

INFORMATION ASYMMETRY AND THWARTING SPAM

Marshall Van Alstyne, *Professor, University of Michigan School of Information*

Theodor Loder, *PhD Student, University Of Michigan Electrical Engineering and Computer Science*

Rick Wash, *PhD Student, University of Michigan Center for Information Technology Integration*

Summary

Due to its low cost, speed, and freedom from geographical constraints, email has become a ubiquitous and arguably essential means of communication. Unfortunately, the same properties that make it so useful, combined with its openness and trusting design, enable unscrupulous marketers to broadcast email to untargeted audiences. The result is unnecessary and unwarranted costs for recipients. Recent estimates indicate more than 50% of email is now spam, the volume of spam is growing rapidly, and worldwide costs exceed \$20 billion annually [4].

Legislative and technological solutions continue to be the primary means pursued to stop or limit spam. No fewer than eight bills [2] have been introduced into Congress over the last several years, and President Bush signed the CAN-SPAM Act into law in December, 2003. More than half the states [8] have enacted laws to regulate email. Concurrently, the technology industry is mobilizing to provide products and services intended to give back some measure of control to mailbox owners. In 2002, at least \$54.4 million was invested in anti-spam startups, up 65% from the previous year [5]. Using principles of information economics, we develop an alternative to the popular mechanisms for filtering and banning communications as well as to challenge-response systems which verify sender identity. Based on a very simple model, we show that use of the right mechanism - one that facilitates communication rather than blocks it or bans it - can improve social welfare. Our mechanism encourages selective targeting of messages, has dynamically adjusting prices (accounting for a recipient's value of time), depends on ex-post verification (not ex-ante classification) of content so that deceitful subject lines do not matter, transfers wealth directly to

recipients and so requires neither rebate mechanisms nor government oversight, and it is incentive compatible. Recipients have reason to adopt it, not only to manage their incoming messages, but to receive wealth from those who would ask for their attention. Last, the mechanism is bi-directional, allowing both signaling and screening.

Existing and Proposed Solutions

The popular approaches to managing spam fall roughly into two categories, technological and regulatory, with market mechanisms also explored but not yet implemented in any significant fashion. Popular *technological* approaches include challenge-response, rule-based filters, Bayesian filters, and community classification. The intent of the *legal* approaches is to regulate email communications. Various laws have been proposed to tax spam, force identification tags or labeling, create do-not-spam lists, and impose criminal charges on behavior outside prescribed guidelines. Although a detailed analysis of CAN-SPAM is outside the scope of this paper, the national law strategy may fail for several reasons: a high cost of enforcement, a lack of incentive compatibility, issues of jurisdiction (spammers are already overseas), and the nebulous definition of what constitutes spam. From another perspective, the one-size-fits-all approach has the potential to halt fruitful, mutually-desired exchange. A few articles have explored *market-based* mechanisms for allocating receiver attention [6, 3, 11]. Such mechanisms include stamps, surcharges on communication, and auctions. These might work by shifting the burden of screening from recipients to senders who know more about message content. Shortcomings of focusing on senders include voluntary participation in surcharge mechanisms and also the ability to lie about content ex ante in order to elevate interest.

An Economic Approach

We extend early proposals in several ways. First, we introduce a formal model that allows incentive analysis and welfare comparisons across proposals, including the ability to explore different recipient policies regarding interruptions. Second, we extend the mechanism to make it bidirectional, that is, we allow welfare transfers in both directions, which

further enhances the creation of markets for attention. By permitting screening and signaling, recipient choice or sender choice can help designate high value messages. Third, we compare this not just to the baseline case of no intervention, but to a 'perfect' filter, which we define as a filter that is costless to operate and makes no mistakes (no false positives or false negatives). Although no such filter exists, we use this as a proxy for any kind of filtering or banning technology. We then show that situations exist where an economic solution creates greater welfare and remains incentive compatible.

Our Key Intuition. The pure technological and regulatory approaches limit unwanted communications by blocking or banning them. This goes against a classic principle of economics: *In terms of individual and aggregate social welfare, a system that facilitates valuable exchange and side payments will generally dominate a system that grants only unilateral veto power to either party.* Our primary assumption is that the person who composes a message knows more about its content than a person who has not yet read it. This private information favors the sender, and standard mechanisms exist for screening out informed parties that would take advantage of uninformed parties. These include reputations and warranties. We therefore propose a screening mechanism that allows recipients to discriminate between classes of high and low quality senders or conversely a signaling mechanism that allows high quality senders to rise above the noise.

Attention Bond Mechanism (ABM). In the case of any sender who has a prior relationship with a recipient, reputation systems work well. Such persons can simply be "whitelisted" and their messages passed through unchallenged. These lists could also be created for recipient inboxes based on the recipients own outbox or through "letters of introduction" based on the CC: field of known contacts. In the case of strangers, the warranty mechanism is more suitable. Analogous to a standard bond mechanism, delivering email to an inbox requires an unknown sender to place a small pledge into escrow with a third party. In the case of screening, recipients determine the size of this bond, which they can dynamically adjust to their opportunity costs. The email is delivered only after the recipient receives suitable confirmation that the bond has been posted. When the recipient opens the email, she may act solely at her discretion to seize the pledge. Taking no action releases the escrow after a period of time. If a recipient expects further

communication with a particular sender and wishes to remove the bond requirement, they can add the sender to the whitelist, whereupon messages from the sender will pass through the screen unencumbered. The idea is simply to cause those who would misuse communication to signal their intention by their willingness to incur risk. Senders of valuable communications bare little exposure.

The Model. Intuition follows from a graphic representation (Figure 1) of sender and recipient gains (or losses) due to acts of communication. For simplicity, let there be arbitrary maximum and minimum values to a message, which are positive and negative respectively. These represent the range of value from welcome and unwelcome communication. Under sender choice, messages will not be sent when the sender's value is negative. This eliminates messages to the left of C_s (the marginal cost of composing and sending a message). Increasing sender costs, for example by taxing senders, provides one means of curtailing low sender value messages, and would be rejected in Figure 1 by moving C_s to the right. Importantly, however, total value of communicating is the sum of sender and receiver value. Total surplus increases in the positive direction on both the "s" and "r" axes, in the region northwest of the welfare line W . Assuming that a filter stops all messages for which the value does not justify the cost of reading, the filter would eliminate messages south of C_r . Relative to the no-intervention case, this reduces recipient losses. It also, however, eliminates a region of positive social surplus. The triangular region below C_r but above W represents possible gains from trade. Within this region, unrecognized but legitimate marketing organizations, political campaigns, charities, persons seeking interviews, and remote contacts of one's social network might offer value in return for a recipient's attention. In economic terms, recapturing and dividing this surplus represents an opportunity for both parties. In general, it will not be possible to recover this area perfectly. First, a recipient cannot know the value of a message from an unknown sender before seeing it. Second, realized value to senders and receivers can be private information, implying that the amount of surplus is unknown and subject to misrepresentation. Third, negotiating an acceptable division of surplus is complicated by the difficulty that the act of communication is itself the subject of the negotiation. A mechanism to substantially reclaim a measure of this surplus represents the bulk of the larger paper.

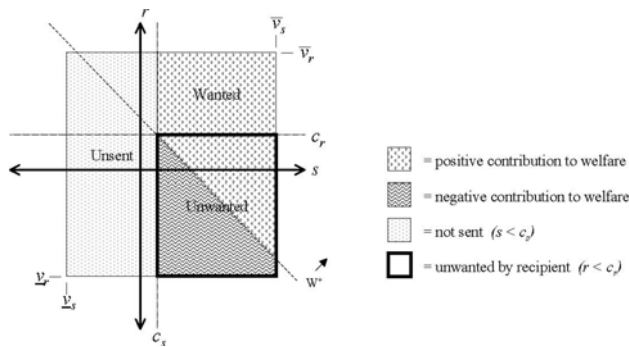


Figure 1: Distribution of Email

Screening Mechanism. Standard solutions to information asymmetry problems, where one party has information the other does not, include reputations [7], sunk costs of signaling [9], and bonding or warranties [1]. In the context of communication from unknown senders, the absence of established reputation is a primary obstacle in dealing with spam. Pseudonymous sending is itself the problem. Sunk costs, in turn, are expensive and preferably avoided as socially inefficient. Thus we focus on bonding or warranties as an effective screen. We also focus on the potential for a valuable relationship with a given sender as distinct from the potential for a specific valuable message. This avoids the considerable technical difficulty of classifying a particular message ex-ante and instead relies on ex-post verifiability of a sender's type.

For a detailed description of our model and a set of formal proofs, please refer to our working paper on the web at, <http://ssrn.com/abstract=488444>.

Caveats

While we believe the use of attention bonds represents the best approach to spam, we have several caveats. First, the mechanism does not dominate the perfect filter with all value distributions. It is particularly strong when the bulk of the distribution is of negative value and there is significant sender surplus to be transferred to the recipient, but it can do harm for primarily desirable distributions. Due to risk aversion with some senders, some email that is potentially valuable to the recipient but of little perceived value to the sender will not be sent (in favor of attention bonds, a positive correlation between values will reduce this loss). Organizations often have sales or marketing related inquiry addresses, suggestion drop boxes (anonymous tip lines), or may otherwise value inbound information to a degree where creating any restriction or additional barrier for the sender is unacceptable. The ABM may not be appropriate for

these addresses. While we do not specifically analyze transaction costs - the bond exchange process is assumed to have negligible costs - the model can easily accommodate them. Alternatively, a constant cost or a function of bond size can be split between sender and receiver in some proportion. The results are qualitatively unchanged - if transaction costs are non-zero, then additional fees reduce the size of the region where the attention bond dominates the perfect filter. Adoption, protocol, and infrastructure issues are likely to be significant, but in the interest of brevity are not discussed here.

Social Benefits

The main benefits of the Attention Bond Mechanism are the ability to cause those who would misuse communications channels to reveal their intentions and the ability to improve to social welfare. Those individuals intending to send spam are unlikely to warrant that their messages are not spam. As modeled, the screening mechanism offers a strict Pareto improvement relative to no intervention and a range of potential improvement relative to even an ideal or perfect filter. By making markets instead of foreclosing them, however, the ABM has several benefits beyond the scope of the analytic model.

Availability of Contact Information. Spiders and web crawlers mine web pages for legitimate addresses in order to send them spam [10]. If communications from strangers are mostly wasteful, then recipients may prefer to hide their contact information. In contrast to the alternatives, a successful screen raises the expected value of communication from unrecognized senders, motivating email users to publish their contact information. This reduces search costs and facilitates valuable interaction among strangers.

Generality of Mechanism. The ABM is a general economic mechanism for allocating attention and should be applicable to many forms of interrupt-capable communications media, such as email, telephone, instant messaging, and SMS (mobile phone) messaging. By brokering transactions, the telco's switches effectively become a marketplace.

Cheaper Channel Costs. The economics of the Attention Bond Mechanism compare favorably with those of competing communications channels. Direct marketing through traditional mail, for example, incurs costs of printing and postage.

Reduced Arms Race. Even as filters improve, clever misspellings, unrelated subject lines, text hidden in images, and programs designed to thwart

challenge-response systems all show the escalating resilience of spam technologies to spam filters. Rather than relying on ex ante classification of message type, the proposed incentive mechanisms relies on ex post verification of a message's value. The ABM scales with little social efficiency loss to accommodate more diverse distributions.

Political Speech. One little explored consequence of filtering technology is its effect on speech. Successful filters have the potential consequence of indiscriminately eliminating the good along with the bad and the tasteless. We grant that political speech represents one form of communication where the no intervention case may be more attractive although it may mean incessant spam. Relative to filters, however, attention bonds offer less disruption in valuable communication.

Individual Tailoring. Although not modeled explicitly, it is easy to see that the size of screen and the seize policy can be functions of other variables. This allows the mechanism to adjust dynamically to individual tolerance of interruptions, opportunity costs of time, sizes of social networks, a desire to inconvenience the fewest senders, etc. Such external factors must be learned in the case of filters and are difficult to incorporate in the case of certain taxes and other spam proposals.

Conclusions

Our principal finding is that for a wide variety of plausible conditions, signaling and screening mechanisms dominate mechanisms whose chief purpose is to block or ban email exchanges. In particular, welfare can improve both collectively and for those recipients that filters and legislation are designed to protect. The mechanism works by forcing unrecognized senders to act on their private knowledge of their own distribution, valuable or wasteful, as it applies to their intended communication. For communication of low value to recipients, the mechanism enforces a wealth transfer from senders to recipients. Communication of high value to recipients is delivered with little exposure to the sender. The net result is that well-targeted communications behave analogously to direct mail advertisements. An added benefit, however, is that resources consumed in the physical mail channel as transaction costs - marginal costs of printing and posting bulk mail - are instead captured as value to recipients. This is incentive compatible for recipients while bi-directionality allows information and wealth transfers in either direction, helping to promote valuable transactions rather than veto them.

The mechanism also depends on ex post verification not ex ante classification so it suffers less from deceitful, and costless, signals in subject lines. Finally, the mechanism allows dynamically adjusted prices, accounting for recipients' value of time.

For more information on this topic, please visit our working paper on the web at: <http://ssrn.com/abstract=488444> or the MIT Center for eBusiness website: <http://ebusiness.mit.edu/>.

References

- [1] G. A. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500, August 1970.
- [2] I. Ayres and B. Nalebuff. Want to call me? pay me! In *Wall Street Journal*, October 8 2003. 2003.
- [3] S. E. Fahlman. Selling interrupt rights: A way to control unwanted e-mail and telephone calls. *IBM Systems Journal*, 41(4):759–766, 2002.
- [4] S. Hansell. Diverging estimates of the costs of spam. In *New York Times*, August 28, 2003.
- [5] J. Hopkins. Anti-spam start-ups cash in as junk e-mail grows. In *USA Today*, July 21, 2003.
- [6] R. Kraut, S. Sunder, J. Morris, M. Cronin, and D. Filer. Markets for attention: Will postage for email help? In *ACM Conference on CSCW*, pages 206–215, 2003.
- [7] C. Shapiro. Consumer information, product quality, and seller reputation. *The Bell Journal of Economics*, 13(1):20–35, Spring 1982.
- [8] www.spamlaws.com/state/index.html .
- [9] M. Spence. Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374, August 1973.
- [10] T. Tompkins and D. Handley. Giving e-mail back to the users: Using digital signatures to solve the spam problem. *First Monday*, 8(9), September 2003.
http://firstmonday.org/issues/issue8_9/tompkins/index.html .
- [11] T. V. Zandt. Information overload in a network of targeted communication. RAND (forthcoming).

CENTER FOR EBUSINESS MISSION

Founded in 1999, the Center for eBusiness is the largest research center in the history of the MIT Sloan School. Our research is supported by the National Science Foundation and corporate sponsors. We fund more than 45 faculty and more than 60 research projects. Our mission is to be the leading academic source of innovation in management theory and practice for eBusiness.

Examples of Current Focused Research Projects:

- Theory T: Trust-Based Marketing
- Implications of e-Commerce for New Services and Structure of Logistics Systems
- How Do Intangible Assets Affect the Productivity of Computerization Efforts?
- Wireless and Mobile Commerce Opportunities for Payments Services
- Two-Tier Support Business Models
- The Impact of the Internet on the Future of the Financial Services Industry
- Pricing Products and Services in the High-Tech Industry

The Center for eBusiness has recently entered into Phase II, adjusting its agenda to focus more explicitly on business value, while at the same time including technologies beyond the Internet in its purview. The early period of exploration and experimentation is coming to an end and there is now the opportunity, and the necessity, to focus more explicitly on using digital technologies to deliver measurable business value. Amidst all this change, the business fundamentals of investment, revenues, expenses, profits, and satisfying customers have only grown more important. At the same time, a broader, inter-related set of technologies is at our disposal. While the Internet has been an important catalyst, related digital technologies are often at least as relevant.

We are co-located with MIT Sloan's Center for Information Systems Research initiative and the Center for Coordination Science to facilitate collaboration. We also collaborate with the Media Lab and the Program on Internet and Telecoms Convergence.



The Center for eBusiness gratefully acknowledges the support and contributions of its Sponsors.

CENTER FOR EBUSINESS SPONSORSFounding Sponsors

BT
 General Motors
 Hewlett-Packard
 Intel
 MasterCard International
 PricewaterhouseCoopers
 Suruga Bank
 UPS

Research Sponsors

CSK
 France Telecom
 Nortel Networks
 Qwest Communications

Member Sponsors

Amazon
 Bank of Tokyo-Mitsubishi
 Cisco
 Citigroup
 GEA
 Publicis Technology
 SAS

CONTACT INFORMATION

Center for eBusiness at MIT
 MIT Sloan School of Management
 3 Cambridge Center, NE20-336
 Cambridge, MA 02142
 Telephone: 617/253-7054
 Facsimile: 617/452-3231
<http://ebusiness.mit.edu/>

Glen L. Urban, Chairman
 Erik Brynjolfsson, Director
 David Verrill, Executive Director
 Steve Buckley, Communications & IT
 Robynne DeCaprio, Program Administration Associate