Advanced Filtering

Tobias Eggendorfer

Advanced Filtering Fails Too

Tobias Eggendorfer

Overview

- Not so advanced Filtering
- Advanced Filtering
- Prevention
- Identification

„Classic“ Filtering

„Classic“ Filtering

- Black- & Whitelists

„Classic“ Filtering

- Black- & Whitelists

„Classic“ Filtering

- Black- & Whitelists

„Classic“ Filtering

- Black- & Whitelists
- Simple Content Analysis
  - Bad Word Filter

Workarounds

- IP-Blacklists
- Open Relay
- Open Proxy
- Bot-Net

Bad Word Lists

Bad Word Lists

- VIAGRA becomes \/ I @ 6 R /-\

Bad Word

- VIAGRA becomes \/ I @ 6 R /-\

Bad Word Lists

- VIAGRA becomes \/ I @ 6 R /-\
- Favourite

B	U	Y			
V	I	A	G	R	A
C	H	E	A	P	

CAUTION!

THIS MACHINE HAS NO BRAIN USE YOUR OWN

Advanced Filtering

- Spammers learn their workarounds
- We learn how to work around their workarounds
- They learn how to work around our workarounds for their workarounds
- ...

What‘s on the menu?

- Web-Bug-Analysis
- URL-Blacklisting
- OCR for Pictures
- Picture-Analysis
- Greylisting
- ...

Heuristics

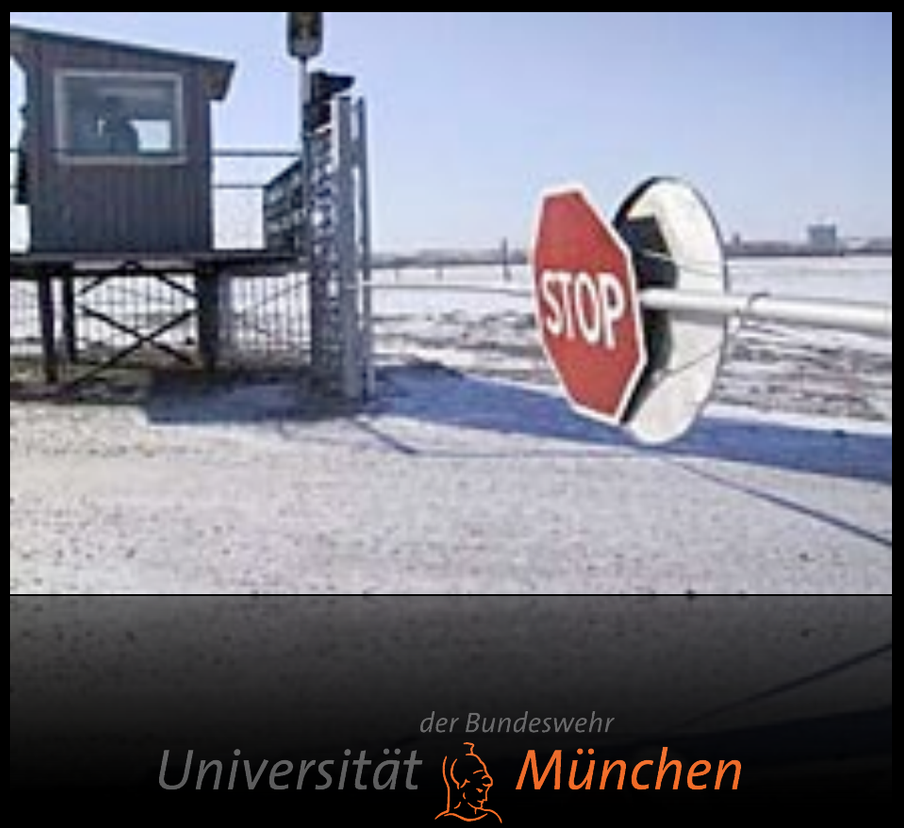
- Put the results together
- Stir them
- Taste it

Heuristics

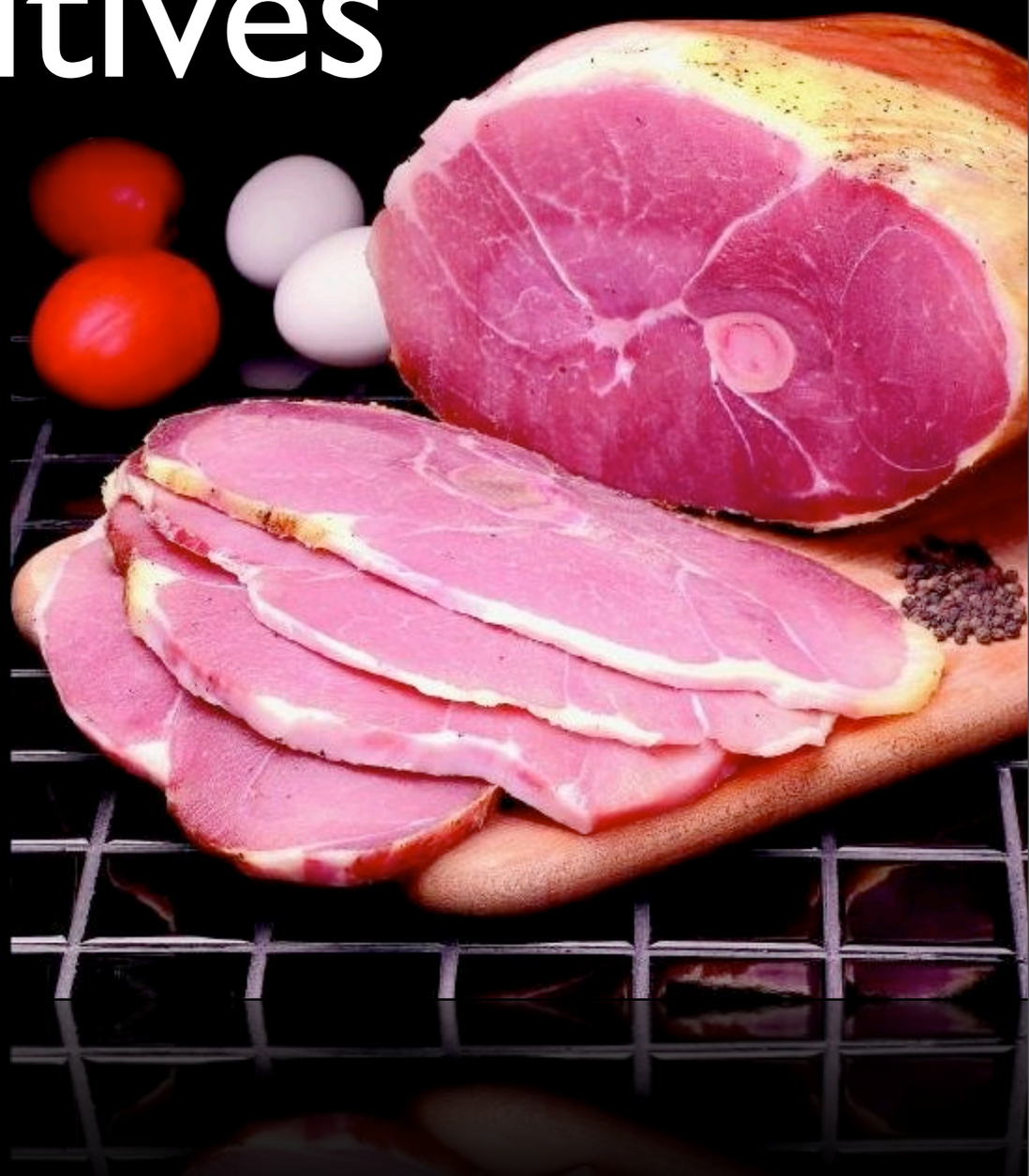
- Put the results toge
- Stir them
- Taste it

Spam filters‘ limits

- Heuristics = qualified guess
- Guess != knowledge
- Guess → false positives & false negatives



False Positives



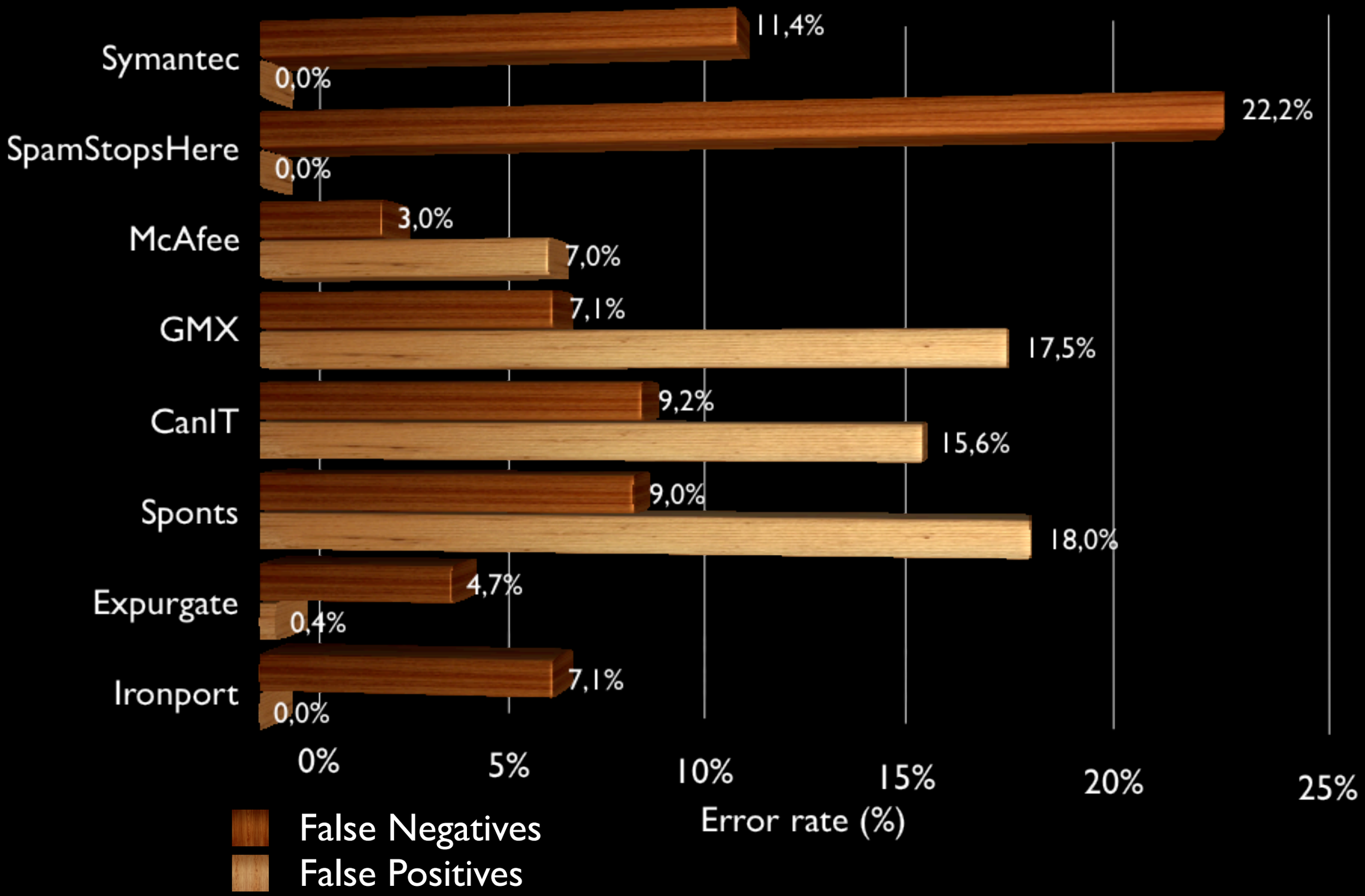
- Important message
  - deleted,
  - denied or
  - delayed
- Loss of revenue
- Liability for not-served contract

False Negatives

- Spam mistaken for ham
  - → manual filtering based on
    - Subject
    - Body
- Human false positives

Usual ratio

- High rate of false negatives
  → „relaxed“ filter
  → low rate of false positives
- Low rate of false negatives
  → more aggressive filtering
  → high rate of false positives


Symantec
11,4%
0,0%
SpamStopsHere
22,2%
0,0%
McAfee
3,0%
7,0%
GMX
7,1%
17,5%
CanIT
9,2%
15,6%
Sponts
9,0%
18,0%
Expurgate
4,7%
0,4%
Ironport
7,1%
0,0%
0%
5%
10%
15%
20%
25%
Error rate (%)
False Negatives
False Positives

The truth is...

- Known ratio is true for
  - Blacklists
  - Content-Filters
- Comparative filters break the rule

But...

- Newsletters?
- Multiple recipients
- Some lost interest
  ➔ mark it as spam

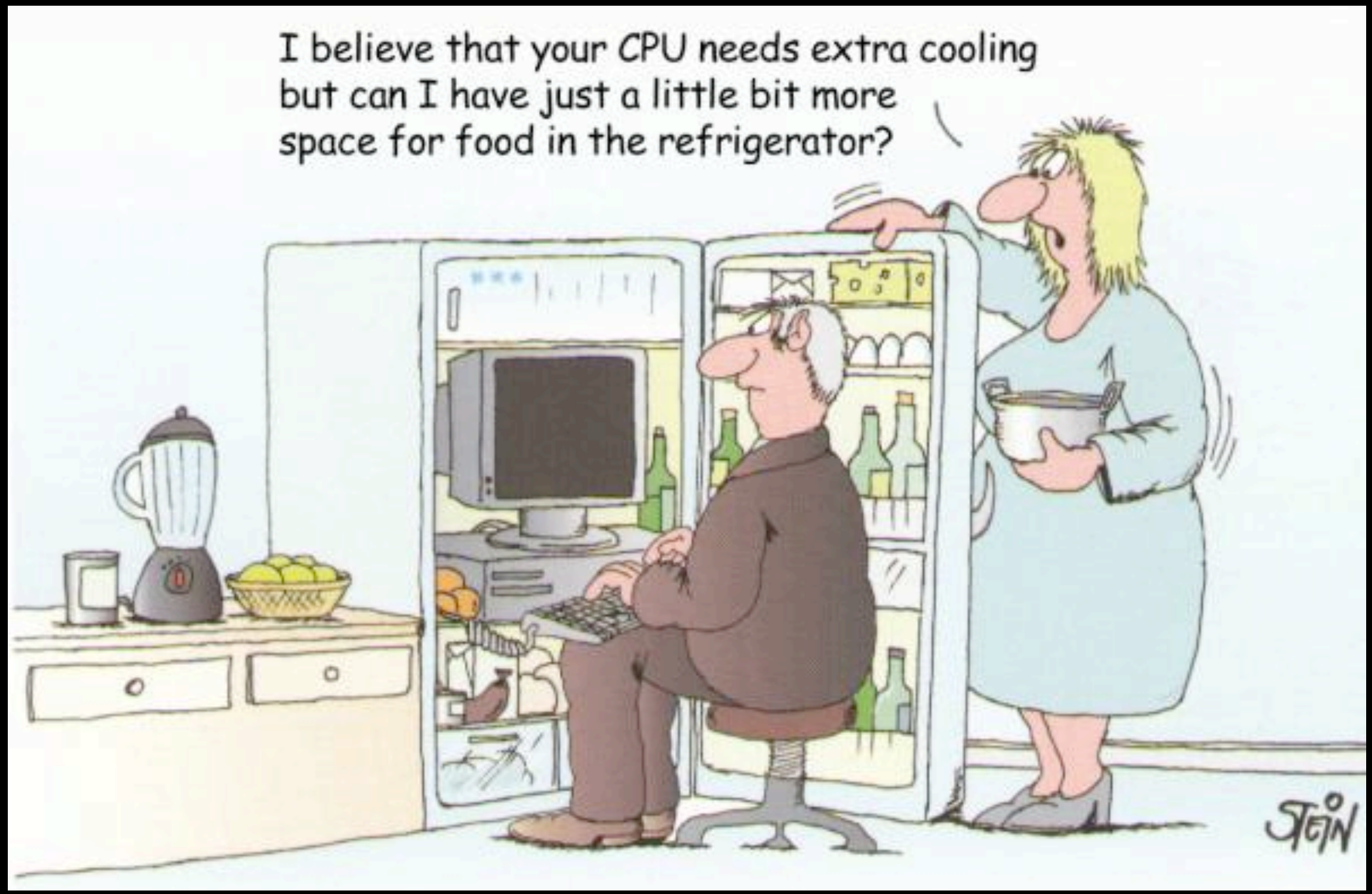
Whitelist them!

- Whitelist all mailinglists?
- I run a few
  - Who knows them?
  - Who would whitelist them?
  - Why would they whitelist me?

Computing power wasted

- Dual Xeon, 4 GB RAM
- For a spam filter?

DoS? Bandwidth?

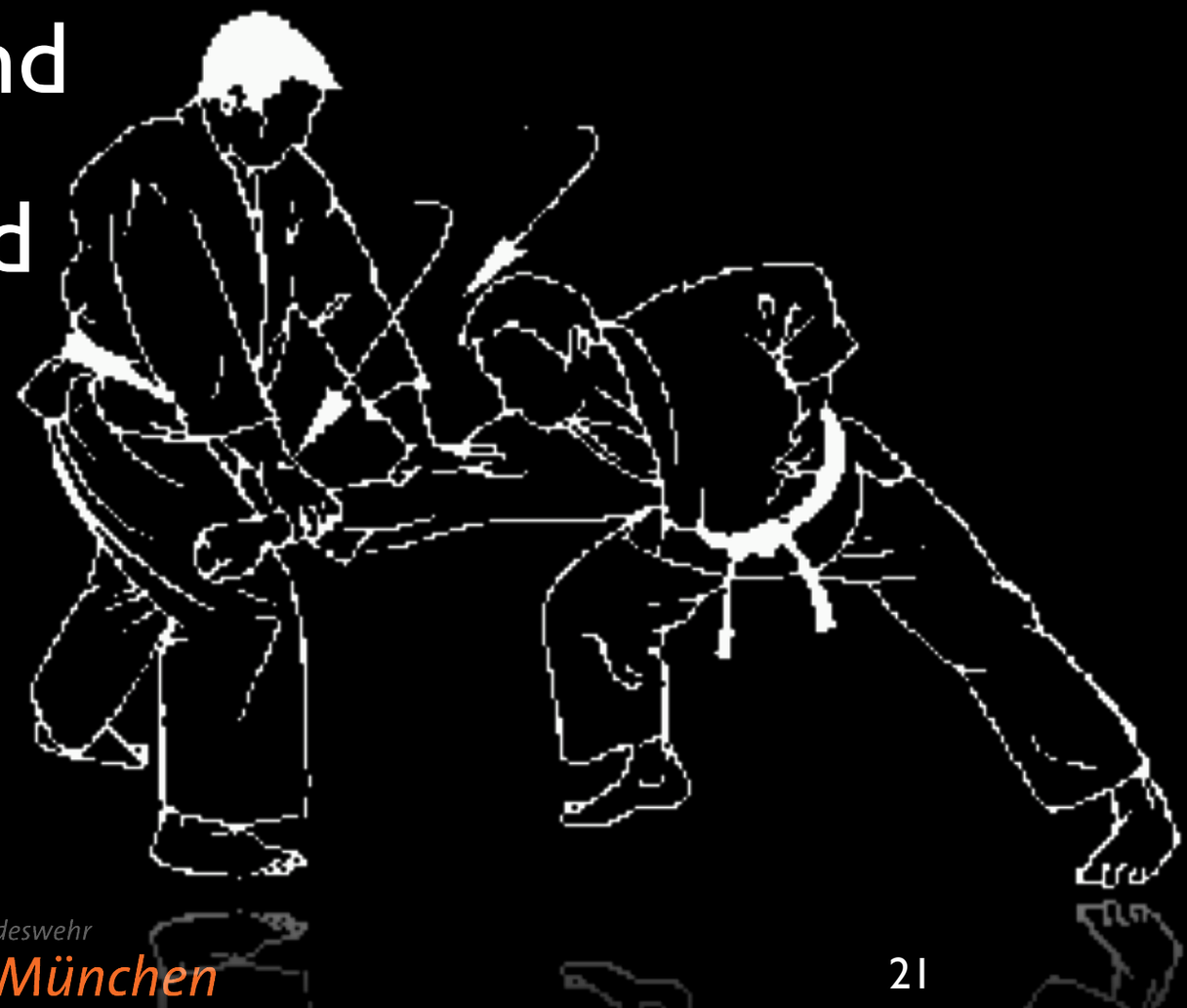


Filtering fails

Spam filtering
is
Russian Roulette
for
communication

Why?

- It‘s reaction.
- Reaction = step behind
- Action = take the lead



Prevention

- Economic approach
- Prevent address collection

It‘s about money

- When would you take a risk?
  - Rewards / Risk relation
  - Money / Risk relation

What risk would you accept for

What risk would you accept for

and for

There might be other motivators...

Some things money can‘t buy.

Earnings

- 5 000 - 10 000 US $ per day
- 25 000 - 50 000 US $ per week
- 1 000 000 - 2 000 000 US $ per year (incl. a few weeks off)

Risks for spammers

- Major threads
  - Jail
  - Liability

Money / Risk Ratio

- Risk of being caught
  - Using bot nets
  - Anonymous payment
  - Intangible goods
- Very low

That is...

- Increase the risk
- But how to?
  - Anti spam forensics are very limited
  - A few ideas later...

No email ➔ No spam

- Prevent address collection
- Obfuscation of addresses
- HTTP tar pits

No brute force

- Address testing with SMTP

RCPT TO: <user@example.com>
250 Recipient ok

No brute force

- Address testing with SMTP

RCPT TO: <user@example.com>
250 Recipient ok

How to prevent this?

- Lie: User unknown
- But when?

Anti-Brute-Force

- Max. 3 bad PINs
- Card withdrawal
- Delay

But in real life

- No legitimate bad attempts
  - at least not 10‘000
- Card withdrawal is final
  - IP block is not
- One ATM & one card
  - No parallel waiting

If there is a delay...

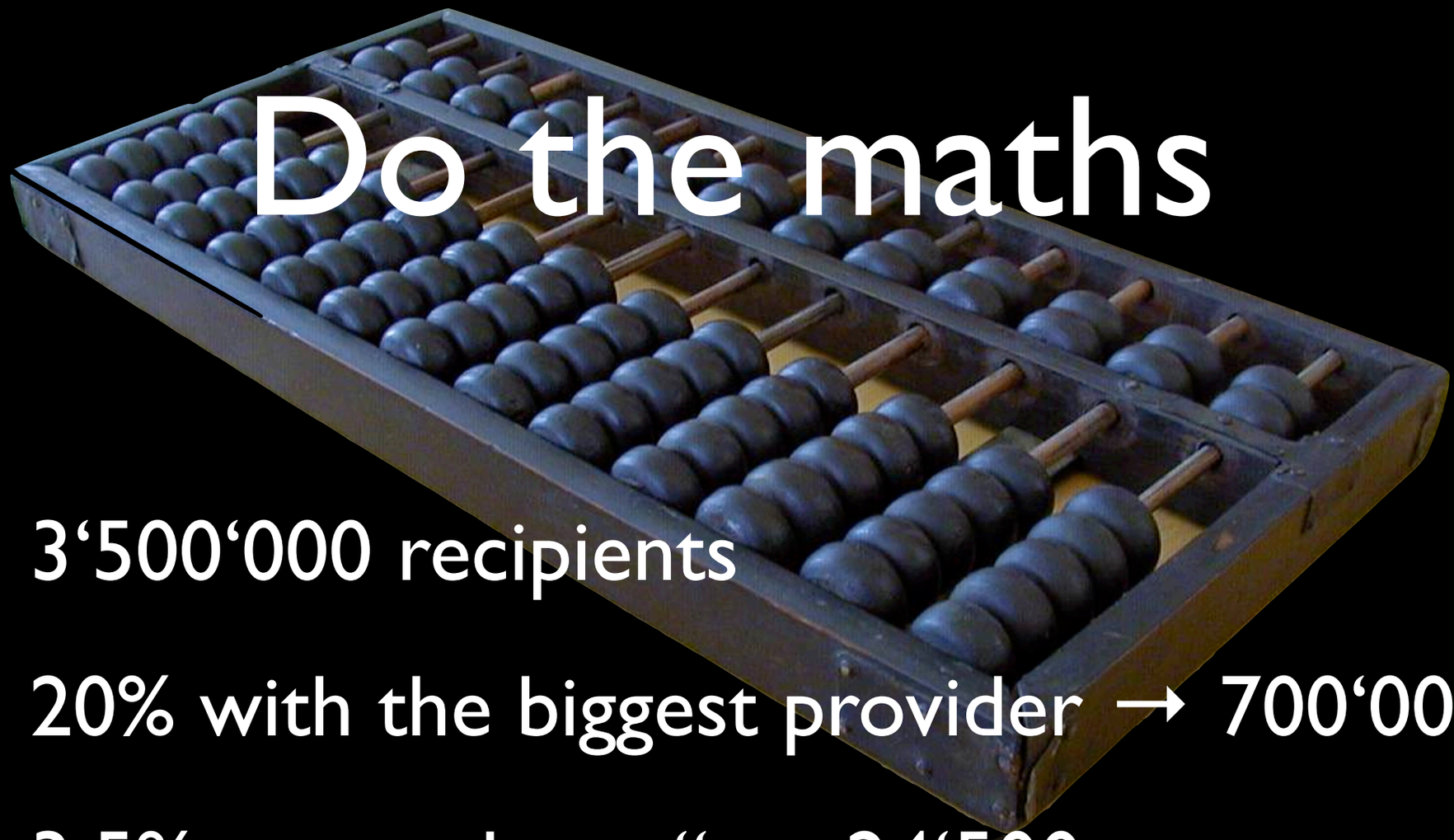
- ... delay the connection setup too
- SMTP_GREET_DELAY with sendmail
- Greet delay > bad response delay

Limit recipients?

- MAX_RCPTS_PER_MESSAGE in sendmail
- Think large
- 3.5 million recipients for a newsletter
  20% with the biggest provider
- You really want a limit?

Constant delay?

- BAD_RCPT_THROTTLE in sendmail
- 1 second delay each
- 3.5 million recipients
  3.5% of those are bad
- Feasible?



Do the maths

- 3‘500‘000 recipients
- 20% with the biggest provider → 700‘000
- 3.5% are „unkown“ → 24‘500
- 1 second delay per recipient → 8 days

Be flexible

- $t_{delay} = \lambda * (bad / total) * t_{wait}$
- Why?
  - Spam: (bad / total) ≈ 1
  - Others: (bad / total) ≈ 0

Prevention‘s Downsides

- Like vaccacination
- Takes time
- No immediate effect
- 50% less spam after 6 month

Immediate pain relief

- SMTP tar pit simulator
- 80% less spam
- No false positives yet

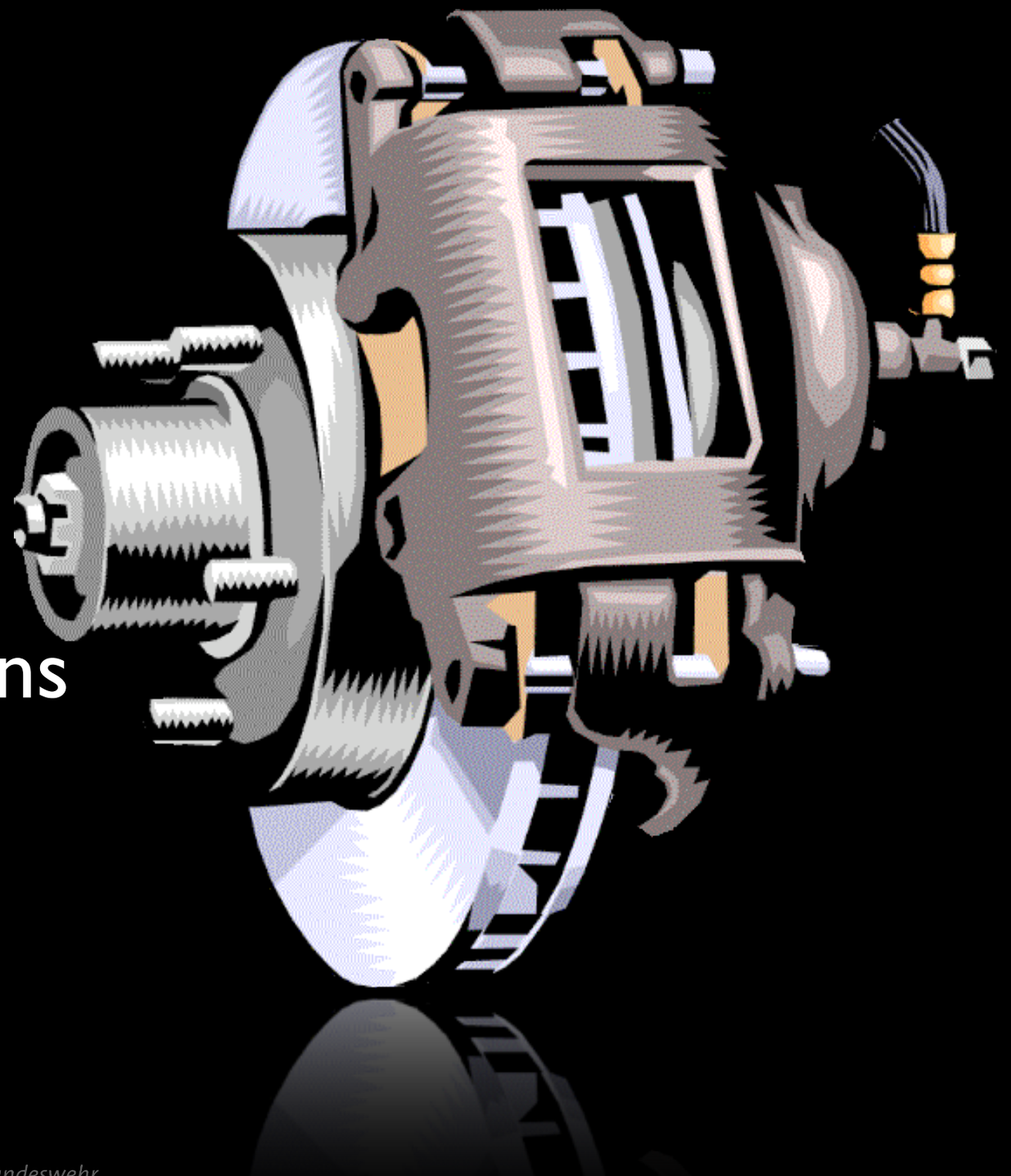
SMTP tar pit

SMTP tar pit

```
220 mail.example.com ready
EHLO spammer.com
250-Welcome spammer.com
250-We do not like spam here.
250-Take your time to read this
250-Commands available
250-...
250 RCPT
```

On its own: useless

- Idea: Delay spam run
- But:
  - Multiple connections
  - Terse time outs



What happens?

- Spammer disconnects quickly
- Addresses not tested
- Spam not delivered
- Maybe: Server blacklisted with spammers

What the simulator does:

What the simulator does:

- Stutter n bytes

What the simulator does:

- Stutter n bytes
- Then: Open up to full speed
- Filters: 80%

Details

- Bulkmailer disconnects after 60 < n < 120 bytes
- Best delay: 1 sec

Optimisation

- Random stutter bytes
  $70 < n < 200$
- Random wait time
  $0.75 \text{ sec} < t < 1.5 \text{ sec}$
- Remember wait time / stutter bytes per IP

Work around?

- Wait longer
  But how long?
- Ignore tar pits
  But be trapped then...

Hunting them down

Jagdszene

The business

- Participants
  - Spammer
  - Vendor
  - Address-Vendor
  - Rent-a-Bot
  - Bullet-Proof Hoster

The process

- Address acquisition
- Send spam
- Product provisioning
- Online Shops & Payment
- Product delivery

Known methods

- Analyse email, identify sender
- Observe bot nets
- Who buys what
- Hoster
- Payment process

Email analyses

- Most headers are forged
- Most traces lost in bot nets
- Low quality proof
  (Investigator might have forged it)

Bot net observation

- Observe a zombie
  - Who controls it?
  - Who rents it?
- Problem:
  - Anonymous usage
  - Security issues

Buying products

- How to supervise:
  - Intangible goods
  - Faked products
  - Generic products
- Stock spam is easy

Buying products

- How to supervise:
  - Intangible goods
  - Faked products
  - Generic products
- Stock spam is easy

New concept

- Identify address traders
- Increase their risk
- Increase prices for addresses
- Reduce possible profit
- Change rewards / risk ratio

Identify address collectors

- Publish traceable email addresses:
  192.0.2.15.20080124123000@example.com
  C000020F124344AA5778@example.com
- Disadvantages:
  - Time
  - International IP tracking

In flagranti

- Identify the harvester in flagranti
- Idea:
  - Distributed HTTP tar pit network

Tar pit network

- Shares
  - Client-IP = Harvester-IP
  - Access time
  - Access frequency

Advantages

- Investigation starts as early as possible
- Harvesters often don‘t use bot nets
  → Identification of address trader
- Tested system to prevent address collection

But

- Crawling tar pits is not illegal
- Therefore:
  - Publish individual, traceable email addresses
  - Receive spam
  - Know before who is going to be the sender

Advantages

- Distributed HTTP tar pit network identifies harvester
- Crafted email addresses prove spamming
- Identify address traders
- Higher risk → Higher price per address
  → Higher costs → Reduced revenue

Conclusion

- Current anti spam techniques are limited
- Prevention is better
- Economic approaches are better
- Criminal investigation and prosecution are better

tobias@eggendorfer.info

tobias@eggendorfer.info