

Spam – Solving it Economically?

by
Omar Zia Khan

Latest News

- 4 million emails received a day
- 1 department devoted to elimination of spam
- 10 emails per day manage to slip through

References

- In addition to assigned papers
- www.wikipedia.org for definition of spam
- www.whatis.com for history of spam
- “The Economics of Spam: the Spam Business Isn't Always What You'd Think” by Joe St Sauver
- “The Economics of Spam” by Stephen Cobb
- “Spam Economics, Bonds, and Restoring Valuable Communication” by Thede Loder, Marshall Van Alstyne, Rick Wash, Mark Benerofe

Outline

- Spam (Definition & Issues)
- Existing Solutions and their Inadequacy
- Parasitic Economics of Spam
- Economic Solutions for ESPs
- Economic Solutions in general
- Issues with Economic Solutions
- Conclusion

Spam – Defined

- Definition
 - Act of sending unsolicited electronic messages in bulk
- Other names
 - Unsolicited Commercial Email (UCE)
 - Unsolicited Bulk Email (UBE)
- Historical Perspective

“The term is said to derive from a famous Monty Python sketch (“Well, we have Spam, tomato & Spam, egg & Spam, Egg, bacon & Spam...”) that was current when spam first began arriving on the Internet. SPAM is a trademarked Hormel meat product that was well-known in the U.S. Armed Forces during World War II.”

Some Basic Stats (2002)

- Estimated loss of 10 Billion?
- $\geq 60\%$ of all email is spam (Brightmail)
- 30% users concede curtailing use of email due to spam in a survey
- 40% list it as the worst IT problem in another survey

Some Non-Obvious Issues with Spam

- Attention Grabbing Spam
- Hoaxing Spam
- Fraudulent Spam
- PrOn Dialers
- Virus/Worm Infection
- Identity Theft
- Money-Making

Existing Solutions

- Legislative
 - Labeling
 - Opt-out
- Technological
 - Filtering (Rule Based or Bayesian)
 - Challenge Response (Quasi-Turing Tests or HIP)
 - Authentication Based

Issues with Legislative Solutions

- Definition of Spam
- Jurisdiction and Enforceability
- Lack of Incentive Compatibility

Issues with Technological Solutions

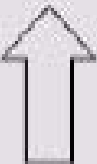
- Filtering
 - “*The Dog ate my Homework!!!*”
 - False positives and negatives
 - Passive measure
- Challenge Response
 - Automated Email Usage
 - Loss of Human Time
 - Inexpensive HR available
 - Use of compromised machines
- Authentication
 - Acquisition of new identities

Economical Solutions

- Computational Challenges
- Use of eStamps
- Escrow Services

The Economics of Postal or Snail Mail (small)

Cost per message



To snail recipient

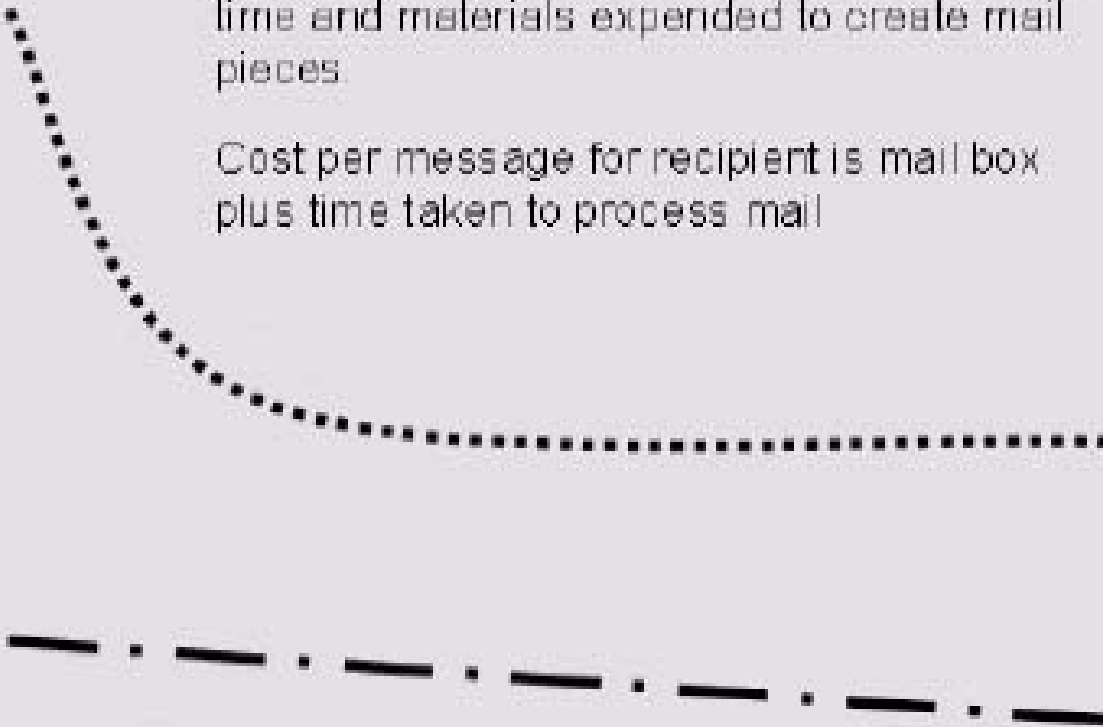


To snail sender



Cost per message for sender is postage plus time and materials expended to create mail pieces

Cost per message for recipient is mail box plus time taken to process mail

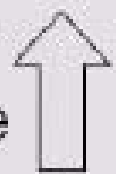


Number of messages sent



The Parasitic Economics of Spam

Cost per message



To spam recipient



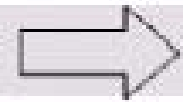
To spam sender



Cost per message for recipient is Internet connection and mail box charges, plus time taken to process mail

Cost per message for sender is charges for Internet connection plus time expended to create mail pieces

Number of messages sent



Some Basic Arithmetic

- Value of Product is \$49.95
- Marketing commission per sale is \$19
- Snail Mail
 - \$1 per brochure mailing cost for 5000 brochures
 - 5.26% response rate needed to pay marketing commission
($263 \times 19 = 5000$)
- Email
 - \$100 per million messages (with dialup!)
 - 3.5 million messages gave 81 sales at response rate of 0.0023% within first week and a \$1500 revenue
 - Task performed by a regular email marketing company
- Costs paid
 - Bandwidth
 - Opt out option

ESPs and Spammers

- Why Spammers use ESPs
 - Avoid blackhole lists
 - Save bandwidth
 - Avoid rate limiting on port 25 data
 - Obscurity
- Why ESPs want to stop Spammers
 - Abuse of resources
 - Cost of response to complaints
 - Damage to reputation of ESP
 - Risk of getting in a blackhole list

Basic Economic Model

- Account Creation Costs
 - Use of HIPs and limiting of per day emails allowed
 - Premium accounts have yearly fees
- Basic Model
 - Cost of creation of account, C
 - Messages per day allowed, D
 - Probability of a user complaining, p
 - Delay b/w complaint receipt & action, L
- Spammer can send at most $1/p$ messages
- Cost per message is Cp
- No. of messages allowed doesn't matter
- Probability of user complaining is a critical factor

Simplified Model

- Failure of existing techniques
 - Free Signup means $C=2$ cents (HIP)
 - Probability of user complaining $p= 1/1000$
 - Cost per message $C_p = 0.002$
 - Lowest price spammers charge is 0.0025

Simplified Model + Delay

- Allow delay in spam reporting and blocking
- DXL messages can be sent during delay
- Chance of complaint on a given day,
$$q=1-(1-p)^D$$
- If D is small, $q=pD$ and expected messages sent, E
$$E=LD + D/q \approx LD + D/(pD) = LD + 1/p$$
- For small values L and D , E approaches $1/p$ not 0!
- No. of spam messages sent is independent of messages per day

Per Message Challenging

- Trivial signup costs doesn't help
- Per message costs (C for n messages)
- C/n per message cost can be prohibitive
- Cost can be
 - Computational cycles
 - HIP solution
 - Monetary costs

Limited Initial Challenging

- Charge for only first n messages k times
- Can spammers get away by sending initial nk good messages?
- Optimal strategy for spammer is to spam as much as possible initially
- Mathematical notations shows
 - Number of messages per day approach 0
 - Lowering messages per day doesn't affect
 - Equivalent to charging for every n messages

Limited Initial Challenging (Cont'd)

- Equivalent to requiring a high signup cost
- HIP usage favors limited initial challenging for lesser user annoyance
- Computational challenges may favor high sign-up cost through pre-payment
- Issue of zombie machines for computational computation

Increased Limits

- Account termination replaced by resetting account payment schedule
- Limit exhaustion may lead to resetting account payment schedule
- Concept of multiple streams and tokens
- Termination of a stream rather than account incase of a complaint

Complaint Procedure Standardization

- Probability of a complaint is most dominant factor in economic prevention of spam
- Manual process in place currently
- Forwarding of mail required
- Issue of list cleansing
- Setting up of honey pots

Summarizing for ESPs

- Low sign-up costs and per day email limits are not sufficient deterrents
- Probability of a spam being reported plays the most important role
- Sender must help in eliminating spam by promptly reporting it
- Model penalizes sender very harshly based solely on receiver feedback

Moving beyond ESPs

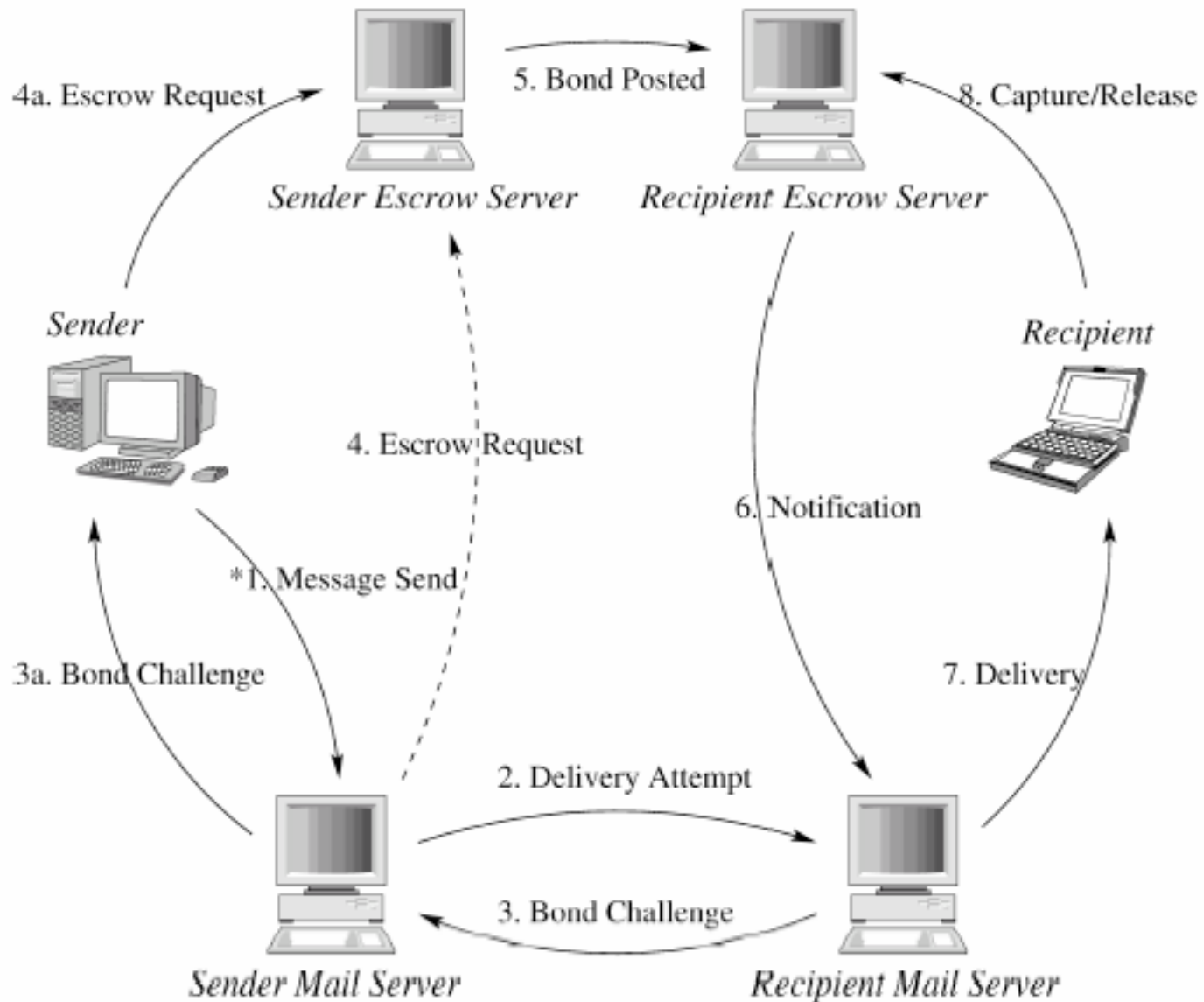
“I expect that eventually you'll be paid to read unsolicited e-mail. You'll tell your e-mail program to discard all unsolicited messages that don't offer an amount of money that you'll choose. If you open a paid message and discover it's from a long-lost friend or somebody else who has a legitimate reason to contact you, you'll be able to cancel the payment. Otherwise, you'll be paid for your time.

When this day comes, spam will cease to be a problem because people will be able to decide what their time is worth and advertisers will have to pay significant sums to reach people”.

Attention Bond Mechanisms

- Sender possesses private information
- Use of reputations for acquaintances
- Use of warranties for strangers
- Conversion of warranty into a reputation

Attention Bond Mechanisms



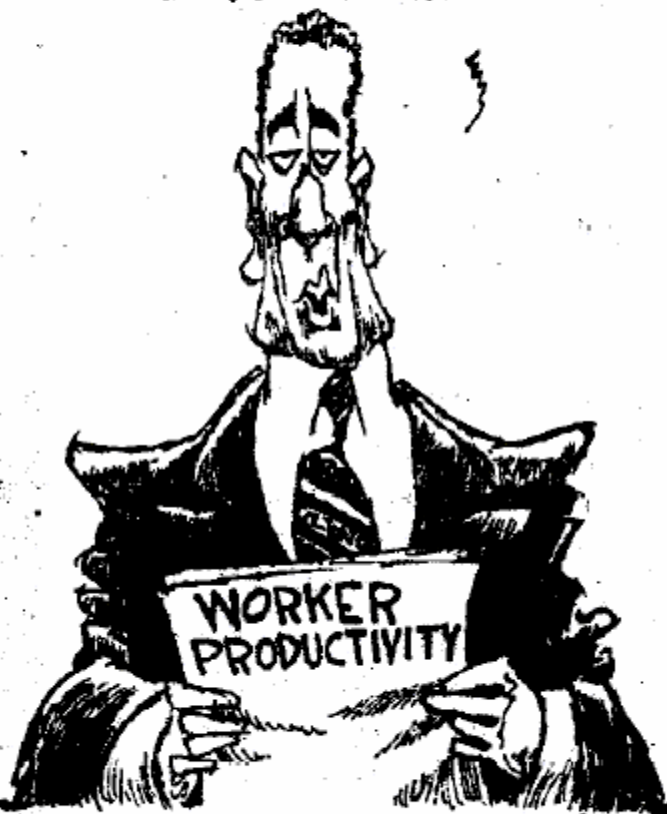
Issues with Economic Solutions

- Overhauling of whole email system needed
- Technological loopholes are not always handled by economic solutions
- All users are never rational

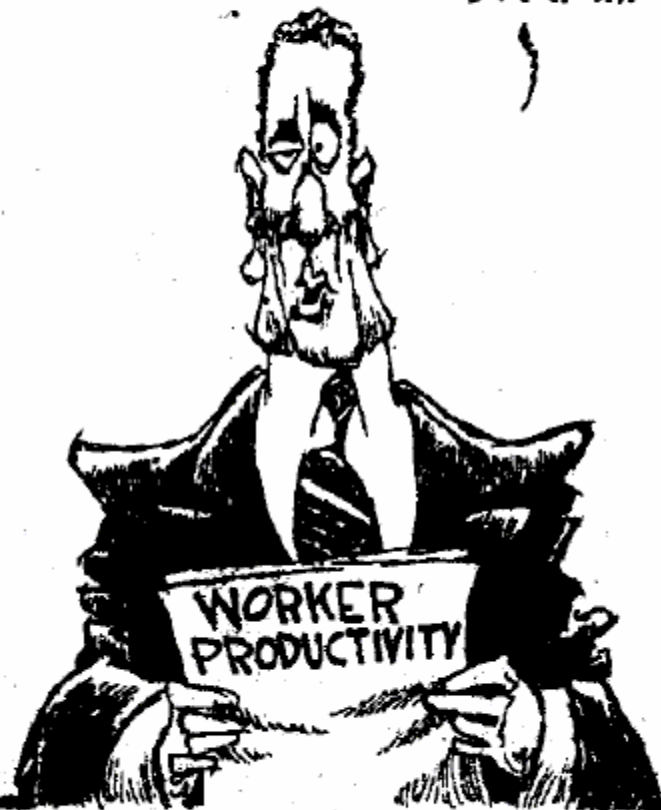
Email and Economic Solutions

- Why is Email popular?
 - Low (no) Cost of Communication
 - Asynchronous and Fast
 - Single Hop (No Third Party Infrastructure)
- Would it remain so after this?
- Would it be worse than spam?
- Trade-off

AMERICANS PRODUCED
THE SAME OUTPUT IN
FEWER HOURS LAST
QUARTER....



UNFORTUNATELY, THE
REMAINING HOURS
WERE DEDICATED
TO DELETING "SPAM..."



Nick Anderson

The Louisville Courier-Journal

Washington Post Writers Group