Yahoo! SpamGuard

Miles Libbey
Anti-Spam Product Manager
Yahoo! Mail

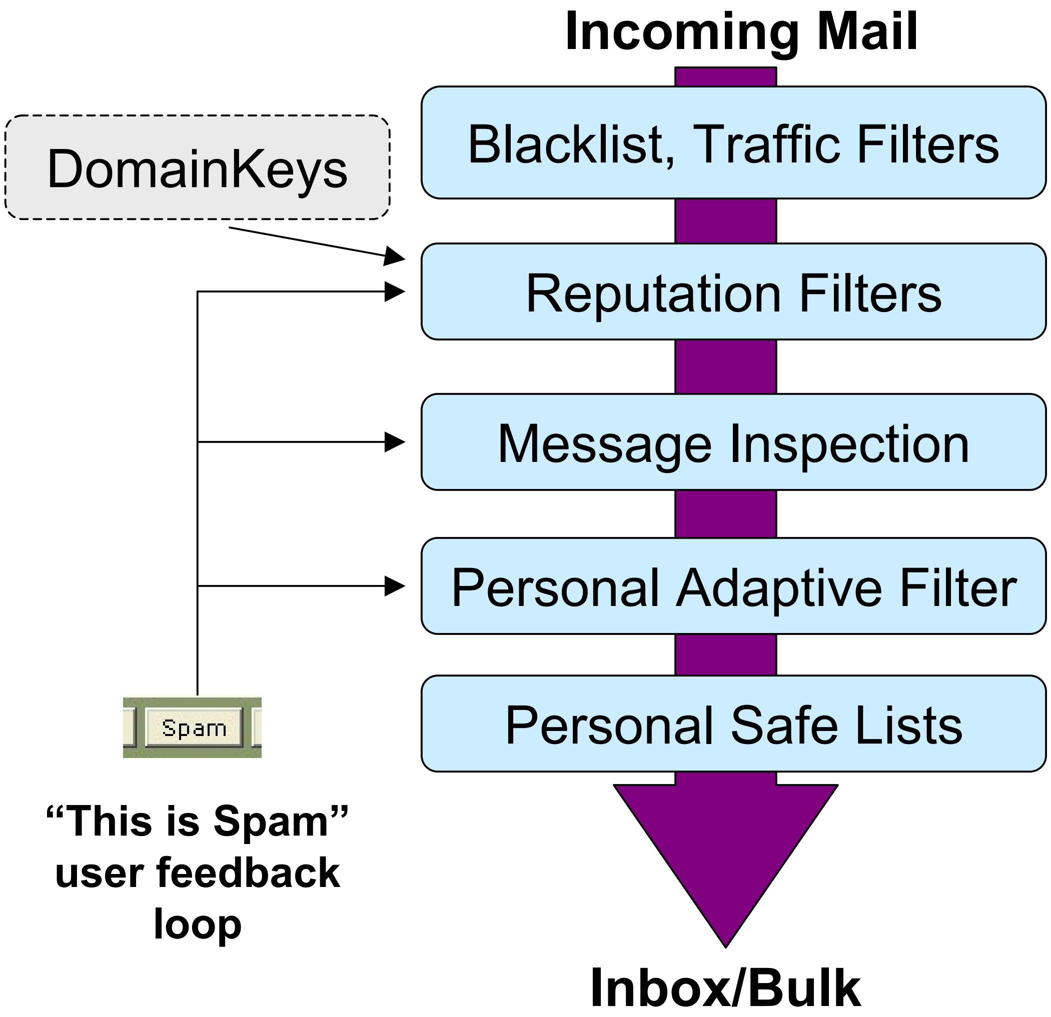
February 17, 2004

http://public.yahoo.com/~miles/nist.pdf

About Yahoo! Mail

- Largest US Web mail system
  - Tens of Millions of users worldwide
- Target rich audience for spammers
  - Novice users
  - Free account can be used as throwaway
  - Spammers can easily test their effectiveness
  - “Namespace” well utilized
- Control over UI and MTA
  - Network knowledge
  - Everyone on same version
  - “This is spam” reports
- Tons and tons of data
  - User spam reports
  - Billions of spam messages each day

Quick Overview of SpamGuard



Incoming Mail

Blacklist, Traffic Filters

Stops dictionary attacks and mail from compromised sources – proxies, trojans, etc.

DomainKeys

Reputation Filters

Flags known good/bad sources

Message Inspection

Analyzes each message for spam/ham characteristics

Personal Adaptive Filter

User-specific spam/ham characteristics

Personal Safe Lists

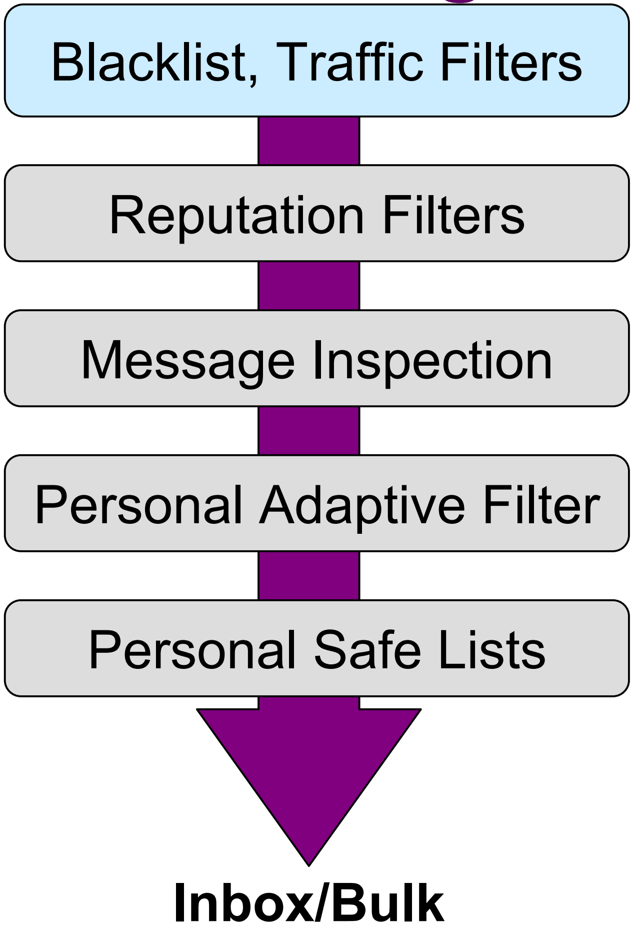
Override spam decisions

Spam

“This is Spam” user feedback loop

Inbox/Bulk

Blacklisting and traffic filters



- Open proxy, other compromised machine detection and blocking
- Traffic filters de-prioritize dictionary attacks, bad mailing list practices, detect traffic abnormalities
- Feedback from traffic filters for compromised candidates
- SMTP error message contains URL to visit to retest
- >1.5M known compromised IPs today
- List still growing as fast as ever
- Most effective in early to mid ‘03

We estimate that spammers have access to more computing resources than the world’s top 25 supercomputers combined

Identity in Email

Problem: lack of authentication in Mail protocols makes it trivial to forge FROM addresses, and as a result:

- We cannot trace spammers by their stated domain
- Users can be fooled into thinking an email is from the wrong party

Currently email identity is IP address

- Maintenance issues – companies constantly add/change IPs
- ESPs – several companies per IP
- Forwarding services break identity
- Spoofing – can only spoof IPs that you control (not an issue)

Possible solutions:

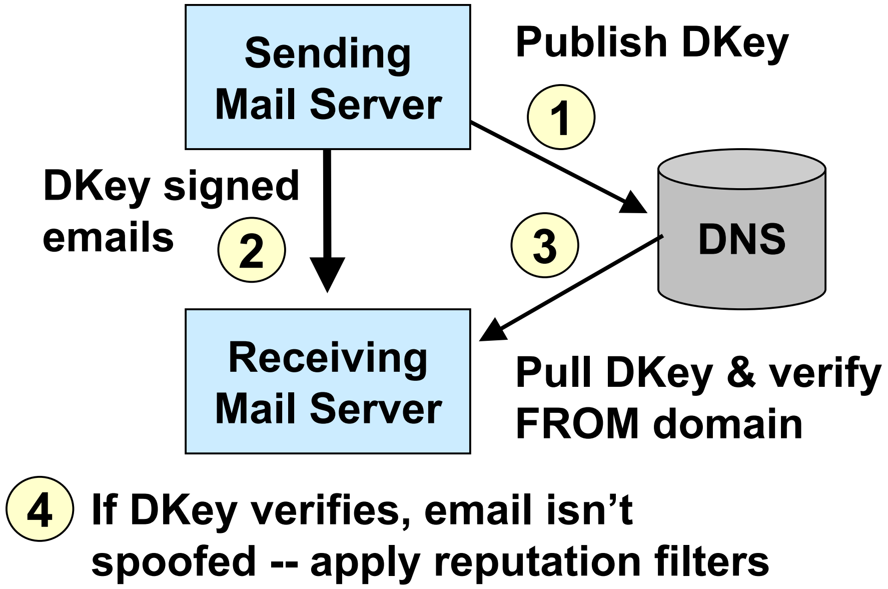
- SPF, and other IP-based solutions (LMAP, DS, RMX, etc.)
- DomainKeys

Identity & DomainKeys

DomainKeys

DomainKeys leverages sending server email signing to begin to solve this problem

- Sign using self-generated keys
- Publish public-key in DNS
- Signature put in headers
- Protect/prove user-view of author's identity



Sending Mail Server

Publish DKey

1

DKey signed emails

2

DNS

3

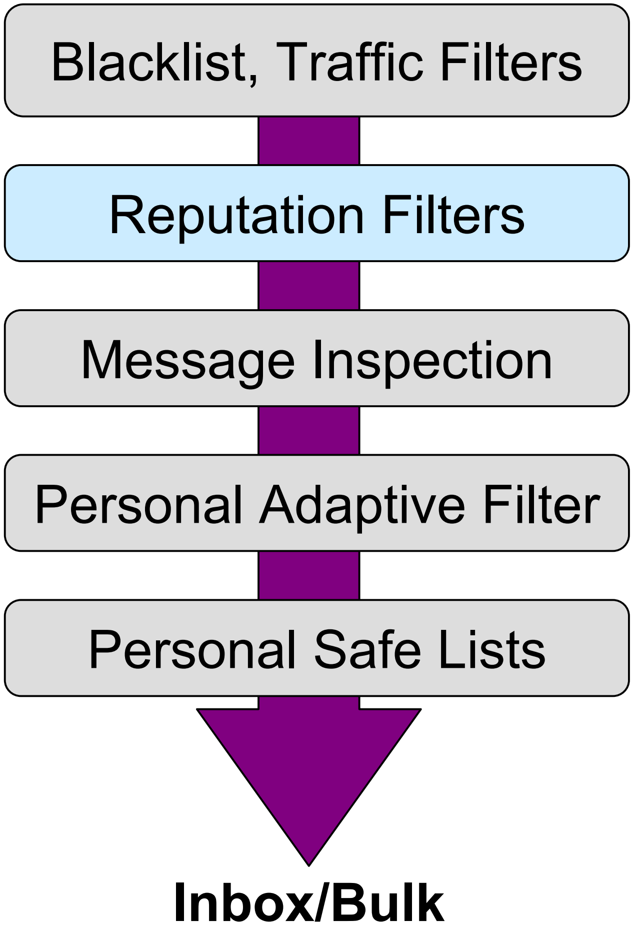
Receiving Mail Server

Pull DKey & verify FROM domain

4 If DKey verifies, email isn't spoofed -- apply reputation filters

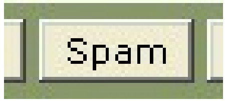
Yahoo! is currently working with major email providers and industry groups to bring DomainKeys to market

Reputation Filters



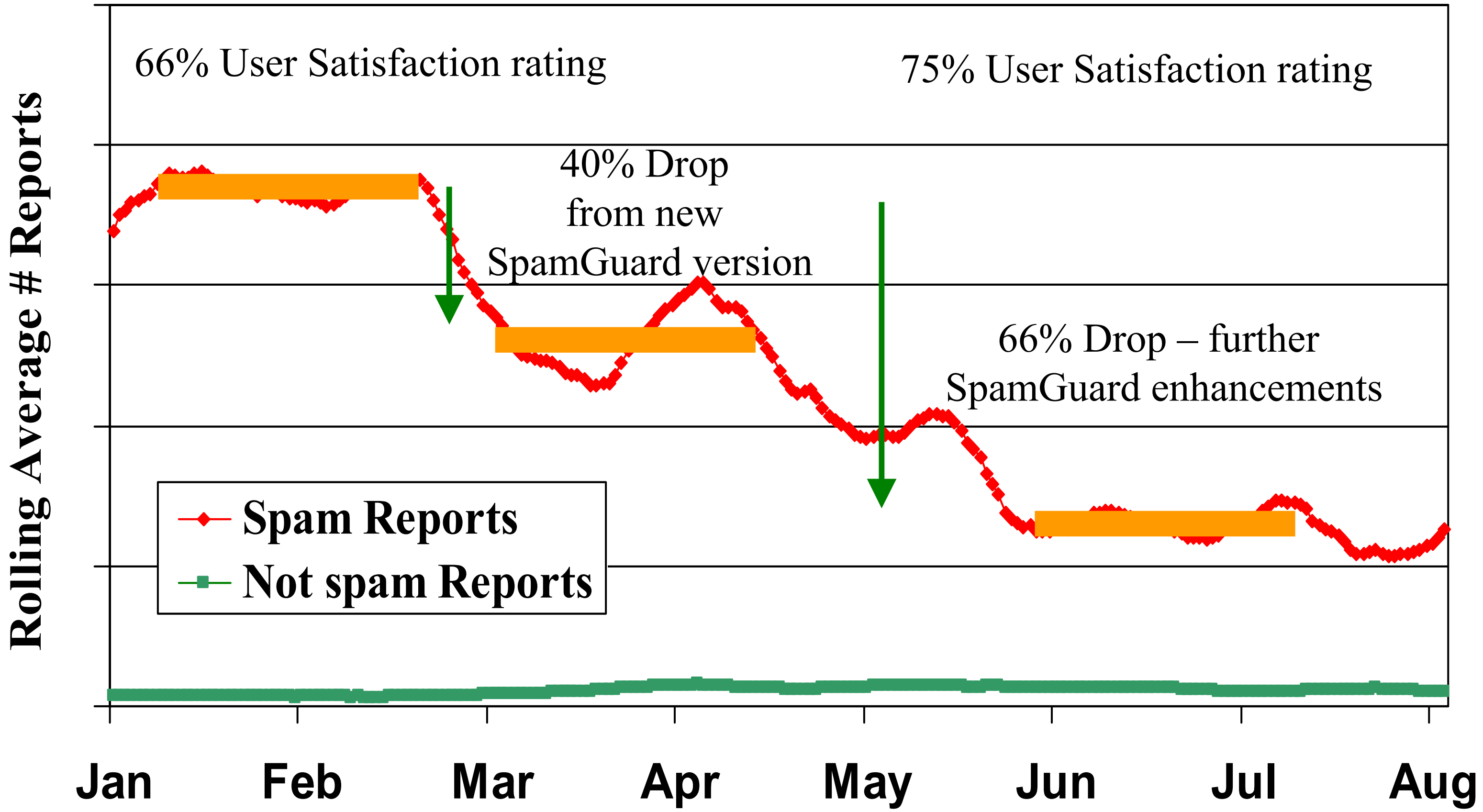
- Constantly measuring user's view of IP's reputation
  - Good – directly to inbox
  - Bad – directly to bulk folder
  - Unknown or Fuzzy – more filters
- Feedback from users
  - This is spam; This is not spam
  - Gaming issues
- Reputation can only be as good as identity system

‘This is spam’ user feedback

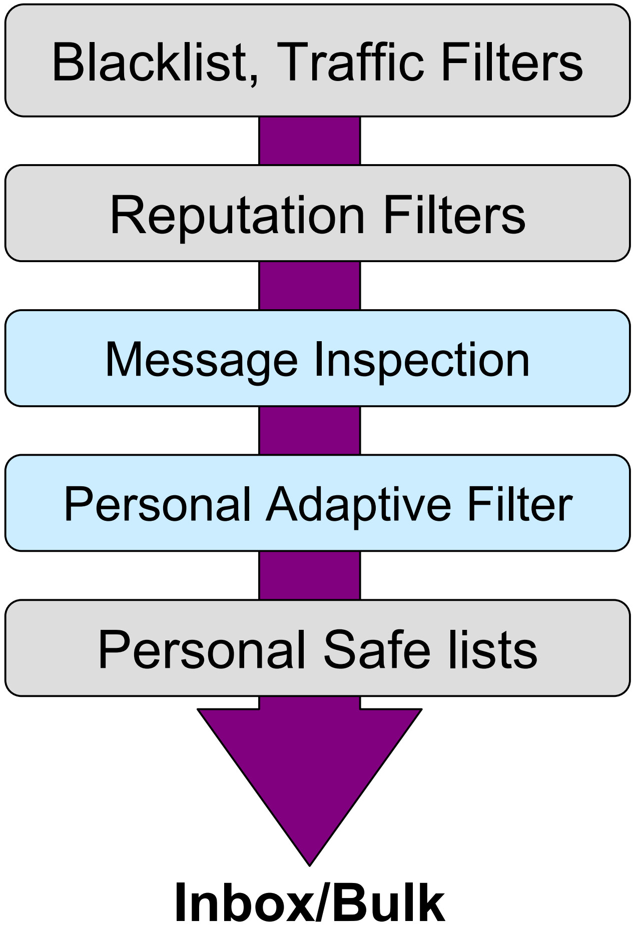


- Real-time feedback from users about anti-spam efficacy
- Anti-spam goal: Minimize total number of reports
- Measure community’s view of IP’s reputation
- Find compromised machines candidates
- Find new spammer tricks

User feedback helps us understand and react to new spammer behavior



Message Inspection



- Designed to automatically detect current spam attacks
- Drop especially virulent viruses (Sobig, myDoom etc)
- Cocktail of approaches
- Personal Adaptive filter adapts to user specific view of spam

Takeaways

- No silver bullet against spam – need cocktail of approaches
- User feedback is immensely helpful
- Spammers rapidly evolve – constant improvement necessary
- Good news:
  - Industry working more closely together than ever before
  - Spammers increasingly have to cross criminal bar (fraud, identity theft, etc.) to enter market

Do You Yahoo!™ ?

http://public.yahoo.com/~miles/nist.pdf