The Effect of Filters on Spam Mail

Eran Reshef*and Eilon Solan†

June 14, 2005

Abstract

We provide a model to study the effect of improving the quality of the filters that users use, and the increase of the cost of mailing spam messages, on the total number of messages that spammers send.

We show that increasing the cost of mailing spam messages have the unequivocal effect of reducing the total number of spam messages sent by spammers. On the other hand, improving the quality of the filters have an ambiguous effect: users who uses the improved filter receive less spam messages, but the total number of messages sent by spammers might increase, and other users, who keep their old filters, might receive more spam messages.

1 Introduction

Bulk electronic mail, also known as spam, is a serious problem, which becomes graver every year. Spam volume has jumped five fold since 2003, and according to Ferris Research estimates, spam will cost about $50 billion globally in 2005.

Many authors have suggested different ways to fight spam mail using various techniques (see, [1]-[15] among others).

*Blue Security Inc., www.bluesecurity.com

†School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel, and MEDS Department, Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208-2001. e-mail: eilons@post.tau.ac.il

Some authors have developed models to analyze some aspects of spam mail. For example, [6] examines the efficiency of various approaches to regulate spam mail, and emphasize legal aspects; the effects of pricing mail messages are examined in [14] and [7].

The most common way to fight spam messages is by installing a filter, which tries to identify spam messages and filter them out. The filter that a user uses indeed reduces the number of spam messages that the user receives, but it does not entirely eliminate the problem. Moreover, filters have inherent problems. For example, they are difficult to maintain, senders of spam mail adapt to their strategies, and they sometimes filter innocent mail. On top of that, once filters improve, to reach their potential customers, spammers, who send the bulk electronic mail, have to increase the amount of spam messages they send, so it is not clear what is the overall effect of the improved filter on the user. Finally, since the number of spam messages increases, the internet itself is more loaded, and its efficiency decreases.

In the present paper we present a model that allows us to study the effects of improving the quality of the filters on the amount of spam messages spammers send, and on the amount of messages that users receive.

At a first approximation we suppose that there are two types of users: users who use high-quality filters, and users who use low-quality filters. The first group of users may include rich users, users who purchase commercial filters, and poor users whose internet provider provides a high-quality filter. The second group may include users who do not purchase commercial filters and whose internet provider does not supply a high-quality filter. We call the first type of users *rich users*, and the second *poor users*.

We study in details the optimization problems of the spammers: what is the optimal number of messages the spammers should send to the users. Often in optimization problems, especially when several decision makers are involved, there are several optimal solutions, and therefore a question arises concerning which solution will be chosen by society. We prove that in *all* optimal solutions the total number of messages that are sent to each user is the same, so that, from the point of view of users, it does not matter which optimal solution is chosen by the spammers.

We go on to study qualitative properties of the solution. There are two intuitive ways to reduce the number of spam messages. First, one can increase the cost of sending those messages, and second, one can improve the quality of the filters.

We show that increasing the cost of sending spam messages has the un-

equivocal effect of reducing the total number of spam messages sent by spammers. This implies that the number of messages that passes the filters decrease as well.

However, improving the quality of the filter has an ambiguous effect: the number of messages that the user whose filter is improved indeed decreases, but the total number of messages sent by spammers might increase. This implies that (i) other users, whose filters do not improve, might receive more spam messages, and (ii) the internet might becomes less efficient. It turns out that this ambiguous effect happens when the cost of sending spam messages is not too high. When this cost happens to be very high, the effect of improving filters is unequivocal. We provide a simple condition that indicates the minimal cost above which improving filters is beneficial for the whole society.

The main conclusion from our results is that to effectively fight spam messages the cost of sending spam messages must increase. Without such a measure, the huge investment in improving filters only causes flooding the internet with more spam messages, thereby reducing its efficiency, and flooding poor users, who do not have the state-of-the-art filters, by more and more spam messages.

2 The Model

We now present the formal model that we study.

2.1 The population

There are M *users*, who are divided into two groups - *rich users* and *poor users*. We denote by β the percentage of rich users among all users.

Most users delete spam messages the moment they receive them. However, some users read the spam messages they receive, and sometimes purchase the product that is being advertised, termed a spam product. We call these users *potential buyers*, and denote the percentage of potential buyers among all users by s. We assume that potential buyers are evenly distributed among the population. That is, the percentage of potential buyers among both poor users and rich users is s.

We assume that all users have a spam filter. However, rich user have better filters than poor users. The quality of a filter is measured by the

percentage of spam messages it allows to pass. We denote the quality of the filter of poor users and rich users by q and r respectively. Thus, a poor users's filter filters out $1-q$ percents of the spam messages, while a rich user's filter filters out $1-r$ percents of the spam messages. Setting $q=1$ effectively means that poor users do not have a filter.

As rich users have better filters, we have

$$r < q. \tag{1}$$

As we said before, a potential buyer may purchase spam products. To simplify the analysis, we suppose that each potential buyer may spend an amount of T dollars on spam products every year, and he or she does so in a single purchase.

We denote the probability that the potential buyer will not purchase the product that is advertised in any given spam message by p. That is, when a potential buyer receives a spam message, with probability p he or she deletes it, and with probability $1-p$ he or she spends T dollars on the spam product. Once this user spends his or her T dollars, he or she will not purchase any other spam product that year.

2.2 The spammers

There are N *spammers*. The decision problem of each spammer is how many spam messages to send to each user. We assume that spammers do not distinguish between users, so that a spammer mails the same number of messages to all users. We also assume that the users do not distinguish between the spammers, so that the percentage of potential buyers who purchase from any given spammer is given by the percentage of spam messages that user receive from the spammer. That is, if a given spammer, say Spade, sends x messages to a given user, while all other spammers send y messages to that user, then, provided the user purchases a spam product, the probability he or she will purchase it from Spade is $x/(x+y)$.

Spammers have two types of cost - a fixed cost and a per-message cost. The fixed cost of a spammer is denoted by D dollars per year. The cost of each message is denoted by d dollars per message.

The goal of each spammer is to maximize his or her expected gain.

2.3 The Payoff of Spammers

We now analyze the decision problem of a specific spammer. To this end, we calculate the expected gain of a spammer.

Denote by x the number of messages per user that a specific spammer, say Spade, sends, and by y the total number of spam messages per user sent by all other spammers.

Thus, the total number of messages sent to each user is $x + y$, while the number of (non-filtered) spam messages he actually receives depends on his or her filter: poor users receive $q(x + y)$ messages, while rich users receive $r(x + y)$ messages.

Spades's cost is

$$D + xMd. \tag{2}$$

The expected number of potential buyers who actually purchase a spam product is

$$Ms\left((1-\beta)\left(1-p^{q(x+y)}\right)+\beta\left(1-p^{r(x+y)}\right)\right). \tag{3}$$

Indeed, the number of potential buyers is Ms. Among those buyers, the percentage of rich users is β. Each such buyer receives $r(x + y)$ messages, so the probability he or she purchases a spam product is $1 - p^{r(x+y)}$. Similarly, the probability that a potential buyer who does not have a filter purchases a product is $1 - p^{q(x+y)}$.

The proportion of spam messages sent by Spade among all spam messages is $\frac{x}{x+y}$. Therefore the percentage of buyers who purchase from Spade is $\frac{x}{x+y}$. As each potential buyer spends T dollars when purchasing a spam product, the total expected payoff of Spade is

$$W(x,y) = -D - xMd + \frac{x}{x+y}TMs\left((1-\beta)\left(1-p^{q(x+y)}\right)+\beta\left(1-p^{r(x+y)}\right)\right). \tag{4}$$

If $W(x, y) < 0$, Spade has a negative payoff, and will go out of market. Otherwise, Spade makes a profit.

For the mathematical analysis it is more convenient to assume that x, the number of spam messages the spammer sends to each user, need not be an integer, but can be any non-negative real number.

2.4 Stable configurations

A vector $\vec{x} = (x_1, \ldots, x_N)$, which indicates the number of spam messages each spammer sends, is termed a spam configuration.

Definition 1 *A* spam configuration *(or simply a* configuration*) is a vector* $\vec{x} = (x_1, \ldots, x_N)$*, where for each* i *the quantity* x_i *is the number of spam messages per user sent by spammer* i*.*

The market reaches a stable state if no spammer finds it profitable to change the number of spam messages he mails to each user. Usually, decision makers adapt to small changes in the environment by slightly changing their behavior, and they do not consider significant changes that may increase their profit. We therefore define a spam configuration to be stable if no spammer can profit by slightly changing the number of spam messages he or she sends.

Definition 2 *A spam configuration* $\vec{x} = (x_1, \ldots, x_N)$ *is* stable *if for every spammer* i

$$\frac{\partial W}{\partial x}\left(x_i, \sum_{j \neq i} x_j\right) = 0.$$

2.5 Direct cost and direct gain

The *direct cost* of mailing a single spam message is d. We define the *direct gain* of a single spam message by

$$g = Ts(-\ln p)((1-\beta)q + \beta r). \tag{5}$$

We now explain the motivation of this definition. Consider the first message that is sent. With probability s the user who receives it is a potential buyer, with probability $(1-\beta)q + \beta r$ the message is not filtered, and with probability $1-p$ it causes the user to purchase a spam product. Therefore, a single message sent to a random user has probability $s(1-p)((1-\beta)q + \beta r)$ to make that user purchase a spam product. In practice, p is close to 1, and therefore $(-\ln p)$ is close to $1-p$. As the amount spent in a spam purchase is T dollars, g represents the expected income from the first message that is sent to a random user.

As we show below, the behavior of the market is different when $g > d$ and when $g < d$.

3 Results

The first result states that if the direct gain is lower then the direct cost, there will be no spam.

Theorem 3 *Suppose that $g < d$, and let $\vec{x} = (x_1, \ldots, x_N)$ be any configuration. If $\sum_{i=1}^{N} x_i > 0$ then there is i such that (i) $x_i > 0$, and (ii) $W\left(x_i, \sum_{j \neq i} x_j\right) < 0$.*

The theorem says that if the direct gain is lower then the direct cost, then if some spammers do mail spam messages, at least one of those spammers suffers a loss. Thus, in this case sending spam messages is not profitable, and there will be no spam.

The second result states that a stable configuration exists as soon as the direct gain exceeds the direct cost.

Theorem 4 *If $g > d$ a stable configuration exists.*

In many models where stable points exist, there exist many such stable points, and then a question arises concerning which stable point the market will reach, and which factors that were not taken in the analysis influence the outcome. The next result states that from the point of view of users, all stable configurations are equivalent - each user receives the same amount of spam messages in all stable configurations. Thus, even though there might be several stable configurations, in which the total amount of spam is divided in different ways among the spammers, the users are indifferent among those stable configurations.

Theorem 5 *Let $\vec{x} = (x_1, \ldots, x_N)$ and $\vec{x}' = (x'_1, \ldots, x'_N)$ be two stable configurations. Then $\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} x'_i$.*

After we established the general existence and uniqueness of stable configurations, we turn to study qualitative properties of those configuration. Namely, how changes in the environment affect the total number of messages that are sent by spammers or received by users.

For every p, q, r, β, and d denote by $z(p, q, r, \beta, d)$ the total number of messages sent by spammers in all stable configurations (by Theorem 5 this quantity is well defined).

Two ways to fight spam messages are (i) increasing the cost of sending those messages, and (ii) improving the quality of the filters. The following two theorems establish that these two methods indeed reduce the amount of spam messages that a user receives. Theorem 6 shows that as the direct cost increases, fewer messages are sent by spammers. Theorem 7 shows that as filters improve, users receive less spam messages.

Theorem 6 *For every fixed p, q, r, and β, the function $d \mapsto z(p,q,r,\beta,d)$ is monotonic decreasing.*

Theorem 7 *For every fixed p, q, β and d, the function $r \mapsto r \cdot z(p,q,r,\beta,d)$ is monotonic increasing. For every fixed p, r, β and d, the function $q \mapsto q \cdot z(p,q,r,\beta,d)$ is monotonic increasing.*

Observe that there is a significant difference between Theorem 6 and Theorem 7. Whereas Theorem 6 states that increasing the direct cost reduces the total number of messages spammers *send*, Theorem 7 says that improving the filter reduces the number of spam messages that the user *receives*. Theorem 7 does not say that the total number of spam messages sent decreases as well. Thus, it might happen that improving the quality of the filter of rich users indeed reduces the total number of spam messages that rich users receive, but the total number of messages sent by spammers increase. This has the effect of (i) making the internet less efficient, and (ii) increasing the number of spam messages that poor users receive (since the quality of their filter did not change).

We explore this issue now.

Theorem 8 *The function $r \mapsto z(p,q,r,\beta,d)$ is monotonic decreasing if and only if*

$$(1-\beta)\left(qe^{-qN/r} + r\frac{N-1}{N}(1-e^{-qN/r})\right) + \beta r\left(e^{-N} + \frac{N-1}{N}(1-e^{-N})\right) > \frac{dN}{Ts(-\ln p)}. \tag{6}$$

Similarly, the function $q \mapsto z(p,q,r,\beta,d)$ is monotonic decreasing if and only if

$$(1-\beta)\left(qe^{-N} + \frac{N-1}{N}(1-e^{-N})\right) + \beta\left(re^{-rN/q} + \frac{N-1}{N}q(1-e^{-rN/q})\right) > \frac{dN}{Ts(-\ln p)}. \tag{7}$$

Thus, suppose that the function $r \mapsto z(p,q,r,\beta,d)$ is monotonic decreasing. This means that as r decreases (i.e., filters improve) z increases (i.e., more spam messages are sent). Since the right-hand side in (6) depends linearly on d, this result implies that if the direct cost of sending a single message is not too high, improving the quality of the filters has the undesirable effect of increasing the total spam messages sent by spammers.

Observe that Theorem 8 provides a simple condition that should be satisfied so that improving the filters is beneficial to all users, and to society as a whole.

4 Existence and uniqueness of a solution

In the present section we prove Theorems 3-5.

Before diving into the proofs we calculate the derivative of W.

$$\begin{aligned}\frac{\partial W}{\partial x}(x,y) &= -Md+\frac{x}{x+y}TMs(-\ln p)\left((1-\beta)qp^{q(x+y)}+\beta rp^{r(x+y)}\right)\\ &\quad+\frac{y}{(x+y)^2}TMs\left((1-\beta)(1-p^{q(x+y)})+\beta(1-p^{r(x+y)})\right). \quad (8)\end{aligned}$$

Proof of Theorem 3. Assume w.l.o.g. that $x_i > 0$ for every i (otherwise, drop the spammers who do not send spam messages from the market). Denote $z = \sum_{i=1}^{N} x_i$, and recall that

$$W(x_i, z-x_i) = -D - x_i Md + \frac{x_i}{z}TMs\left((1-\beta)(1-p^{qz})+\beta(1-p^{rz})\right). \quad (9)$$

Summing Eq. (9) over $i = 1, \ldots, N$, and using Lemma 10 we obtain

$$\begin{aligned}\sum_{i=1}^{N} W(x_i, z-x_i) &= -ND - zMd + TMs\left((1-\beta)(1-p^{qz})+\beta(1-p^{rz})\right)\\ &< zM\left(-d+\frac{Ts}{z}\left((1-\beta)(1-p^{qz})+\beta(1-p^{rz})\right)\right)\\ &< zM\left(-d+Ts\left((1-\beta)(-\ln p)q+\beta(-\ln p)r\right)\right)\\ &= zM(g-d).\end{aligned}$$

Hence, if $g < d$ we have $\sum_{i=1}^{N} W(x_i, z-x_i) < 0$, so that at least one spammer receives a negative payoff. ■

Proof of Theorem 4. Substituting $(x,y) = \left(\frac{1}{N}z, \frac{N-1}{N}z\right)$ in (8) we obtain:

$$\begin{aligned}\frac{\partial W}{\partial x}\left(\frac{1}{N}z, \frac{N-1}{N}z\right) &= -Md+(-\ln p)\frac{TMs}{N}\left((1-\beta)qp^{qz}+\beta rp^{rz}\right)\\ &\quad+\frac{(N-1)TMs}{Nz}\left((1-\beta)\left(1-p^{qz}\right)+\beta\left(1-p^{rz}\right)\right).\end{aligned}$$

One can easily verify that

$$\lim_{z\to\infty}\frac{\partial W}{\partial x}\left(\frac{1}{N}z, \frac{N-1}{N}z\right) = -Md < 0. \quad (10)$$

By (16)

$$\lim_{z\to 0} \frac{\partial W}{\partial x}\left(\frac{1}{N}z, \frac{N-1}{N}z\right) = M(g-d). \tag{11}$$

Since $g > d$, the limit in (11) is positive. Since the function $z \mapsto \frac{\partial W}{\partial x}\left(\frac{1}{N}z, \frac{N-1}{N}z\right)$ is continuous, it follows from (10) and (11) that there is $z_* > 0$ with

$$\frac{\partial W}{\partial x}\left(\frac{1}{N}z_*, \frac{N-1}{N}z_*\right) = 0. \tag{12}$$

Define the configuration $\vec{x} = (x_1, \ldots, x_N)$ by

$$x_i = \frac{z_*}{N}, \quad \forall i.$$

By (12) $\vec{x}$ is stable. ■

Define

$$\begin{aligned} F(p,q,r,\beta,z) &= (-\ln p)\left((1-\beta)qp^{qz} + \beta r p^{rz}\right) \\ &\quad + \frac{N-1}{z}\left((1-\beta)\left(1-p^{qz}\right) + \beta\left(1-p^{rz}\right)\right). \end{aligned}$$

As we will see below, this function is related to the derivative of W.

By Lemma 11, and since the function $z \mapsto p^{qz}$ is monotonic decreasing, we have the following.

Lemma 9 *For every p, q, r, β, and z, $\frac{\partial F}{\partial z}(p,q,r,\beta,z) < 0$.*

We are now ready to prove Theorem 5, which states that in all stable configurations the total amount of spam messages that are sent is the same.

Proof of Theorem 5. Let $\vec{x} = (x_1, \ldots, x_N)$ be a stable configuration, and denote $z = \sum_{i=1}^{N} x_i$. Then

$$\begin{aligned} 0 &= \sum_{i=1}^{N} \frac{\partial W}{\partial x}(x_i, z - x_i) \\ &= -NMd + TMs(-\ln p)\big((1-\beta)qp^{qz} + \beta r p^{rz}\big) \\ &\quad + \frac{N-1}{z} TMs\big((1-\beta)(1-p^{qz}) + \beta(1-p^{rz})\big) \\ &= TMs\left(-\frac{dN}{Ts} + F(p,q,r,\beta,z)\right). \end{aligned}$$

In particular, z, the total number of spam messages sent to each user, satisfies

$$F(p,q,r,\beta,z) = \frac{dN}{Ts}. \tag{13}$$

By Lemma 9 the function F is monotonic decreasing in z, hence for every p, q, r and β there can be at most one solution z to (13), and the result follows. ■

5 Qualitative Properties of the Solution

In this section we prove the qualitative findings, regarding the effects of changes in the market on the number of spam messages sent by spammers or received by users.

Proof of Theorem 6. Fix p, q, r, β and d. By Eq. (13) and Theorem 5 the total number of spam messages sent to each user in a stable configuration, $z(p,q,r,\beta,d)$, is the unique solution of Eq. (13). By Lemma 9 the function $z \mapsto F(p,q,r,\beta,z)$ is decreasing. Therefore, by increasing d we increase the right-hand side of (13), so that the left-hand side of (13) must increase as well. But this implies that the solution decreases. ■

Proof of Theorem 7. We prove only the first claim. The proof of the second claim is analogous.

Set $u = rz$. u is the total number of spam messages that bypass the filter of a rich user. Set

$$H(p,q,r,\beta,u) = F\left(p,q,r,\beta,\frac{u}{r}\right).$$

Then,

$$\begin{aligned} H(p,q,r,\beta,u) &= (-\ln p)\big((1-\beta)qp^{qu/r} + \beta r p^u\big) \\ &\quad + \frac{r(N-1)}{u}\big((1-\beta)(1-p^{qu/r}) + \beta(1-p^u)\big). \end{aligned}$$

Since the function $u \mapsto p^{cu}$ is monotonic decreasing for every $c > 0$, lemma 11 implies that

$$\frac{\partial H}{\partial u}(p,q,r,\beta,u) < 0. \tag{14}$$

Also,

$$\begin{aligned}\frac{\partial H}{\partial r}(p,q,r,\beta,u) &= (1-\beta)q(-\ln p)^2\frac{qu}{r^2}+(-\ln p)\beta p^u \\ &\quad +\frac{N-1}{u}\big((1-\beta)(1-p^{qu/r})+\beta(1-p^u)\big) \\ &\quad -\frac{r(N-1)}{u}(1-\beta)(-\ln p)\frac{qu}{r^2}p^{qu/r}.\end{aligned}$$

By Lemma 10 the third term is larger than the fourth term, and therefore $\frac{\partial H}{\partial r}(p,q,r,\beta,u)>0$.

Let $u(p,q,r,\beta,d)=r\cdot z(p,q,r,\beta,d)$ be the number of spam messages that bypass the filter of a rich user. Then

$$H(p,q,r,\beta,u(p,q,r,\beta,d))=\frac{dN}{Ts}.$$

Set

$$G(r)=H(p,q,r,\beta,u(p,q,r,\beta,d)).$$

By the chain rule

$$\begin{aligned}0 &= G'(r) \\ &= \frac{\partial H}{\partial r}(p,q,r,\beta,u(p,q,r,\beta,d))+\frac{\partial H}{\partial u}(p,q,r,\beta,u(p,q,r,\beta,d))\frac{\partial u}{\partial r}(p,q,r,\beta,d).\end{aligned}$$

Since $\frac{\partial H}{\partial r}(p,q,r,\beta,u(p,q,r,\beta,d))>0$ while $\frac{\partial H}{\partial u}(p,q,r,\beta,u(p,q,r,\beta,d))<0$ it follows that $\frac{\partial u}{\partial r}(p,q,r,\beta,d)>0$, as desired. ■

Proof of Theorem 8. Define

$$G(r)=F(p,q,r,\beta,z(p,q,r,\beta,d)).$$

By Eq. (13) $G(r)=dN/Ts$ for every r, and therefore $G'(r)=0$. By the chain rule

$$0=G'(r)=\frac{\partial F}{\partial r}(p,q,r,\beta,z(p,q,r,\beta,d))+\frac{\partial F}{\partial z}(p,q,r,\beta,z(p,q,r,\beta,d))\frac{\partial z}{\partial r}(p,q,r,\beta,d).$$

By Lemma 9 $\frac{\partial F}{\partial z}(p,q,r,\beta,z(p,q,r,\beta))<0$, and therefore

$$\frac{\partial z}{\partial r}(p,q,r,\beta,d)>0 \quad\Leftrightarrow\quad \frac{\partial F}{\partial r}(p,q,r,\beta,z(p,q,r,\beta,d))>0.$$

The derivative of F w.r.t. r is

$$\begin{aligned}\frac{\partial F}{\partial r}(p,q,r,\beta,z) &= (-\ln p)\beta p^{rz} - (-\ln p)^2 rz\beta p^{rz} + (N-1)\beta(-\ln p)p^{rz} \\ &= \beta(-\ln p)p^{rz}(N-(-\ln p)rz).\end{aligned}$$

Hence, the derivative is negative if and only if $z(p,q,r,\beta,d) > \frac{N}{(-\ln p)r}$.

Since F is monotonic decreasing, $z(p,q,r,\beta,d) > \frac{N}{(-\ln p)r}$ if and only if

$$\frac{dN}{Ts} = F(p,q,r,\beta,z(p,q,r,\beta,d)) < F\left(p,q,r,\beta,N/(-\ln p)r\right)).$$

We will now check when this inequality holds.

Observe that

$$p^{1/(-\ln p)} = \exp\left(\frac{1}{-\ln p}\ln p\right) = \exp(-1) = e^{-1}.$$

Hence

$$\begin{aligned}F\left(p,q,r,\beta,N/(-\ln p)r\right)) &= (-\ln p)(1-\beta)\left(qpe^{-qN/r} + r\frac{N-1}{N}(1-e^{-qN/r})\right) \\ &\quad +(-\ln p)\beta r\left(e^{-N} + \frac{N-1}{N}(1-e^{-N})\right).\end{aligned}$$

The result follows. ■

6 Final Comments

The model we presented is simplistic, and further research is needed to improve our understanding of the behavior of spammers in real life.

For example, some of our assumptions can be weakened. We assume that each potential buyer makes at most one purchase per year. The model can be enriched by allowing users to make more than one purchase. We assume that rich users and poor users spend the same amount of money each year. This assumption can be weakened as well. We also assume that all spammers send the same number of messages to all users. However some spammers target themselves on a specific subgroup of users.

Another point we ignore is the behavior of the anti-spam companies. For example, anti-spam companies adapt their filters to the spammers. A

spammer who increases the number of messages he sends has a higher chance of being identified by the anti-spam companies, and having the filters learn how to identify his messages. This effect may reduce the number of spam messages sent by spammers.

References

[1] S. Ahmed and F. Mithun (2004) Word Stemming to Enhance Spam Filtering Proceedings of the 1st Conference on Email and Anti-Spam (CEAS 2004), Mountain View, CA, USA.

[2] I. Androutsopoulos, G. Paliouras, V. Karkaletsis, G. Sakkis, C.D. Spyropoulos and P. Stamatopoulos (2000) Learning to Filter Spam E-Mail: A Comparison of a Naive Bayesian and a Memory-Based Approach. In H. Zaragoza, P. Gallinari, and M. Rajman (Eds.), Proceedings of the Workshop on Machine Learning and Textual Information Access, 4th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD 2000), Lyon, France, pp. 1-13.

[3] Z. Chuan, L. Xianliang and X. Qian (2004) A Novel Anti-spam Email Approach Based on LVQ. Proceedings of the 5th International Conference, PDCAT 2004, Singapore. Eds. Liew K.-M., Shen H., See S.

[4] J. Clark, I. Koprinska and J. Poon (2003) A Neural Network Based Approach to Automated E-Mail Classification. Proceedings of the IEEE/WIC International Conference on Web Intelligence, p.702.

[5] A. Gray and M. Haahr (2004) Personalised, Collaborative Spam Filtering. Proceedings of the 1st Conference on Email and Anti-Spam (CEAS 2004), Mountain View, CA, USA.

[6] D.W.K. Khong (2004) An Economic Analysis if Spam Law. Erasmus Law and Economics Review, **1**, 23-45.

[7] R.E. Kraut, S. Sunder, R. Telang, and J. Morris (2005) Pricing Electronic Mail To Solve the Problem of Spam. Human-Computer Interaction Vol. 20, pp. 195-223.

[8] B. Leiba and N. Borenstein (2004) A Multifaceted Approach to Spam Reduction. Proceedings of the 1st Conference on Email and Anti-Spam (CEAS 2004), Mountain View, CA, USA.

[9] K. Li, C. Pu and M. Ahamad (2004) Resisting SPAM Delivery by TCP Damping. Proceedings of the 1st Conference on Email and Anti-Spam (CEAS 2004), Mountain View, CA, USA.

[10] E. Michelakis, I. Androutsopoulos, G. Paliouras, G. Sakkis and P. Stamatopoulos (2004) Filtron: A Learning-Based Anti-Spam Filter. Proceedings of the 1st Conference on Email and Anti-Spam (CEAS 2004), Mountain View, CA, USA.

[11] P. Pantel and D. Lin (1998) SpamCop– A Spam Classification & Organization Program. Proceedings of AAAI-98 Workshop on Learning for Text Categorization.

[12] I. Rigoutsos and T. Huynh (2004) Chung-Kwei: a Pattern-discovery-based System for the Automatic Identification of Unsolicited E-mail Messages (SPAM). Proceedings of the 1st Conference on Email and Anti-Spam (CEAS 2004), Mountain View, CA, USA.

[13] M. Sahami, S. Dumais, D. Heckerman and E. Horvitz (1998) A Bayesian Approach to Filtering Junk E-Mail. Proceedings of AAAI-98 Workshop on Learning for Text Categorization.

[14] T. Van Zandt (2004) Information Overload in a Network of Targeted Communication. RAND Journal of Economics, **35**, 542-560.

[15] W. Yerazunis (2003) Sparse Binary Polynomial Hash Message Filtering and The CRM114 Discriminator. Proceedings of 2003 MIT Spam Conference.

Appendix

Here we prove some technical results we need regarding the function $(1 - p^x)/x$.

Lemma 10 *For every $p \in (0,1)$ and every $x > 0$ one has*

$$x(-\ln p)p^x < 1 - p^x < (-\ln p)x. \tag{15}$$

Proof. Since all terms in Eq. (15) vanish at $x = 0$, it is sufficient to compare their derivatives. That is, we need to prove that

$$(-\ln p)p^x - x(-\ln p)^2 p^x < (-\ln p)p^x < (-\ln p).$$

However, this inequality holds since $p \in (0,1)$ and $x > 0$. ■

By L'hopital's rule one obtains the following.

$$\lim_{x \to 0} \frac{1 - p^x}{x} = -\ln p. \tag{16}$$

Lemma 11 *For every $p \in (0,1)$ and every $q > 0$ the function $f(z) = \frac{1-p^{qz}}{z}$ is monotonic decreasing over $(0, \infty)$.*

Proof. The derivative of f is

$$\begin{aligned} f'(z) &= -\frac{1}{z^2}(1 - p^{qz}) + \frac{1}{z}(-\ln p)qp^{qz} \\ &= \frac{1}{z^2}\big(qz(-\ln p)p^{qz} - (1 - p^{qz})\big). \end{aligned}$$

By Lemma 10 we have $1 - p^{qz} > qz(-\ln p)p^{qz}$, and the result follows. ■