Anti-spam filtering techniques

Stéphane Bortzmeyer
AFNIC (“.fr” registry)
bortzmeyer@nic.fr

ITU, 19 january 2006

Background on this work

This work started in the french Working Group on spam fighting, created by the government. The group includes actors from many sides. I manage the technical subgroup.

Background on this work

This work started in the french Working Group on spam fighting, created by the government. The group includes actors from many sides. I manage the technical subgroup.

This work was then sent to the OECD Task Force on spam (`www.oecd-antispam.org`) and later refined. It will be part of the OECD Anti-Spam Toolkit.

"The OECD Anti-Spam Tookit is a first step in a broader initiative to help policy makers, regulators and industry players orient their policies relating to spam solutions and restore trust in the Internet and e-mail."

Background on this work

This work started in the french Working Group on spam fighting, created by the government. The group includes actors from many sides. I manage the technical subgroup.

This work was then sent to the OECD Task Force on spam (`www.oecd-antispam.org`) and later refined. It will be part of the OECD Anti-Spam Toolkit.

The final french report should be out this month. This talk is mostly a summary of the technical part of the report.

The work and its results

We focused on **incoming** spam. Outgoing spam is a different problem (to be addressed later).

The work and its results

We focused on **incoming** spam. Outgoing spam is a different problem (to be addressed later).

The emphasis is on **practical** advices. We care for the poor system administrator, overwhelmed by work and who is asked to fight spam too.

The work and its results

We focused on **incoming** spam. Outgoing spam is a different problem (to be addressed later).

The emphasis is on **practical** advices. We care for the poor system administrator, overwhelmed by work and who is asked to fight spam too.

There are many anti-spam solutions. We keep only the good ones.

The work and its results

We focused on **incoming** spam. Outgoing spam is a different problem (to be addressed later).

The emphasis is on **practical** advices. We care for the poor system administrator, overwhelmed by work and who is asked to fight spam too.

There are many anti-spam solutions. We keep only the good ones.

We fully recognize that spam requires a multi-thing approach: technical, legal and social solutions are necessary and discussed in the full report. We just focus in this talk on the technical part.

Background on security

Spam is just a branch of the vast domain of **network security**. It raises exactly the same issues as other security problems: tradeoffs between efficiency and cost, collateral damages, mix of technical and social issues, etc.

Background on security

Spam is just a branch of the vast domain of **network security**. It raises exactly the same issues as other security problems: tradeoffs between efficiency and cost, collateral damages, mix of technical and social issues, etc.

Security is a process, not a product

Do not think you will be safe because you bought Anti-Spam Platinum Gold 3.0

Background on security

Spam is just a branch of the vast domain of **network security**. It raises exactly the same issues as other security problems: tradeoffs between efficiency and cost, collateral damages, mix of technical and social issues, etc.

Security is a process, not a product

Do not think you will be safe because you bought Anti-Spam Platinum Gold 3.0

Security is a tradeoff

Yes, we could suppress **all** the spam. Just shut off the computers.

Background on security

Spam is just a branch of the vast domain of **network security**. It raises exactly the same issues as other security problems: tradeoffs between efficiency and cost, collateral damages, mix of technical and social issues, etc.

Security is a process, not a product

Do not think you will be safe because you bought Anti-Spam Platinum Gold 3.0

Security is a tradeoff

Yes, we could suppress **all** the spam. Just shut off the computers.

In the real world, the question is not "How to suppress spam?" but "How to limit spam without killing email?"

The perfect solution

The perfect solution

1. No false negative: catches all the spam

The perfect solution

1. No false negative
2. No false positive: catches only the spam

The perfect solution

1. No false negative
2. No false positive
3. Low cost: small price, runs on small computers, blocks the spam before transmission, saving bandwidth

The perfect solution

1. No false negative
2. No false positive
3. Low cost
4. No change: installs on the current systems, require no change in habits

The perfect solution

1. No false negative
2. No false positive
3. Low cost
4. No change
5. Based on open standards: no vendor lock-in, ability to understand what it does, preferably free (as in free speech) software

The perfect solution

1. No false negative
2. No false positive
3. Low cost
4. No change
5. Based on open standards

This solution does not exist

But this checklist is a good method to evaluate imperfect solutions.

State of the art

Today, almost no spam gets in the default mailbox, if everything is configured with state-of-the-art tools.

State of the art

Today, almost no spam gets in the default mailbox, if everything is configured with state-of-the-art tools.

Yes, today, the spam never reaches users

But it has a cost: computers, bandwidth, engineers. And it requires to use tools choosen by the experts, not snake-oil sold by salesmen.

State of the art

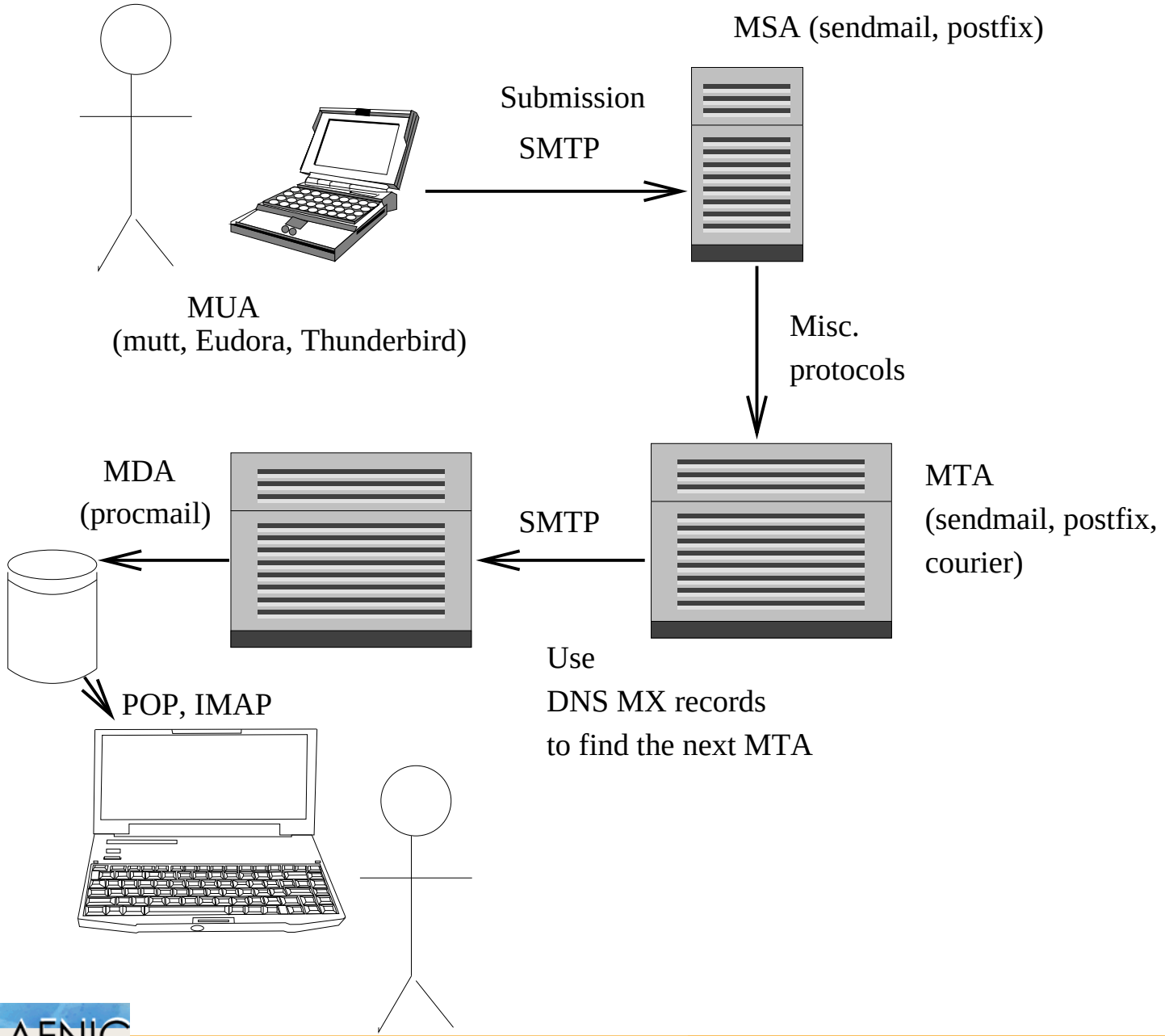
Today, almost no spam gets in the default mailbox, if everything is configured with state-of-the-art tools.

Yes, today, the spam never reaches users

But it has a cost: computers, bandwidth, engineers. And it requires to use tools choosen by the experts, not snake-oil sold by salesmen.

Also, we cannot be sure it will stay that way in the future: the research must go on.

Reminder: email architecture



Good practices

Good practices

1. Greylisting `http://www.greylisting.org/`

Good practices

1. Greylisting
2. Heuristic filters, with scores computed against real spam and ham $\Rightarrow$ SpamAssassin `http://www.spamassassin.org/`

Good practices

1. Greylisting
2. Heuristic filters, with scores computed against real spam and ham $\Rightarrow$ SpamAssassin
3. Bayesian filters on the user's desktop like bogofilter
   `http://www.bogofilter.org/`

Good practices

1. Greylisting
2. Heuristic filters, with scores computed against real spam and ham $\Rightarrow$ SpamAssassin
3. Bayesian filters on the user's desktop

This group of three kills almost all the spam with very little false positives. It "just" requires big machines (spam is a big plague in the countries of the South).

Greylisting

Deliberately returns a **temporary** error when receiving email from a new machine.

The typical spammer software does not retry. A legitimate MTA does.

Surprisingly effective and very simple to deploy.

Heuristic filters

Starts from spamicity tests:

- Attempts to disguise the word 'viagra'
- HTML has very strong "shouting" markup
- Claims you can opt-out
- ...

A score is then computed **automatically** (humans can be wrong on the spamicity of something) for each test.

Tests are applied to the message and a total "spam score" is produced.

The better known one is SpamAssassin, from the Apache Foundation. Many anti-spam appliances use it.

Bayesian filters

Starts from nothing: they have no prejudice.

Human users trains the filter by giving it spam and ham.

The filter learns the vocabulary of each.

Then, it can calculate a spam score for the message, using Bayes statistics.

Proper training is important

So there is a user interface and user training issue.

They are the most efficient filters today. But no solution is perfect alone: you need to combine several techniques.

bogofilter in practice

Analysis of one message with `bogofilter -v` :

```
                          n    pgood     pbad      fw
"$99"                     8  0.000275  0.002554  0.901823
```

pbad = "spamicity". fw = probability of being a spam. The string 99$ is a good spam mark.

You can also display the whole database with `bogoutil` :

```
Viagra 88 0 20041116
```

Viagra was in 88 spams and no ham (a doctor would have different results: this is my personal database).

Less good practices

There are other methods but either unrealistic, harmful or questionable. Since we emphasize **practical** advices, we mention shortly these techniques in the report.

Not specific technique mentioned here, to be reserved for discussion.

The issue of authentication

Authentication is not an anti-spam technique by itself (spammer can have a passport, too).

But it may help:

1. Better accountability may deter spammers,
2. Whitelisting cannot work without authentication.

Authentication techniques

A lot of unsolved questions: what to authenticate? (Which identity?)

Also, there is no common identity service (the Internet has no government).

Authentication techniques

A lot of unsolved questions: what to authenticate? (Which identity?)

Also, there is no common identity service (the Internet has no government).

1. SPF, Sender Policy Framework: the sender indicates in the DNS which machines can send mail on its behalf. Since it is the DNS, it authenticates a domain, not an user.

Authentication techniques

A lot of unsolved questions: what to authenticate? (Which identity?)

Also, there is no common identity service (the Internet has no government).

1. SPF, Sender Policy Framework
2. DKIM, Domain Keys Identified Mail, IETF Working Group security/dkim: the sender cryptographically signs the headers. The key is typically a domain key, not an user key. The key can be retrieved by various means, including the DNS.

Authentication techniques

A lot of unsolved questions: what to authenticate? (Which identity?)

Also, there is no common identity service (the Internet has no government).

1. SPF, Sender Policy Framework
2. DKIM, Domain Keys Identified Mail, IETF Working Group security/dkim
3. PGP, Pretty Good Privacy: very good system for user authentication, only deployed in limited communities.

What to do with spam?

Once it is detected by the heuristic filter or the bayesian filter, what do we do with spam?

Technical hint: rejection during the SMTP session is an interesting solution because you never took responsability for the mail. Not always easy to do and has its own problems.

What to do with spam?

Once it is detected by the heuristic filter or the bayesian filter, what do we do with spam?

1. Tell the sender? **No**, most spams are joe jobs (the address is forged).

Technical hint: rejection during the SMTP session is an interesting solution because you never took responsability for the mail. Not always easy to do and has its own problems.

What to do with spam?

Once it is detected by the heuristic filter or the bayesian filter, what do we do with spam?

1. Tell the sender? **No**, most spams are joe jobs (the address is forged).
2. Drop silently? Harsh but probably necessary, if the end user accepted it.

Technical hint: rejection during the SMTP session is an interesting solution because you never took responsability for the mail. Not always easy to do and has its own problems.

What to do with spam?

Once it is detected by the heuristic filter or the bayesian filter, what do we do with spam?

1. Tell the sender? **No**, most spams are joe jobs (the address is forged).
2. Drop silently? Harsh but probably necessary, if the end user accepted it.
3. File in a spam mailbox: probably the best default solution. Copies are a good idea, specially at the beginning, because they allow later screening.

Technical hint: rejection during the SMTP session is an interesting solution because you never took responsability for the mail. Not always easy to do and has its own problems.

A few hints about outgoing spam

(Not part of the report)

To fight spam as its source:

A few hints about outgoing spam

(Not part of the report)

To fight spam as its source:

1. Try to stop MS-Windows machines to be recruited as **zombies**. “A botnet is comparable to compulsory military service for Windows boxes.” (Stromberg)

A few hints about outgoing spam

(Not part of the report)

To fight spam as its source:

1. Try to stop MS-Windows machines to be recruited as **zombies**. "A botnet is comparable to compulsory military service for Windows boxes." (Stromberg)
2. Rate-limit outgoing mail (but you need exemptions because some users host mailing lists) and/or block outgoing SMTP (you also need exemptions or you are no longer an Internet access provider).

The problem of collateral damages

Like medical drugs...

...anti-spam solutions have secondary effects. You can limit them, you cannot suppress them (do not believe the ads).

The problem of collateral damages

Like medical drugs...

...anti-spam solutions have secondary effects. You can limit them, you cannot suppress them (do not believe the ads).

1. False positives: legitimate messages are refused

The problem of collateral damages

Like medical drugs...

...anti-spam solutions have secondary effects. You can limit them, you cannot suppress them (do not believe the ads).

1. False positives: legitimate messages are refused
2. High human and machine costs (fighting the spam is a full-time job for some)

The problem of collateral damages

Like medical drugs...

...anti-spam solutions have secondary effects. You can limit them, you cannot suppress them (do not believe the ads).

1. False positives: legitimate messages are refused
2. High human and machine costs (fighting the spam is a full-time job for some)
3. Loss of trust: some people switch away from email

The problem of collateral damages

Like medical drugs...

...anti-spam solutions have secondary effects. You can limit them, you cannot suppress them (do not believe the ads).

1. False positives: legitimate messages are refused
2. High human and machine costs (fighting the spam is a full-time job for some)
3. Loss of trust: some people switch away from email
4. Loss of freedom: more filters, more rules, less authorized things

Deeper changes ahead

Deeper changes ahead

The anti-spam struggle sometimes change the architecture of the Internet in a bad way.

Deeper changes ahead

The anti-spam struggle sometimes change the architecture of the Internet in a bad way.

The Internet is a continuum between the operators and the end-users, with all sort of people in-between. There is room for the ordinary user, the savvy user, the university, the bank, the big operator in a northern country, the small operator in Africa...

Deeper changes ahead

The anti-spam struggle sometimes change the architecture of the Internet in a bad way.

The Internet is a continuum between the operators and the end-users, with all sort of people in-between. There is room for the ordinary user, the savvy user, the university, the bank, the big operator in a northern country, the small operator in Africa...

Some anti-spam proposals try to make it a binary network: only operators (all from the North) and end-users, pure consumers.