

OECD - workshop on SPAMMING
2-3 February 2004



**Measuring SPAM:
a Government perspective**

Direction du Développement des Médias
M. Aufrant - J.M. Nivlet

The case of SPAMMING



- The DDM is a Prime Minister office department in Paris, its duty is to set the French regulatory framework on mass media activities such as : the press, radio and television. It is also in charge of questions related to content in Information Society.
- A new mission entrusted to the DDM by the French Government Commission for Information Society.

What is spamming about ?

- Spamming is condemned as **a nuisance** by internauts.
- Spamming is condemned as **a nuisance** by advertisers and electronic marketing businesses.
- Spamming is condemned as **a nuisance** by Internet Access Providers.
- But spamming also is an economic opportunity: **the anti-spamming business.**

A draft review



- Many professions and organisations concerned by spamming have published figures and surveys: anti-spam solution publishers, Internet Service Providers, consultants or marketing businesses, specialists in privacy protection, non-profit organizations, the authorities (FTC in the US, CNIL in France...)
- So far, about **thirty papers** have been identified from the Internet.

The intensity of spamming (1)

- **Spamming boxes**, that of the FTC, of the Belgium Commission for Privacy, of the CNIL, count absolute numbers of spams; they indicate flows in **ten-thousands or hundred-thousands of spams**.
- But in comparison to the total flow of received E-mails, these fluxes only represent **0.002%** of total E-mail in the US and **0.050%** in France.
- 2 spams out of 100 000 E-mails are denounced in the US; 50 spams out of 100 000 E-mails are denounced in France.

The intensity of spamming (2)

- All other assessments (made using automatic spam filters or using surveys on internauts' opinion) directly **measure the proportion of spams in received E-mails.**
- Presently, results concentrate around the 50% level, with extreme estimates ranging from 30% to 75% in the US:
estimates leading to **billions of spams per day.**

The statistical difficulty



- Between estimates made using spam boxes and those made using automatic filters or opinion surveys, the gap is astronomic: the order of magnitude in their difference is in a factor 1 to 100 000.
- Our aim: to understand the origins of that gap.
- In average, internauts would not complain for more than **one e-mail out of hundred thousand** received? Such a result then requires to seriously question the reality of the nuisance and identify all the characteristics of the nuisance measurement.

The nature of nuisance

- **Under reporting:** on that matter, the impact of the new legal framework will have to be assessed.
- The **tangible** nuisance versus the **individual feeling** of nuisance: this subject has a **quantitative aspect** (counting E-mails and measuring their characteristics), as well as a **qualitative aspect** (opinion, perception and behaviour when facing a nuisance).
- Whether at home or at work, the population of internauts is **not equally** exposed to **the risk** of nuisance.

Present insufficiencies



- A lack of objective and tangible elements: we still don't know the content composition of internauts' mail boxes.

- **"Black-and-white" approaches.**

A standardised spam survey protocol:

How many E-mails do you usually receive per day?

What is the usual proportion of spams?

A two-stage survey on spamming



- 1. Sampling of internet service subscribers.
- 2. Then : **statistical analysis of E-mails** contained in the sampled internaut's box.

E-mail classification



Personal message; professional message;
B to B e-mailing; automatic reply, including
order confirmation and tracking; UCE or spam;
canvassing or fidelity marketing message;
newsletter message; *mailing list* message;
news group message.

Personal and non personal E-mail



- The second stage of the survey would exclude personal E-mail which in terms cannot be considered as spam.
- **Non personal E-mail:**
 - *marketing E-mail;*
 - *newsletters E-mail;*
 - *other non personal E-mail,*
in which " illicit " mail (spams).
- OK if at home, to think about if at work place.

Possible topics to be surveyed



- E-mail equipment.
- Composition and the processing of E-mail.
- Commercial transactions.
- Communication and Mass media.
- Awareness of the spam threat.

A steering group



- Identify the needs; validate the method and the analysis of data.
- Open it to representatives of government agencies and business associations.
- A more ambitious objective : a "benchmark survey" intending to synthesize private and public approaches.