Spam and Anti-Spam Measures: A Look at Potential Impacts

Christopher Lueg
University of Technology, Sydney, Australia

lueg@it.uts.edu.au

Abstract

The proliferation of unrestricted Internet access has brought the community spam which has become a serious problem costing companies billions of dollars per annum. Typical anti-spam measures, such as filtering and blocking techniques, exist but focus on solving the spam problem on the message transportation level. Using such techniques may have impacts beyond the realm of spam-filters and block lists. In this paper we argue that implementing typical anti-spam measures means that computers are assigned the power to assess legitimacy of email. This means, for example, that legitimate email might be rejected because the sender used the 'wrong' mail server or the wrong terminology. In this paper, we describe some of the core problems and discuss alternatives.

Keywords: spam, spam filter, blacklists, information filtering, legitimacy, solicitation

Introduction

The advent of the Internet has enabled numerous novel forms of business and leisure activities (e.g., Castells 2001). Popular examples are e-business companies like amazon.com and online communities, respectively. Proliferation of unrestricted Internet access also enabled new forms of interacting with customers. Kania (2001) outlines that "the Web [one of the Internet's most popular services] allows two-way communication [...] a dialogue in which buyer and seller learn more about one another." Aaker and Joachimsthaler (2000) argue that "to understand the Web, an experience-based model such as a theme park or a retail store is a better metaphor than passively received advertising."

Apart from providing numerous benefits proliferation of unrestricted Internet access has enabled new forms of unwanted communication. The problem of receiving unwanted email is known quite some time (e.g., Denning 1982). "Spam" however has become a problem just recently. Spam is a term denoting the unsolicited sending of typically commercial email. Although relatively young spam has started to threaten the unrestricted information exchange on the Internet which many consider one of the building blocks of the Internet.

In this paper we focus on impacts of spamming on business communication. This does not mean that spamming does not impact other areas. There are findings suggesting that impacts on consumer communication may be even stronger as consumers may prefer services free of charge over services they have to pay for. Due to their specific business models "free" services may tend to pay less attention to ensuring reliable mail delivery than services targeting business customers.

From a business perspective spam sent to employees was regarded as something annoying but spam was rarely considered a serious problem that businesses need to address in a systematic way. Typically it was assumed that deleting a spam message popping up in the inbox takes only a few seconds.

Paper Accepted as a Regular Paper

A lack of investigations of the economical impacts of spam might have contributed to the underestimation of the looming spam problem.

One of the first investigations of the spam load corporate networks have to cope with was conducted by Cranor and LaMacchia (1998). They investigated the mail intake at selected AT&T and Lucent mail domains and found that in April 1998 only 2.5% would classify as spam. They conclude that a 30% spam figure mentioned by AOL may not be typical for businesses in general as AOL users might attract an unusually high number of spam messages. Cranor and LaMacchia (1998) also report however that between April 1998 and the end of the study in August 1998 the amount of spam sent to the AT&T and Lucent domains under observation doubled to 5%.

Just a few years later the situation looks quite different. Market researchers Aberdeen Group (2002) estimate in their "2003: Predictions for Security and Privacy" report that spam in corporate networks is around 25% of all emails and will reach 50% within the year's time.

Investigations by Ferris Research estimate the damage caused by spam at 8.9 billion USD for U.S.-American companies and around 2.5 billion USD for European companies. The damage for U.S.-American and European ISPs (Internet Service Providers) is estimated to be around 500 million USD. When calculating the loss of productivity caused by spam Marten Nelson of Ferris Research assumed that deleting single spam-messages would take around a second; further time may be needed if spam is not immediately recognized as such and if emails classified as spam (so-called false-positives) have to be researched in corporate mail archives. Assuming in average 4.4 seconds work per spam mail adds up to 4 billion USD annual loss of productivity for U.S.-American companies. This means in particular that from a business perspective the wide-spread assumption that removing spam from an inbox takes only insignificant time is wrong.

Considering these figures it is not a big surprise that companies are increasingly investing in anti-spam measures. New companies have specialized on developing anti-spam tools specifically designed for companies and their networks. Typically, these tools are targeting spam on the mail transportation level. This means these tools focus on filtering and/or blocking email.

From a business perspective however there are a number of issues that concern side-effects and wider impacts of anti-spam measures with regard to corporate communications. After all, properly working email is mission-critical and inappropriate handling of customer emails may result in the loss of customers and other damages.

In this paper we argue that implementing typical anti-spam measures means that computers are assigned the power to assess legitimacy of email. We proceed as follows. In the next section we briefly summarize some background information about spam: definitions of spam, the important difference between "solicited" and "unsolicited", products typically advertised by spamming, and three points indicating why spamming could become such a serious problem. Then we outline the most common anti-spam measures and argue that the specific characteristics of these measures demand careful consideration of side-effects. Finally we present conclusions and outline future research directions.

Background: What is Spam?

In what follows we briefly discuss definitions of spam, the difference between "solicited" and "unsolicited", products typically advertised by spamming, and a summary of why spamming could become a serious problem.

Definitions of Spam

In the context of computer networks the term "spam" originally referred to a specific kind of Usenet (NetNews) postings (Baseley 1998). In the meantime using the term for email has become common practice. Mueller (2003) provides a spam definition that covers both Usenet postings and email:

> Spam is flooding the Internet with many copies of the same message, in an attempt to force the message on people who would not otherwise choose to receive it. Most spam is commercial advertising, often for dubious products, get-rich- quick schemes, or quasi-legal services. Spam costs the sender very little to send -- most of the costs are paid for by the recipient or the carriers rather than by the sender.

The Difference between Solicited and Unsolicited

Baseley (1998) explains the crucial difference between "solicited email" and "unsolicited email", starting with the "unsolicited" version:

> Unsolicited email is any email message received where the recipient did not specifically ask to receive it.

However Baseley (1998) also notes that

> Taken by itself, unsolicited email does not constitute abuse; not all unsolicited email is also undesired email. For example, receiving "unsolicited" email from a long-lost friend or relative is certainly not abuse. The reason that it is defined separately is that email abuse takes several forms, all of which begin with the fact that the email received is unsolicited.

Baseley (1998) illustrates the difference by a number of examples. Examples of soliciting email are posting an email request to Usenet, filling out a web form which explicitly mentions email, or subscribing to mailing lists. Examples of acts that do not, by themselves, constitute "soliciting" email are just posting a message to a Usenet newsgroup or any other public forum, chatting in IRC or other chat groups, filling out a survey form at a Web site that does not explicitly say it is for mailings, or posting one's email address on a web page.

Recipients' email addresses are often harvested from online resources: "Email spam lists are often created by scanning Usenet postings, stealing Internet mailing lists, or searching the Web for addresses" (Mueller, 2003). Other methods include "brut force" attacks (systematically generated email addresses) and "dictionary attacks" (e.g., The Spamhaus Project, 2003a). The US-based company Microsoft just filed a lawsuit against people they suspect of having harvested e-mail addresses from their hotmail mail servers (e.g., Festa, 2003).

Products Typically Advertised by Spamming

Stone and Lin (2003) describe spam as "unsolicited messages [...] with misleading subject lines and dubious propositions (from pyramid schemes to porno come-ons)". Brightmail, a company specializing on anti-spam products, provides an annual "Top 10" list of spams they receive on thousands of email "honey pots" and the topics support Stone and Lin's (2003) description. Among the most popular topics of the year 2002 were a credit card scam, the infamous "Nigeria scam" and porn site advertisements.

CAUCE (2003) provides details regarding the typical content of spam messages:

> To understand the problem of UCE [Unsolicited Commercial Email], you must first understand what is most often advertised via UCE. There are many places on the Internet where copies of UCE are reposted by recipients and system administrators in order to help notify the Internet community about where UCE is originating. Surveying mailing lists like SPAM-L and USENET newsgroups in the news.admin.net-abuse.* hierarchy, you will see that there are very few reputable marketers using UCE to advertise goods and services. To the contrary, the most commonly seen UCEs advertise:
>
> Chain letters

Pyramid schemes (including Multilevel Marketing, or MLM)

Other "Get Rich Quick" or "Make Money Fast" (MMF) schemes

Offers of phone sex lines and ads for pornographic web sites

Offers of software for collecting e-mail addresses and sending UCE

Offers of bulk e-mailing services for sending UCE

Stock offerings for unknown start-up corporations

Quack health products and remedies

Illegally pirated software ("Warez")

Why Spamming Could Become a Serious Problem

CAUCE (2003) summarizes the most important aspects as to why "junk mail" could become such a serious problem:

1."Cost-Shifting"

Sending spam is extremely cheap (for the sender). The costs of spamming are paid by others: network maintainers, recipients, etc.

2."Fraud"

Often, spams pretend to be replies or follow-ups to previous inquiries to get people into opening messages.

3."Disguise origin"

Spammers can easily disguise (at least to some extent) the origin of their spam messages. Otherwise it would be just too easy to filter spam and spamming would be rendered useless.

Often spam is "relayed" by third party servers which means that the spam seems to originate from a "neutral" mail server rather than a mail server known to be operated by a spammer.

Anti-Spam Measures

In this section we briefly discuss the two main approaches to fighting spam which are filtering and blocking. Commercial products typically use combinations of different filtering and learning techniques (see Metz, 2003 for a product overview) and may also include blocking techniques. Approaches can be implemented either on the user's desktop or on the respective mail server. There is a business trend to redirect corporate mail to remote servers operated by anti-spam companies. In this case the anti-spam company operates the server and maintains the anti-spam measures applied.

From a technical point of view both filtering and blocking have their specific advantages and disadvantages which will be discussed in the next section.

Content-Based Filtering

Content-based filtering targets the information contained in the body of a mail message (the message 'itself') or the information contained in the message header (mostly 'transport' information but also From: and Subject: headers).

Body-based filtering

A large part of spam messages advertises the same type of products and accordingly tends to use a certain vocabulary.

Based on this knowledge it is possible to design spam filters that filter messages containing terms that are "typical" for above mentioned types of messages. Examples for such terms would be "free porn", "XXX", "Warez" or "Get rich quick".

"Learning" spam filters can be fed new spam instances that have not yet been identified by the software. The software then extracts typical terms and uses them to identify future messages with similar characteristics as spam.

A rather powerful content-based approach is "SpamNet" (Cloudmark 2003) which is a distributed spam-recognition approach that is based on a "community consensus model". SpamNet is based on P2P technology and allows users to share "unique identifiers" of spam they received. Those identifiers can then be used to filter the spam at other sites that have not seen this particular piece of spam yet.

Origin-based filtering

Often spam is sent through mail servers that are known to be operated by spam-friendly companies. Spam messages may also show features that are characteristic for spammers trying to disguise the origin of their messages. Such characteristics can be used to filter spam based on its origin. The following example of actual porn spam received by the author helps illustrate the point:

> From: "hallo unbekannter" <lust_auf_mich2000@yahoo.de>
>
> Received: from mx0.gmx.de (HELO mx0.gmx.net) (213.165.64.100)
>
> by [the mail server used by the author] with SMTP; 14 Jan 2003 16:42:10 -0000
>
> Received: from acb84c8a.ipt.aol.com (HELO mail.mailer.de)
>
> (172.184.76.138)
>
> by mx0.gmx.net (mx029-rz3) with SMTP; 14 Jan 2003 16:42:51 -0000

[extract of header information]

The message was part of a series of more than 1000 (in words: thousand) porn spam messages. According to the header information all porn spam messages were sent though AOL. In the case of the spam described above the header information says the spam originates from AOL host acb84c8a.ipt.aol.com. In the case of this particular message a spam filter might consider the message's origin and the (probably faked) "HELO" information when computing the likelihood that the message qualifies as spam. The HELO information is part of the email transmission protocol and is easily faked; it is not unlikely that the "mail.mailer.de" header information is actually inserted by a specific software used by the spammer.

Blocking

A more radical approach to fighting spam is blocking mail delivery during the transport process when suspicious mail servers are trying to deliver messages. Knowing that a particular mail server is often used for sending spam it is possible to configure the own mail server such that mails from the suspicious server are no longer accepted. Such blocking approaches can use IPs addresses, domain names (if they do a reverse DNS lookup) and other information provided during the "handshake" between mail servers.

Lists of "spammer-friendly" mail servers are shared on the internet. Services such as Spamhaus Block List (SBL) allow other servers to check suspicious mail servers in realtime by making an inquiry to SBL over the internet when a suspicious server tries to deliver mail.

The Spamhaus Project (2003b) summarizes the approach as follows:

> The Spamhaus Block List (SBL) is a realtime database of IP addresses of static spam-sources, including known spammers, spam operations and spam support services. SBL listings are based in part on Spamhaus' Register of Known Spam Operations (ROKSO) database, spammers which Spamhaus believes are responsible for 90% of all American and European spam.

Impacts of Anti-Spam Measures

In the previous sections we have outlined that spam causes significant costs. Anti-spam measures such as filtering incoming mail or blocking suspicious mail servers can be used to significantly reduce the amount of spam messages. Still companies investing in anti-spam technology need to consider that deploying anti-spam technology is not a "free ride". By this we mean that implementing such measures always involves coping with specific advantages and disadvantages.

The advantage of spam filters is that the sheer amount of spam received by companies and their employees is significantly reduced. Typical measures for the success of anti-spam measures such as the number of false-positives (genuine mail classified as spam) indicate that these measures work reasonably well (see Metz 2003 for an overview).

The disadvantage is that a properly working corporate mail system is no longer a question of technically efficient and reliable mail transportation but it also becomes a question of "appropriate" content.

Baseley (1998) discusses the important difference between "solicited email" and "unsolicited email" but these differences are fuzzy and to some extent in the eye of the observer. Most people would agree that the infamous 'Nigeria' scam is spam that deserves to be blocked but opinions may be divided in the case of political messages. The latter may be just as 'unsolicited' as the Nigeria scam but many might be willing to receive the message. Similarly jokes may be received unsolicited but according to the author's experience many simply like to receive them.

It is questionable if filters will ever be able to reliably distinguish between legitimate mail and advertisements. In the words of Hall (1997): "User-written rule sets are doomed to lose an `arms race' against clever human junk mailers."

A recent spam example may illustrate the point:

> Date: Sat, 29 Mar 2003 09:15:07 -0800
>
> From: Kate <beatwar@mega781.com>
>
> To: [author's email address]
>
> Subject: Should America be at War?
>
> Message-ID: <155195208-1463747838-1048958107@b.Mega78.com>

In the context of the political situation end of March 2003 (war against Iraq just started) it is more than likely that quite a few concerned people would be willing to look at the email as it fits their emotional situation:

> With the current situation in the Middle East, we're conducting a United States Consumer survey about the WAR WITH IRAQ and we'd like to invite you to help us out.
>
> Let your opinions be heard amongst Americans and the world!

> After answering the question above, you qualify to receive a complimentary* USA T-shirt. Your opinion counts!
>
> [...]

First of all however the spam seems to exploit for address collection purposes the willingness of ordinary citizens to participate in public opinion formation processes. Most likely the link to click (http://c.Mega78.com/maaaXyBaaWZy2abIHBMc/) is "personalized" such that clicking the link would reveal which spam message sent to which email address triggered the clicking.

Verbal descriptions of spam, such as "Internet spam is one or more unsolicited messages, sent or posted as part of a larger collection of messages, all having substantially identical content." (Monkeys 2003) can be produced. Still it is a long way from verbal descriptions to technical operationalizations of these definitions.

So far there is no precise *technical* definition of (email) spam that could be used by mail servers to verify "solicitation" exists. In fact we do not think that such a definition is possible (see below for ways to circumvent the problem). Rather static user profiles can hardly cover all aspects of solicitation, such as having posted deliberate solicitations to Usenet newsgroups, having filled out web forms requesting emails, or having friends and colleagues to send email.

Approaches to filtering spam however are necessarily operationalizing *technical* definitions which means they incorporate such definitions even if they may not be explicit. Similarly, approaches to blocking spam are necessarily based on definitions of what distinguishes "good" and "bad" mail servers.

Without going into too much detail it is clear that the very nature of filtering/blocking based anti-spam measures implies that these measures may impact the distribution of *legitimate* email. This is a worry as for many companies properly working email has become mission-critical.

Two examples may illustrate how content-based filtering methods may prevent delivery of legitimate content or email:

1. Weaknesses of filtering technology in the context of the world wide web were experienced when filtering and blocking tools used to prevent children from accessing inappropriate content blocked legitimate content. An infamous example is the filter software SurfWatch blocking access to the White House home page because the site used the word "couples" (e.g., CNet 1996).
2. In the context of email a similar incident happened just recently when members of the British parliament did not receive messages relating to the "Sexual Offences Bill" under discussion (Heise Online News, 2003). Assumed to be porn the messages had been filtered by anti-spam filters.

Blocking emails on the basis of the mail server that is used to transport the emails is efficient as delivery is prevented. Legitimate mail may be blocked as customers may use ISPs that are 'black-listed' without even knowing about the matter. The Spamhaus Project (2003) addresses the problem as such:

> 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.

Of course, filters may be "individualized", filtering technology is getting smarter and block-lists can be updated but these are gradual improvements; the underlying problem is not addressed.

Alternative Anti-Spam Approaches

As discussed in the previous sections both filtering and blocking measures are inherently problematic as the task of classifying email as legitimate or not is assigned to computers (based on definitions of what constitutes spam).

Proposals for alternative anti-spam measures that are not based on filtering or blocking do exist. We see two major approaches:

Filter/Block-Free Communication Channels

Examples of filter/block-free communication channels are "pricing via processing" and "selling interrupt rights". These approaches are block/filter free in the sense that no further blocking/filtering needs to be applied to messages once they met an "entry criteria". This means that it is not necessary to have computers calculate legitimacy of emails.

"Pricing via processing"

Dwork and Naor (1993) discuss pricing via processing in the sense that senders have to compute complex "pricing" functions in order to get a token that allows them to deliver mail. It is reasonable to assume that spammers would not have the resources to compute a large number of tokens. Accordingly they would be prevented from sending a large number of spam emails.

"Selling interrupt rights"

Fahlman (2002) discusses a similar model requiring senders of messages to buy 'interrupt rights'. It is a reasonable assumption that spammers would not be willing (or able) to pay.

Such approaches could be used to make sure that filter-free and block-free communication channels do exist. Email may still be rejected but then it is not because of message content or message origin but because clearly defined delivery conditions were not met.

The "channel"-based approach as investigated by Hall (1998) could be used as an integrating framework. Of course other media such as phone, fax and snail mail provide communication channels as well but for some reason customers may consider using these channels as inconvenient (which means customers may not want to use them).

Networks of "Trustworthy" ISPs

Another alternative is a network of "trustworthy" mail servers that are allowed to deliver messages to corporate communication channels that are not spam-filtered. All other mail servers would be denied access to these communication channels. ISPs operating trustworthy mail servers would grant access to these mail servers only to trustworthy customers and would act immediately if abuse occurs.

Anecdotal evidence suggests that customers experiencing problems with ISPs that are blocked by the Internet community might be willing to pay for access to reliable mail servers. According to information posted to the newsgroup news.admin.net-abuse.mail, one of the largest ISPs in Italy is offering cheap broadband access but their mail servers get blocked because of a reputation for harboring spammers. Ercolessi (2003) describes how Italian IT managers cope with the situation:

> "The most savvy of IT managers know by now that Interbusiness has no abuse desk, that a large part of the world blocks their IP space, and for this reason they often buy mail services elsewhere, but STILL buy Interbusiness connectivity on purely economical grounds."

Networks of trustworthy ISPs would also face a number of difficulties. First, spammers are infamous for trying to deceive ISPs about the nature of their businesses. Then, there are difficulties to define what constitutes "trustworthy" ISPs and to make sure that a network of trustworthy ISPs is still trustworthy.

Conclusions

The discussion in this paper indicates that although anti-spam measures may provide enormous benefit they also create specific problems. This means that deploying anti-spam measures in corporate environments demands careful reviewing of the pros and cons of anti-spam measures. From a scientific point of view the first international conference on spam filtering at MIT earlier this year (MIT 2003) was an important first step. The next step should be to widen the context in which the spam phenomenon is explored. This means that future conferences on spam should also pay attention to impacts of anti-spam measures on the information distribution on the internet as impacts on corporate communications as discussed in this paper are just one aspect of the problem. Increasing interest in online communities and the trend towards e-government suggest there are quite a few areas in which impacts of anti-spam measures need to be clarified.

References

Aaker, D. and Joachimsthaler, E. (2000). Brand leadership. *The Free Press*. New York, NY, USA.

Aberdeen Group (2002). 2003: *Predictions for security and privacy*. Article available at URL http://www.aberdeen.com/ab%5Fcompany/researchareas/security2003.htm (last access 01/13/03).

Baseley, W. D. (1998). *The email abuse FAQ* (last updated June 25, 1998) Article available at URL http://members.aol.com/ emailfaq/emailfaq.html (last access 01/13/03). The FAQ can also be found in the Usenet newsgroups news.admin.net-abuse.email, news.answers.

Castells, M. (2001). *The Internet galaxy*. Oxford University Press, New York, USA.

CAUCE (2003). *Coalition Against Unsolicited Commercial Email* (CAUCE). The problem. Available at URL http://www.cauce.org/about/problem.shtml (last access 1/10/03).

Cloudmark (2003). *SpamNet* http://www.cloudmark.com/products/spamnet/ (last access 1/10/03).

CNet (1996). SurfWatch to give users more control. *CNet News*. Article available at http://news.com.com/2102-1023-211137.html (last access 02/1503).

Cranor, L.F. and LaMacchia, B.A. (1998). Spam! *Communications of the ACM* August 1998, / Volume 41, No 8, pp. 74-83.

Denning, P.J. (1982) Electronic junk. *Communications of the ACM* Vol 25 No 3, pp. 163-165.

Dwork, C. and Naor, M. (1993). Pricing via processing or combatting junk mail. Lecture Notes in Computer Science 740 (*Proceedings of CRYPTO'92*), pp. 137-147.

Ercolessi, F. (2003). Re: interbusiness.it? Posting to the newsgroup news.admin.net-abuse.email 10 Feb 2003. Message-ID <b296sr$2ivh$1@half.spin.it>.

Fahlman, S. E. (2002). Selling interrupt rights: a way to control unwanted e-mail and telephone calls. *IBM Systems Journal Vol 41*, no 4, 2002, pp. 759-766.

Festa, P. (2003). Microsoft going after hotmail spammers. *CNet News*. Available at http://msnbc-cnet.com.com/2102-1023-985018.html (last access 02/1903).

Hall, R. (1997). *Channels: avoiding unwanted electronic mail*. Available from ftp://ftp.research.att.com/dist/hall/papers/agents/channels-long.ps (last access 19/02/03). A shorter version (without the quote) was published as "How to avoid unwanted email" in Communications of the ACM, Volume 41, No 3, pp. 88-95.

Heise Online News (2003). *Spam-Filter verärgert britische Abgeordnete*. Available at URL http://www.heise.de/newsticker/data/wst-09.02.03-003/ (last access 02/09/03).

Kania, D. (2001). *BRANDING.COM - Online branding for marketing success*. NTC Business Books. Lincolnwood, IL, USA.

Metz, C. (2003). Corporate antispam tools. *PC Magazine*. Available at URL http://www.pcmag.com/article2/0,4149,849390,00.asp (last access 02/16/03)

MIT (2003). *Spam Conference* http://spamconference.org/ (last access 01/13/03).

Monkeys (2003). *Spam defined*. Available at URL http://www.monkeys.com/spam-defined/ (last access 02/16/03)

Mueller, S. H. (2003): *What is spam?* Article available at http://spam.abuse.net/overview/whatisspam.shtml (last access 01/13/03).

Stone, B. and Lin, J (2002). Spamming the world. *Newsweek*, August 19, pp. 40-42.

The Spamhaus Project (2003a). *Spammers grab hotmail and msn addresses*. Available at URL http://www.spamhaus.org/newsdog.lasso?article=114 (last access 01/13/03).

The Spamhaus Project (2003b). *The Spamhaus block list (SBL) advisory frequently asked questions.* Available at URL http://www.spamhaus.org/sbl/sbl-faqs.lasso (last access 01/13/03).

Biography

Christopher Lueg is a Senior Lecturer at the Faculty of Information Technology, University of Technology, Sydney, Australia. His degrees in computer science are from the University of Dortmund, Germany, and his doctoral degree is from the Faculty of Science, University of Zurich, Switzerland. Christopher's research interests are trans-disciplinary at the intersection of computer science, information science and cognitive science. Current research topics are, among others, e-business and the impacts of unrestricted information distribution on the Internet.