Spam
and the Limits of Interpersonal Collaboration

Nathaniel S. Borenstein
IBM Distinguished Engineer
Lotus Chief Antispam Strategist
Spam: How bad is it really?

- #1 concern of messaging buyers surveyed.
- Each cost estimate is bigger than the last.
- Tens or Hundreds of billions?
- 30 cents for each $100 of GNP?

Bottom line: it’s very, very, very, very, bad,
  - and it is absolutely getting worse!
  - and it is fundamentally unsolvable.
  - This talk is about long term strategy.
  - No date/feature/product promises here, sorry!
Spam isn’t just spam

- Spam may be inevitable in open interpersonal electronic communication – the scourge of the Internet era, a hard limit on human collaboration.

- SPIM, SPIT, blog spam, wiki spam, RSS spam, …

- Spam and viruses are inseparable.

- Phishing is a particularly nasty spam subtype.

- An e-mail-only solution, hard as it is, won’t do.

- It will never be eliminated, only kept in check.
Spam is a game, not a puzzle

- Puzzles are solved by algorithms
  - When you edit a document, no one is actively trying to sabotage you
- Games are played by opponents
  - The spam game has very complex rules
- An active adversary changes everything
- Let’s look at some of the rules
How email works (simplified)

How we think of it
How email works (simplified)

We all have ISP’s
How email works (simplified)

Some of us have corporate firewalls & data centers
How email works (simplified)

The rest of us have idiosyncratic service providers
How email works (simplified)

Of course, our ISP's need firewalls
How email works (simplified)

And let’s not forget the email relays…
How email works (simplified)

...which come in very heterogeneous flavors...
How email works (simplified)

...and often conceal malicious and clueless parties
Then there are opaque “routers” we can’t even see…
How email works (simplified)

So we need trusted 3rd parties, like CA’s…
How email works (simplified)

...or someone trusted to track a blacklist...
How email works (simplified)

… or to filter our email based on its content, etc.
How email works (simplified)

And we need to prove WE aren’t spammers!
How email works (simplified)

CA's

Filters

Trusted

Compliance

The spammers ignore any rules, of course…
How email works (simplified)

...as do the “good guys”...
How email works (simplified)

...and the vigilantes...
More complexity

= ?

More vulnerabilities

- Not always, but usually
- The world is NOT getting simpler
- Ever more paths to deception
- Ever more sophisticated protection needed

If there’s an end in sight, I sure don’t see it!
Many Techniques can be Helpful

- No Silver Bullets, though
- Most introduce additional problems and concerns.
- Each has fanatical supporters who disagree.
- The illusion of the best is the enemy of the good.
- Multiple approaches *must* work well together!
A Single Vendor Solution is an Illusion

New innovations are constantly needed

- They can’t always come from a single company
  - (No matter what they claim!)

- Users are best served by an open antispam ecosystem that avoids vendor lock-in

- Vendors should help maximize your productivity, not monopolize your software dollars.
IBM Antispam Strategy

- Lead the creation of a Comprehensive Model of Spam Control
- Build the best platform for integrating multiple cooperating technologies
- Contribute antispam innovations to the community where possible.
- Engage fully in education, standards, and other community efforts.
It’s still a lot harder than it sounds

- Cooperative voting by heterogeneous software, for example

- Not much architectural cooperation is visible among the antispam vendors.

- We’re inviting them to start.
Comprehensive Antispam Model
Draft 0.0.1
Let’s abstract away the details…
Let’s abstract away the details...

\[ S \longrightarrow R \]

... and we have a series of separate steps over time
Bad things can happen at any step,

but in practice, most bad things happen in the sender or receiver’s organization.
Each message transmission can be expressed as a vector through the dimensions of time, trust, and transmission between human minds.
If the world were perfect...
Filters: The Tools you Hate to Love

- Filters were our first line of defense.
- Today they are our last line of defense.
- We’re going to need them for a long time.
- But they will get much more powerful.
A simple filter

S → Filtering Agent → R

Trash
A simple filter

S → Filtering Agent → R

Can occur anywhere,
Can happen more than once,
Typically twice: S and R’s enterprise

Trash
An open architecture requires tight integration to be efficient
An open architecture requires tight integration to be efficient.
Email Authentication Technologies

- Reliable identity is key to identifying spam.
  - But Domain identity suffices!

- There are many good technologies
  - DomainKeys, Identified Internet Mail
  - SPF, Sender-ID, FairUCE
  - S/MIME, PGP
  - Biometrics and more

- Multiple identity technologies must coexist.
Sender appends verification information, in hope of bypassing filters.

“Verification” includes many payment systems.
Recipient-Driven Verification

Includes Challenge/Response, computational challenges, and some payment schemes
Trashing Spam is Too Good for It

Before we trash it, we might:
-- archive it, or
-- report it to some jurisdiction(s)
Humans don’t use protocols

Clarity demands differentiating “Person to Person” and “User Agent to User Agent”
The Passage of Time Means More Complexity

I sell real Estate & Viagra!
The Passage of Time Means More Complexity

There are new spammers out there!

I sell real Estate & Viagra!
The Passage of Time Means More Complexity

S → V1...Vn → Fi → SA1 → SA2 → SA3 → C/R → Fo → J1 → J2 → ... → Jn → T/T → R

I sell real Estate & Viagra!

There are new spammers out there!

I understand PGP but not S/MIME
The Passage of Time Means More Complexity

I understand PGP but not S/MIME

I used to be a spammer but I’ve reformed!

There are new spammers out there!

I sell real Estate & Viagra!
The Passage of Time Means More Complexity

I understand PGP but not S/MIME

I used to be a spammer but I've reformed!

No he hasn't!

There are new spammers out there!

I sell real Estate & Viagra!
Needed: a Distributed Spam-related Data Store
What’s a DSDS?

A shared data model, access protocol, access controls, and spam-related data, including reputation services.
DSDS is one major missing piece
But when all is said and done, educating wetware is what helps most!
And let’s not forget the role of law (and politics)!
The IBM Antispam Strategy (reprise)

• **Lead the creation of a Comprehensive Model of Spam Control**

• **Build the best platform for integrating multiple cooperating technologies**

• **Contribute antispam innovations to the community where possible.**

• **Engage fully in education, standards, and other community efforts.**
Leading in Antispam Innovations

IBM understands the “email commons”

- **Near term focus is architecture**
- **Architecture enables standards**
- **Technology innovations continue**
  - Chung Kwei, Spam Guru filtering algorithms, FairUCE
  - Received-line parsing
  - Multilingual spam control
- **Seeking partners for Consumer Education**
Building the best antispam platform

- Build on/componentize spamguru
- Define & open all relevant API’s
- Integrate with Notes/Domino & IBM Workplace
- Address performance issues
- Work closely with partners
Leading in (many) Antispam Standards

- Pipelined spam processing (plugins, api's)
- Cooperative voting (mechanism, results)
- Identity management (multiple id’s)
- Authentication protocols
- Shared personal & site customization
- Message tracking & tracing
- Advisory message cancellation
- DSDS: Data model, Access protocol, Security
The Future of Spam Control

- Different clients have different needs
- Today’s solutions won’t suffice tomorrow
- Spammers & opponents innovate constantly.
- We’re all in this for the long haul.
Any Questions?

Nathaniel S. Borenstein
nborenst@us.ibm.com