THE STATE OF THE ART OF SPAM AND ANTI-SPAM STRATEGIES AND A POSSIBLE SOLUTION USING DIGITAL FORENSICS

Author and co-authors

Franscois R. van Staden[1], H.S. Venter[2]

[1,2]Information and Computer Security Architectures (ICSA) Research Group Department of Computer Science University of Pretoria South Africa

ruan.vanstaden@up.ac.za

Room 5-1.7, Natural Sciences 2, Hatfield Campus, Pretoria, 0001

+27(12)420-4690

hventer@cs.up.ac.za

Room 4-23, Information Technology, Hatfield Campus, Pretoria, 0001

+27(12)420-3654

ABSTRACT

Electronic communication such as email is an efficient and cost effective communication medium in today's connected world. This paper looks at the strategies employed by spam and anti-spam and shows the co-evolution of these strategies. Anti-spam software makes use of intelligent filtering based on content scanning, block lists, black lists, white lists and mailbox authentication. Spammers have been able to get past anti-spam software by using picture content, mailbox- spoofing and anonymous emailing.

Spammers use the strategy of creating botnets to send spam. Honeypots, systems employed to gather information on unusual system

activity, track and ultimately stop the activities of botnets. The paper looks at honeypots as part of information gathering in a digital forensic process. Digital forensic science has been employed to authenticate email authors and back trace email paths. This paper proposes two strategies for the detection of botnet activity and the tracing of botmasters.

KEY WORDS

Spam, Blacklisting, White listing, Bayes, Chi-squared, Botmaster Botnet, Spam-zombie, Honeypot.

THE STATE OF THE ART OF SPAM AND ANTI-SPAM STRATEGIES AND A POSSIBLE SOLUTION USING DIGITAL FORENSICS

1 INTRODUCTION

Spam is an inconvenience to electronic communication. Before an email can be profiled as spam, the Internet Service Provider (ISP) has to download the spam because anti-spam strategies are implemented either on the user's mailbox or on the company's mail servers. The downloading of spam has a direct impact on the bandwidth use of a company. The harmfulness of spam can be calculated in monetary value. The implementation of anti-spam strategies also has its own cost implications. Spam can clog up the electronic communication lines of a company to such an extent that there is a loss of service and therefore a loss of revenue. According to BBC News, the Microsoft security report for 2008 fourth quarter states that 97% of all email sent through Microsoft email servers, is spam (Waters, 2009). Symantec reported that 73.3% of all email sent in February 2009 were spam (InternetNews, 2009).

The state of the art anti-spam strategy makes use of intelligent filtering. Intelligent filtering is build on all the anti-spam strategies developed to date and includes content scanning, black-listing and white-listing. With each strategy developed there are still situations where intelligent filters gives false positives and false negatives. False positives occur when legitimate email is marked as spam e.g. medical correspondence, including black-listed words. False negatives are when spam is not detected by the spam filter because of picture content or misspelling of black-listed words.

The state of the art spam strategy employed is botnets. Botnets are infected PCs, known as zombies, that work together to aid in cyber crimes. The botmaster controls these botnets from a central point. The problem of eliminating botnets and botmasters is firstly to find zombies in the botnet

and then to trace them to the botmaster. It is possible to trace botnet activity and trace the activity back to the botmaster, by using digital forensics.

The remainder of the paper is constructed as follows; section 2 gives background information on spam, anti-spam and digital forensics, section 3 looks at the current state of the art of spam and anti-spam strategies and section 4 proposes implementation strategies that use digital forensics to augment anti-spam efforts. Section 5 is the conclusion of the paper and also discusses future work needed.

2 BACKROUND

This section discusses spam, anti-spam and digital forensics. The discussions are an overview of the different concepts used in this paper.

2.1 Spam and anti-spam strategies

Spam is defined as unsolicited commercial email (Lueg, et al., 2006) or unsolicited bulk emails (O'Brien, et al., 2003). Anti means to be strongly opposed to a person, action or event (University Press). Anti-spam is defined as an application used by an email user or an email server administrator, to reduce the amount of spam the user receives (Network-Dictionary). Anti-spam is defined by the author, as strategies employed to oppose spam. These strategies can be employed separately or together. In the following sections the author explains the technology strategies employed by anti-spam software and the strategies that spammers use to get past the anti-spam strategies. There are also training and awareness strategies but those are outside the scope of this paper.

2.1.1 Content Scanning

Content scanning is implemented as an application on the email server and as an application addition to the users' email application client (spam-site, 2006), (Mueller, 2009). To give email content a spam probability rating, content scanners use a statistical analysis algorithm such as Bayes or Chi-squared (O'Brien, et al., 2003). The algorithm uses key words and key phrases to calculate the spam probability rating. We also refer to these key words and phrases as patterns. This rating categorises the email as spam, possible spam or non-spam. Spam is stopped at the email server. Possible spam is marked with a spam tag but is still sent to the user. Non-spam is seen as normal email.

To get past content scanners, spammers use techniques like picture content , HTML tag inlay and misspelling of patterns (spam-site, 2006) (Naidoo, 2007). Picture content is a series of pixel values and cannot be scanned the same way as text. When content scanners try to scan text, the scanner ignores the pictures. Email servers can be set so that the ISP will only download picture content if the user gives permission for the action. With the advent of mail clients being able to parse HTML, spammers started using tag inlay like "Viagra", to hide patterns. Content scanners can scan content before and after HTML parsing. Misspelling

patterns, replacing characters but still making it recognizable to the reader e.g. “Vi@gra” or “V1@gra” for Viagra, causes content scanners to ignore the patterns. Users need to add these alternative patterns to the lists of the content scanner to block the mail.

2.1.2 Block list, black list, white list and mail box authentication

A block list is build by individual users (spam-site, 2006). Users block email senders or email domains from the users’ own mail application. Block lists block email from being downloaded to a user’s mailbox in future. A black list is generated at ISP or DNS level where email traffic is monitored for indications of bulk mail originating from a single source (spam-site, 2006), (Lueg, et al., 2006). The ISP or DNS will blacklist email domains suspected of sending bulk mail. White listing is when users set up a list of allowed email accounts and email domains to be downloaded to their email boxes. Mailbox authentication relies on the fact that spammers falsify or spoof the “From” and “Reply-to” tags of an email. A message is sent to the mailbox in the “From” or “Reply-to” tag. If either the “From” or “Reply-to” mailbox does not exist, the mailbox cannot be authenticated and the email is marked as spam.

To get past email block lists and mail box authentication, spammers replace the “From” tag with the “To” tag in the email header (spam-site, 2006), (Mueller, 2009). According to the email profiler, the mail appears to originate from the users’ own mailbox. The spam email bypasses the block list since it is assumed that the user would not block its own mail address or domain. As long as the user’s domain is not blacklisted, the spam email bypasses the black list. The spam email bypasses the white list because the white list automatically adds the user’s address when it creates the list. Since the user’s mailbox does exist, the spam email’s “From” mailbox is authenticated.

Faynberg, et al. (2004) proposed a method for authenticating email. Each gateway and relay server authenticates email by sending a query to the originating mail server, asking if the email received originated from the specified mailbox. Each email forwarded needs to be logged before the email it is forwarded. This log checks if the respective server sent the mail, when there is a query about the sent mail. An addition to the

proposal is to make use of an Authentication, Authorization and Accounting (AAA) server that is trusted to verify an email server. This server certifies that an email server being queried can be trusted to give a true answer. If the AAA server returns with a negative response, the server drops the email regardless of what the sending server's response. When a server drops mail, the server's log notes that the drop action has been performed on the mail.

2.1.3 Intelligent Filters

Intelligent filters are software applications that are installed as part of a user's mail application or as part of a mail server or both (spam-site, 2006), (Mueller, 2009). Intelligent filters use a set of anti-spam strategies to improve the success of the filter. Intelligent filters can be trained to reduce the amount of false negatives and false positives. The idea is that the user or groups of users give input to the filter to train it.

Anti-spam software vendors claim that intelligent filters can be trained to block 99.9% of all spam (spam-site, 2006), (Mueller, 2009). From the claims made it can still be deduced that not all spam can be blocked at all times. Most of the weaknesses, discussed with regard to the other anti-spam strategies, are present in intelligent filters.

2.2 Digital Forensics

Digital forensic science is a relatively new field of study that evolved from forensic science. According to the Oxford Dictionary (University Press), digital forensic science is the systematic gathering of information about electronic devices, which can be used in a court of law. Digital forensic science is more popularly called digital forensics and sometimes also called computer forensics. Palmer (2002) defines digital forensics as " the use of scientifically derived proven methods towards the preservation, collection, validation, identification, analysis, interpretation, documentation, and presentation of digital evidence derived from digital sources for the purpose of facilitation or furthering the reconstruction of events". Palmer's definition describes the digital forensic process. The Oxford dictionary describes digital forensic science. The Digital Forensic Process Model (DFPM) by Kohn, et al. (2009) captures the definition of digital forensic science and states that any digital forensic process must have an outcome that is acceptable by law.

2.2.1 Digital forensics and email

Digital forensics has been used to verify the author of an email or to authenticate a user while the user is using their email application (de Vel, et al., 2001), (Gupta, et al., 2004). Digital forensics has also been used to trace the origin of email messages. Spammers make use of spoofing, open proxy servers and open mail relays, to send anonymous emails.

A proxy server is a computer process that relays a protocol between client and server systems by appearing to be the client to the server, and appearing to be the server to the client (Network-Dictionary, 2009) (Obied, 2006). Proxy servers allow communication between two computer systems by relaying information back and forth between the two connected proxy servers. An open proxy server allows unauthenticated systems to communicate through it. Email cannot be sent straight from one server to the next, it has to pass through a series of Email relay servers (Obied, 2006). Open mail relays are mail servers that are not properly configured to authenticate the origin of an email or to authenticate the email's path. By using an open proxy server or a series of open proxy servers before routing an email through an open mail relay server, the sender of an email can stay anonymous because it appears that the email originated from the last open proxy server the mail was sent through.

2.2.2 Honeypots

Even (2000), states "honeypot systems are decoy servers or systems setup to gather information regarding an attacker or intruder into your system". Spitzner, according to Obied (2006), defines a honeypot as information system resources whose value lies in unauthorized or elicit use of that resource. The author defines a honeypot as a trap set to detect, deflect, or in some other manner counteract attempts at unauthorized use of information systems. The information gathered by the honeypot is used to track where the authorized access originated and what exploits were used. If the honeypot is not accessed, it is of no use. A honeypot logs access information in accordance with digital forensic information gathering techniques. The design, implementation, placement and monitoring of a honeypot is crucial to the effectiveness of the honeypot.

Honeypots have been deployed as open proxy servers and open mail relays, to gather information about the spammers that use them (Obied,

2006). Honeypots have also been employed to gather information on botnets (Obied, 2006). The next section discusses the history of botnets, as well as advances made in the development of new strategies to trace and combat botnets.

3 BOTNETS

According to Network-Dictionary (2009)“a botnet, also known as a zombie army, is a computer connected to the Internet, that has been set up to forward transmissions (including spam or viruses) to other computers on the Internet, without the knowledge of the computer owner.” ESET (2009) defines a botnet as “a group of bot infected PCs that are all controlled by the same command and control center”. According to the author, a botnet can be defined as a group of infected computers or zombies, that are controled from a single controler or botmaster and used to facilitate electronic crime.

Botnets are created by infecting computers with Trojans. Once a computer is infected, the Trojan creates a SMTP (Simple message transfer protocol) account on the local machine. This account is used to send spam and any other electronic content. The Trojans in a botnet used IRC (internet Relay Chat) connections to receive information from the controler. According to InternetNews (2009), the new tactic is for the bots to communicate with each other using Peer-to-peer connections, set up in a family tree fashion to relay information and commands. The IRC method of communication hard codes the address of the controller into the Trojan. The controller's address is extracted from the Trojan during the dissection process. The family tree method of control is when no “child” Trojan knows any of its ancestors other than its direct “parent”. The family tree method makes it harder to find the controller.

Since closing, McColo (News, 2009), a US-based ISP accused of being a major hub for spammer activity, spammers have learned to hide their activity behind the same technology used for secure networking. The biggest botnet, called Sirbizi, closed in late 2008. According to Waters (2009), the infection rates across the world are increasing. Figure 1 shows the number of infected PCs per 1000 for all the world regions. InternetNews (2009), states that MessageLabs is currently monitoring a number of botnets, including Xarvester, Cutwail and Mega-D. Spammers use botnets to create low-volume-high-node-count mail senders. Low-

volume-high-node count means that the nodes are only used to send a small subset of the mails to be able to stay under the radar of bulk mail detectors. Mega-D was detected because it over utilised its bots.

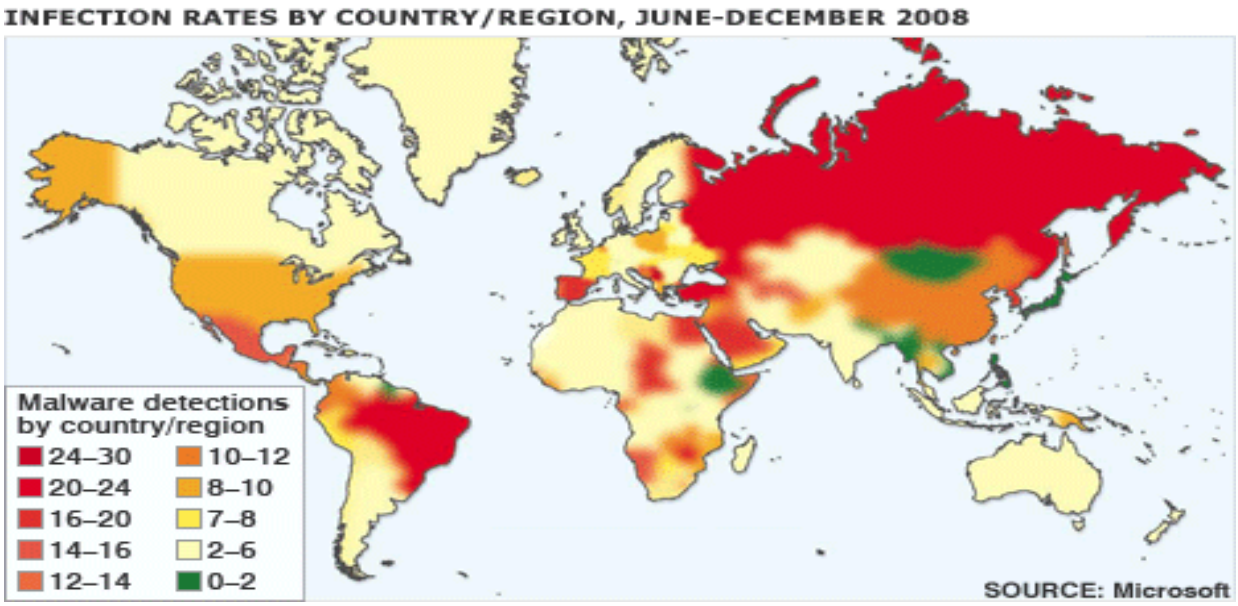

Figure 1 Infections per 1000 PCs for world regions (Waters, 2009)

According to Obied (2006) Microsoft used a zombie machine as a honeypot to detect and trace spam activity. The machine was infected with a botnet's trojan and quarentined. Activity to and from the zombie was monitored. The information gathered by the zombie honeypot helped to track the command and control source of the botnet. This tracking information was used in a lawsuit against 13 spam operations. Using a zombie as a homeypot is only possible if the controller of the botnet is unaware that one of the zombies is being used as a honeypot.

Botnets employing P2P connections between the different zombies makes it harder to use the zombies to track the controler. The next section discusses a proposed state of the art botnet architecture and proposes strategies to combat this state of the art botnet.

4 PROPOSED STRATEGY TO COMPLEMENT ANTI-SPAM USING DIGITAL FORENSIC STRATEGIES

As anti-spam strategies evolve, spammers evolve new strategies to bypass anti-spam. The challenge for anti-spam is to get ahead of the evolution curve and start developing strategies that combat possible future developments in spam strategies. Wang, et al. (2009), suggests that, to effectively protect against new developments in botnet technology and its uses, state of the art botnets should be developed to find ways of combating.

Wang, et al. (2009) presents the design of an advanced hybrid peer-to-peer botnet. The botnet uses advanced techniques to hide its activity by means of encryption and a traffic control algorithm. The botnet uses decentralised control mechanisms to hide the controller and ensure that zombies in the botnet cannot be traced by use of other zombies. The zombies are autonomous. Finding and removing those zombies found, does not impair the rest of the botnet.

The following sections discuss the implementation of an experimental environment. This environment will be used to deploy the botnet and gather information on the working of the botnet. Section 4.1 discusses the implementation of the experimental environment. Section 4.2 discusses the implementation of honeypots, in the experimental environment, as an information-gathering tool. Section 4.3 discusses the creation of a digital forensic profile of the botnet. In the real world, a botnet profile can detect and categorise botnet activity.

4.1 Experimental implementation of botnet

Combating the botnet will require the implementation of the botnet in an experimental environment. The experimental environment needs to consist of open proxy servers, open mail servers, a botnet control machine and a set of workstations used as zombies. Previous studies thought us that spammers use open proxy servers and open mail servers to hide the origin of the spam that they send (Obied, 2006). To enable us to gather the most relevant information about the operation of the botnet we will need to implement the most spammer friendly environment. Figure 2 shows the proposed implementation of the experimental environment.

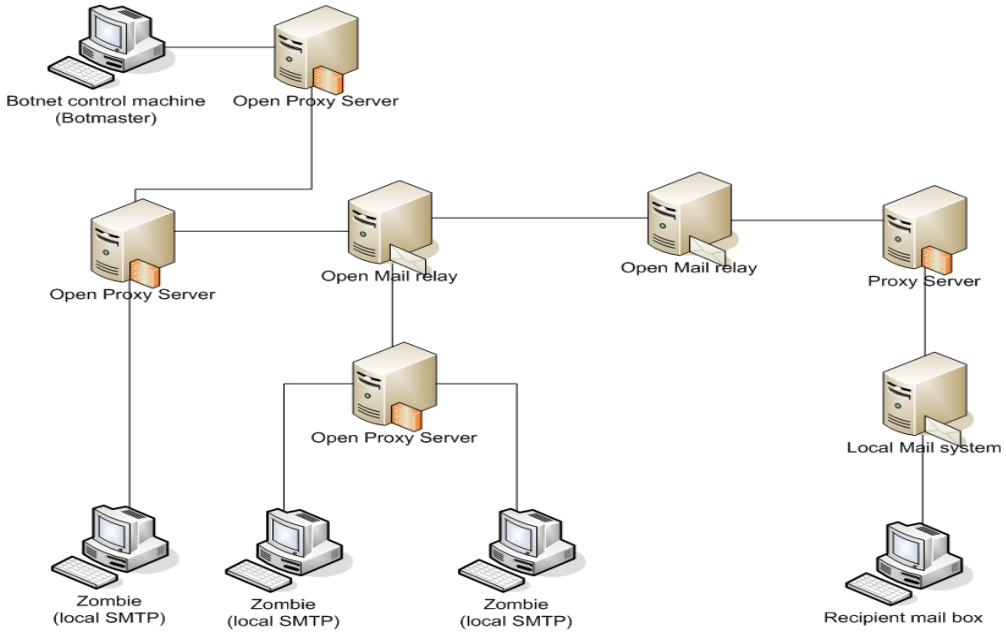

Figure 2 Depiction of the experimental environment implementation

Figure 2 shows the botmaster in its own network connected to the common network with an open proxy server. A second proxy server connects the botmaster to an open mail server. The configuration of the botmaster and open proxy servers creates an open proxy server chain. The open proxy servers are positioned to indicate the borders of the smaller networks within the larger environment. To simulate the real world mail relay environment, two open mail relays are connected to each other and placed in the centre of the experimental environment. The local SMTP service on the zombies is there because of the Trojan infection. The local mail server will have a known set of mail accounts that will receive e-mail from the zombies. The recipient e-mail box will monitor all the mailbox activity on the local mail server.

4.2 Honeypot implementation

Honeypots will be deployed to gather information about the activity inside the experimental environment. The open proxy servers, open mail servers and zombies will be used as honeypots in the experimental environment. The experimental botnet will make use of log-files to capture true activity

of the botnet. The true activity can be compared to the activity recorded with help from the honeypots. Using the comparative information recorded, an effective design and deployment strategy for a production environment, can be determined.

4.3 Profiling

The disadvantage of using honeypots in the experimental environment is that botmasters can create detection methods to detect honeypots and avoid them. Honeypots are used more and more as a general information gathering technique. A new information gathering technique is needed that cannot be detectable by botnets and cannot be bypassed.

Organisations create sub-networks with the use of VPN's over WAN links to connect IT resources together. A botnet creates a sub-network within a greater network, in the same way. A profile of what the botnet sub-network activities might look like is set up, using the experimental information collected during the botnet testing. The creation of a network diagram for all the sub-networks within the greater network will allow for the creation of an activity profile of the sub-networks. By comparing the activity profile of a sub-network with the know activity profile of a botnet, it will be possible to detect the botnet.

The information contained in the profile will consist of normal activity information. This paper defines normal botnet activity as the activity involved with the sending and receiving of control information and the sending of spam messages. The activity of infecting new machines is included in the activity of sending spam messages.

5 CONCLUSION

This paper looks at the strategies employed by spam and anti-spam and shows the co-evolution of these strategies. Anti-spam software makes use of intelligent filtering based on content scanning, block lists, black lists, white lists and mailbox authentication. Spammers have been able to get past anti-spam software by using picture content, mailbox spoofing and anonymous e-mailing.

Digital forensic science has been employed to authenticate email authors and back trace e-mail paths. The latest development in digital forensic information gathering is the use of honeypots. Spammers use botnets to send unsolicited electronic communication that can bypass anti-spam strategies.

This paper proposed two strategies for the detection of botnet activity and the tracing of botmasters. The first strategy consists of an implementation of honeypots to detect botnet activity. The second strategy employs digital profiling to detect the activity of botnets. The challenge for future developments, with regard to anti-spam strategies, will be to improve information gathering, botmaster tracing and botnet detection.

No one can win an evolutionary war. The co-evolution between spam and anti-spam is likely to continue indefinitely. To win the war, anti-spam strategies will need to get ahead of the evolutionary curve and start to develop new ways of detection, information gathering and tracing, proactively.

6 REFERENCES

de Vel, O, et al. 2001. Mining Email Content for Author Identification Forensics. *SIGMOD Record.* December, 2001, Vol. 30, 4.

ESET. 2009. Botnet Definition. *http://www.eset.com/.* [Online] 2009. [Cited: 27 April 2009.] http://www.eset.com/threat-center/threats/botnet.php.

Even, Loras R. 2000. Intrusion Detection FAQ: What is a Honeypot? *http://www.sans.org/resources/idfaq/honeypot3.php.* [Online] SANS Institute, 12 July 2000. [Cited: 27 April 2009.] http://www.sans.org/resources/idfaq/honeypot3.php.

Faynberg, Igor, et al. 2004. *Method and Apparatus for Reducing E-mail Spam and Virus Distribution in a Communications Network by Authenticating the Origin of Email Messages. US2005/0203985A1* United States of America, 29 April 2004. 709/200.

Gupta, Gaurav, Mazumdar, Chandan and Rao, M. S. 2004. Digital Forensic Analysis of Emails: A Trusted Email Protocol. *International Journal of Digital Evidence.* Spring, 2004, Vol. 2, 4.

InternetNews. 2009. Report Says Spam Arms Race Escalating. *http://www.ironport.com/.* [Online] IronPort In the News, 16 March 2009. [Cited: 16 March 2009.] http://www.ironport.com/company/pp_internet_news_03-16-2009.html.

Kohn, Michael, Eloff, J.H.P. and Olivier, M.S. 2009. *UML Modelling of Digital Forensic Process Models (DFPMs).* [Document] Pretoria : Information and Computer Security Architectures (ICSA) Research Group University of Pretoria, 2009.

Lueg, Christopher, Huang, Jeff and Twidale, Michael B. 2006. *Mystery Meat:Where does spam come from,and whydoes it matter?* [Document] Hamberg, Germany : EICAR, EICAR, 29 April 2006. Security in the Mobile and Networked World, Vol. 15.

Mueller, Scott H. 2009. spam.abuse.net. *Fight Spam on the Internet!* [Online] spam.abuse.net, 18 April 2009. [Cited: 21 April 2009.] http://spam.abuse.net/.

Naidoo, Nithen. 2007. *Introduction to Using Spam Methedology to Initiate Proactive Spam Controls.* [Document] Centurion : SensePost, 2007.

Nazario, Jose. 2006. *Botnet Tracking: Tools, Techniques and Lessons Learned.* [Document] Canada : Arbor Networks, 2006.

Network-Dictionary. 2009. http://www.networkdictionary.com/security/b.php. *http://www.networkdictionary.com.* [Online] http://www.networkdictionary.com, 2009. [Cited: 27 April 2009.] http://www.networkdictionary.com/security/b.php.

O'Brien, Cormac and Vogel, Carl. 2003. *Spam Filters: Bayesvs. Chi-squared; Letters vs. Words.* [Electronic] Dublin : University of Dublin, 2003.

Obied, Ahamed. 2006. *Honeypots and Spam.* [Document] Calgary : UniversityofCalgary, 2006.

Palmer, G.L. 2002. *Road Map for Digital Forensic Research.* [Electronic Publication] s.l. : Digital Forensic Research Workshop (DFRWS), Digital Forensic research workshop, 2002.

spam-site. 2006. www.spam-site.com. *www.spam-site.com.* [Online] www.spam-site.com, 2006. [Cited: 03 March 2009.] http://www.spam-site.com.

University Press, Oxford. Possible entries for. *http://www.oup.com.* [Online] [Cited: 22 April 2009.] http://www.oup.com/oald-bin/web_getald7index1a.pl.

Wang, Ping, Sparks, Sherri and Zou, Cliff C. 2009. An Advanced Hybrid Peer-to-Peer Botnet. *TRANSACTIONS ON DEPEDABLE AND SECURE COMPUTING.* Monthly, 2009, Vol. 0, 0.

Waters, Darren. 2009. Spam Overwhelms Email Messages. *BBC NEWS | Technology |.* [Online] BBC, 8 April 2009. [Cited: 8 April 2009.]

http://newsvote.bbc.co.uk/mpapps/pagetools/print/news.bbc.co.uk/1/hi/technology/7988579.stm?ad=1.

www.spam-site.com. 2006. http://www.spam-site.com/your-website-blacklisted.shtml. *www.spam-site.com.* [Online] www.spam-site.com, January 01, 2006. [Cited: March 23, 2009.] http://www.spam-site.com.

7 PERMISSIONS