# The Spam and Attention Bond Mechanism FAQ

Thede Loder, Marshall Van Alstyne, Rick Wash, University of Michigan Mark Benerofe, Vortex Communications

**Summary:** The Attention Bond Mechanism (ABM) is a means of using sender-posted bonds to eliminate spam and facilitate mutually agreeable communication. The ABM can be applied to email and to other communications media.

**Note:** This document is one of a group of related documents that together describe the ABM.

#### The other documents are:

- A <u>one page description</u> that gives an overview of the economics and reasoning behind the design
- An Overview of the ABM Protocol which is a step-by-step walk-though of an how it works
- An academic treatment entitled <u>An Economic Answer to Unsolicited</u>
  <u>Communication</u>, which provides a rigorous formal analysis (subscription to the <u>ACM</u> may be required). An earlier version, entitled <u>Information Asymmetry and Thwarting Spam</u>, is available without subscription from <u>SSRN</u>.

This FAQ is also <u>available as a PDF</u>. The latest version can be found at <a href="http://www.eecs.umich.edu/~tloder/abm\_faq.html">http://www.eecs.umich.edu/~tloder/abm\_faq.html</a>

Version 55.0, last edited 2004-07-06

#### 1 General

- 1.1 Q: What is this FAQ about?
- 1.2 Q: What is spam?
- 1.3 Q: What is the Attention Bond Mechanism?
- 1.4 Q: What is a Bond?
- 1.5 Q: How does the Attention Bond Mechanism work?
- 1.6 Q: Does the sender have to pay for every message?
- 1.7 Q: Will the ABM be complicated to use?
- 1.8 Q: How much does the sender put at risk?
- 1.9 Q: What is a reasonable size for a bond?
- 1.10 O: What is a whitelist?
- 1.11 Q: How can I add senders to my whitelist?
- 1.12 Q: How does the ABM work (in more detail)?
- 1.13 Q: Can you show me a diagram?
- 1.14 Q: In relation to other solutions, where does the ABM fit?

- 1.15 O: How does a sender post a requested bond?
- 1.16 Q: Who gets the payment if a recipient decides the email is spam?
- 1.17 O: Why does the ABM work?
- 1.18 Q: How will use of the ABM change the value of email as a medium?
- 1.19 O: So I get back control of my mailbox?
- 1.20 Q: Paying someone you know to receive your email seems weird. What about etiquette?
- 1.21 Q: Why wouldn't you just seize the bond all the time?
- 1.22 Q: What prevents the recipient from taking the money, regardless of the message value?
- 1.23 Q: Can I keep my existing email address?
- Q: Is the ABM a universal spam cure-all? 1.24
- 1.25 Q: What if the sender does not have an escrow account when they are challenged?
- 1.26 Q: Will I have to check a second mailbox for email that was misclassified?
- 1.27 Q: Will using the ABM require a credit card or a bank account?
- 1.28 Q: Are regulations required to enable the Attention Bond Mechanism?
- 1.29 Q: What are the next steps needed to make the ABM a reality?
- Bond and Warranty
  - Q: What is a bond? What is a warranty?
  - Q: What is the purpose of a sender-posted bond for email?
  - Q: Why a bond and not a warranty?
  - O: How is an Attention Bond different from a product warranty?
  - Q: What are the advantages of allowing customized (per-person) bond sizes?
  - 2.6 Q: What is my optimal bond size? How often should I seize the bond?
- Comparison to other Approaches
  - Q: Why won't filters solve the problem (Summary)
  - O: What about other market-oriented approaches?
  - 3.2 3.3 3.4 Q: Why is the Attention Bond Mechanism a good solution (Summary)?
  - Q: What are other benefits of the ABM?
  - How does the ABM compare to other solutions?
    - <u>3.</u>5.1 How does the ABM compare to regulation or legislative solutions (such as CAN-SPAM)?
    - 3.5.2 How does the ABM compare to filters?
    - 3.5.3 How does the ABM compare to sender-pays solutions?
    - 3.5.4 How does the ABM compare to computational challenges?
    - 3.5.5 How does the ABM compare to community filters?
    - 3.5.6 How does the ABM compare to challenge-response with CAPTCHAs?
    - 3.5.7 How does the ABM compare to Whitelisting?
    - 3.5.8 How does the ABM compare to Blacklists?
    - 3.5.9 How does the ABM compare to a flat tax?
    - 3.5.10 How does the ABM compare to a flat postage fee?
    - 3.5.11 How does the ABM compare to using email stamps (e-stamps)?
    - How does the ABM compare to Brightmail? 3.5.12
    - 3.5.13 How does the ABM compare to Ironport?
    - 3.5.14 How does the ABM compare to Ironport's Bonded Sender Program?

- 3.5.15 How does the ABM compare to Goodmail?
- 3.5.16 How does the ABM compare to Microsoft's Penny-Black?
- 3.5.17 How does the ABM compare to Bankable Postage?
- 3.5.18 How does the ABM compare to Microsoft's "Caller-id"?
- 3.5.19 How does the ABM compare to Domain Keys?
- 3.5.20 How does the ABM compare to SPF?
- 3.5.21 How does the ABM compare to Yahoo's Filtering?
- 3.5.22 How does the ABM compare to Cloudmark?
- 3.5.23 How does the ABM compare to Vipul's Razor?
- 4 Security, Viruses, Honeypots, and Authentication
  - 4.1 Q: What about possibility of fraud or a virus triggering bond payments?
  - 4.2 Q: How does the ABM handle honeypots?
  - 4.3 Q: What is the role of stronger authentication?
  - 4.4 Q: What about privacy?
  - 4.5 Q: Will stronger authentication make email more difficult?
  - 4.6 Q: Why are Certifying Authorities not needed for individual identities?
- 5 Infrastructure and Costs
  - 5.1 Q: Doesn't this require infrastructure?
  - 5.2 Q: Who runs the payment system?
  - 5.3 Q: How will the costs of tracking the warranties be covered? Can it be done cost effectively and or will such costs prohibit the adoption of the system?
  - 5.4 Q: Will existing mail clients need to be changed?
  - 5.5 Q: Does this design require changes to SMTP?
  - 5.6 Q: What happens if warranties are demanded in foreign currencies?
- 6 Compatibility with Existing Uses
  - 6.1 Q: How does this system work with mailing lists that I subscribe to?
  - 6.2 Q: How is this compatible with e-commerce and customer relationships?
  - 6.3 Q: What if I have forgotten my password to a website and need it to be emailed to me?
  - 6.4 Q: What if I meet someone at a conference and they give me their business card with their email address and I write them?
  - 6.5 Q: How will the ABM work for large institutions?
  - 6.6 Q: What does my IT department need to do once we begin using the ABM?
  - 6.7 Q: What about friends and family?
  - 6.8 Q: How does the ABM work with children under 18?
  - 6.9 Q: Many people access their email from Internet cafes and do not have accounts with ISPs. How those without bank accounts, ISP accounts, or credit cards use an escrow account?
- 7 Effects on Participants
  - 7.1 Q: If I am a marketer, will this system stop me from reaching my customers?
  - 7.2 Q: What are the effects on Marketers?
  - 7.3 Q: Will I have to pay taxes on any warranties I claim?
- 8 Adoption
  - 8.1 Q: Why would anyone participate?
  - 8.2 Q: How can user adoption be encouraged?
  - 8.3 Q: Why should marketers participate?

- 8.4 Q: What are the benefits to ISPs?
- 9 Social Implications
  - 9.1 Q: Will this cause a loss of freedom of speech?
  - 9.2 O: If everyone starts using the ABM, what happens to anonymous email?
- 10 Other
  - 10.1 Q: Is this really a new idea?
  - 10.2 Q: So what's your contribution?

#### 1 General

#### 1.1 Q: What is this FAQ about?

This FAQ describes the Attention Bond Mechanism (ABM), a solution to email spam. It contains a vision for a competitive industry structure, a description of the operation and protocol of the ABM, comparisons to existing solutions and its relationship to them, and next steps. Please address any comments, or suggested edits to abm\_project –atumich.edu.

#### 1.2 Q: What is spam?

A: We define spam as *any email that you would rather have not received*. It is important to note that with this definition, we assume you know whether or not something is spam only *after* you have received it.

An alternative definition is "any email for which after reading the email, you feel that if you could charge the person sending it to you for wasting your time, you'd do it."

The above definition excludes things like email that represent a bill or a death in your family. In the case of a bill, despite the fact that you would rather not have to pay it and hate being reminded of it, you typically still want to pay since the alternatives to not paying are usually worse.

#### 1.3 Q: What is the Attention Bond Mechanism?

A: The Attention Bond Mechanism is a means to improve the value of communication between two parties and increase the likelihood that such communication is mutually desired. In other words, it is a means of eliminating spam.

The basic protocol of the ABM can be used for many different communications media, including email, phone, SMS, and instant messaging. This FAQ focuses primarily on its application to email and spam.

#### 1.4 Q: What is a Bond?

A: A bond is a sum of money (or other exchangeable good) which one party in a transaction sets aside with a third party *before* the primary transaction occurs, as a show of good faith.

A bond is similar to a warranty, but not the same. The difference between a bond and a warranty is that a warranty is a *promise* to pay, but is not paid or set aside in advance. If the second party to a transaction (the recipient, in the case of spam) is dissatisfied, with a warranty she must request payment from the first party. With a bond, she instead requests payment from the third party, which holds the bond.

The third party plays the role of an escrow. The bond is placed with the third party, who provides the service to both primary parties of holding the bond and releasing it when certain events occur.

#### **Example Bonds:**

For better or for worse, perhaps the most widely known example of a bond is a "bail bond". In the situation where a suspected criminal offender captured, they are sometimes required to post a bond in order to leave custody of authorities. The money set aside as the bond constitutes a promise to return to a trial for formal interaction in person. Should they not return on the promised date, the bond money is forfeit.

It is important that the bond is large enough such that showing up in court is actually in the best interest of the suspect offender. Bonds are also used as investment instruments, in housing purchases, and for providers of contracting services.

#### 1.5 Q: How does the Attention Bond Mechanism work?

A: Email from a sender who is pre-approved or whitelisted (See Q: What is a whitelist?) goes directly to the recipient. A sender who is not pre-approved by the intended recipient is required to post a small sum of money (a bond) in order for his email to be delivered. In effect, the sender warranties the content of his email and pre-pays the warranty in the form of the posted bond.

Assume, for example, that a sender sends an email and receives notification that he is required to post a bond to ensure delivery. Assume also that he then posts the bond (see "Q: How does the sender post a requested bond?"). When the bond is posted, the email is delivered to the recipient.

Once the email is in possession of the recipient, the recipient can choose whether or not to claim the bond. If the recipient claims it, bond money is transferred to her and the sender loses it. If instead the recipient is satisfied with the email, she allows the bond to be returned to the sender unclaimed and no money changes hands.

#### 1.6 Q: Does the sender have to pay for every message?

A: The sender does not have to pay anything if they are sending email to people with which they correspond regularly, for example to friends, family, co-workers, other professionals, or mailing list subscribers. These senders will typically be whitelisted by recipients (see "What is a Whitelist?"). The sender may have to pay if he sends email to someone with whom he has never had prior communication, or to someone who does not want the communication.

#### 1.7 Q: Will the ABM be complicated to use?

A: From the standpoint of the average user, the ABM will be simple to use. The complexity of the underlying system, including the details of its operation, various parties involved, and protocols used can be hidden from the end user, just as the complexity of the underlying Internet or the stock market is hidden from users today. Existing mail clients can be adapted to hide the complexity and provide a very simple set of options. ISPs and corporations will be able to automate the setup of the accounts required by the ABM for their users. Users of email who want greater control of their accounts (or greater visibility) or want customized options will be able to set up their own accounts or have access to services that provide greater levels of control.

#### 1.8 Q: How much does the sender put at risk?

A: The sender must post a bond at least as large as the size required by the recipient. It could be pennies or more than \$5.00. See What is a reasonable size for a bond?

#### 1.9 Q: What is a reasonable size for a bond?

A: For most people, a good size bond is likely to be somewhere between a few cents and less than \$1.00.

If a recipient demands too large a bond, marketers will avoid them altogether or only rarely take the risk of sending an email. This might be fine for some users of email, but a recipient's friends might choose to avoid sending to them as well. With a very high bond, the recipient risks alienating young or less developed relationships, particularly if the senders are risk-averse. A sender may not want to risk losing a large sum of money when the relationship with the recipient is tenuous or where an expectation of future interaction is limited.

On the other hand, if a recipient demands too small a bond, she will end up receiving spam. It is likely that for each person, there is a unique optimal size. Over time, more sophisticated recipients will learn what is most appropriate for their own social and business relationships, and set the bond accordingly. To make it easy initially, recipients can simply take a default size provided by their ISP or escrow agency, or may use the default in whatever software tool they use for mail. Most users may never need to change the default.

#### 1.10Q: What is a whitelist?

A: A whitelist is a list of identities of approved senders. Whitelists are used to route email from selected senders around filtering, screening, or additional challenges and allow the email to be placed directly into the recipient's mailbox. A requirement for an efficient whitelist is that the identity of the sender of the message is cheaply and easily recognizable, yet hard (or computationally expensive) to fake.

See also "How does the ABM compare to Whitelists?" and "How does the ABM compare to Blacklists"

#### 1.11Q: How can I add senders to my whitelist?

A: There are several ways to add senders to your whitelist. One way is to use your mail tool when a message from a new sender arrives. You could do this simply by clicking on a button. Another way is to use your email software to manually add a sender's identity by typing it in. Since this could be a nuisance if you are adding many senders, you could make use of a software feature that automatically adds your list of contacts or scans your archived mail folders and your outbox for identities of those who are already acceptable to you. This would need to be done only once during setup, though there may be reasons to manage the whitelist later.

When sending email, your mail system can automatically add the identity of the recipients you have specified if they are not already whitelisted. This way, when they reply to your email they will already be on your whitelist and the reply will not be blocked.

A third automatic method is analogous to a "letter of introduction" commonly used in the last century. If someone you know "cc's" a third party they want you to know, then the system can automatically grant that person a temporary waiver until you can evaluate whatever they send.

#### 1.12Q: How does the ABM work (in more detail)?

A: The inbox of the recipient is protected by a whitelist. The recipient has control of which identities are on the whitelist. Any email sent from a sender listed on the whitelist is immediately delivered to the recipient's inbox. Email from a sender not on the whitelist is blocked and the sender is sent a challenge message.

Up to this point, the design of the ABM is quite similar to existing challenge-response systems, but hereafter the approach differs.

With the ABM, the challenge message contains a request for the sender to post a bond to an escrow account controlled by the recipient (see "Q: What is a reasonable size for the bond?" for a discussion of bond size). The recipient's escrow account can be set up in advance or when first required. The recipient can set it up with a third party escrow

service, or it could be set up on behalf of the recipient by their ISP, a bank, or by their employer.

The challenge reaches the sender in both human readable and machine-readable formats, which a sender to manually respond to the challenge, but also lets software operated by the sender automatically respond. If the sender chooses to accept the challenge, they authorize posting of the requested bond. Once the bond is posted with the recipient's escrow service, the recipient's mail system receives a notification. This notification triggers delivery of the original email into the recipient's inbox.

When the recipient opens the email, she can see the associated bond and may decide whether or not to claim it. She might think, for example, "no, this is spam, I'm going to collect the bond", or think, "This is a direct marketing offer for a product I am not interested in. I will collect the bond as compensation for having to review it." Or they may say to themselves "this is interesting," then review, keep and/or respond to the mail and not collect the bond, as they would if the email is from a friend, family member, or business associate.

By default, if the recipient does not explicitly choose to keep the bond, the bond is automatically returned to the sender after a few days. If the recipient chooses to keep the bond, she notifies her escrow service by clicking on a button. The value of the bond is then deposited into her own general-use account.

To avoid the extra work of having to respond to challenges by hand, a sender may choose to activate a policy, similar to the following: "If I ever send someone an email and they respond with a challenge, as long as the amount of the challenge asks for less than \$0.50, it's ok to automatically post the bond. For any requested bond larger than \$0.50, return the challenge to my inbox so that I may review it and decide." Such a policy could be implemented in software on the sender's mail servers, and would reduce the number of times direct intervention is required. With such a setup, most of the time the posting of bonds would be transparent and automatic.

#### 1.13Q: Can you show me a diagram?

A: For a set of simple diagrams that describes the various parties, their interactions, and the steps involved see the How it Works document.

#### 1.14Q: In relation to other solutions, where does the ABM fit?

A: The ABM is essentially a combination of several simple and well-known technologies:

- 1) stronger identities (authentication Sender ID, Domain Keys, S/MIME, GPG)
- 2) Whitelists (authorization Access Control Lists)
- 3) A means of posting, claiming, and transferring payment (electronic payments)

4) Simple Agent-based negotiation (Challenge Response and execution of policies)

One of the key pieces is email authentication. The big ISPs are already collaborating on this. Authentication makes effective whitelists possible. Once you have an effective whitelists, you can provide a means for people to get around the whitelist (or "through" it) by taking a risk.. If you make the risk refundable, based on the recipient's decision, you have the ABM. The agent-based negotiation allows the cost of both the challenge and response to be handled cheaply by computers.

See the section on Comparison to other Approaches.

#### 1.15Q: How does a sender post a requested bond?

A: The sender posts a bond by authorizing their escrow agency to move the amount of money requested as a bond into an account controlled by the intended recipient. This can be done with a mouse click when the sender receives the challenge, or they may set things up so that it can be done automatically by establishing a policy with their own mail servers. See "How does it work (in greater detail)"

## 1.16Q: Who gets the payment if a recipient decides the email is spam?

A: The recipient is the main benefactor. However, a small portion will be necessary to cover the costs of the escrow agencies and to provide the right incentives for the ISPs or enterprises that operate the recipient's and sender's email accounts. This size of this portion should be low (perhaps a few percent) due to competition between service providers.

#### 1.17Q: Why does the ABM work?

A: It works because it lets both the recipient and the sender cheaply negotiate the terms under which they both want communication. The low costs are possible because software on both sides can handle the first few rounds of communication – no human involvement unless they want to be involved.

It also works because it lets mailbox owners levy a penalty on spammers without penalizing everyone else. The bulk of the penalty paid by the spammers goes to the mailbox owner.

## 1.18Q: How will use of the ABM change the value of email as a medium?

A: It will help restore the value of email by making it reliable again. But the improvement goes beyond restoration of the pre-spam value of email. The ABM expands

the use of email by allowing payments, which makes it easier for mailbox owners and companies to maintain relationships, and gives mailbox owners a way to effectively bill for their time. It increases the value of the information marketplace by allowing marketers, first-time senders and recipients to fine-tune their communications, letting each side tell the other what they want and how much they are willing to pay for it.

#### 1.19Q: So I get back control of my mailbox?

A: The ABM puts the mailbox owner in control. No more spam, only email you want. Email you do not want will have enough money attached so that you're still happy you received it. And if you never want to receive email from someone unknown to you, you can simply use the whitelist portion of the ABM by itself, and disable payments.

## 1.20Q: Paying someone you know to receive your email seems weird. What about etiquette?

A: Etiquette will evolve around the ABM, just as etiquette has evolved for the telephone, voicemail, email, and instant messaging. The design of the ABM assumes honesty, but will ultimately punish those who abuse it. If you improperly collect the bond or require a bond that is too big, senders will take you off their list. See <a href="What is a reasonable size for a bond?">What is a reasonable size for a bond?</a>

#### 1.21 Q: Why wouldn't you just seize the bond all the time?

A: Nothing in the ABM directly prevents you from always seizing the bonds posted by senders. However, by always seizing bonds you could hurt yourself more than you hurt others. If you just try to make money off other people, then they will stop communicating with you; marketers will quickly blacklist recipient identities that always seize the bond, or achieve the same effect by tracking and managing lists of those senders that regularly release it. See the "honeypot" question.

When you collect bonds more often, you increase the costs to the senders. Consequently, less email will be sent to you, and you probably do not want to lose email from you friends or work associates. But when spammers are the senders, collecting from them is ok. So, rather than keeping the bonds all the time, you will want to keep bonds posted by spammers but not by your friends.

## 1.22Q: What prevents the recipient from taking the money, regardless of the message value?

A: The ABM enables the recipient to seize the money solely at her discretion. Nothing, other than perhaps etiquette and good judgment, prevents claiming a bond.

Why is this ok? The ABM puts the burden on the sender to think carefully about to whom they are sending email. If it really is to a friend or acquaintance, the sender should

have no fear of sending an email and, if necessary, posting the bond. See <u>Q: Why</u> wouldn't you just seize the bond all the time?

#### 1.23Q: Can I keep my existing email address?

A: Yes you can. There is no reason to lose your present email address if your ISP or web-mail provider supports a system that implements the ABM. Even if you regularly access your mail with a traditional client such as Eudora or Outlook (rather than through the Web, as with Hotmail and Yahoo), you should be able to keep your address. However, if you switch providers, you might not be able to take your address with you. However, this is no different than the way it works today.

#### 1.24Q: Is the ABM a universal spam cure-all?

A: No. There are certain types of mailbox uses for which the ABM would be inappropriate. For example, suggestion drop boxes, sales and info email addresses, and any situation where the recipient wants no barriers, what so ever, to be placed in front of the sender.

For example, the New York Times, CNN, and many local newspapers and broadcasters have an email address that people can use to send anonymous tips of breaking stories. Any additional barrier on such addresses address, aside from the unavoidable costs to the sender of taking the initiative, is too much. In cases such as these, recipients can set the threshold to zero cents, and may want to use an appropriate filtering solution. Fortunately, due to the small number of people who would get submitted emails, spammers may not even be justified in sending to these addresses.

## 1.25Q: What if the sender does not have an escrow account when they are challenged?

A: If the sender does not already have an escrow account setup, they can set one up before they respond to the challenge. The challenge email can contain instructions as to where to go to get started. Alternatively, they can contact the recipient through other means, and ask to be added to their whitelist.

To aid the adoption, if the recipient chooses they can optionally include in a challenge message references alternatives other than bond payment that do not require account creation with an escrow agency. During early days of adoption, few people will have escrow accounts, so this alternative can make things a little easier for the typical sender until the infrastructure becomes widespread.

For example, the alternative challenge could be a CAPTCHA (like a Turing Test), where the sender is required to take a simple test that is designed to prove they are human, similar to what is used today by challenge-response systems. See "Q: How does the ABM compare to Challenge-Response?"

## 1.26Q: Will I have to check a second mailbox for email that was misclassified?

This is an option, but not a requirement. The recipient's mail server can be instructed to deliver email from a non-whitelisted sender (a suspect sender) into a separate mailbox (commonly called a 'grey mail' box). The recipient can then view the contents of this mailbox at any time. When notification is received from the recipient's escrow agency that a sender has posted a bond for a particular message, that message is 'promoted' out of the 'grey mail' box and placed the recipient's primary mailbox.

## 1.27Q: Will using the ABM require a credit card or a bank account?

A: In order to post bonds or collect them, users of ABM will need an escrow account. Escrow accounts can be set up in advance for users by their ISPs or by their employers. Alternatively, users can set up an escrow account directly through an escrow agency on their own. Users should be able to support multiple email addresses per escrow account if this is desired

To post a requested bond will require the sender to have an account that has a high enough balance. The funding of this account could be accomplished through a credit card payment, or through other payment means. In the case where the account is setup by an ISP, the funding for this escrow account balance could come from an additional payment on top of that for the first month of service.

If a recipient claims bonds, over time the sum of claimed bonds will accumulate unless they are spent paying bonds to others. At some point, if the balance gets high enough, the recipient may want to transfer money out of the account to a general-purpose bank account. The ACH checking network would be suitable for these transfers, if the size of the transfer was set large enough to offset the transaction cost.

## 1.28Q: Are regulations required to enable the Attention Bond Mechanism?

A: The only regulations required may already be adequately addressed by the existing regulations covering consumer banking and inter-banking networks (further investigation necessary). It is possible that the escrow agencies and underwriters will need appropriate regulation.

## 1.29Q: What are the next steps needed to make the ABM a reality?

A: We believe the next steps are as follows:

- 1) Ensure that high level decision makers are informed that a welfare-favorable solution exists that has the right incentives for all parties (except the spammers). This is to help avert any laws that make things worse and have to be undone. We've started this with the FTC, and hope to do more shortly with the ITU and future FTC related meetings, like the upcoming authentication meeting.
- 2) Explain the value chain to the business and technology communities. Marketers put the funds in at the top, most flows to the recipients, but the escrow agencies and the ISPs/enterprises that provide the connectivity take a cut. We'd like to open a dialog with folks at the DMA, for example.
- 3) Get the ABM in front of the major spam solution providers that exist now. Many are already in good position to roll out the enterprise/ISP/organization-focused server software. If they see the open standards on the immediate horizon, that will light a fire under them to get the standards implemented first in their own products.
- 4) Involve the development/infrastructure communities, design and publish the standards. Get the word out so that the developers of existing email software (clients and servers) can upgrade their products.
- 5) Set up the first escrow agencies. Aside from open source credit union-like non-profits, go to big transaction players (maybe citibank?) who have the trusted consumer brands and customer relationships. Similarly, go to the major ISPs who might also want to run their own escrow agency for their member bases.

We'd like to get the idea out there and heard, and a public standard (for challenge messages, inter-escrow agency payments, and authentication formats ala x509 or openPGP) created. Adoption will not be a problem as long as the standard is open and straightforward, and the appropriate community is convinced that this approach will get rid of spam and still enable email use to have great value. Escrow agencies can range from marketer-facing to consumer facing (integrated with ISP, stand alone, or like a credit union for people who like particular privacy policies).

### 2 Bond and Warranty

#### 2.1 Q: What is a bond? What is a warranty?

A: In lay terms, a bond is a risk that a person on one side of a transaction takes to prove to the other party that he will find the completed transaction valuable. Technically, this is described as a "contingent liability with an expiration date". A warranty is a promise to pay compensation for dissatisfaction after the transaction completes as opposed to setting aside compensation in advance as with a bond. See the previous section, Q: What is a Bond?

#### 2.2 Q: What is the purpose of a sender-posted bond for email?

A: The purpose is to make spam too expensive for spammers to send. The bond is required only for first-time communication from a stranger, or for subsequent communication if a mailbox owner has removed the sender from her whitelist.

The sender who believes his message is not spam is willing to "put up" that money - to risk it - to signal his belief that if the recipient reads the email, she will agree that it is not spam. Spammers, in general, cannot afford to take this risk.

Of course, the bond can also represent a pledged amount of money a legitimate marketer is willing to spend to reach a person in exchange for his or her time. Again, the recipient can decide to accept it or not.

#### 2.3 Q: Why a bond and not a warranty?

A: For tangible products and services, a higher quality warranty (one that is likely to be more costly to provide if requested) implies that the goods offered will satisfy. Someone offering a poor quality product or service will find warranties more expensive than someone offering higher quality because they will need to spend money fixing their broken promises. The good guys incur much lower warranty costs. In the digital world, such a warranty can be used to imply "even though you don't know me, my message is worth your time."

The trouble is, in the online world, it is possible for the party offering a warranty to "split the scene" after the primary transaction completes, and just disappear. Since it costs so little to change identities, there is no way to track them down, and no way to enforce the claim on the unpaid warranty. Warranties cannot be used. The Internet is like the Wild West, so it's wise to collect cash up front, before providing spending additional effort.

It is for this reason we propose a bond, paid up front, rather than a warranty. The bond goes to a third party before the transaction completes. The third party, who requires both a good reputation and significant sunk costs to retain and attract customers, will not disappear after the transaction, even if the sender does. Therefore a claim on the bond can be effective, and the recipient gets paid the amount of the bond.

## 2.4 Q: How is an Attention Bond different from a product warranty?

A: There are two primary differences. The first difference is that with the Attention Bond, the money for the bond is collected up front, yet with a warranty, it is only promised.

Attention Bonds also differ from standard product warranties in that recipients of email get to choose the size of the bond, allowing customization. Usually, in the non-electronic world, the cost of customizing a warranty to each individual buyer of a product is too great and so it is rarely done.

## 2.5 Q: What are the advantages of allowing customized (perperson) bond sizes?

A: The advantage is primarily economic. Different people have different value for their time, have different tastes, and may want to receive more or less communication from outsiders. The best bond size will be somewhat different for each person. A fixed bond has little flexibility and cannot accommodate these differences.

## 2.6 Q: What is my optimal bond size? How often should I seize the bond?

A: You should be guided by three factors. First, if your time is very valuable, you want to set a higher warranty level. If high enough, any strangers are unlikely to bother you. Second, if new information is valuable to you, you should set a lower warranty level. This encourages more people to contact you with new ideas, observations, and information. Finally, you need to decide when to force senders to honor their warranty. If they're bothering you with useless sales pitches, then go ahead and claim it -- they've wasted your time. If someone is your long lost high school buddy, then you're glad they reached you and you don't need to bother claiming it. Pick the factors that are most important to you. See Q: What is a reasonable size for a bond? and Q: Why wouldn't you just seize the bond all the time?

### 3 Comparison to other Approaches

#### 3.1 Q: Why won't filters solve the problem (Summary)

A: Technology-based filters are required in many situations and are useful, but they don't solve the problem and they have two terrible side effects. First, they create a technical "arms race" - with spammers on one side and filter makers and Internet service providers on the other. You, the recipient, are stuck in the middle. This kind of battle only increases the volume and variety of junk the spammers send to try to get their message through. Second, while some filters claim to get closer to 99.8% effectiveness - false positives are still a problem. The incidence of false positives (a false positive is good email that is incorrectly classified as spam) increases with filters as they are modified to include more screening criteria and rules.

For many, the occurrence of these false positives is intolerable. For example, AOL is constricting its filters so tightly that some moms coordinating after-school activities can no longer rely on email and are returning to the phone. Admissions departments are having email acceptances simply "dropped" as bulk mail, which means students don't receive them. Emails that include website addresses are often determined by the filters to be "spam" even between people with a long history of communications between their two addresses

Even spam filter experts at the January 2004 MIT conference on spam (www.spamconference.org) report losing "good mail" from friends and business associates - including valuable contracts - in their bulk mail and spam folders because they were misclassified.

One commonly implemented means to address the problem of false positives it to provide a 'grey-mail' box, in which suspect messages are placed for later review. Unfortunately, although use of a grey-mail box reduces the likelihood of missing an important email, there are two problems. Unless you check your 'grey-mail' box as often as you do your regular inbox, delays are introduced. And, if you have to periodically go through the grey-mail box to check to see if something important was misclassified, you still end up having to look at all the spam.

#### 3.2 Q: What about other market-oriented approaches?

A: As compared to email-stamps, a flat tax on email, or per-email postage fees, the Attention Bond Mechanism has the advantage that charges are levied against the sender only when the recipient claims the bond, as opposed to every time an email is sent. Between parties who have whitelisted each other, the cost sending email via the ABM is the same as it is today, nearly indistinguishable from zero. Solutions that advocate a per-email charge raise the costs for everyone - spammers and non-spammers alike. The result of these approaches would be the loss of the cost advantage of the medium.

With computational challenges, where the sender's computer is forced to compute the answer to a suitably difficult problem, the cost incurred by the sender is non-recoverable. It is a pure loss. Literally, the sender has been forced to pay with time and energy and to heat the environment with their CPU. (While a sophisticated scheme of sharing or tracking compute cycles and distributing problems of general social value would be possible to set up, the overhead is unlikely to be worthwhile, and computation is not nearly as flexible or tangible as money from the perspective of the recipient.)

Worse, in order for a computational solution to really work, the total cost of providing the CPU time would need to be comparable to the cost of a bond – perhaps several cents. At around \$0.00001 per CPU-second, this could mean several thousand seconds, introducing an unacceptable delay in delivery. If computed in advance, the work effort must be tracked, similar to the accounting required for cash.

Another proposed bond approach is to require companies to post a single large bond with a trusted entity, usually the maker or service provider of anti-spam filtering systems (see <a href="How does the ABM compare to Ironport's Bonded Sender Program?">How does the ABM compare to Ironport's Bonded Sender Program?</a>) Similar to community filtering, if enough recipients indicate that an email is spam, the sender is required to forfeit the bond. Unfortunately, this solution only works if the filters work, and filters have problems with false positives. As it is a form of community filtering, this approach has the 'tyranny of the minority' problem in that others end up deciding what you get to read. Since the value of the forfeit bond is given to the filtering company itself or to some otherwise uninvolved third party (perhaps a non-profit of the providers

choice), the single bond approach lacks the incentive compatibility of the ABM. The recipient does not get compensated for their lost time or annoyance.

## 3.3 Q: Why is the Attention Bond Mechanism a good solution (Summary)?

A: Unlike pure technological approaches, such as all types of filtering, the ABM is less subject to an arms race. It will not result in an increase in spam email traffic. It short-circuits the expensive cycle of spammers and filter makers endlessly trying to out smart each other. Specifically, it punts the arms race into the realm of cryptography, where carefully controlled communication is well studied and has known outcomes.

False positives are no longer a problem; important email can be sent reliably without concern for loss. If email does not get through to the recipient, there is no way for them to collect the bond, and legit mailbox owners will not want to automatically claim bonds without reading the email first. Without false positives, there is no longer a requirement for recipients to scan a list of suspect email to prevent accidental deletion, saving time and money.

The ABM gives each individual the ability to have their own definition of spam, accounting for their unique tastes. In contrast, community based filtering designs, such as Cloudmark or Yahoo's filters, may delete email that you want because it was thought to be spam by other community members. One could consider this problem to be another form of false positive, or a tyranny of the vocal minority.

The ABM facilitates exchange instead of shutting it down. It allows sender and recipient to cheaply negotiate the terms and conditions under which both can be satisfied and gain. Filters, which cut off communication, stop this process from happening and therefore stop potentially gainful exchange and relationships.

The ABM allows email from automated systems to go through as long as they are from whitelisted addresses or the sender has posts the bond. Challenge-response systems, with reverse-Turing tests as the only means of validating the sender, stop the flow of email from automated systems, which are very cost efficient and their use is often desired by both sender and recipient.

Unlike banning and labeling laws, the ABM is not subject to problems with multiple jurisdictions, lack of incentive compatibility, and does not have the costs of enforcement and adjudication.

#### 3.4 Q: What are other benefits of the ABM?

A: There are many other benefits, starting with support for quality direct marketing. Many Fortune 1000 companies, legitimate small businesses, and others have shied away from email for fear of being viewed as a "spammer" – something that could compromise

the integrity of their brands and their hard-earned reputations. Attention Bonds let these legitimate marketers back into the medium at a lower cost than the alternatives.

Next, use of sender bonds will result in a reduction in search costs (the overhead of finding someone's address). With sender bonds, recipients have the incentive to publish their email address rather than obscure it. There is no more fear of having your email address "spidered" by search software, viruses or a screen scraper. If you are sent an email, it is a chance to earn a payment. This ultimately adds value to the information marketplace and communities created by it, because legitimate people and businesses can find you easily and put up the bond to reach you.

The use of sender bonds will also allow mid-sized ISPs to continue to compete with major ISPs. The biggest players (AOL, Microsoft-MSN/Hotmail, Earthlink, etc.) are very focused on spam because of the added infrastructure costs to carry the traffic of billions of messages each day. Spam now constitutes more than 60% of email. These trafficking costs are huge. But, as hard as it is for the large scale ISPs, it is even harder for the mid-size ISPs, which have fewer resources. Unless the spam problem is addressed in a cost-effective way, small and mid-sized ISPs will keep dropping out, creating less competition and ultimately less choice and value for the consumer, since only the larger ISPs can support and suffer the added costs over time.

#### Other benefits include:

- The Attention Bond Mechanism is a general mechanism; it can also be applied to SMS messaging, telecom, and instant messaging. The same account used for posting the bond in one medium can be used for the others and have the same effect.
- A system that associates payment with communications traffic creates the opportunity for funds transfers on a very large scale. Such a system could facilitate interactions between legitimate buyers and sellers for other forms of exchange; the financial infrastructure will already be in place.
- Individual Tailoring. The ABM is better than "one-size" fits all; information exchanged with the ABM is ruled neither by majority or minority vote. Individuals can express their preferences through their bond level and their actions. This allows the marketplace to evolve, so senders and recipients can provide each other better and more customized information.
- Signaling of interests. When recipients don't claim the bonds, it is immediate feedback to the marketers (senders) about a recipient's interests. The information marketplace can evolve to a higher level, and better accommodate its members' preferences and needs.
- Adjustable screening. Businesses can set their minimum bond sizes for different addresses at different levels, letting them adjust to the needs of different customer groups, suppliers, buyers, and associates. For example, mailboxes that are for customer service or sales leads might have a bond of zero no bond required, employee mailboxes set at another level, and senior executives still another level.

• Political Speech. Relative to a system that is completely costless, any system that introduces some friction results in less communication overall. Yet frictionless communication creates spam and with it all the noise that makes many messages meaningless. The ABM increases the ability of any party who really cares about an issue to rise above the noise. You can win attention by simply and literally insuring that your message is worth reading. Does this mean you might pay for getting news out? Yes, it does. But even with this risk, the ABM is likely to be preferable to alternatives. First, it will likely be substantially cheaper than other forms of political speech including print, TV, and radio media. For non-profits and issues of general concern, people are unlikely to collect bond money. Second, sending costless communication isn't really costless. It merely transfers processing costs from senders to receivers, as spam clearly shows.

#### 3.5 How does the ABM compare to other solutions?

There are many of other solutions, implemented and proposed. Some specific solutions and solution classes discussed below.

## 3.5.1 How does the ABM compare to regulation or legislative solutions (such as CAN-SPAM)?

A legal definition of spam is one-size fits all, rather than individualized; essentially a blunt instrument. In order to make the threat of fines credible, law enforcement agencies would need bigger budgets for enforcement. Finding and bringing spammers to trial is an expensive process, and the ease of offshore Internet access creates difficulties with jurisdiction. The spammer's evasion costs, like the enforcement costs, are an unnecessary deadweight loss.

A flat tax would include mandated technologies, enforcement costs, collection costs, international coordination, and has the unfortunate property of increasing the cost of the medium for everyone, rather than for just the spammers.

#### 3.5.2 How does the ABM compare to filters?

The ABM will not suffer from false positives and false negatives. No periodic visits to a grey-mail box are needed to catch misidentified email.

Filters use ex ante classification (i.e. they look at the content before the email is delivered). A filter can be considered an externalization of the recipient's preferences. The filter judges message quality, based upon its contents, in behalf of the recipient. However, the preferences are not exact, but are simply an approximation, subject to the form and limitations of how they are represented in the filter and the knowledgebase available to it. The ABM uses ex-post verification. After the email is delivered, it is the recipient themselves that decides if it is spam. The recipient's assessment of the email is certain to be better than that of a filter (or at least for the foreseeable future).

With most filters, when an email is sent to the grey-mail box or erased, the sender is not given any notification. This can trigger time-consuming losses to both parties. With the ABM, the sender of a blocked email is sent a challenge. The challenge serves as notification that the original delivery did not go through, allowing them to take action.

Filtering is subject to an expensive arms race. Filter makers and spammers continue to spend time and effort to outwit each other. Since the ABM makes use of stronger identities and can ultimately use cryptography for message classification, it avoids the arms race

With the ABM, the sender runs the risk of paying the recipient the amount of the posted bond. With filtering, the recipient has no chance of getting any compensation.

The popularity of filters leads to an increase in the overall amount of spam sent, since senders may try multiple permutations of the same basic message to get a message through. This leads to an increase in traffic, most of it junk. While the ABM does require some additional communication to occur beyond just the delivery of the original message, the extra transmissions can be relatively small. Since second attempts are no more likely to pass through than the first, the incentive to try multiple messages disappears, and with it the excess bandwidth consumed.

(See Why won't filters solve the problem?)

#### 3.5.3 How does the ABM compare to sender-pays solutions?

Sender-pays is a general term which describes a class of anti-spam solutions that are designed to pass some of the recipient's costs to the sender, thereby raising the cost to sender. The primary effect is that a sender who bears more costs will be more careful with his targeting (rather than indiscriminate) when sending. If the costs are sufficiently high, a sender will choose not to send at all.

There are several forms of sender pays, including a flat tax, a per-message delivery fee (postage), computational challenges, and bonds. Although the sender clearly pays when required to take a CAPTCHA (with their effort/time), challenge response is not usually considered sender pays. The ABM is similar to sender-pays, but is more accurately "sender-risks".

(see <u>How does the ABM compare to a flat tax?</u>, <u>How does the ABM compare to computational challenges?</u>, <u>How does the ABM compare to a flat postage fee?</u>, <u>How does the ABM compare to Ironport's Bonded Sender Program?</u>, <u>Q: How does the Attention Bond Mechanism work?</u>)

#### 3.5.4 How does the ABM compare to computational challenges?

A system that makes use of a computational challenge, like the ABM, is usually a hybrid of a whitelist, challenge-response, and a means of forcing the sender to perform an

expensive calculation. Proposed designs include Hashcash, Cam-Ram, and Microsoft research's Bankable Postage.

An often-promoted advantage of computational challenges systems is that third parties are unnecessary; the sender's computer and the recipient's computer are enough. However, use of computational challenges has several drawbacks. The first is that the computation introduces a delay in delivery. If the delay is made short by simplifying the computation, the cost to the sender is diminished along with its screening effect. Since compute cycles are comparatively cheap, particularly when compromised machines are available to harness, the delay would need to be significant in order to block spam. While pre-computation can be introduced to eliminate the delay, pre-computation requires a third party tracking system, similar in complexity to the ABM.

A second disadvantage is that the recipient does not get the benefit of the sender's effort (and nor does anyone else). There is no 'transfer' of value. The computation is a deadweight loss.

Proposed extensions, such as Bankable Postage, address this 'money burning' problem by allowing the work of the computation to be exchangeable. The additional infrastructure can also be used for pre-computation, eliminating the problem of delay. Yet if generic, the same infrastructure can be used to transfer real money, in the form of bonds, which do not require exchange to be useful.

Finally, unless the computations have some social value (rather than just CPU heating), the use of money is more efficient (in the economic sense).

#### 3.5.5 How does the ABM compare to community filters?

Community filters such as Yahoo's filters, Cloudmark, and Vipul's Razor, like other filters, have the generic problems of filtering. See <u>How does the ABM compare to filters?</u> and <u>Q: Why won't filters solve the problem (Summary)</u>

Since community filters harvest the decision making of your neighbors (in that such filters record the classification made by others in the community), your neighbors end up deciding what you read. If other members think and ad is spam, but it is something that you would have liked to know about, you lose. In contrast, with the ABM, each person can have their own definition of spam.

Community filters do reduce the amount of effort for each community member, since they share the burden of classification. Yet, there is still the burden of classification some of the time, and recipients have no potential for being paid for their inconvenience.

## 3.5.6 How does the ABM compare to challenge-response with CAPTCHAs?

The key advantage the ABM has over traditional challenge-response is the use of automatic posting and return of the response (in the form of a bond). This removes the requirement for direct human interaction and does not exclude email that is automatically generated and sent.

Since the human task of solving a CAPTCHA can be outsourced to a place where labor is inexpensive, bulk senders can reduce their costs, yet individual senders with a higher cost of labor pay comparatively more.

#### 3.5.7 How does the ABM compare to Whitelisting?

While a whitelist can do a good job, especially when coupled with stronger sender identities, a whitelist alone is an incomplete solution. The reason is that there are often unknown senders from whom you want email. In order for a conversation to start with these unknown senders, there must be a way for them to contact you. This means one of two things: you force them to use another medium, such as the telephone, fax, or snailmail (non-electronic mail), or you must provide a grey-mail box or alternative address that you check on a regular basis. The first approach creates more costs for the sender and may screen out contact initiation from those you would like to hear from. The other creates more costs for you.

Unlike existing challenge-response solutions, the ABM allows relationships to be established entirely via email without requiring the sender to use any other medium such as the telephone, the web, or instant messaging.

#### 3.5.8 How does the ABM compare to Blacklists?

Blacklists are lists of sender addresses and source mail servers that have been known to send spam in the past. Information from many sources (available to receiving mail servers) help to indicate where (and whom) an email is from. By deleting email that arrives from those named on the blacklist, spam can be reduced.

If senders were confined to using the same machines (with the same IP addresses) to send all their spam, blacklists would work quite well. Unfortunately, spammers are able to acquire new machines and new addresses from which to send mail with low costs and relative ease. New email identities can be created programmatically in email headers. Access to compromised machines can be purchased on the black market or compromised directly by viruses and Trojan horses created by spamming organizations. Since email source identities are cheap, blacklists are largely ineffective, and like filters, subject to an arms race.

The ABM uses stronger identity for the senders, and uses a whitelist. Emails sent from someone not on the whitelist are challenged, rather than deleted or just quarantined in a grey-mail box. The sender, if he includes a return address, can be notified immediately that his mail did not complete its journey, and can choose to authorize a bond to ensure delivery.

## 3.5.9 How does the ABM compare to a flat tax? 3.5.10 How does the ABM compare to a flat postage fee?

- 1) Postage stamp systems or flat taxes charge a uniform price for all people a 'blunt instrument'. Some mailbox owners might want to erect higher barriers to spam than others. With the ABM, the mailbox owner decides how much sender must risk to ensure delivery. (Individualized)
- 2) With postage stamp systems, the sender pays the postage costs regardless of whether or not the recipient wanted the email. This artificially raises the costs of all communication. With the ABM, the sender only pays if the recipient explicitly claims the bond. (allows zero cost initiation of relationships, entirely within the medium)
- 3) The proceeds from a postage stamp system go to the issuing company or to the government. With the ABM the recipient receives the bulk of the posted bond, in direct proportion to the number of times she is the target of spam. Other stakeholders, such as ISPs and the escrow agencies, get a cut in the form of a small transaction fee, so they get a benefit too.

#### 3.5.11 How does the ABM compare to using email stamps (e-stamps)?

A: With the ABM, you don't pay if you're not sending spam. Email can remain free whenever it is between two parties that want to communicate -- even when they are initially strangers. With e-stamps, you pay every time, all the time, regardless of intent or use. Valuable communication will be lost.

A flat tax (or postage fee) must to be collected. Even if the government subsidizes the cost of implementing and operating a collection system, the proceeds do not directly go to recipients, but rather to the government or a private agency that could be authorized do to the collection. In contrast, with the ABM, the recipient receives the bulk of the financial benefit as a transfer from the sender.

A flat tax or per-email postage also has the disadvantage of being 'one size fits all'. For some mailbox owners, the default flat tax may not be high enough to eliminate spam. Individual preferences for quality and content are possible with the ABM, due to the ability of the mailbox owner to set the size of their bond.

#### 3.5.12 How does the ABM compare to Brightmail?

Brightmail's solution is a form of filter, where the filter rules are generated dynamically at a central facility and then published out to all the mail servers of its customers. As a filter, it is subject to the problems of false negatives, false positives (very few in their particular implementation), and includes a grey mail box. Brightmail has the operational overhead of running an operations center to create new rules, due to the arms-race nature of filtering. The costs for this are then passed down to the subscribers.

#### 3.5.13 How does the ABM compare to Ironport?

Ironport produces a hardware appliance for processing email. Their appliance makes use of Brightmail's dynamic rule creation and distribution system. See <a href="How does the ABM">How does the ABM</a> <a href="Compare to Brightmail?">Compare to Brightmail?</a>

## 3.5.14 How does the ABM compare to Ironport's Bonded Sender Program?

With Bonded Sender, recipient mailboxes are protected by the Ironport filters. Marketers who wish to reach these recipients run the significant risk of having their emails filtered, unless they are members of the Bonded Sender program. Compliant marketers (members of the program) who want to be sure their email is delivered are required to post a bond. The bond will be forfeit in sizable increments (\$20 at the time of this writing), for each complaint received by mailbox owners.

Like the ABM, Bonded Sender makes use of a bond, but it is a slightly different kind. The bond in Bonded Sender is a 'bulk bond' or a single bond for multiple recipients. In contrast, the ABM uses a per-recipient bond, of a size set by the recipient, rather than by the service's operators. In a sense, using Bonded Sender is analogous to paying protection money: 'If you pay us, we won't block your spam'.

Bonded Sender has several significant drawbacks:

- 1) The capital required for posting the bulk bond, combined with the overhead of establishing the relationship with Bonded Sender may be prohibitive for smaller businesses. Although a \$500 minimum bond is available to non-profits (as of this writing), the ABM requires no more than the size of a single recipient's required bond to be posted, which we expect to be as a little as \$0.05.
- 2) Recipients do not get any of the value of the claimed bond. Although a portion of the collected bond is to be given to charities, the exact portion (as of this writing) is not published. In contrast, with the ABM, the recipient gets the bulk of the collected bonds. Should they wish to donate these funds to charity, they can still do so, and donate to one of their own choice.
- 3) Bonded Sender is closed rather than an open standard. It is available only to those protected by Ironport appliances.
- 4) The bond size is fixed, rather than set by the mailbox owner. This ignore individual preferences and value of time.
- 5) Recipients must complain (through an intermediary), but there is no direct benefit (compensation) for complaining. If the cost and effort of complaining is high, few people will make the effort and recipients will continue to be spammed.

#### 3.5.15 How does the ABM compare to Goodmail?

to be completed:

- 1) Trusted email class (protection money to ensure delivery)
  - a. Paid piecewise, as opposed to up front with Bonded Sender
- 2) Relies on filtering
- 3) Recipients do not get proceeds
- 4) Like a flat tax or sender-pays, only proceeds go to private company and ISPs.
- 5) See comparison to Bonded Sender

#### 3.5.16 How does the ABM compare to Microsoft's Penny-Black?

Penny Black is the name of a set of spam reduction projects at Microsoft Research, and is not a specific solution to spam. (See <u>How does the ABM compare to Bankable Postage?</u>)

#### 3.5.17 How does the ABM compare to Bankable Postage?

Microsoft's Penny Black group's proposal for Bankable Postage is similar to the ABM, with the exception that the ABM allows transfers of utility (money) between sender and recipient, while Bankable Postage requires some additional transfer steps to achieve some aspects of the ABM. Also, with Bankable Postage, the currency of transfer is sunk computational costs, and while possible, it does not include recipient chosen variable bond sizes. See <a href="How does the ABM compare to computational challenges?">How does the ABM compare to computational challenges?</a>

# 3.5.18 How does the ABM compare to Microsoft's "Caller-id"? 3.5.19 How does the ABM compare to Domain Keys? 3.5.20 How does the ABM compare to SPF?

Caller-Id, SPF and Domain Keys are a proposed means of increasing the strength of the sender's identity in an email.

For example, with SPF the source mail server of the message (domain) can be verified as valid for the sender listed in the message. Microsoft and others propose that senders of email register the possible source addresses of their mail servers with the Internet's Domain Name Service (DNS). In general, strengthening identity is needed for many of the proposed solutions, including the ABM.

Note: When coupled with a whitelist, stronger identities can be very effective at blocking spam. However, whitelists and identity alone do not address the problem of first contact. If you block everything from those you do not already know, they are forced to contact you through some other means. This means added cost (even to senders you want to hear from), and delay. The ABM solves this problem. See <a href="How does the ABM compare to Whitelisting?">How does the ABM compare to Whitelisting?</a>

#### 3.5.21 How does the ABM compare to Yahoo's Filtering?

Yahoo's anti-spam system uses a combination of regular filters and community filtering (although it may use other technologies as well). Drawbacks to their approach include false negatives and false positives, need for a bulk mail folder (grey-mail box), and the 'tyranny of the vocal minority' of the community filters. (See <a href="How does the ABM">How does the ABM</a> compare to community filters?, and <a href="Q: Why won't filters solve the problem">Q: Why won't filters solve the problem</a>?)

#### 3.5.22 How does the ABM compare to Cloudmark?

Cloudmark is a community filtering system.

See How does the ABM compare to community filters?

#### 3.5.23 How does the ABM compare to Vipul's Razor?

Vipul's Razor is a community filtering system. See How does the ABM compare to community filters?

# 4 Security, Viruses, Honeypots, and Authentication

## 4.1 Q: What about possibility of fraud or a virus triggering bond payments?

A: There are several types of possible fraud. (see also <u>Q: How does the ABM handle Honeypots?</u>). It might be possible for someone to write a malicious virus that causes a mail program used by many people to send messages to addresses owned by the virus writers. The virus writers could attempt to claim and keep the value of the bond.

Proper safeguards will be important, but as with any financial network, it may be impossible to completely eliminate the risks. A depleted escrow account would certainly serve as an indicator that something is wrong and the machine or account has been compromised. However, liability, at maximum, would be limited to the current balance in the compromised person's escrow account. In addition, escrow agencies could provide limits on liability, similar to the way credit card owners have limited liability if their card is stolen. Due to the tracking of payments and the need for a means of getting money out of the system for use, covering one's tracks may be more of a challenge.

#### 4.2 Q: How does the ABM handle honeypots?

(In the honeypot scenario, a person sets up hundreds or thousands of mailboxes, sets a positive (maybe high) warranty, and then does everything they can to make the addresses for the dummy mailboxes known. Unsuspecting marketers send to these addresses, and if they post bonds, the honeypot creator profits by collecting the posted bonds.)

A: There is nothing in the ABM itself that prevents this from happening. We'll argue that this is not a bad thing. If marketers stand to lose money when sending email to unknown addresses, they have the incentive to find out ahead of time if these addresses actually belong to a consumer that might buy their product. The very threat of this type of thing will stop frivolous untargeted broadcast spam. It is quite likely that a secondary industry will appear to help marketers determine the validity of new email accounts in advance of their 'first contact' campaigns (like a credit reporting agency or a reputation system in the abstract).

As with banking and credit agencies, reputation agencies will emerge to serve the marketing community. Marketers will target campaigns at those identities with proven track records of making purchases, while new and unproven identities will have 'unknown credit' and a correspondingly higher risk. As a result, informed senders will not send at all or will only be willing to pay a very small bond due to the high risk of dealing with an unknown recipient.

In summary, the presence of honey and of honeypots causes senders to do the right thing.

#### 4.3 Q: What is the role of stronger authentication?

A: Stronger authentication (not necessarily cryptographically strong) allows whitelists be effective. This makes it essential, as whitelists allow the cost of email to remain low for the bulk of communication. The adoption of a stronger means of authentication is an important aspect of any realistic spam solution, including the ABM.

Whitelists that rely solely on the sender addresses and other easy-to-fake header information allow a spammer to guess which identities may be on the whitelist, then fake an indentity to get their email through. For example, it is likely that millions of people would have the sender address "orders@amazon.com" on their whitelist, and so without stronger identity whitelists will be ineffective.

Work is being done to create stronger authentication. See <u>How does the ABM compare</u> to Microsoft's "Caller-id"?.

#### 4.4 Q: What about privacy?

With the ABM, the identity of the sender need not have any correlation with his Internet domain or even his real world identity. It just needs to be hard enough (expensive enough) for someone else to fake the identity (since being able to fake an identity on the whitelist makes it ineffective).

Also, if a sender's identity is not whitelisted and the sender is required to post a bond, all that matters for the bond mechanism to function is that the identity of the *bond underwriter* is recognized in the sender's subsequent response. As long as the bond underwriter is trusted, the recipient can accept the email and bond regardless of the sender identity. The underwriter will issue payment if the bond is claimed.

This is very important property, as it preserves sender privacy while still allowing the email through (e.g. the bond can be purchased with anonymous digital cash).

## 4.5 Q: Will stronger authentication make email more difficult?

A: Prior to spam, there wasn't enough of a reason for most people to switch to an authenticated format. Adoption of stronger authentication has been hindered due to poor user interfaces. Most mail tools were (and many still are) clumsy at best for the non-expert to use for signing messages and managing identities. This technology *can* be made simple for people to use, and there is opportunity for organizations that manage to do it well.

## 4.6 Q: Why are Certifying Authorities not needed for individual identities?

A: Certifying Authorities are needed when you wish to tie a real-world identity to a digital identity, thus leveraging the real-world reputation. With email between individuals, reputation can be established over time, entirely within the context of the medium. While some people may wish to have their identities signed by a CA, it is not necessary. What is important is that the identity is difficult for someone to fake. Several commonly used public-key cryptographic systems have this property today.

#### 5 Infrastructure and Costs

#### 5.1 Q: Doesn't this require infrastructure?

A: Absolutely. But the spam wars already require a huge amount of infrastructure and it is growing exponentially as the technical arms race continues. ISPs are hit hard, corporations are hit hard, and so are consumers. It costs time. It costs money. It costs brainpower. Any real solution will require infrastructure. So that money should be spent on the right solution and the right infrastructure, which will put an end to the escalating costs. In this case, that means using either bearer bonds or an escrowed "wire transfer" to facilitate the exchange of warranties. Mail gateways that service large user bases (Hotmail, Yahoo and major corporations and ISPs for example), will need to provide their users with whitelist management and identity based blocking at their gateway, or at least allow clients to implement a means of doing so on their own.

#### 5.2 Q: Who runs the payment system?

A: There will likely be several parties involved. Each sender and recipient will need a relationship with an escrow agency, where the funds they make use of for posting

warranties are kept. For most users, the agency could be their current ISP. Marketers will maintain relationships with special escrow agencies that are focused on their unique needs. One or more underwriters of the electronic payments will be necessary, independent of the escrow agencies that maintain a relationship with customers (senders, as individuals or as marketers, and recipients).

# 5.3 Q: How will the costs of tracking the warranties be covered? Can it be done cost effectively and or will such costs prohibit the adoption of the system?

A: Ultimately, the senders pay for the cost of the infrastructure, analogous to the way users of the U.S. mail system pay for its operation. The companies that provide the infrastructure, escrow agencies and underwriters, may get revenue from at least two sources: transaction fees and account holder float.

The majority of this income will be from transaction fees - a small commission on any warranties that are claimed by recipients. Knowledgeable marketers consider the lifetime-value of acquiring a customer when deciding whether or not to target them. If this lifetime value is high enough, the marketer can justify significant acquisition costs – and spend money to gain a new customer. The ABM provides a means for companies to reach new customers, as well as maintain relationships with existing ones. If a marketer wants new customers and sends out bulk emails to establish new relationships, many recipients will choose to keep the warranty. Nevertheless, for many marketers, it will remain profitable to do the campaign, and the responses they get back over time will let them target their service offers and market smarter.

Meanwhile, average individual senders (non-marketers), whose warranties are rarely seized, will suffer little in the way of losses yet will be able to accumulate money due to claiming the bonds of direct markers if such email is received. The result is that most individual users will be able to make use of the system for free, and they may even profit. If individuals accumulate money in their accounts, it can be dispersed via a wire transfer when it reaches a certain size so as to keep the overhead cost of the wire transfer to a minimum. The system as a whole will operate similarly to the way long distance companies do their tracking and accounting today. Telcos track sub-penny transactions across multiple providers for a single call, and do so for a month at a time (there are phone companies that bill by the second, at \$0.03 a minute). At the end of the month they send a single bill.

Escrow companies can bring in addition revenue by making use of 'float', via the practice of lending out currently held money in bulk. Most individual account holders, or at least those with no prior relationship with an escrow company, will be required to prefund their accounts to some minimum (\$5 for example). These accounts may very well accumulate funds during use from claimed warranties and have an even higher value. With many millions of account holders, the float on the sum of their accounts will be significant, and will help offset the operational costs.

The bulk of the costs are essentially subsidized by the direct marketers, and the costs are still lower to them than traditional means (contrast losing 10 cents every time for sending an email vs. a loss of 35 cents to send you an advertisement in the mail). The ISPs and enterprises, which provide email accounts to individual users, can share some of the revenue provided by the float and commissions to cover some or all of their costs.

#### 5.4 Q: Will existing mail clients need to be changed?

A: To provide the smoothest and most integrated user experience, existing clients will need to be modified, but this is not a requirement. Those using POP or IMAP to access their mail (some users of Outlook, Eudora, and web mail interfaces such as Yahoo and Hotmail) can initially use an Internet hosted proxy service. However, over time, makers of these clients and providers of mail services will want to update their user interfaces to provide management of the users' whitelist and escrow accounts, and to give users the ability to adjust and personalize some characteristics such as the size of their minimum warranty. Essentially, they will want to update their software and interfaces to support this solution, just as they have made modifications to support the existing filtering and blocking systems now in use.

#### 5.5 Q: Does this design require changes to SMTP?

A: No changes are necessary. All information needed to make the system work can be wrapped inside a regular message, facilitated by existing standards such as S/MIME.

## 5.6 Q: What happens if warranties are demanded in foreign currencies?

A: When a sender is required to provide a warranty for a recipient in a country with a different currency, the warranty must be exchanged into the native currency of the recipient. The exact details of this need to be worked out and are outside of our expertise.

Sender's may wish to establish escrow accounts managed in foreign countries for foreign currencies. During account setup, when the account holder funds the account, they could pay with a credit card and the funds would be exchanged at that time at the current exchange rate. Account funds would be held subsequently in the native currency.

### 6 Compatibility with Existing Uses

## 6.1 Q: How does this system work with mailing lists that I subscribe to?

A: Subscribers would need to add the mailing list's source address to their white list at the time they signup (this can be facilitated by a link on the lists' webpage). When subscribing to a list, the list operator may send the subscriber a challenge (a demand for a warranty), which subscriber would have to post. If a subscriber removes the mailing list from their whitelist, it can have the same effect as unsubscribing. The next time an email comes from the mailing list, it will be bounced with a challenge back to the list operator, at which point they can automatically remove the subscriber from their list and prevent future mailings, or elect to post the warranty.

## **6.2** Q: How is this compatible with e-commerce and customer relationships?

A: For new customers, when signing up on a website, the website operators should notify the customer what identities will be used to send them email. The customer can add the identity to their whitelist, a simple process that can be facilitated by use of a special link embedded in the e-commerce site's web page. If clicked, the link triggers an update to the users' whitelist.

In the case that the customer has not modified their whitelist, when the ecommerce company sends them an email (for example, if the email contains a confirmation of an order), the company can decide how it responds if challenged. The company might respond by posting a warranty as requested, which the customer/recipient would be able to claim and the company accepts as a cost of doing business. However, the content of the email can be customized with an incentive NOT to claim the warranty (a coupon or discount for future use) that at the same time encourages the recipient to shop again. Also, any time the customer does claim the warranty, the company will have a record of it and can use this in making decisions for future interactions with the customer.

The company might decide that demanded warranty is too high, in which case the email is undelivered and the customer will need to log into the web site to get their information. A notification of the current status of the whitelist can be placed in the account management section of the marketer's website, reminding the customer they can increase the ease of communication by modifying their whitelist.

For existing customers, if the customer already has the e-commerce company on their whitelist, no warranty will have to be posted and communications will continue as they have before.

## 6.3 Q: What if I have forgotten my password to a website and need it to be emailed to me?

A: Typically, you would already have the website's identity on your whitelist, so the email with your password would go to your mailbox without being blocked and challenged. If it is not, the form you use to request the password can give you the correct identity to add to your whitelist before you make the request.

As an alternative, the form used to request your lost password can ask that you provide an additional passpharse (other than your password) as input. You add this second passphrase to your whitelist, and the email sent from the website containing your password would contain the second passphrase in its subject line. Your whitelist would recognize the passphrase and allow the email through rather than issue a challenge to it. In effect, your whitelist enforces this policy: "allow any email with the passphrase 'xxxxxx' in the subject line to pass through to my inbox."

# 6.4 Q: What if I meet someone at a conference and they give me their business card with their email address and I write them?

A: You would post the warranty (depending upon its size), but assume that the recipient would not collect it when they see the note that says "nice to meet you at the conference, I am following up to discuss..." If they do collect the bond, then you know they aren't interested in talking to you. We expect that etiquette will develop quickly in this area.

#### 6.5 Q: How will the ABM work for large institutions?

A: The needs of large institutions, such as universities, non-profits, and large businesses are often different than those of individuals. If the corporation is providing the email accounts, advanced whitelist management capabilities will be needed to implement appropriate policies. For example, it may be important for members of a company, such as its executive team or managers to be able to broadcast email to employees to make announcements or create mailing lists. Also, it should be easy to whitelist entire domains or specific administrative groups within or associated with the institution, so divisional, department-wide, or global whitelists are required.

For productivity or other policy reasons, an institution may want to set or enforce certain minimum and maximum warranty sizes (again by division, group, office, person), and be able to track the flow of any warranty payments collected or posted by company employees. Email will be managed as part of a company's overall communication strategy; the corresponding server software must be as sophisticated as its environment.

## 6.6 Q: What does my IT department need to do once we begin using the ABM?

A: Servers that perform whitelist enforcement and management, manage other user preferences, generate and store of identities, facilitate the posting and claiming of warranties, hold inboxes, and communicate with an off-site escrow facility will all be required somewhere in a company's overall service architecture. Updates to existing email client software may also be desirable, but not required. It will be possible to outsource many of these activities, including regular operations. The entire system could

be implemented as a web-based email service, implemented entirely locally (even the escrow accounts), or some combination in between. As the market for such software develops, an IT department will be able to choose their level of operational involvement.

Unfortunately, like any spam solution rolled out system-wide, some integration will be necessary if a company chooses to operate one or more of the components in-house. However, there is no reason why the ABM should be any more difficult than other solutions in this regard.

#### 6.7 Q: What about friends and family?

A: Email from friends and family on your whitelist comes through to your mailbox as it did before. A friend or family member not on your whitelist will need to be added to the whitelist in advance (by giving you their email in person or by phone), or they can send you an email with a warranty and you simply accept the email and do not collect the warranty, then place them on the whitelist for the future.

#### 6.8 Q: How does the ABM work with children under 18?

A: The Attention Warranty System helps families, parents and children get control of kids' mailboxes. Setting the warranty high enough on kids' email accounts will block porn and other unwanted spam emails, and limit marketers from reaching children. Children can add new friends to their white list on their own if permissioned to do so by their parents (major ISP software such as AOL already provides for various levels of parental permissioning and delegation of authority regarding email, instant messaging, etc.), or see their parents if required. Children using email as sub accounts on major ISPs (AOL, Earthlink, MSN, United, cable companies and telcos) that already have sophisticated parental/minor controls will be able to add escrow accounts and allocations at the discretion of the master account or screename holder who manages the account billing. This is already the case for other fee-based services from these ISPs. Freemail accounts such as yahoo, hotmail, and others, will require credit cards that will be subject to the same governance issues they are today for online shopping and fee-based services.

# 6.9 Q: Many people access their email from Internet cafes and do not have accounts with ISPs. How those without bank accounts, ISP accounts, or credit cards use an escrow account?

A: The Internet cafes can provide an escrow account for their patrons. Senders using cafes can pay in cash in their local currency to the cyber cafe owner, who can have one or more escrow accounts for use of their patrons. If recipients claim any warranties put up by a sender, the cafe can keep track and debit from the sender's local account.

### 7 Effects on Participants

## 7.1 Q: If I am a marketer, will this system stop me from reaching my customers?

A: No, in fact it improves your ability to reach them. Existing customers can simply white list the marketer, so that communication comes through to them, providing the service and quality they expect. Prospects can do the same, and/or set a threshold or "price" for their attention, which gives the marketer a good sense of the recipient's interest, value and they can measure their marketing cost against that information and ongoing feedback based on response rates from recipients to various campaigns.

#### 7.2 Q: What are the effects on Marketers?

A: This system helps legitimate marketers. Those who have retreated from email marketing for fear of tainting their brand and products can re-enter legitimately and smartly. Those skilled in database marketing may have an added advantage in that they can better refine their target lists. And because it is an economic system that allows the recipients to "signal" and provide information in terms of value and interest back to marketers (the senders), ultimately it makes the marketers smarter and more efficient about how to successfully reach the right targets.

#### 7.3 Q: Will I have to pay taxes on any warranties I claim?

A: If you are earning significant amounts of money through collected warranties, it is likely that you will be taxed. The specifics are not yet determined. We expect that for most people, the sum of money earned in collected warranties will be below the \$600 threshold necessary for reporting. However, in the case where an account owner has earned more, their escrow company can require the owner to furnish their taxpayer ID an can issue a 1099 statement for tax purposes.

### 8 Adoption

#### 8.1 Q: Why would anyone participate?

A: Users of the ABM will get greater value from their email and will not be troubled by spam. Strangers simply can't reach you if they're unwilling to pledge that they will not waste your time. Corporations will see their costs of trafficking spam greatly reduced, as will ISPs.

#### 8.2 Q: How can user adoption be encouraged?

A mail tool plug-in or a proxy service can be offered for free that includes a traditional challenge-response system along with means of handling warranties. This offers immediate benefit for those most troubled by spam, even before warranties are regularly risked. As the installed base grows, it will become cost effective for marketers to communicate with the participants, and paying the warranties in bulk will be simple. Because the ABM offers recipients the opportunity to profit from any unwanted interruptions, users have a greater incentive to stay with it than other spam solutions.

#### 8.3 Q: Why should marketers participate?

A: The ABM allows marketers to reach their potential customers, and is it will be competitive to other means - cheaper than regular mail via the post office.

#### 8.4 Q: What are the benefits to ISPs?

A: ISPs and web mail providers, aside from ultimately realizing the cost savings of reduced spam, can gain a valuable revenue stream. They serve a channel to the consumer and as such can obtain a portion of the money shared by all the infrastructure parties in the value chain. The marketers fund the top of the chain, and the ISP acts as an agent, helping their customers manage flows of information and funds. The escrow companies and underwriters (some of which might also be ISPs) also stand to make significant profits in the form of transaction fees.

### 9 Social Implications

#### 9.1 Q: Will this cause a loss of freedom of speech?

A: No. People are still free to start their own newspaper, television station, radio station, and do their own desktop publishing offline. They can still send emails to everyone they know and have been whitelisted by, and encourage recipients to forward them on to others. They can set up a website, list it on search engines, and email random addresses if they are willing to post a warranty, and if the recipient finds the speech of value, then the recipient can read it and not collect the warranty. If the recipient views it as spam, they can collect the warranty and delete the email. The right to be left alone is as fundamental as freedom of speech. Consumers can unlist their telephone number, or sort their mail based on bulk postage vs. first class and throw out the bulk mail. This is no different. In fact, if the warranty mechanism helps you rise above the noise of people shouting for attention, it could increase the ability to reach people you care about.

## 9.2 Q: If everyone starts using the ABM, what happens to anonymous email?

A: Anonymous email without a warranty can still be sent, no different than today. Mailboxes that do not challenge unknown senders will still be able to receive email from anonymous sources. However, since many mailbox owners will require a sender to post a warranty if they are unrecognized, email without payment will be blocked (if it didn't work this way, spam would be rampant). To address this, several options will be available. The standard escrow agencies can offer secondary addresses for which they agree not to reveal the identity of the sender unless there is a court order. These accounts can be funded through a regular account or a separate payment, but would appear as 'anonymous' to any recipient. You as a sender would have to be doing something illegal in order for your identity to be divulged. This provides the same level of anonymity afforded by websites and webmail companies that let people post and send without revealing their true identity.

Those who require anonymity beyond this level will still have options, though at increasing costs and less convenience. For example, companies could offer to receive payment by anonymous digital cash (or cash paid in person or by mail), and email could be sent through these services by accessing them from a cyber-caf.

In general, the cost of discovering a sender's identity can be made arbitrarily large, but as the level of anonymity increases, so will the cost to the sender who wishes to remain hidden.

#### 10 Other

#### 10.1Q: Is this really a new idea?

A: The concept of charging the sender, as applied to email, has been floating around the Internet for many years. In fact, the general concept of paying for the attention of someone else has been around as long as professional services have existed. Accountants, lawyers, doctors and engineers have long charged for their time, either by the hour or by project duration. The fees involved were high enough to cover the costs of the accounting. For example, a professional might bill at \$200 an hour for a job, \$20 of which is lost to the accounting process as overhead. If you are not a professional who can charge these rates, and say, make \$20 an hour, it's hardly justifiable to use an expensive accounting system.

The ABM is possible because the costs of the accounting have dropped dramatically due to information technology. Since the costs are lower, the ability to perform careful accounting is becoming available to everyone, regardless of hourly rate. Your computer can keep track for you. When an advertiser wants some of your time, they pay for it (or at least risk the payment) in the form of the bond. The ABM just gives people the ability to cost-effectively account for and charge for their time, even if the time is measured in seconds.

#### 10.2Q: So what's your contribution?

A: As far as we know, we're the first to formally prove that all parties can potentially be better off than even a perfect filtering technology. The idea is that promoting the right kinds of valuable electronic interaction among friends, acquaintances, business associates, marketers and strangers can make everyone happier than simply shutting off communication. For the academically inclined, the proof can be found here — <a href="http://ssrn.com/abstract=488444">http://ssrn.com/abstract=488444</a>.

This has implications for how to approach solving the spam problem, what kinds of regulation are fruitful and what should be avoided, and how standards and technologies can work together to make it happen. We have focused on how to maintain (and to the degree that the email medium has already been damaged by spam, restore) the unique properties and value of email. Most importantly, we have tried to find a system where everyone has the incentive to improve the quality of the email information marketplace, so that innovation can continue, and to present it as a workable solution. This system focuses the costs and penalties on those (spammers) who are abusing it, does not unfairly penalize good-citizen users by constricting flow with filters or imposing tax burdens and costs, and provides a valuable way for marketers to connect with existing and potential customers (recipients), while giving the customers control over what is theirs – their mailbox.