Stopping Spam by Extrusion Detection

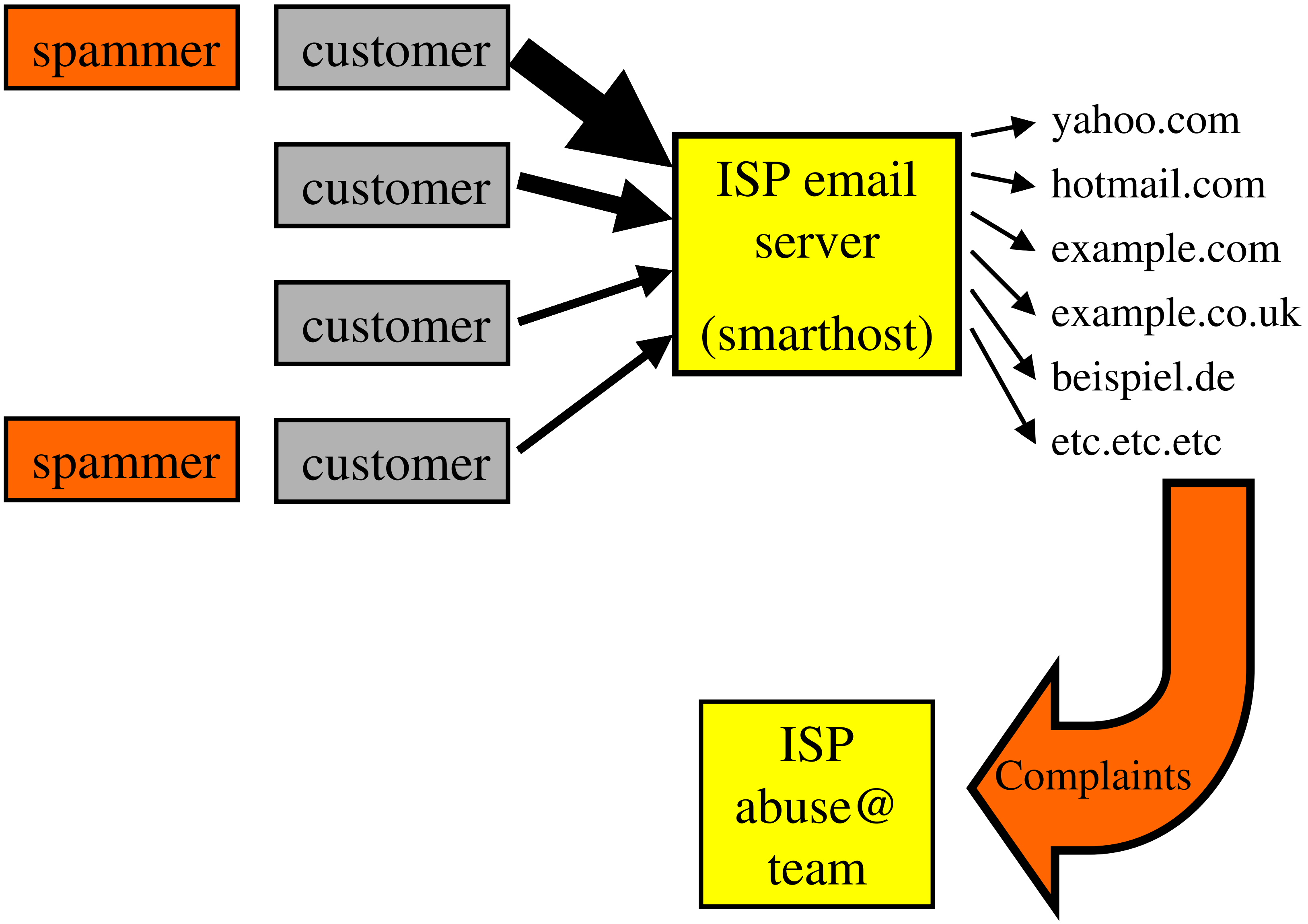
Richard Clayton

CEAS

31st July 2004

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Demon

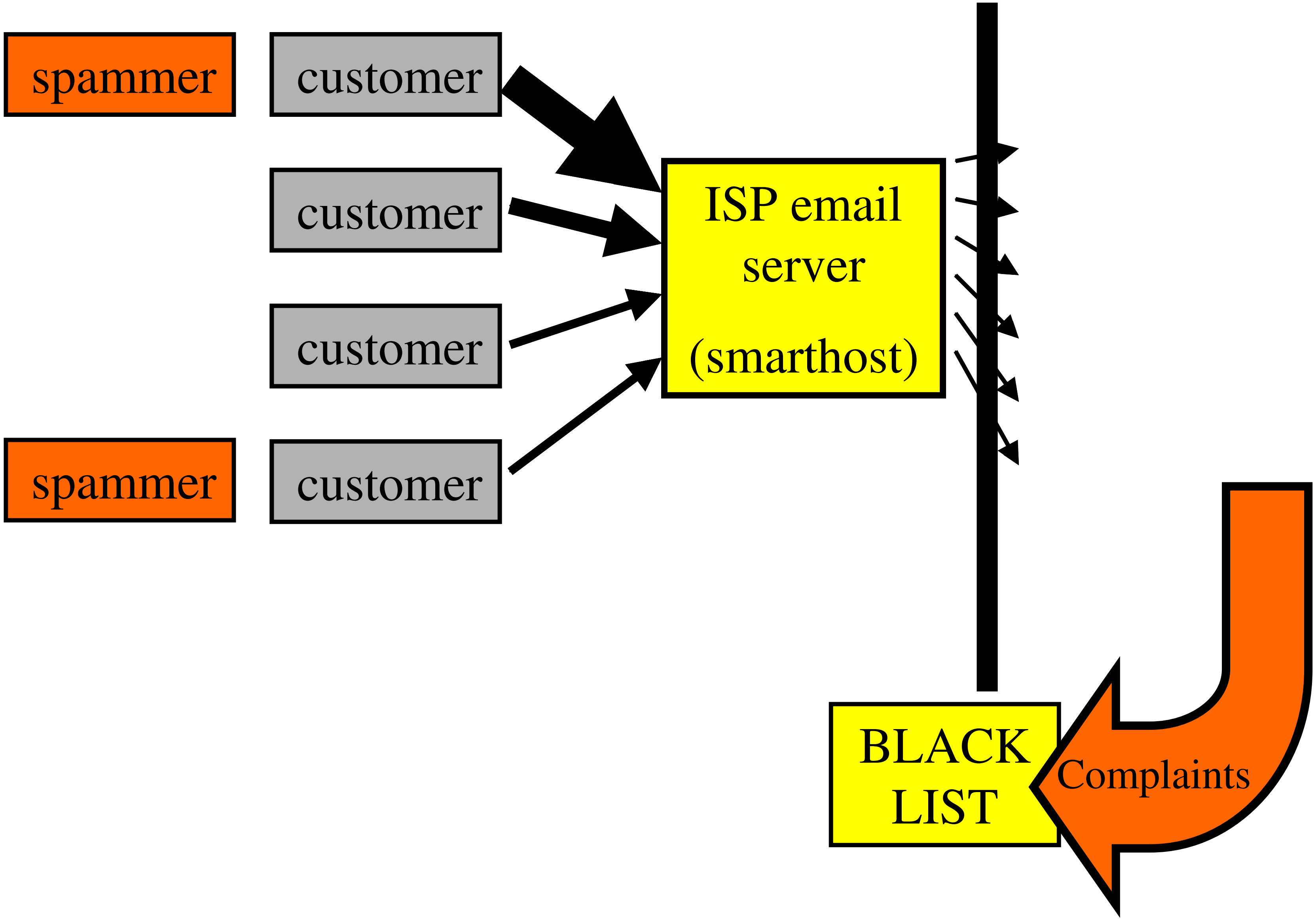

spammer
customer
customer
customer
spammer
customer
ISP email server (smarthost)
yahoo.com
hotmail.com
example.com
example.co.uk
beispiel.de
etc.etc.etc
Complaints
ISP abuse@ team

Current (Jul 04) problems for ISPs

- ⚐ Insecure customers
  - very few real spammers in the UK!
- Open proxies
  - mainly “trojans on non-standard ports”
- SMTP AUTH
  - Exchange “admin” accounts + *many others*
- Systems still insecure “out of the box”
  - brand new XP is compromised before secured

ISP's Real Problem

- Blacklisting of IP ranges & smarthosts
  - `listme@listme.dsbl.org`
- Rapid action necessary to ensure continued service to all other customers
- But reports may go to the blacklist and not to the ISP (or will lack essential details)



spammer

customer

customer

customer

spammer

customer

ISP email server (smarthost)

BLACK LIST

Complaints

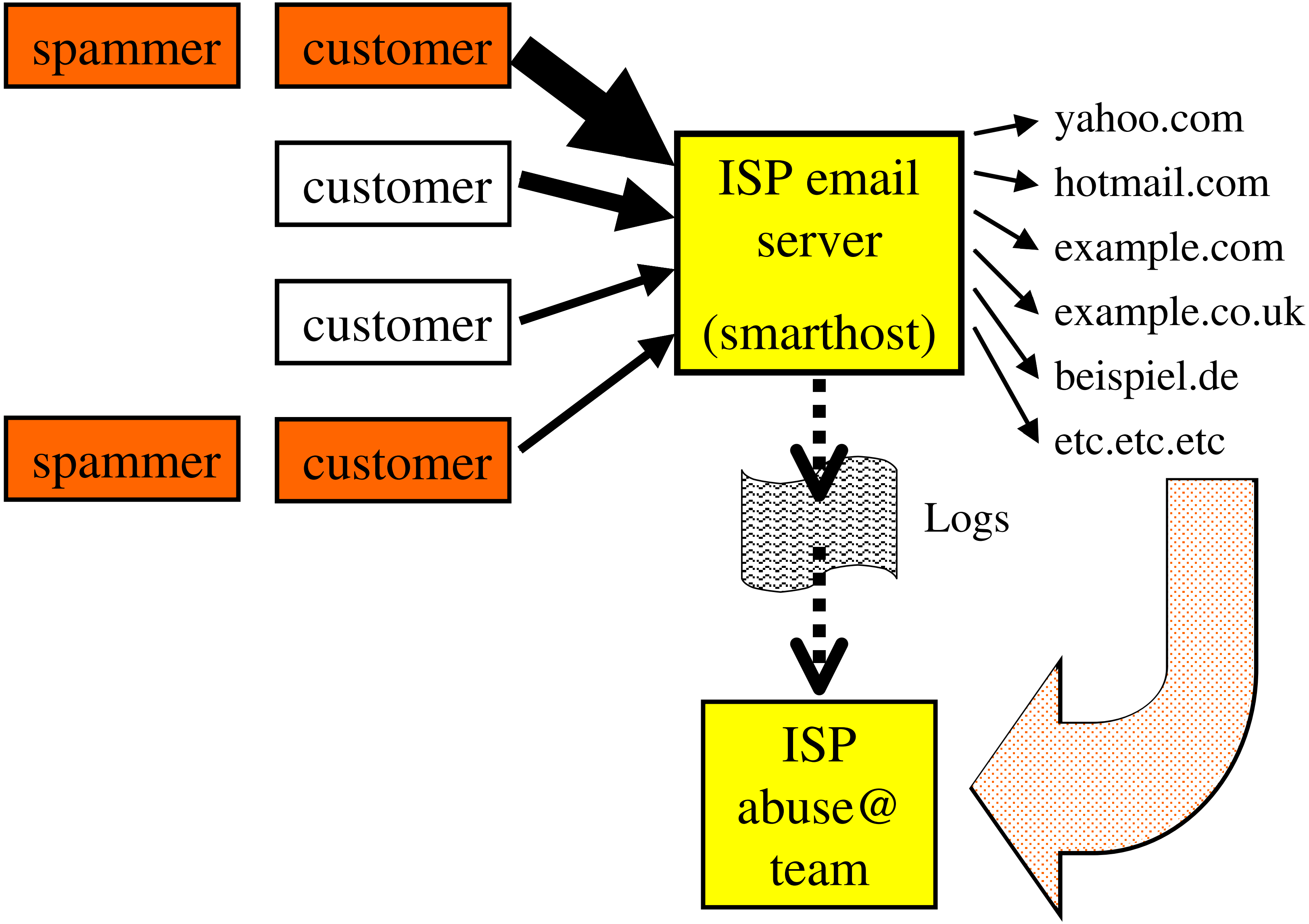
Why Spotting Spam is Hard

- Expensive to examine outgoing content
- Legal/contractual issues with blocking
  - and “false positives” could cost you customers
- Volume is not a good indicator of spam
  - many customers with occasional mailshots
- “Incorrect” sender doesn’t indicate spam
  - many customers with multiple domains

Key Insight

- Lots of spam is to ancient email addresses
- Lots of spam is to invented addresses
- Lots of spam is blocked by remote filters

- Can process server logs to pick out this information. Spam has delivery failures whereas legitimate email mainly works


spammer
customer
customer
customer
spammer
customer
ISP email server
(smarthost)
yahoo.com
hotmail.com
example.com
example.co.uk
beispiel.de
etc.etc.etc
Logs
ISP abuse@ team

My Log Processing Heuristics

- ⚐ **Report “too many” failures to deliver**
  - more than 40 works pretty well
- Ignore “bounces” !
  - have null “< >” return path, these often fail
  - detect rejection daemons without < > paths
- Ignore “mailing lists”
  - most destinations work, only a few fail
  - more than one mailing list is a spam indicator!

Bonus! Also Detects Viruses

- Common for mass mailing "worms" to use address book (mainly valid addresses)
- Recent trend towards scanning the browser cache and (Swen) accessing Usenet servers
  - so many addresses now invalid or badly formed
  - plus remote sites may reject incoming malware
- **So virus infections are also detected**

Evaluation at Large UK ISP

- 28 day period (1-28 March 2004)
- No public holidays (ie 20 working days)
- 85K active customers (of 200K total)
- 33.4 million emails (51.8 million destinations)
- System had been in production 6 months
  - hence there are no edge effects (initially was spotting dozens of problems per day)
- No major virus events occurred

Evaluation Methodology

- Manually check all reports from system
  - spamming patterns are very obvious
- False positive occurs when report is wrong!
- False negatives assessed by comparison of results with manual inspection of results from a far more sensitively tuned version.
  - also examined all other reports of viruses etc

Results (total over 28 days)

Abuse Type	total detected	false positive	false negative
Real Spammers	0	0	0
Open Servers	56	69	10
Virus Infection	29	6	4
Email loops	14	3	0

Looking More Closely

Abuse type	total	False+ve	False -ve
Open Servers	56	69	10

FALSE POSITIVES:

36 customers running multiple genuine mailing lists

22 customers with >40 delivery failures during one day

11 assorted other reasons (see paper)

FALSE NEGATIVES:

7 (of the 10) were one “cutecandy” spammer (using a fixed sender string & remote sites accepted a dictionary attack)

Future Work

- Spammers will evolve!
  - Spam resembling bounces will be hard to spot
  - Valid MAIL FROM will be harder to detect
  - Reducing the volume will be harder to spot
- Viruses will evolve!
  - Changing HELO isn’t doing them much good
  - May begin to avoid nonsense destinations

Conclusions

- Spammers & viruses that hide a pattern at the destination make a pattern at the source
- Some simple heuristics currently spot these patterns : with delivery failures being key
- False positives mainly caused by software & users that are being especially clueless ☹

Stopping Spam by Extrusion Detection

http://www.cl.cam.ac.uk/~rnc1/

THE END : Any questions ??

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Demon