E-Mail Security as Cooperation Problem

Yacine Gasmi and Joerg Schneider[1]

Communication and Operating Systems, Technische Universitaet Berlin,
Berlin, Germany
gasmi@tu-berlin.de, komm@cs.tu-berlin.de
home page: https://www.kbs.tu-berlin.de

Abstract. E-mail communication still has to cope with certain security problems. The most visible result is the mass of unsolicited messages outnumbering the regular e-mails in magnitudes. The technical reasons for this unfavorable situation are manifold, e.g., unreliable sender authentication, loose and ad-hoc coupling between the involved servers, and only few ways to complain about misbehavior of users of foreign systems. In this paper, we will show that resolving these problems requires a deeper analysis of the related economic aspects. We will demonstrate in a simplified game-theoretical model that the individual choices of the involved parties will lead to a suboptimal state, where less effort is put into the security of e-mail communication. As a consequence, some kind of cooperation between the involved parties is needed to implement functions which provide a social benefit rather than an individual one.

Keywords: game theory, spam, filtering, mail server infrastructure

1 Introduction

E-mail is designed as an open infrastructure. It is very easy to add your own server for sending and receiving e-mails. However, this open design leads to the key problem of e-mail usage nowadays: Spam. These unsolicited messages can enter the e-mail infrastructure anywhere and fill up the mail boxes of all users.

Most users ignore the spam messages. However, a study by Kanich et al. showed that 0.00001% of the monitored spam mails lead to a sale of the advertised product [3]. Combined with a very large number of messages send out for each campaign, this results in an estimated net income of $7,000 per campaign.

An e-mail on its way from the sender to the recipient passes multiple systems. In a simplified model, these systems belong to four domains: the sending user, the senders mail server operator, the receiver's mail server operator, and the addressed end user (see Figure 1). In each of these domains, anti-spam measures can be implemented [10]. In general, the sending user has a low level of interest in the protection. On the other hand, the receiving user has the highest interest in fighting the problem as he is facing most of the costs generated by spam [7]. In this paper, we are going to concentrate on the other two players, i.e., the mail server operators. Both can filter the e-mails during transmission and reject unsolicited mails.

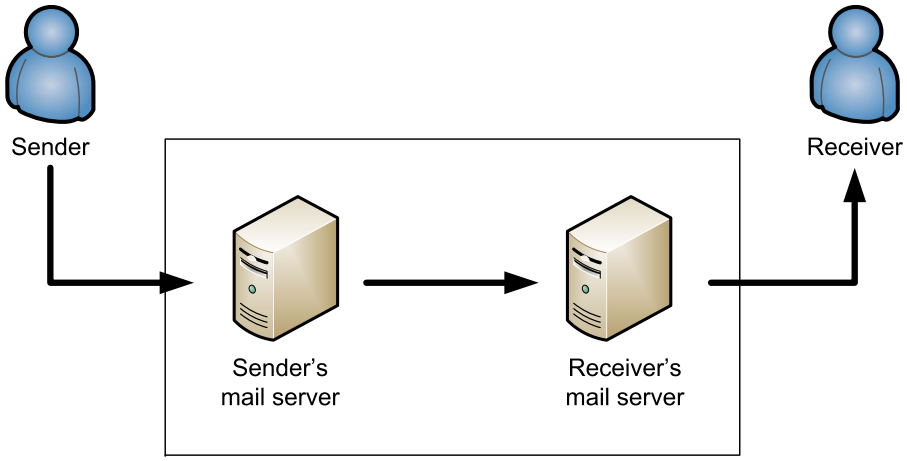

Fig. 1. Four parties are involved in each e-mail communication. In this paper we concentrate on the relation between the mail server operators.

Filtering incoming e-mails on the receiver's mail server seems to be in the interest of mail server operators. Since clients can base their purchase decisions on the spam protection provided and, hence, choose an operator that provides good spam protection. Filtering outgoing messages already at the sender's mail server has benefits for the whole e-mail infrastructure [10]. Nevertheless, as the positive effect is rather small for the operator performing the filtering, it does not enjoy great popularity among mail server operators compared to the use of incoming filters.

In this paper we model the decisions of the mail server operator in a simplified game-theoretical model. Using the model we will show that mail server operators tend to provide a rather low level of filtering outgoing mails (compared to incoming filtering).

In the remainder of this paper, we start with a discussion of established spam protection and filtering mechanisms and show related work on modeling the spam market economics. Then, we introduce our economic model and discuss our findings for an uncoordinated market. Before concluding, we include a reputation system based on black lists in the model and show the enhancement for the general spam protection level.

2 Related work

According to Lessig, the operation of technical infrastructures depends on four regulators: Law, norms, market, and architecture [5]. In this paper we will examine the economic aspects of the spam problem in mail infrastructures.

The law aspects were discussed by Moustakas et al. [8]. They compared the anti-spam laws in the European Union and the United States and identified

the different approaches. However, they concluded that anti-spam laws can be effective only if complemented by technical enforcement and international cooperation.

There are a number of technical approaches to change the mail system architecture in order to impede the distribution of unwanted messages. They are usually applied either at the receiving mail server or at the receiver's computer.

There are two prominent initiatives to authenticate the real sender of a message to be able to trace offenders of social norms and laws. One is the Domain Keys Identified Mail[1] (DKIM) system. Here, each mail server in the transport chain cryptographically signs for the correctness of the sender's identity. A user receiving an offending mail can then identify at least the mail server where the message entered the mail infrastructure.

The other identification system is called Sender Policy Framework[2] (SPF) [2]. SPF is based on the idea, that most users send their e-mails using the mail server of their own organization. Thus, each organization names distinguished mail servers authorized to send mail in their name. The receiving server can now validate, whether the mail with a given sender address entered the mail infrastructure using a server authorized by the owner of the address or not. However, traveling users have to use their authorized mail server at home and cannot use one nearby.

Another widely adopted technical strategy against spam is to include message filters in the receiving server's or client's software. Very effective and widely spread are Bayesian filters [1]. These filters are trained with old messages — wanted and unwanted — and can then identify spam messages based on the learned word count statistics. However, Bayesian filters have certain shortcomings. They can be tricked with manipulated messages and may also block legitimate messages.

Additionally, mail server operators could try to identify misbehaving parties in the infrastructure and filter messages based on the reputation of the sending mail server. DNS blackhole lists (DNSBL) are used to propagate this reputation. The blackhole lists are based on the domain name system to facilitate the established caching infrastructure. However, one does not query for the real sender domain name but for a special sub domain within the DNSBL domain. If a server was caught sending spam, e.g., by receiving a message at a honey pot address or due to user complains, it is added to the blackhole list. Most blackhole lists provide an unlisting timeout, i.e., the server is removed if no other incident is recorded. In the second part of this paper, we extend our model to include the effects of a reputation based blackhole list.

The economic aspects of main systems have also been widely discussed. However, most paper concentrate on the relationship between the sending and receiving user. For example, Khong and Building examined different strategies to regulate main infrastructures [4]. They identified three categories of approaches: opt-out, filtering and blocking, and opt-in. They showed that sending unsolicited

[1] http://www.dkim.org/
[2] http://www.openspf.org/

messages can also have social benefits and concluded that only the opt-in approach leads to no net social loss. Loder et al. discussed monetary approaches to reduce spam [6]. Four cases were compared: a baseline case without any control mechanism, a case where the user has a perfect filter, a flat tax for each message, and an Attention Bound Mechanism (ABM). The ABM protocol allows the sender and receiver to negotiate with low overhead the value of the message for both and, thus, establish an individual price. The authors showed that implementing the ABM reduces unsolicited messages and keeps at the same time the communication channel open for strangers to initiate a new connection. Other mechanisms like opt-in and filtering block such communication attempts by strangers.

In contrast to these economic works we concentrate on the relationship between the main server operators. The game-theoretical model we present in this paper is based on the work done by Hal R. Varian [9]. Varian showed how system reliability as a public good that depends on the single efforts of the involved agents results in a free-rider problem.

3 An economic model

In this section we introduce a simple game-theoretical model that shows how sending and receiving mail servers invest in filter technologies. In a second step, we extend our model to reflect the influence of blocking misbehaving mail servers as it is done with the DNS blackhole lists.

For simplicity reasons we consider a network of two mail server operators A and B that exchange e-mails of their respective clients with each other. All e-mail traffic originates from and is delivered to clients associated to one of the two operators. In our model A will play the role of sending out messages to operator B. When we talk about B, in turn, we generally mean the recipient in the communication between A and B.

In the model we consider two basic mechanisms to prevent the delivery of unsolicited mail to the clients. First, each mail server operator makes use of anti-spam techniques to prevent incoming spam from being forwarded to its clients. Second, outgoing spam filters are used to block spam originating in one of their own clients from being transmitted to the other mail server operator.

The details of the applied anti-spam techniques are abstracted away in the following considerations, as we focus on economic aspects. By the term *outgoing filter* we mean any mechanisms that the sending provider can apply in order to prevent spam from being transfered to the other provider. *Incoming filter* in addition stands for mechanisms that prevent spam from reaching the associated clients.

3.1 Notation

Both mail server operators A and B decide how much effort to invest in incoming and outgoing filters. We define x_i and x_o as the efforts of A put into incoming and

outgoing filters. The efforts of recipient B are defined as y_i and y_o, respectively. Each provider has specific costs for applying the filtering measures which are expressed in the model through the constants c_{i_A}, c_{o_A}, c_{i_B} and c_{o_B}.

A and B receive values v_A and v_B for successful operation of their mail services. These values are reduced by spam harming the e-mail communication of their clients and, hence, depend on the filtering efforts exerted.The effect of exerted efforts on the values is reflected by the probability function $P(F(x, y)$. We assume that P is differentiable, increasing in F and concave at least in the relevant part.

3.2 Simple case

To see how much sending operator A would like to invest into outgoing filters, we consider the simple case where P depends on the incoming filtering effort of the recipient and the outgoing filtering effort of the sender. We define $F_A = x_i + y_o$ and $F_B = y_i + x_o$ and get the following utility functions for A and B (both acting as sender and recipient):

$$U_A = P(x_i + y_o)v_A - x_i c_{i_A} - x_o c_{o_A}$$

$$U_B = P(y_i + x_o)v_B - y_i c_{i_B} - y_o c_{o_B}$$

It is obvious that the best choice for A is to put no effort x_o into outgoing filters at all. Because x_o only affects the spam traffic reaching mail provider B. Hence, A has no incentive to invest effort in outgoing filters as it would incur additional cost without benefits.

3.3 Including blacklisting

The simple case above does not consider an important aspect of todays fight against unsolicited e-mails: A central instrument in incoming filters is to blacklist *bad* e-mail server operators (e.g. through DNSBL) and to deny all e-mails originating from those operators (at least for a certain period of time). For our model this means that operator B can decide to block all messages from operator A if the latter is considered to send out too much unsolicited mails. This gives A an incentive to put some effort x_o into outgoing filters. The following will show how big this effort might be.

To include blacklisting into our model we introduce another probability function R to the utility function U_A of the sending e-mail provider A:

$$U_A = [P(x_i + y_o) - R(y_i - x_o)]v_A - x_i c_{i_A} - x_o c_{o_A}$$

R reflects the negative effect on A's utility through having legitimate e-mail from A's clients being blocked by B.[3] On one hand R depends on B's effort

[3] There is a similar negative effect for the receiving mail provider B when legitimate e-mails do not reach its clients. But this effect can be treated as already included

y_i in filtering incoming e-mail traffic: The more restrictive B filters incoming messages, the more likely A might be blocked if spam leaves its server. On the other hand the outcome of R depends on the effort of A in preventing spam from leaving its mail server.

We assume that $R(y_i - x_o) = 0 \forall (y_i, x_o) | y_i \leq x_o$. Thus, there is no negative effect on A's utility as long as it exerts at least as much effort on outgoing filters as B does in incoming filters. In turn, the negative effect increases with an increasing difference between B's effort y_i and A's effort x_o: For $y_i > x_o$ we assume that R is positive, differentiable and increasing, at least in the relevant parts.

To see, how much effort A and B would invest in outgoing and incoming filter mechanisms under these conditions, we have to solve the terms

$$\max_{x_o}[P(x_i + y_o) - R(y_i - x_o)]v_A - x_i c_{i_A} - x_o c_{o_A}$$

$$\max_{y_i}[P(y_i + x_o) - R(x_i - y_o)]v_B - y_i c_{i_B} - y_o c_{o_B}$$

For A we get

$$-R'(y_i - x_o)(-1)v_A - c_{o_A} = 0$$

which can be transformed into the equation

$$R'(y_i - x_o) = \frac{c_{o_A}}{v_A}$$

For B we receive

$$P'(y_i + x_o) = \frac{c_{i_B}}{v_B}$$

Let $G_{R'}$ and $G_{P'}$ be the inverse functions to R' and P'. Then we have

$$y_i - x_o = G_{R'}(\frac{c_{o_A}}{v_A})$$

$$y_i + x_o = G_{P'}(\frac{c_{i_B}}{v_B})$$

Now we define

$$\bar{x} = G_{R'}(\frac{c_{o_A}}{v_A})$$

$$\bar{y} = G_{P'}(\frac{c_{i_B}}{v_B})$$

The following two equations

$$x_o = y_i - \bar{x}$$

$$y_i = \bar{y} - x_o$$

in P, as it affects the same two input variables y_i and x_o (and x_i and y_o for A, respectively).

give us the reaction function for A to B's choice on the effort exerted in y_i

$$f_A(y_i) = y_i - \bar{x}$$

and the reaction function for B to A's choice on the effort exerted in x_o

$$f_B(x_o) = \bar{y} - x_o$$

The resulting reaction functions f_A and f_B are depicted in Figure 2.

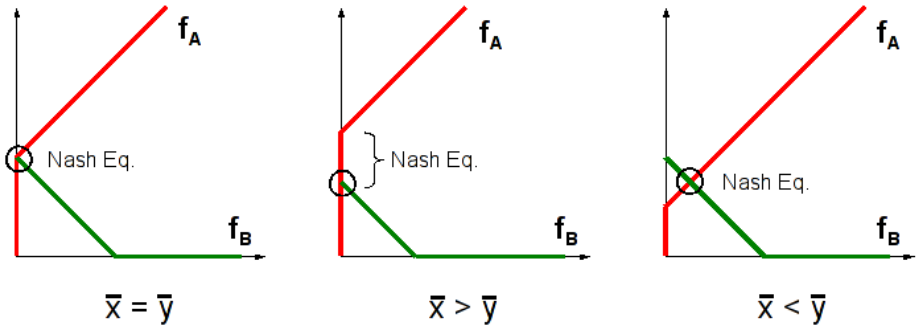

Fig. 2. Reaction functions f_A and f_B and corresponding Nash equilibria

Now that we have the reaction functions for both, sending operator A and receiving operator B we would like to find the existing Nash equilibria, i.e. the stable states where neither of both would want to change the made decision on invested efforts, as no change could increase utility for the respective party. To find the existing nash equilibria we have to solve the equation

$$x_o = y_i - \bar{x} \wedge y_i = \bar{y} - x_o$$

Doing so leads us to the following statements: Sending operator A would exert effort x_o into outgoing filters according to the equation

$$x_o = \frac{1}{2}(\bar{y} - \bar{x})$$

The corresponding equation for determining the effort y_i of receiving operator B is

$$y_i = \frac{1}{2}(\bar{y} + \bar{x})$$

.

Having the definitions of $\bar{y}$ and $\bar{x}$ in mind, we see that the decisions depend on the cost-benefit-ratio of the two operators. Due to the vague definitions of the functions P and R we cannot precisely determine the outcomes. However, we can derive some interesting results from the obtained equations.

From the definitions of P and R we can derive $\bar{y} > 0$ and $\bar{x} > 0$ and, consequently $y_i > x_o$. This means that the mail server operators will always exert more effort into incoming filters than into outgoing filters, independent from the benefit-cost-ratio of the respective operators and efforts. Thus, even if the benefit-cost-ratio would favor the use of outgoing filters more efforts would be put in incoming filters, which would lead to a socially suboptimal state.

This result is not surprising, as the sending operator has no benefit from investing more into outgoing filters than is demanded by the receiving operator in order to not being blocked. However, the result shows that the only case where the efforts are equally distributed is the trivial case where neither of them invests any effort at all. As soon as the receiving operator exerts any effort, the sending operator will exert less effort than its counterpart.

To render this result more precisely, we distinguish two different cases depending on the values of $\bar{x}$ and $\bar{y}$.

1. $\bar{x} \geq \bar{y}$:
   In this case we get $x_o = 0$ and $y_i = \bar{x} = \bar{y}$, i.e. B exerts all the effort and A no effort at all.
2. $\bar{x} < \bar{y}$:
   Here, we get $x_o > 0$ and $\bar{x} < y_i < \bar{y}$. So, A exerts some effort and B's effort is somewhat smaller than $\bar{y}$.

The second case can hardly be interpreted due to the missing details on R and P. However, the first case shows that unless the sending operator's benefit-cost-ratio regarding outgoing filters falls below a certain level with respect to the receiving operator's benefit-cost-ratio regarding incoming filters, the sending operator will exert no effort at all and free-ride at the expense of the receiving operator.

For a more detailed analysis of the underlying relations between efforts and benefit-cost-ratio of the involved parties, a more precise definition of the respective probability functions is needed.

4 Conclusion

Unsolicited messages are a major threat for the open e-mail infrastructure. Not only can they carry malicious software to the user, but they also cost the receiving end user time and money to distinguish between wanted and unwanted messages [7]. In a simplified model, e-mail communication is handled by four parties: the sender, the sender's mail server operator, the receiver's mail server operator, and the receiver (see Figure 1). While previous economic research concentrated on the relation between both end users, we modeled the relation between the mail server operators. Both mail server operators can apply filters to reduce the amount of unsolicited messages within the system. However, filtering outgoing messages already at the sender's mail server has benefits for the whole e-mail infrastructure [10].

In a first step, we developed a game-theoretical model. The model considered the costs and benefits of filtering. Our analysis showed, that the sender's mail server operator has no incentive to perform outgoing mail filtering. The whole filtering work load has to be done by the receiver's mail server.

In a next step, we extended our model by a reputation system as it is used in the DNS blackhole lists (DNSBL). Mail servers caught forwarding spam mails may be blocked by other mail servers. As all users of the blocked server are unable to communicate — not only those responsible for the spam — the server operator has an incentive to filter outgoing e-mails. Using the extended model, we identified two settings depending on the benefit-cost ratio and the lost function for being blocked. We showed for the first setting that if the risk to get caught and the lost due to unsatisfied legitimate senders are too small, the mail server operators do not invest in outgoing filter. In the second setting mail server operators have an incentive to invest in outgoing mail filtering — namely to avoid being blocked by other mail servers. However, our model showed that an operator's investments in outgoing filters are always bounded by the investment on incoming filter at the communication partner.

In future work, we are planning to analyze other means of cooperation between mail server operators. Furthermore, we want to derive more properties that can be incorporated into the probability function used in the model. A more precise and fine-grained probability function would allow to get further and better results from the model presented in this paper.

References

1. Androutsopoulos, I., Koutsias, J., Chandrinos, K.V., Spyropoulos, C.D.: An experimental comparison of naive bayesian and keyword-based anti-spam filtering with personal e-mail messages. In: SIGIR '00: Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval. pp. 160–167. ACM, New York, NY, USA (2000)
2. Görling, S.: An overview of the Sender Policy Framework(SPF) as an anti-phishing mechanism. Internet Research 17(2), 169–179 (2007)
3. Kanich, C., Kreibich, C., Levchenko, K., Enright, B., Voelker, G., Paxson, V., Savage, S.: Spamalytics: An empirical analysis of spam marketing conversion. In: Proceedings of the 15th ACM conference on Computer and communications security. pp. 3–14. ACM (2008)
4. Khong, D., Building, W.: An economic analysis of spam law. Erasmus Law & Economics Review 1, 23–45 (2004)
5. Lessig, L.: Code and other laws of cyberspace: version 2.0. Basic Books (2006)
6. Loder, T., Alstyne, M., Wash, R.: An Economic Response to Unsolicited Communication. The BE Journal of Economic Analysis & Policy (1) (2006)
7. McAfee, ICF International: The carbon footprint of email spam. http://resources.mcafee.com/content/NACarbonFootprintSpam (2009)
8. Moustakas, E., Ranganathan, C., Duquenoy, P.: Combating spam through legislation: a comparative analysis of US and European approaches. In: Proceedings of the Second Conference on Email and Anti-Spam. Citeseer (2005)
9. Varian, H.R.: System reliability and free riding. In: Economics of Information Security. pp. 1–15. Kluwer Academic Publishers (2004)

10. Yoke, H.K., Tan, L.: Curbing spam via technical measures: An overview. In: Proceedings of ITU WSIS Thematic Meeting on Countering Spam (2004)