposed in the literature and many of them are currently implemented in commercial anti-spam toolkits. For example, detection methods can be categorized into two: content based & metadata

based. The former uses content to analyze spam and hence is computationally intensive and more reliable whereas the latter only uses metadata i.e. links or url or email headers and is relatively fast but comparatively less reliable. Both approaches rely on data mining techniques which can either be supervised, semi-supervised or unsupervised. Supervised methods require a labeled data set for spam classification whereas unsupervised do not [26]. Some anti-spam methods are language dependent and hence may not be able to apply to non English language spam, which can be a problem.

3.2 Detection Approach: Users Flag Content as Spam

This approach is a subset of the detection strategy, where the end users are involved in helping to fight spam. This feature is commonly seen in free email services like Yahoo Mail, Hotmail or Gmail etc., where the users have the option to select an email and tag it as Spam. Lately this feature has also become popular in blogs and forums. This feature is very good, as it can help the anti-spam detection algorithms to build up a spam data set; however, the downside of this approach is that spammers can equally use this feature to tag genuine content as spam. So studying the effectiveness of this method is very interesting. Currently there are no publicly available results to show whether this strategy is working [27-28].

3.3 Prevention Approach: CAPTCHA

Prevention approach was developed to defeat spambots by requesting spambots to go through an online test named as CAPTCHA. The aim of this test was to distinguish human users from spambots. The test requires the user to type in unclear, curvy or ghostly characters from in an image to a registration form. Most users should be able to surpass this test easily but bots would fail even if they use optical character recognition techniques. CAPTCHA is used in almost all commercial emails sites (Yahoo, Google), online forum, blogs, and social networking sites to prevent automated registrations. CAPTCHA also helps bloggers in dealing with comment spam. However there are some drawbacks e.g.:

⇒ it relies on human visibility, hence it is inconvenient for users with bad vision.
⇒ at times it is even very difficult for normal users to decipher the CAPTCHA.
⇒ free CAPTCHA servers incur longer delays in processing
⇒ many spammers have developed OCR techniques to automatically read CAPTCHA.
⇒ as Optical Character Recognition (OCR) techniques improve, CAPTCHA's images become harder and hard to decipher even by humans. This damages the typical users web experience
⇒ as computers get more powerful, they will be able to decipher image and voice CAPTCHA requests similar to humans.

3.4 Attack Approach: Poisoning Spammers Database

This is a relatively new approach to address spam. The basic idea is to infiltrate spammers' database and poison it with fake email address, with an aim to reduce the

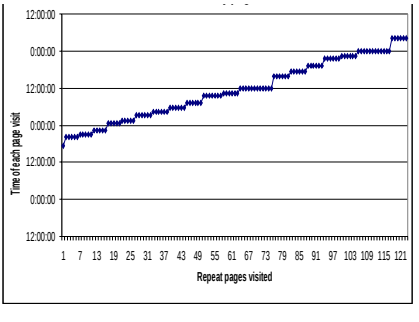
effectiveness of spamming campaigns. So instead of waiting for the spammers to attack, this method takes an active approach towards attacking spammers. This method generates random email addresses and waits for bots to index them. Once the spammer realizes that their database is full of invalid information, it will reduce the effectiveness of their spam campaigns. However the main concern is that whether the list of random emails generated are really invalid or they do belong to someone. An additional concern is that spammers may not bother if their database is poisoned, they may just send spam to all the emails addresses they have, since they are not using their resources any way. WPOISON and SUGARPLUM are examples of such services available on the Internet.

4 Experimental Results

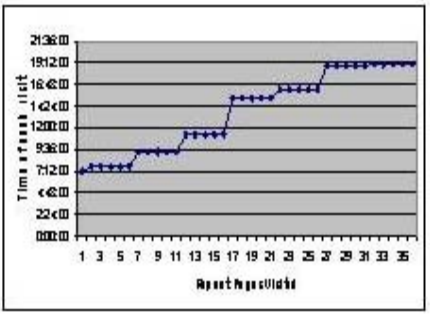
In order to monitor & analyze the spambot behavior we conducted an experiment by setting up a honeypot which had public access to the Internet. We decided to use Pligg as a web application to assess the spambot behavior because Pligg's anti-spam features are very weak. We modified Pligg source code to integrate it with our user navigation tracking (UNT) functionality to track spambot navigation behavior. We tracked: *url visited, session identification, user agent, referral link, start time, last login time, total active time, total time per visit, average time user spends on each session and total active time*. To advertize the honeypot we listed the URL of this honeypot on several sites and we got sufficient spambots. We now explain our 4 main observations from this experiment.

4.1 Observation 1: Spambots Failed Attempts

Over a period from 16th May 2009 to 9th July 2009 we got 412 unique spambot visiting our site. Overall there were 599 visits from different bots i.e. repeat bots. From the total registered users 9 spamuser's attempted to add content on this site. Fig. 1(a) & 1(b) shows failed attempts by two spambots to add comments to Pligg Story Page, since we had changed the HTML code to trap the bot to study its navigation



(a) Bot 1 attempted 21 times to add a comment



(b) Bot 2 attempted 7 times to add a comment

Fig. 1. Spambots failed attempts to add comments to Pligg Story Page

patterns. We found that spambots were programmed to follow a standard path when attempting to add content and that was as follows: Homepage →Show Story →Show Story →Show Story →Show Story.

4.2 Observation 2: Navigation Pattern Detection

When we activated content submission some other spambots started submitting content and had the following navigation pattern User Page → Submit News → Submit News → Submit News → Submit News → Upcoming News. Fig. 2 shows the similarity between navigation behaviors of two spambots. We also created an interface to analyze each record to help us further discover spambots behavior.

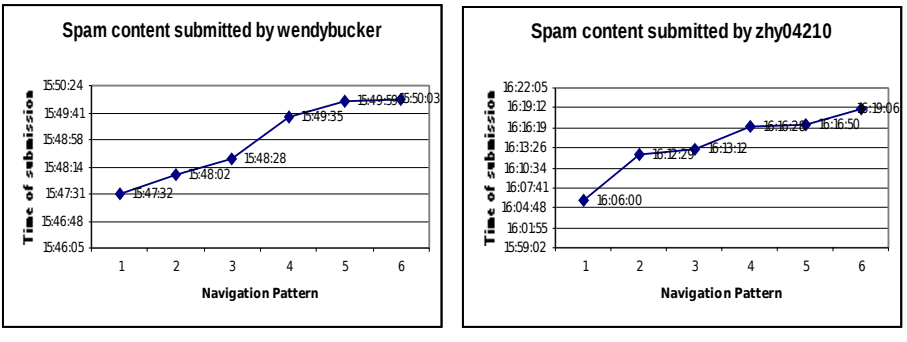


Fig. 2. Bot 3 & Bot 4 submitting spam on the Pligg website following similar pattern

4.3 Observation 3: Normal Navigation Pattern

Some spambots used 8 different user agents or modified the header info i.e. browsers and operating systems to access the Pligg website. The normal navigation pattern for almost all users was as follows: Register →Register →Submit →Submit →Upcoming News.

4.4 Observation 4: Motivation of Navigation Pattern

We assume this pattern of navigation shows that the spambot first wants to ensure that their account has been created, then they want to ensure that their content has been submitted and finally checking whether that content is live or not. If the content is live they consider their job to be complete and do not visit the website for a predefined interval of time.

5 Discussion and Conclusion

We have seen that the current anti-spam filtering techniques are not effective in combating webspam. Purely on a technical front addressing spam problem is going to be a difficult challenge. We have reached a stage where the technical battle between

spammers and anti-spam providers has reached a tipping point. We need to look at how we can solve the spam problem by relying on non-technical measures e.g. increasing awareness amongst the community about spam could be a good alternative. However the effectiveness of such an idea needs to be investigated. Australian media and communication authority has a website on spam, scam and fraud, but how many people actually visit this site is a different question. We believe people would visit the website once they become aware that they have been stung by a spammer, by then it is too late. Alternatively should we be looking at some management or legal solutions to combat spam? May be introducing tough money laundering legislation may be a good choice however the problem arises when spammers (or spam servers) are physically located in a different country where there is no spam legislation. At this moment, spammers are exploiting these legislative loop holes. Other than this, would it be possible to change the monetary equation such that spamming becomes an unattractive job? If some strategies can be investigated along these lines it would be very promising. The issue becomes even worse if money is not in the equation, some research indicates that spammers may be politically or even religiously motivated, in which case their motivations may be similar to computer virus creators. So the question now lies whether spam will infiltrate other communication mediums such as Digital TVs? Radios? Wireless sensor networks?

References

1. Hayati, P., Potdar, V.: Evaluation of spam detection and prevention frameworks for email and image spam: a state of art. In: Proceedings of the 10th International Conference on Information Integration and Web-based Applications & Services, ACM, Linz (2008)
2. Aaron, W.: Ending spam's free ride. netWorker 7(2), 18–24 (2003)
3. Fazlollahi, B.: Strategies for Ecommerce Success, p. 300. IGI Publishing (2002)
4. Chris, K., et al.: Spamalytics: an empirical analysis of spam marketing onversion. In: Proceedings of the 15th ACM conference on Computer and communications security, Alexandria, Virginia, USA, CM (2008)
5. Moheeb Abu, R., et al.: A multifaceted approach to understanding the botnet phenomenon. In: Proceedings of the 6th ACM SIGCOMM conference on Internetmeasurement, ACM, Rio de Janeiro (2006)
6. Group, e., Spam: By Numbers (June 2003)
7. Neal, L.: Vendors Fight Spam's Sudden Rise. Computer 40(3), 16–19 (2007)
8. Nucleus, R.: Spam: The silent ROI Killer. Research Note D59 (2003), 14 July 2009 [cited; Available from: `http://www.nucleusresearch.com`
9. Rockbridge, A.I.: National Technology Readiness Survey: Summary Report 2005 (2004)
10. Vrhnjak, S., Staff, C.: Spam is a big polluter in more ways than one (2009)
11. Yao, Z., et al.: BotGraph: large scale spamming botnet detection. In: Proceedings of the 6th USENIX symposium on Networked systems design and implementation, USENIX Association, Boston (2009)
12. Husna, H., et al.: Behavior Analysis of Spambotnets. In: 3rd International Conference on Communication Systems Software and Middleware and Workshops, 2008. COMSWARE 2008, Bangalore, pp. 246–253 (2008)
13. [cited 13] (July 2009), Available from: `http://www.joewein.net/dl/bl/from-bl.txt`

14. Antispam. [cited 13] (July 2009) Available from: http://antispam.imp.ch/swinoguri-rbl.txt
15. Joewein. [cited 13 July 2009] Available from: http://www.joewein.net/dl/bl/dom-bl.txt
16. Juniper. [cited 13 July 2009]; Available from: http://www.juniper.net/security/spam/
17. Lowd, D., Meek, C.: Good Word Attacks on Statistical Spam Filters. In: Second Conference on Email and Anti-Spam (CEAS), Palo Alto, CA (2005)
18. Cunningham, P., et al.: A Case-Based Approach to Spam Filtering that Can Track Concept Drift. In: Ashley, K.D., Bridge, D.G. (eds.) ICCBR 2003. LNCS, vol. 2689, p. 3. Springer, Heidelberg (2003)
19. Yinglian, X., et al.: Spamming botnets: signatures and characteristics. In: Proceedings of the ACM SIGCOMM 2008 conference on Data communication, ACM, Seattle (2008)
20. Workathome. [cited 13 July 2009]; Available from: http://www.workathomeforum.in/online-adplacing-homejob.htm
21. Leiba, B., Borenstein, N.: A multifaceted approach to spam reduction. In: First Conference on Email and Anti-Spam, CEAS (2004)
22. Cobb, S.: The Economics of Spam (2003)
23. Rich, L.L.: Internet Legal Issues: SPAM (1999)
24. Schwartz, E.I.: Spam Wars (2003)
25. Halprin, R.: Dependent CAPTCHAs: Preventing the Relay Attack, 26 (2009)
26. Hayati, P., Potdar, V.: Toward Spam 2.0: An Evaluation of Web 2.0 Anti-Spam Methods. In: 7th IEEE International Conference on Industrial Informatics (INDIN 2009), Cardiff, Wales, ' (2009)
27. Sheffield, M.: 'Flag Spam,' the Preferred Tool of the Left's Web Censors. 2008 [cited 14,]; 2008/10/07/flag-spam latest-tool-censors-left (July 2009), Available from: http://newsbusters.org/blogs/matthewsheffield/
28. userscripts.org. Flagging Content Feature. [cited 14 July 2009]; Available from: http://userscripts.org/topics/1362
29. Hayati, P., Potdar, V.: Spammer and Hacker, Two Old Friends. In: 3rd IEEE International Conference on Digital Ecosystems and Technologies (DEST 2009), Istanbul, Turkey (2009)
30. Hayati, P., Potdar, V.: Toward Spam 2.0: An Evaluation of Web 2.0 Anti-Spam Methods. In: 7th IEEE International Conference on Industrial Informatics (INDIN 2009), Cardiff, Wales (2009)
31. Hayati, P., Chai, K., Potdar, V., Talevski, A.: Behaviour-based web spambot detection by using Action Time and Action Frequency. In: The 2010 International Conference on Computational Science and Applications, Springer, Heidelberg (2010)
32. Hayati, P., Potdar, V., Chai, K., Talevski, A.: Web Spambot Detection Based on Web Usage Behavior. In: The International Conference on Advanced Information Networking and Applications, AINA 2010 (2010)
33. Ridzuan, F.H., Potdar, V., Talevski, A.: Key Parameters in Identifying Cost of Email Spam. In: The 2010 International Conference on Computational Science and Applications, Springer, Heidelberg (2010)
34. Ridzuan, F.H., Potdar, V., Talevski, A.: Key Parameters in Identifying Cost of Spam 2.0. In: 24th IEEE International Conference on Advanced Information Networking and Applications, AINA 2010 (2010)

35. Sarencheh, S., Potdar, V., Yeganeh, E.A., Firouzeh, N.: Semi-Automatic Information Extraction from Discussion Boards with Applications for Anti-Spam Technology. In: International Conference on Computational Science & its Applications (ICCSA 2010), Springer, Heidelberg (2010)
36. Hayati, P., Chai, K., Potdar, V., Talevski, A.: HoneySpam 2.0: Profiling Web Spambot Behaviour. In: Yang, J.-J., Yokoo, M., Ito, T., Jin, Z., Scerri, P. (eds.) PRIMA 2009. LNCS, vol. 5925, pp. 335–344. Springer, Heidelberg (2009)