ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/229048601

Mystery Meat: Where does spam come from, and why does it matter?

Article · May 2002

READS
35

3 authors:

Christopher Peter Lueg
University of Tasmania
110 PUBLICATIONS **460** CITATIONS
SEE PROFILE

Jeff Huang
National Taiwan University
6 PUBLICATIONS **59** CITATIONS
SEE PROFILE

Michael B. Twidale
University of Illinois, Urbana-Champaign
112 PUBLICATIONS **1,877** CITATIONS
SEE PROFILE

Available from: Christopher Peter Lueg
Retrieved on: 10 May 2016

Mystery Meat: Where does spam come from, and why does it matter?[1]

Christopher Lueg, Jeff Huang & Michael B. Twidale

About Author(s)

Dr. sc. nat. Christopher Lueg is a Professor of Computer Science in the School of Computing at the University of Tasmania. Contact Details: c/o School of Computing, Private Bag 100, Hobart TAS 7001, Australia, e-mail christopher.lueg@utas.edu.au.

Jeff Huang is an M.S. Student in the Department of Computer Science at the University of Illinois Champaign, IL 61820 U.S.A., U.S.A. e-mail: huang6@uiuc.edu.

Michael B. Twidale, PhD is an Associate Professor in the Graduate School of Library and Information Science, University of Illinois, 501 E. Daniel Street, Champaign, IL 61820 U.S.A. e-mail: twidale@uiuc.edu

Keywords

email, spam, filtering, blocking,SMTP, digital divide, digital redlining,

1Proceedings of the 15th EICAR Annual Conference "Security in the Mobile and Networked World". Hamburg, Germany, 29 April - 2 May 2006.

Abstract

Unsolicited commercial email or spam is recognized as a problem disrupting email communication and costing the community dearly. In order to protect recipients from receiving spam, anti-spam measures building on technologies, such as filters and block lists, have been deployed widely. There is some evidence that certain anti-spam measures based on the purported origin of the spam cause unintended consequences which relate to issues of equity of access which we term digital redlining. Spammers have an interest in bypassing such measures by obscuring the real origin of their messages. Investigating these effects means we need to determine the true origin of spam, despite the efforts of spammers to confuse us and spam filters. The aim to find the true origin of spam is different from the objective of most anti-spam developers who are mainly interested in identifying spam when it knocks on their front door (mail server). In this paper we discuss why the difference between originator and delivery host matters when investigating digital redlining. We also highlight some of the difficulties we are facing when trying to determine the originating host as opposed to the delivering host.

Introduction

E-mail is widely regarded as one of the most important services provided by the Internet. In the U.S., for example, almost all Internet users use email which seems to be rapidly becoming more used than the telephone (Haythornthwaite and Wellman, 2002, p. 6). The term "spam" is used as a colloquial and convenient substitute for "unsolicited commercial email" (UCE). Spam has become a major problem causing significant costs to the community (e.g. McKusker, 2005).

In order to protect recipients from receiving spam, technical anti-spam measures have been deployed widely. Developing accurate anti-spam measures is technically challenging as there is no such thing as a precise, "technical" definition of spam; by definition, the nature of unsolicited/unwanted commercial email aka spam depends on the recipient's attitude towards receiving respective messages (Lueg, 2003; 2005). E.g., in a report on the spam problem and how it can be countered, the Australian National Office for the Information Economy defined spam as "unsolicited electronic messaging, regardless of its content" (NOIE, 2002 , p. 7). The report explicitly mentions that "arriving at an agreed definition of spam is a potentially contentious issue, as the direct marketing industry, ISPs, spammers, blacklisters and privacy and consumer groups have their own interests and views."

Spammers are effectively engaged in an arms race with spam recipients and providers of spam filtering technologies. Each measure by one side provokes a counter-measure by the other (e.g., viruslist.com; also see Hulten et al., 2004 for examples). However, innocent third parties can get caught in this crossfire, in particular perfectly legitimate email that gets mis-classified as spam and ignored or never even seen by its intended recipient. In this paper we focus on two related issues.

The first issue is that secondary effects of certain anti-spam mechanisms, in particular blacklisting of mail servers, may not only have the unfortunate effect of falsely declaring some email as spam, but in so doing disproportionately affect the emails of certain disempowered groups, an effect that we call digital redlining. Certain blocks of IP addresses may be redlined because mail servers operating in these net blocks have shown to be used for spamming and they are suspected to be used again in the near future. The most extreme case of digital redlining is blocking whole countries as

proposed e.g., by SpamStopsHere. The interesting and ironic point is that there is evidence that in certain cases, blacklists created in Western countries are redlining servers located in other –often economically disadvantaged– parts of the world due to spam that actually originated from those same Western countries. Rather than resolving the spam problem originating in their own countries, the Western blacklist filters discriminate against foreign servers that were (ab)used for sending spam around the world and back into the West.

The second issue we address is the problem of collecting supporting evidence for the above observations. The problem is that email headers can easily be manipulated which means it is difficult to extract reliable data about the originating host as opposed to the delivering host. This difficulty of determining the true origin and provenance of an email is of course both a challenge for research in this area, and may also contribute to designing more effective spam filters that can more equitably discriminate between legitimate and illegitimate email.

A related research challenge is that to our knowledge, there is no comprehensive conceptual framework available that could be used as a scaffold for the kinds of analyses we are interested in. Such a framework should support analyses of information equity issues in a systematic way. This lack of conceptual work had been noted by Lueg (2001) in the context of more general information distribution issues. In the related context of forensic computing, Slay, Turner & Hannan (2004) argue for establishing a framework supporting the use and also the positioning of different disciplinary approaches as this would help overcome the current focus of forensic computing on securing and analyzing the information "at hand", typically in a reactive way. In this sense, we see striking similarities between challenges in forensic computing and challenges in spam filtering research which appears to be focusing too much on immediate mail delivery issues and paying little attention to wider impacts of the technologies deployed.

Intended effects of anti-spam measures: protecting users and resources

The primary, intended effect of deploying anti-spam measures is a reduction of the amount of spam that end users have to deal with. Reducing the amount of net traffic devoted to passing around emails that recipients do not want to receive is also desirable, but end user attention is by far the most scarce resource and the one that should be optimized.

Most filtering targets the information contained in the body of a mail message (the part of a message users normally see) but can also consider information contained in the message header (mostly information used to transport the message but also From: and Subject: information). Typical examples of characteristics used to sort out spam messages are terms such as "free porn", "XXX", "Warez" or "Get rich quick". Many spam filters use a probabilistic approach known as Bayesian filtering, effectively awarding each email a number of ‘points’ for features that have been found to correlate with known spam, such as certain words in the content and headers, or a large proportion of HTML in the body. If the message’s points exceed a certain threshold, it is deemed to be spam. In this paper we focus on the header information and how it can be used in spam filtering, but of course in reality, this information is combined with the analysis of the body of the message to make an overall decision.

Spam filters may target the origin of a message. If a message is sent through a mail server that is known to be operated by a spam-friendly company, then this raises the probability that the message is itself spam, and for some filters this may be sufficient for the spam/notspam decision. Of course

spammers are aware of the existence of spam filters and know much about their algorithms (certainly as much as is in the public domain, which is all the information that we have available for this analysis). Therefore spam messages may be designed to disguise their origin or provenance.

Related to origin-based filtering is the more radical approach of blocking. Blocking means a mail server simply refuses to accept any mail from certain servers, often according to blacklists shared on the Internet. Blocking approaches can use IP addresses, domain names and other information provided by the mail exchange.

Lists of alleged spam servers are shared on the Internet. Services such as the Spamhaus Block List allow mail servers to check in real time if a server trying to deliver email has earned a spammer reputation. The Spamhaus Project (2003), the organization hosting the above mentioned list, describes their block list as follows: "The Spamhaus Block List (SBL) is a real time database of IP addresses of static spam-sources, including known spammers, spam operations and spam support services."

Anti-spam products often implement combinations of different filtering, blocking and learning techniques (see Metz, 2003 for an overview). Kaspersky Anti-Spam ISP Edition, for example, uses a combination of linguistic analysis, formal analysis of message characteristics and blocking based on blacklists and whitelists.

Discussing the details of specific spam filtering technologies would exceed the scope of this paper. See e.g., the CEAS conference proceedings (URL http://www.ceas.cc) for further information.

It is also worth noting that in certain countries telecommunication legislation may prevent or restrict usage of specific spam filtering technologies. For example, preventing servers from receiving certain spam messages by blacklisting suspicious mailservers is basically illegal in EU countries because employees have the right to see these messages. See Gattiker (2005) for details.

Unintended consequences of anti-spam measures: digital redlining

The objective of using blocking techniques is to reduce a mail server's intake of messages likely to be classified as spam anyway. Messages may not be checked thoroughly since they may be assumed to be spam because they originate from alleged spam sources. This saves processing time, but of course means that such a coarse filter opens up the method to false positives.

Typically, blockings are based on information coming from blacklists shared on the Internet. There is anecdotal evidence that blocking not only reduces the spam intake but also may cause ‘collateral damage’ undermining reliability of email communication. The Spamhaus Project (2003) being a major host of such a black list notes the possibility of adverse effects as follows: *"Can the SBL block legitimate email? The SBL's primary objective is to avoid 'collateral damage' while blocking as much spam as possible. However, like any system used to filter email, the SBL has the potential to block items of legitimate email if they are sent from an IP under the control of a spammer or via IPs belonging to spam support services. The chances of legitimate email coming from such IPs are slim, but need to be acknowledged."*

From a technical point of view, blocking is an effective way to reduce a server's spam load. Anecdotal evidence suggests that blocking is also perceived as a way to suggest to mail server

operators to protect their servers against misuse by spammers. A discussion in the newsgroup news.admin.net-abuse.blocklisting, comprising more than a hundred statements, indicates blocking is a double-edged sword though as there are quite different opinions regarding accountability and reliability of blocklists (see Blue, 2003 for details). There is an abundance of requests to be removed from block lists posted to the net-abuse related newsgroup news.admin.net-abuse.blocking (see, for example, Rodriguez, 2003). Often, these requests are coming from businesses renting IP blocks from major ISPs. Even though these businesses do not spam themselves, they are blocked because their ISPs have a spam history or previous owners of their net blocks were spammers. Gaudin and Gaspar (2001) quote figures suggesting that using one particularly unreliable blacklisting service was found to cause 34% false positives, i.e., genuine email classified as spam. Cole (2003) provides a detailed overview as to why mail servers located in the net block (or IP range) of 'innocent' businesses may become 'collateral damage' and how these businesses may address the situation. Cohn & Newitz (2004) of the Electronic Frontier Foundation (EFF) discuss impacts of spam filters from the point of view of mailing list operators.

Varghese (2003a) reports the following incident: *"AOL says it is blocking email from Telstra's BigPond users because it has received complaints from its subscribers about spam being sent to them from BigPond addresses. Company spokesman Nicholas Graham said AOL had been [...] essentially compiling a whitelist of IPs from which mail would be allowed to reach AOL users."* BigPond is a major Australian ISP. This means AOL blocked a significant part of a whole continent's email users! Of course Telstra is a large company, and when it became a victim of this draconian anti-spam algorithm, it complained vociferously and was able to get the decision overturned within a week (Varghese, 2003b). Which leads to an interesting question of ethics and power in the digital world: are big, rich, powerful, western ISPs and servers better able to reverse presumably unfair blacklisting than their peers in poorer regions of the world? If so, the policy, which may appear solely to be about spam, and which all of the stakeholders may agree is a problem, may actually raise serious issues of equity across countries and levels economic and political power.

Hosted spam-filtering service SpamStopsHere mentions on their web site: "[i]t is well known that a huge amount of spam originates in China, Taiwan, Brazil and Argentina. It is clear to us that there are businesses in these countries that primarily just send spam. South Korea has a huge number of "open relays" which are computers that have been hacked and are controlled by spammers. Therefore, blocking email from countries notorious for sending spam is an effective filtering method. Which countries to block, if any, is a business decision. Click here for a discussion of why China, Taiwan and South Korea should be blocked." The idea that whole countries should be blocked because a lot of undesirable activity originates there (or appears to originate there – see later) raises many obvious questions about access and fairness. The similarities to the historic practices of certain US banks in discriminating in loan applications against areas of cities deemed poor economic risks (and not coincidentally frequently occupied by African Americans, so contributing to rising problems in those areas) inspired us to use the provocative term 'redlining'.

"Forensic" spam tracking challenges

Knowing where spam really comes from is useful in many kinds of approaches to prevent it, limit it or mitigate its consequences. As noted, the provenance of an email message can be a contributory piece of evidence in a judgment about its candidacy to be treated a spam by various filtering algorithms. Unfortunately spammers are aware of this and so take measures to try and conceal a

message's true origins.

Although our interests as researchers in identifying spam are similar to those of anti-spam developers, our concerns are broader, extending to issues of the efficacy and equity of anti-spam measures undertaken. This requires gaining a richer picture of the true origins of spam. This is different from the objective of most anti-spam developers who are mainly interested in identifying spam at the time it knocks on their front door – their mail server. In terms of provenance, server's spam filters mostly focus on what is known about the last server to relay the message to them (see e.g., Goodman, 2004). Consequently, blacklists merely list prospective spam delivering hosts but do not attempt to provide information about originating hosts.

Given the ease of fabricating the earlier history of the message's progress, this is an understandable simplification, but, we believe, an unfortunate one. In what follows we discuss why the difference between origin and delivery host matters when investigating digital redlining and highlight some of the difficulties we are facing when trying to determine the true originating spam host as opposed to the delivering host.

One of the challenges in the context of understanding digital redlining is identifying what header information is actually reliable. The way header information is forged is also of interest as it may tell who the spammer is trying to blame in the case of joe jobs. A 'joe job' means hiring a spammer to spam under the name of another person's domain, or web pages. The intended effect is that lots of people complain to the Internet service provider (ISP) hosting the domain or the web page advertised in the spam as they mistakenly assume they know the source of the spam. Further it is necessary to take into account information from other non-technical sources to understand how spamming is affecting the networked world. This means understanding spamming requires methods from a range of disciplines, from computer science and information science to social sciences and possibly political sciences. It also means it would be nice to have a conceptual framework supporting the use and the positioning of different disciplinary approaches as required. As mentioned earlier, a similar argument was put forward by Slay et al. (2004) in the context of forensic computing (McKemmish, 1999).

Identifying Fake Received Headers

The immediate technical challenge we are facing is that email is relayed from one server to another, and each server tacks on a Received: header indicating where it got the email from. These are simple unprotected text strings, listing the route the mail has taken, with the most recent server first. Thus it is possible for a spam-sending computer to falsely claim that it received the email from another computer, by inventing and adding prior Received: headers. These faked headers may be from real or nonexistent servers. Since there is no direct way of knowing whether or not a Received: header is legitimately added or not, it is generally accepted that there is no fool-proof method to find the origin of an email. What can be done instead is to identify the last external server that handles the email before passing it to an internal server, (Goodman, 2004). Assuming you can trust your internal servers, this IP cannot be falsified because it was reported by an internal server, which you have control over. Because of the lack of a fool-proof method to find the originating IP address of an email, the last external server is often used for tracking and reporting purposes instead. This may be good enough for spam filtering but not for analysis. So the origin of an email is generally not retrieved.

If a faked Received: header is found in the list, this indicates that one of the servers claiming to have subsequently received the message intentionally tried to foil attempts to identify the source of the email. Typically the guilty server is the actual originator the email.

For example imagine that we receive an email with header information including the following:

```
Received: from Echo <plus more info>
Received: from Delta <plus more info>
Received: from Charlie <plus more info>
Received: from Bravo <plus more info>
Received: from Alpha <plus more info>
```

The first line is the last stop on the way to us; Echo, our own internal server which we trust. If we can somehow determine that Bravo is a fake entry, and we trust Echo, then we must suspect that either Charlie or Delta are the true origin and that they faked Bravo (and also Alpha, which looks plausible to us).

Each received line usually includes ('from') the host name reported by the sending server through the HELO command, followed by a pair in parenthesis -- a host name gotten by the receiving server by performing a reverse lookup on the IP address, followed by the actual IP address of the sending server. Following this is the host name of the receiving server ('by'), and a ('with') token followed by some additional information about the SMTP server and an ID for tracking purposes, and finally, some time information (following a semi-colon).

```
Received: from relay7.cso.uiuc.edu (relay7.cso.uiuc.edu
[128.174.5.108]) by expms6.cites.uiuc.edu (MOS 3.4.8-GR) with
ESMTP id BDH13397; Sat, 21 Jan 2006 02:22:39 -0600 (CST)
```

In practice however, SMTP servers do not always use this format. It is possible that a forgery is done well enough that it is impossible to tell whether or not a header sequence is fake. Fortunately, many attempts to forge Received: headers are imperfect and can be detected. There are several ways a false Received header can be identified:

Illegal IP addresses

IP Addresses contain 4 bytes; each byte is a number between 0 and 255. Thus, any time a computer claims that it was passed email coming from an IP where a byte is greater than 255 is obviously false. Illegal IP Addresses are often generated by the programs that create the false headers, which simply string together 3 digits for each byte, so it is often possible to see an IP address such as 729.493.230.588.

IP addresses with leading zeros

A smarter program that generates random IP addresses will restrict the first digit to 0-2 and/or ensure that the second and third digits are 0-5 if the first digit is a 2. This will create valid IP addresses, but there can still be a clue that indicate the IP address was made by a program rather than one that was reported by an SMTP server. The giveaway is the leading zero in the IP address.

For example, the IP address 4.235.08.96 would most likely have been generated by a program because an SMTP server would report it as 4.235.8.96.

Reserved IP addresses

There are several blocks of IP addresses which are not valid internet IP addresses. Most of these IPs are reserved for use by the IANA (Internet Assigned Numbers Authority), the organization that oversees allocation of IP address and documented in an RFC (RFC 3330). Blocking reserved IP addresses is a popular method of spam filtering, used by many spam filters including Spamassassin. These reserved IP addresses slowly change over time at IANA's discretion, usually because a certain block gets allocated to an entity or a regional internet registry, so to use this method to identify fake headers also requires information about when the email was sent as well. Fortunately, this time information is also in each Received lines.

Broken path

It is also possible to discover that a fake header has been inserted if there is a broken link in the chain of Received headers. For example,

```
Received: from exserver.chgh.org.tw ([203.69.196.131]) by
expansionpack.xtdnet.nl (8.11.6/8.9.3) with ESMTP id fA9LD9m04845
     for <foo@bar.com>; Fri, 9 Nov 2001 22:13:10 +0100

Received: from mcpeely.concentric.net (0-1pool5-38.nas2.los-
angeles1.ca.us.da.qwest.net[63.233.5.38])    by internation.co.uk
(8.11.0/8.9.3) with SMTP id gAJEYgl19984     for <foo@bar.com>;
Fri, 9 Nov 2001 20:29:03 GMT
```

Most likely, the second Received header is fake because there is no Received header to show how the email was moved from internation.co.uk to exserver.chgh.org.tw. Therefore, we suspect the second Received header to be fake and deem exserver.chgh.org.tw as the (spoofing) origin of the email. However, one must watch for legitimate mismatches caused by internal relaying (see below), but that is unlikely in this example.

Suspect Received Headers

Sometimes, even though it is impossible to tell definitively whether a Received header is fake, it might be possible to spot certain suspicious attributes that contribute to the probability that the header might have been tampered with. With an accumulation of such suspicious attributes, we may choose to make a determination that the header and so the whole message is fake, but we must be prepared to acknowledge that perfectly legitimate messages may also have these features.

HELO/EHLO mismatch

When a email client connects to an SMTP server, it is usually required to identify itself by its hostname using the HELO or EHLO (for extended SMTP) command. Originally, email clients were trusted to provide the correct hostname. However, now SMTP servers often also perform a reverse

lookup on the IP address of the computer connecting to itself. Both the reported hostname name and then in parentheses the looked up name are reported in the Received header in the format below.

```
Received: from hostname.reported.by.helo.command.com
(actual.hostname.from.lookup.com [63.233.5.38]) ...
```

Theoretically, the hostname from the reverse lookup and the hostname given by the connecting client should be the same. However, since the hostname given by the HELO/EHLO command is up to the client, the hostname could be fictitious if the client wants to be evasive. Therefore, a mismatch between the hostnames could be seen as suspicious and headers beneath that header should be analyzed carefully. One caveat is that the hostnames could be different because of internal relaying and/or forwarding. For example, Microsoft hotmail might identify itself as mail.hotmail.com but the reverse lookup could report omc1-s7.bay6.hotmail.com.

SMTP path analysis, geographical path analysis

Using an IP to City database (e.g., geobytes, ip-2-location), it is possible to track down the geographical path of an email. Software that does this already exists (e.g., visualroute). It is possible to use this to analyze the path of the spam, and determine the rationality of the path of the email. If the Received headers of an email indicate that it has traveled from the United States to Belgium to New Zealand, this is much more suspect than an email that has traveled only domestically, or only changed countries once. This will not determine the veracity of a Received header but merely indicates that a set of Received headers is suspect. It is possible for email messages to zig-zag their way across the world, exploiting variations in traffic density, but messages that oscillate between countries can also be a way of creating a spoof provenance.

Email that travels to multiple countries is very rare. The most complicated email path today occurs when an email is relayed from a set of sending servers to the destination set of servers. Thus, even if an email travels through different regions or cities, it is more suspect. It used to be fairly popular among US spammers though. For example:

```
Received: from ms1.hinet.net (root@168.95.4.10) by
amadeus.ifi.unizh.ch with SMTP; 7 Dec 2000 07:32:24 -0000
Received: from power.jyinfo.com.tw (root@[203.75.22.178]) by
ms1.hinet.net (8.8.8/8.8.8) with ESMTP id PAA26963; Thu, 7 Dec
2000 15:22:45 +0800 (CST)
From: ieowirut@centrum.cz
Received: from epost.de (1Cust41.tnt1.bloomington.il.da.uu.net
[63.27.139.41]) by power.jyinfo.com.tw (8.9.3/8.8.7) with SMTP id
PAA13339; Thu, 7 Dec 2000 15:18:53 +0800
```

Assuming the header lines are genuine, the email originates from a uu.net host in the US and was relayed to Switzerland (amadeus.ifi.unizh.ch) via Asian mail servers.

Dynamic hostnames in the path

Most SMTP servers have static IP addresses that do not change very often. On the other hand, a consumer's internet access, via say cable modem, often has a dynamically assigned IP address and

respective hostname. Therefore, if a set of Received headers shows a dynamic hostname in the middle of the path, some of the headers may be fake.

Finding the "True" Originator

Algorithmically, the process of finding the originator of the email is fairly straightforward. Starting at the top of the set of Received headers, process down until either the end of the list, or a false or suspect header is found. However, as described earlier, determining what information is correct (or at least trustworthy) is a challenge and will remain a challenge even though authentication techniques are being introduced in a number of countries.

To gain an impression of trends in spamming we manually investigated a few hundred spam messages collected in December 2000, December 2001 and December 2005. The first two sets were collected in Zurich, Switzerland, the last set in Sydney, Australia.

December 2005		
straightforward	223	63%
dubious	46	13% (most relayed)
manipulated	80	23%
December 2001		
straightforward	48	42%
dubious	33	29% (most relayed)
manipulated	33	29% (most advertising.com spam)
December 2000		
straightforward	45	39% (incl. 8 surecom.com spams)
dubious	60	53% (most relayed)
manipulated	8	7%

Even though the data collection is not intended to be representative, the data does show a trend also reported in the literature: less third party relaying and more direct delivery. By direct delivery we mean that the delivering host is also the (alleged) originator:

```
Received: from haticeg (p169.net220148067.tnc.ne.jp
[220.148.67.169]) by filter.it.uts.edu.au (Postfix) with SMTP id
40214DF379; Thu, 1 Dec 2005 05:31:46 +1100 (EST)
```

The reason for less third party relaying is that most open relays were shut down (e.g., Hoffman, 2002) because of abuse by spammers. Direct delivery of spam points towards hosts that are infected by bots that are remotely controlled by a bot-master. As Kaspersky Lab state on their web site virus.com:

"Using open relay and open proxy servers is [...] time consuming and costly. First spammers need to write and maintain robots that search the Internet for vulnerable servers. Then the servers need to be penetrated. However, very often, after a few successful mailings, these servers will also be detected and blacklisted."

As a result, today most spammers prefer to create or purchase bot networks.

"In 2003 and 2004 spammers sent the majority of mailing from machines belonging to unsuspecting users. Spammers use malware to install Trojans on users' machines, leaving them open to remote use. [...] Anyone who has the client part of a program which controls the Trojan that has infected a victim machine controls the machine or network of victim machines. The resulting networks are called bot networks, and are sold and traded among spammers."

From the digital redlining point of view, the question will be where infected machines are located and whether there is a link between infection rate, blacklisting and economic strength of the respective geographic area. One might expect that computers in economically weaker regions are easier to infect due to the cost of keeping software and hardware up-to-date. However, the most infected operating systems seem to be Windows 2000 and Windows XP systems i.e., relatively new and resource-intense operating systems more likely to be found in economically stronger regions than in economically weaker regions. Looking into blacklists we actually found a bias towards economically stronger regions and bot infections may explain the pattern.

It will also be interesting to see to what extent concepts, such as SPF or domain-based key authentication, will help ease the spam problem. Both are technologies that may help identify a forged sender. DomainKeys is an email authentication system that verifies the domain of an email sender and the message integrity. SPF is a system that allows domain name owners to use special DNS records to publicize which computers are authorized to send email from that domain. However, DomainKeys and SPF are not anti-spam technologies. They prevent spammers from spoofing addresses from domains that use the technology, but spammers can still spoof domains that don't publish SPF/DomainKeys or they can use their own domains. Also, even if there is a valid SPF or DomainKeys signature, it does not indicate that the message isn't spam. Also see URL http://security.weburb.dk/frame/show/news/3917 for a discussion of the economics of domain-based email authentication.

Summary and Future Research

In the long term, anti-spam measures deployed in Western countries to protect them against spam produced in these countries may contribute to excluding those located in economically less developed countries unable or unwilling to prevent spamming. As anti-spam legislation is getting tougher in Western countries, such as the U.S., E.U. and Australia, there is a good chance that spam will increasingly be sent from or relayed in developing countries. This paper provides three main contributions to the discussion of the impact of spam and anti-spam technologies.

First, we have chosen to focus on the provenance of spam; where it originates and the route it takes to get to the recipient. By contrast, the majority of analysis work by spam filter researchers has looked at the content of spam. Of course content is important, and perhaps is the most important part. But we believe that provenance can play at least a vital contributory role in understanding how spam and spammers operate and hence how best to develop measures to counteract it effectively and equitably.

Secondly we have noted that where provenance is used, due to the difficulty of obtaining accurate information, it is often used somewhat crudely in existing filters, focusing almost exclusively on the last hop of the message, and rendering that particular mail-bearer the subject of suspicion. This has clear consequences for efficiency and effectiveness of spam filters. However, far more importantly, we wish to raise the point that while providing considerable benefit, certain anti-spam measures, such as filtering and blocking, particularly when they use crude measures of provenance, can also create problems of access, specifically of equity of access where message from certain sources may be blocked excessively and unfairly more than from other, more privileged sources. We believe this risk of digital redlining needs more careful study.

Last but not least we see this work as a contribution to the discussion of the role of conceptual frameworks in information distribution settings and forensic computing. Frameworks should support the integration of different disciplinary methods ranging from computer science and information science to social sciences and possibly political sciences.

Acknowledgments

The authors are grateful to the anonymous reviewers for their insightful comments.

References

Blue, T.M. (2003). Blocklists Accountability, Standards, Who is Policing Blocklists. Posting to the Usenet newsgroup news.admin.net-abuse.blocklisting on 12 Aug 2003. Message-ID: <65ab995e.0308121542.4bccaede@ posting. google.com>

Cohn, C. & Newitz, A. (2004). Noncommercial Email Lists: Collateral Damage in the Fight Against Spam. Retrieved 30 January 2006 from URL http://www.eff.org/wp/?f=SpamCollateralDamage.html

Cole, W.K. (2003). Blacklists, Blocklists, DNSBL's, and survival: How to Survive as a Non-Combatant Emailer in the Spam Wars. A collection of frequently asked and too-often poorly answered questions. Retrieved 2 January 2004 from URL http://www.scconsult.com/bill/dnsblhelp.html.

Gattiker, U.E. (2005). Information Security This Week (EU-IST). Editor Urs E. Gattiker. Published 27 September 2005. ISSN1600-1869.

Gaudin, S. & Gaspar, S. (2001). The Spam Police. Tactics Used by Self-Appointed Spam Fighters Come Under Fire. Network World, 09/10/01 Retrieved 30 December 2003 from URL http://www.nwfusion.com/research/2001/0910feat.html.

Goodman, J. (2004). IP Addresses in Email Clients. Proceedings of the Conference on Email and Anti-Spam, July 2004.

Haythornthwaite, C. & Wellman, B. (2002). The Internet in Everyday Life. An Introduction. In: Wellman, B. and Haythornthwaite, C. (eds.) The Internet in Everyday Life. The Information Age Series. Blackwell Publishing, Malden, MA.

Hoffman, P. (2002). Allowing Relaying in SMTP: A Series of Surveys. Internet Mail Consortium Report: UBE-RELAY IMCR-016, http://www.imc.org/ube-relay.html

Hulten, G., Penta, A., Seshadrinathan, G. & Mishra, M. (2004). Trends in Spam Products and Methods. Clients. Proceedings of the Conference on Email and Anti-Spam 2004, July 2004.

Kaspersky (2003). Kaspersky Labs Now Battling Spam at the ISP Level. Product announcement released 12/30/2003.

Lueg, C. (2001). Towards a Framework for Analyzing Information-Level Online Activities. Proceedings of the 2nd Australian Information Warfare & Security Conference (IWAR 2001) Perth, WA, Australia, November 2001.

Lueg, C. (2003). On the Relevance of Spam and Anti-Spam Measures to Information Security Management. Proceedings of the 1st Australian Information Security Management Conference (InfoSECURITY 2003), Perth, WA, Australia, November 2003.

Lueg, C. (2005). From Spam Filtering to Information Retrieval and Back: Seeking Conceptual Foundations for Spam Filtering. 69th Annual Conference of the American Society for Information Science and Technology. Charlotte NC, USA, 28 October-2 November, 2005.

McKemmish, R. (1999). What is forensic computing? Aust. Institute of Criminology, Canberra.

McCusker, R. (2005). Spam: nuisance or menace, prevention or cure? Trends & issues in crime and criminal justice No. 294. March 2005. ISBN 0 642 53875 1; ISSN 0817-8542.

Metz, C. (2003). Corporate Antispam Tools. PC Magazine. Retrieved 16 February 2003 from URL http://www.pcmag.com/article2/ 0,4149,849390,00.asp.

NOIE (2002). Final Report of the Australian National Office for the Information Economy (NOIE) review of the spam problem and how it can be countered. Retrieved 3 May, 2003 from http://www.noie.gov.au/projects/confidence/Improving/Spam/ Interim_Report/contents.htm.

Rodriguez, G. (2003). SPEWS blocking a range with mine included , how to get out? Posting to the Usenet newsgroup news.admin.net-abuse.blocklisting on July 22 2003. Message-ID <b7efc7c3.0307220944.3e1a4d00@ posting.google.com>

RFC 3330. Special-Use IPv4 Addresses. IANA. September 2002. Retrieved 23 January, 2003 from http://www.ietf.org/rfc/rfc3330.txt.

Slay, J., Turner, P. & Hannan, M. (2004). Developing Forensic Computing Tools and Techniques within a holistic framework: an Australian Approach. 2004 IEEE Information Assurance Workshop, Westpoint, NY, June 10-12, 2004.

SpamStopsHere. http://www.spamstopshere.com/antispam_howitworks.aspx

The Spamhaus Project (2003). The Spamhaus Block List (SBL) Advisory FAQ. Retrieved 13 January, 2003 from URL http://www.spamhaus.org/sbl/sbl-faqs.lasso.

Varghese, S. (2003a). AOL blocking BigPond mail because of spam. The Sydney Morning Herald 29 April 2003.

Varghese, S. (2003b). Telstra problems with AOL resolved. The Sydney Morning Herald 30 April, 2003.

virus.com (n.d.) The Evolution of Spam. Retrieved 15 January 2006 from http://www.viruslist.com/en/spam/info?chapter=153350530