# spamHINTS update

**PI:** **Prof. Ross Anderson**

**Researcher:** **Dr. Richard Clayton**

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

# Summary

- Best Practice Document
- Very early results from sFlow data
- Latest results from sFlow data

# Sharing traffic data

- Want to encourage ISPs to share email log info about incoming spam & viruses
- Send report to host ISP indicating:
  - source IP address (and of course time)
  - source email address (probably forged)
  - destination email address
  - metadata (size, HELO message, filter results)
  - diagnosis of problem

# Lawyers!

- Reporting is straightforward except…
- … email addresses are personal data
  - Information Commissioner quite clear on this
- Much is of course forged, but amongst this may be some real email, and source/ destination details could be sensitive
  - so must meet legal obligations

# Legitimate processing

- Asking another ISP to take action to prevent their user sending spam/virus traffic can be seen as legitimate processing
- So jump through correct hoops & all OK
  - inform customers of processing
  - (try to) inform senders of processing
  - ensure processing covered by privacy policy
  - address any promises of confidentiality

# Best Practice document

- Would also be desirable for processing to be in line with industry Best Practice!

- Hence recent draft of:

  *Best Practice for reporting abuse issues based on traffic data*

# Components of Best Practice

- Reports based on traffic data must only be sent by prior agreement
- Reports should not be unduly repetitive
- The evidence on which the report is based must be clearly given (& accurately timed)
- Needs warning about personal data
- Must keep customers informed (as above)

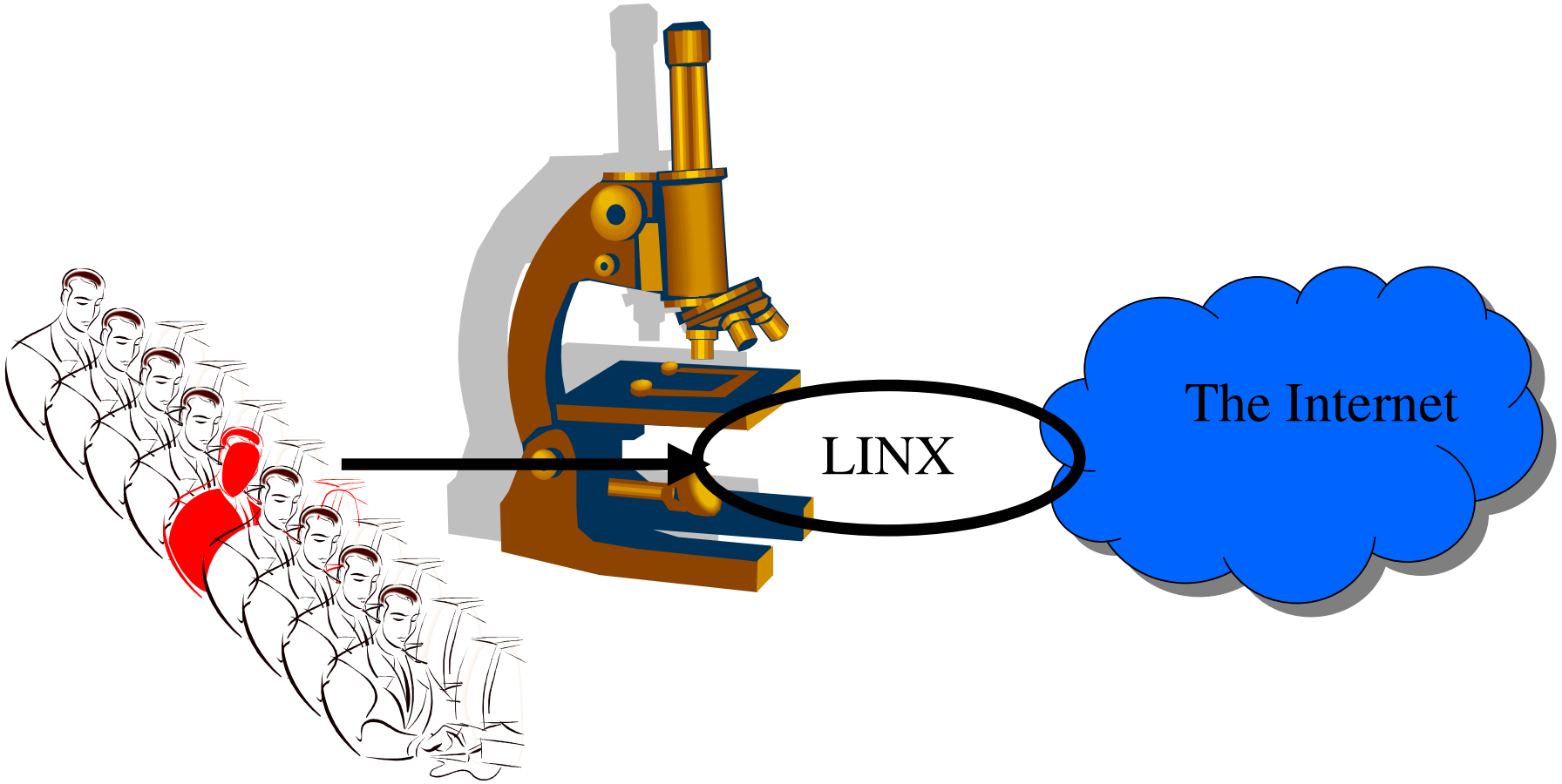
# Approval of BCP?

- Document presented at LINX52
- Some relatively minor revisions
- No comments on mailing lists
- Don’t believe it is controversial…

  …so please could we approve it!

**http://www.spamhints.org/TrafficDataBCP.pdf**

# spamHINTS @ LINX



# Latest sFlow results

- Processed 25 hours of data from Foundry switches (Wednesday last week) *thanks Ivan!*
- Very limited analysis so far... (and note very carefully the caveats on coverage – no Extreme data, no private peering data, no data where peering is at other IXPs!)
- Can get picture of who is sending email
- Can correlate with info from Demon logs

# Raw statistics

- 9,672,926 SMTP packets detected
  - SYN 413,902 (4.2%)
  - SYN/ACK 194,649 (2.0%)
  - RST 72,235 (0.7%)
  - FIN 387,433 (4.0%)
- 611,563 unique IP addresses involved
  - 378,457 client only
  - 170,648 server only
  - 62,458 both client & server

| | | | |
|---|---|---|---|
| 18,794 | 646 | 296 | Turkey |
| 17,724 | 5,867 | 2,905 | UK |
| 16,558 | 1,556 | 407 | Korea |
| 11,248 | 1,239 | 420 | Spain |
| 10,731 | 1,871 | 409 | Italy |
| 9,897 | 866 | 196 | Korea |
| 7,701 | 489 | 64 | France |
| 6,092 | 594 | 404 | Portugal |
| 5,854 | 979 | 526 | UK |
| 5,258 | 105 | 130 | India |
| 4,982 | 966 | 185 | Germany |
| 4,503 | 569 | 416 | Middle East |
| 4,416 | 36 | 2 | China |
| 4,200 | 137 | 63 | India |
| 4,072 | 420 | 161 | France |
| 3,814 | 742 | 467 | UK |
| 3,737 | 477 | 232 | Malasia |
| 3,480 | 5,183 | 3,422 | UK |
| 3,291 | 7,123 | 1,296 | USA |
| 3,267 | 125 | 45 | Vietnam |
| 3,188 | 171 | 10 | Brazil |
| 3,185 | 10 | 4 | India |
| 2,975 | 1,224 | 654 | Ireland |
| 2,825 | 160 | 29 | Africa |
| 2,750 | 467 | 455 | UK |
| 2,719 | 752 | 47 | Europe |
| 2,683 | 407 | 122 | Israel |
| 2,618 | 1,328 | 311 | Germany |
| 2,604 | 24 | 5 | India |
| 2,586 | 425 | 113 | Israel |

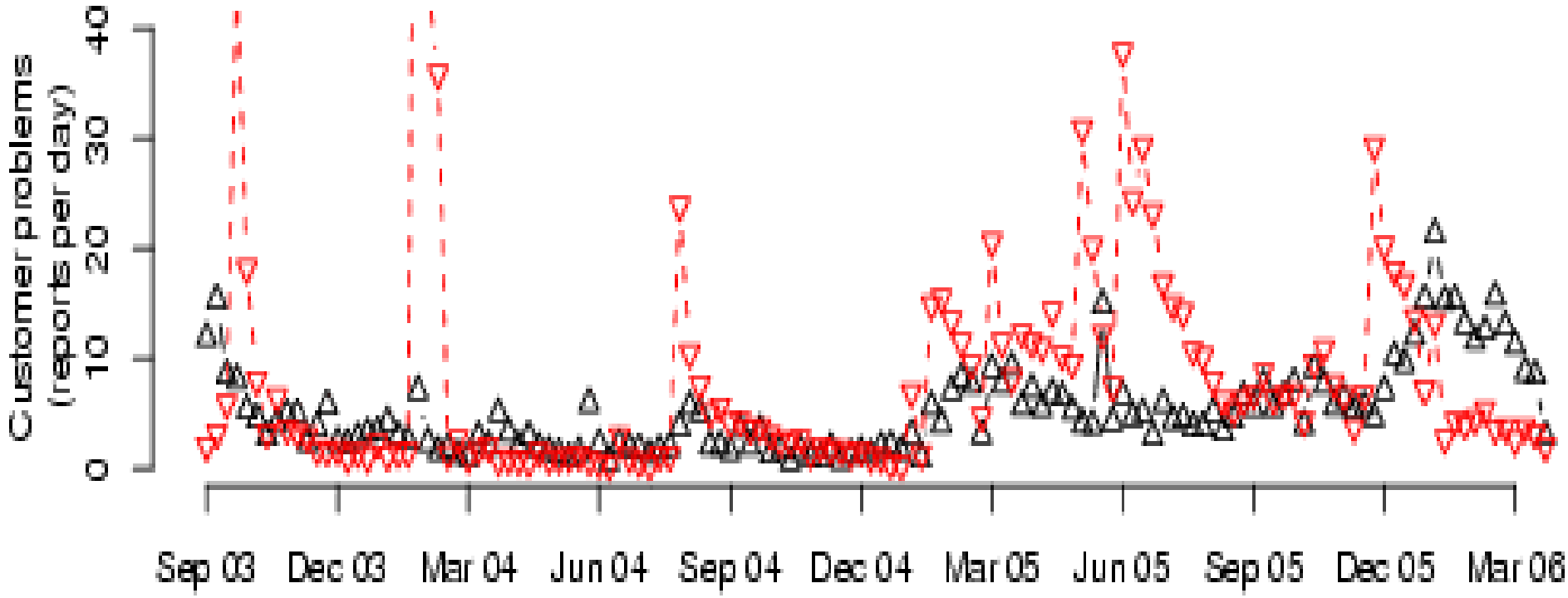
# Client AS "top 30"

columns are:

- clients
- servers
- both

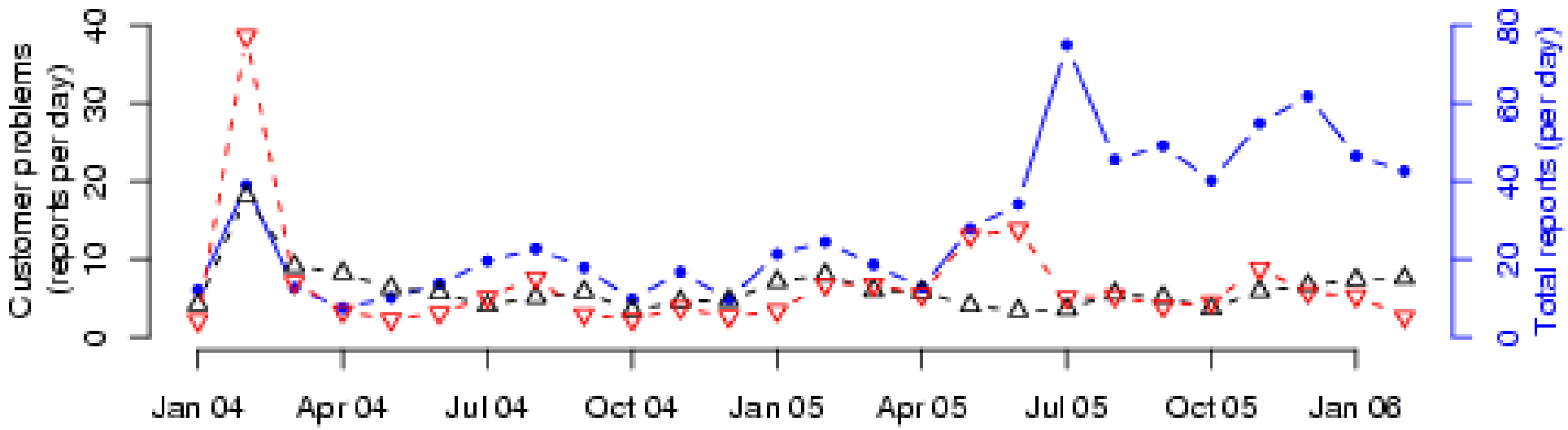
NB: 100th place has 855 clients…

# Email log processing @ Demon



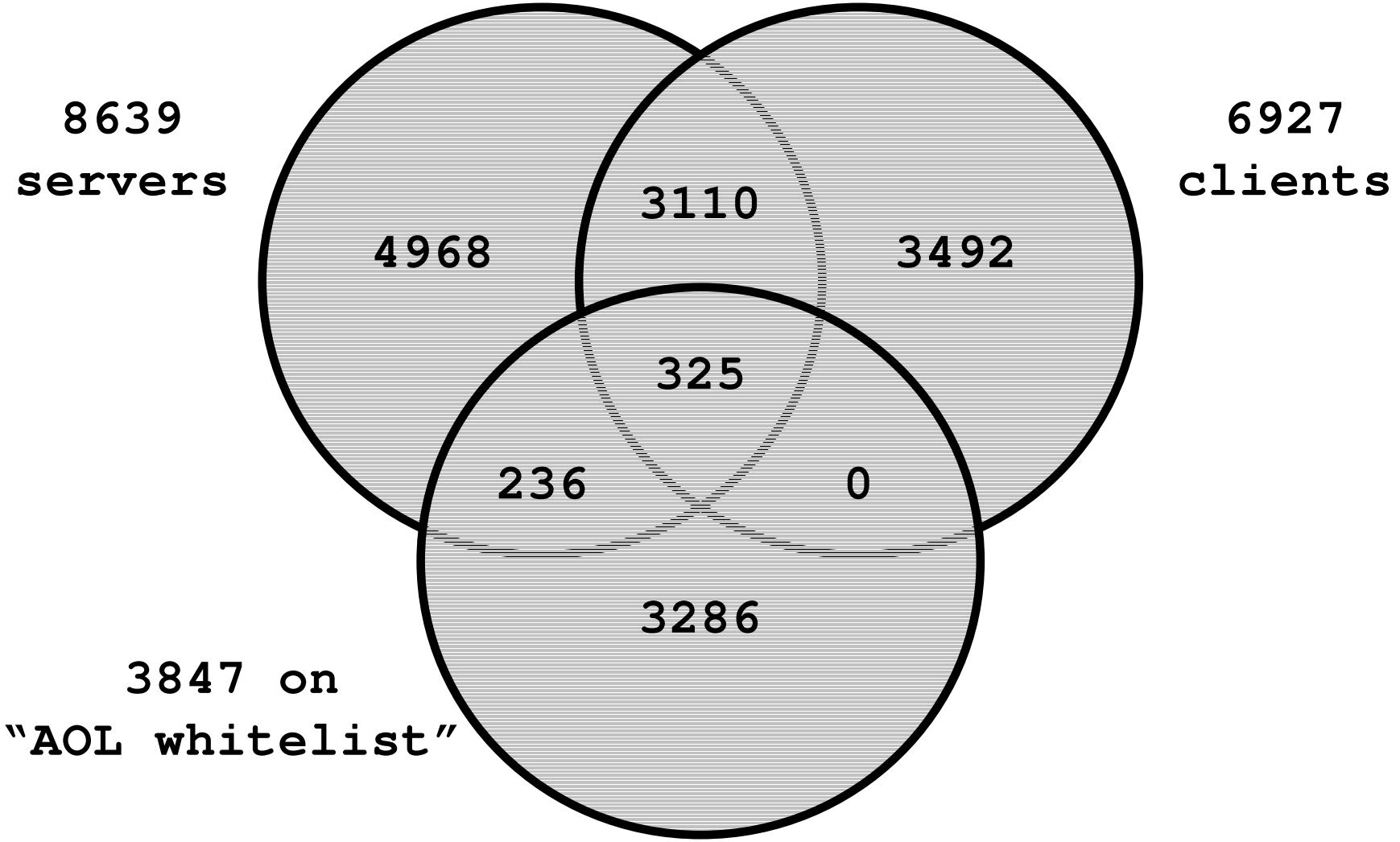

Detection of spam (black) and viruses (red)

# Incoming reports (all sources)




spam (black), viruses (red), reports (blue)

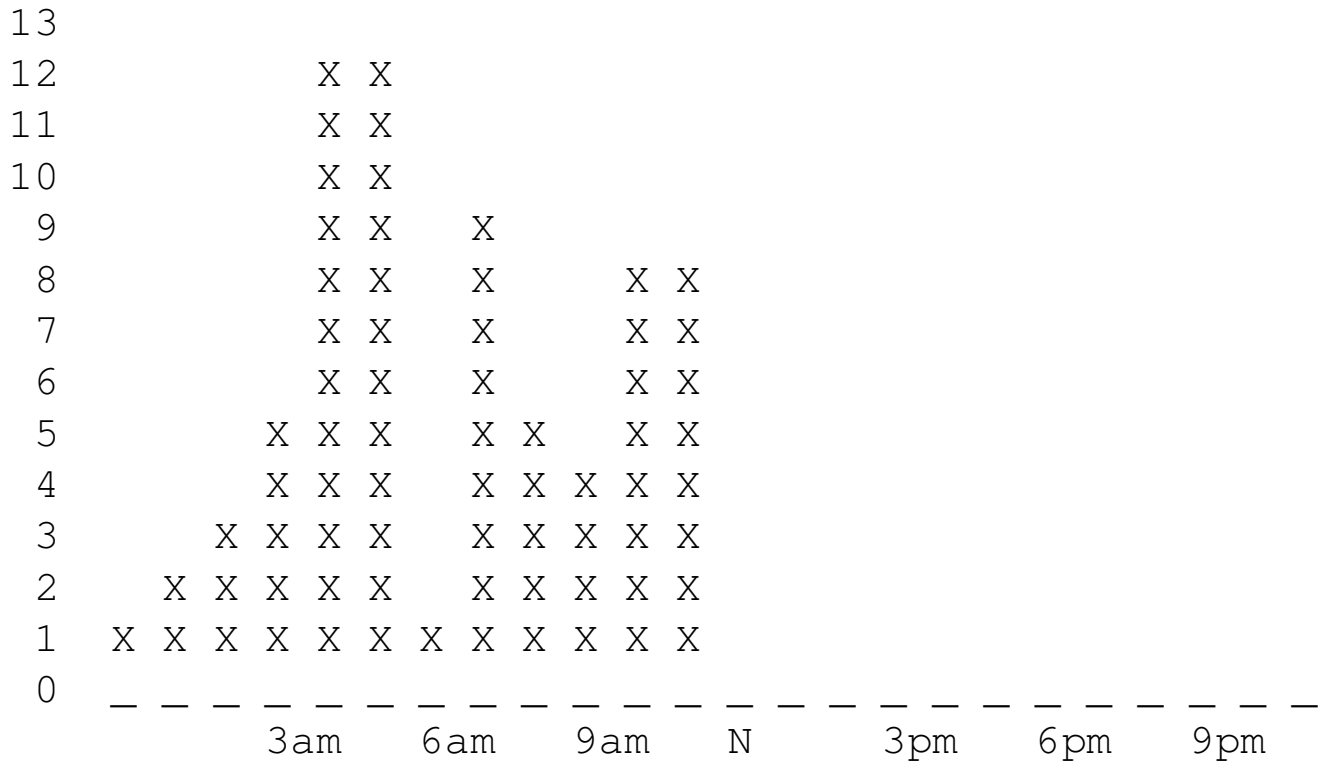
# sFlow Results : Demon customers



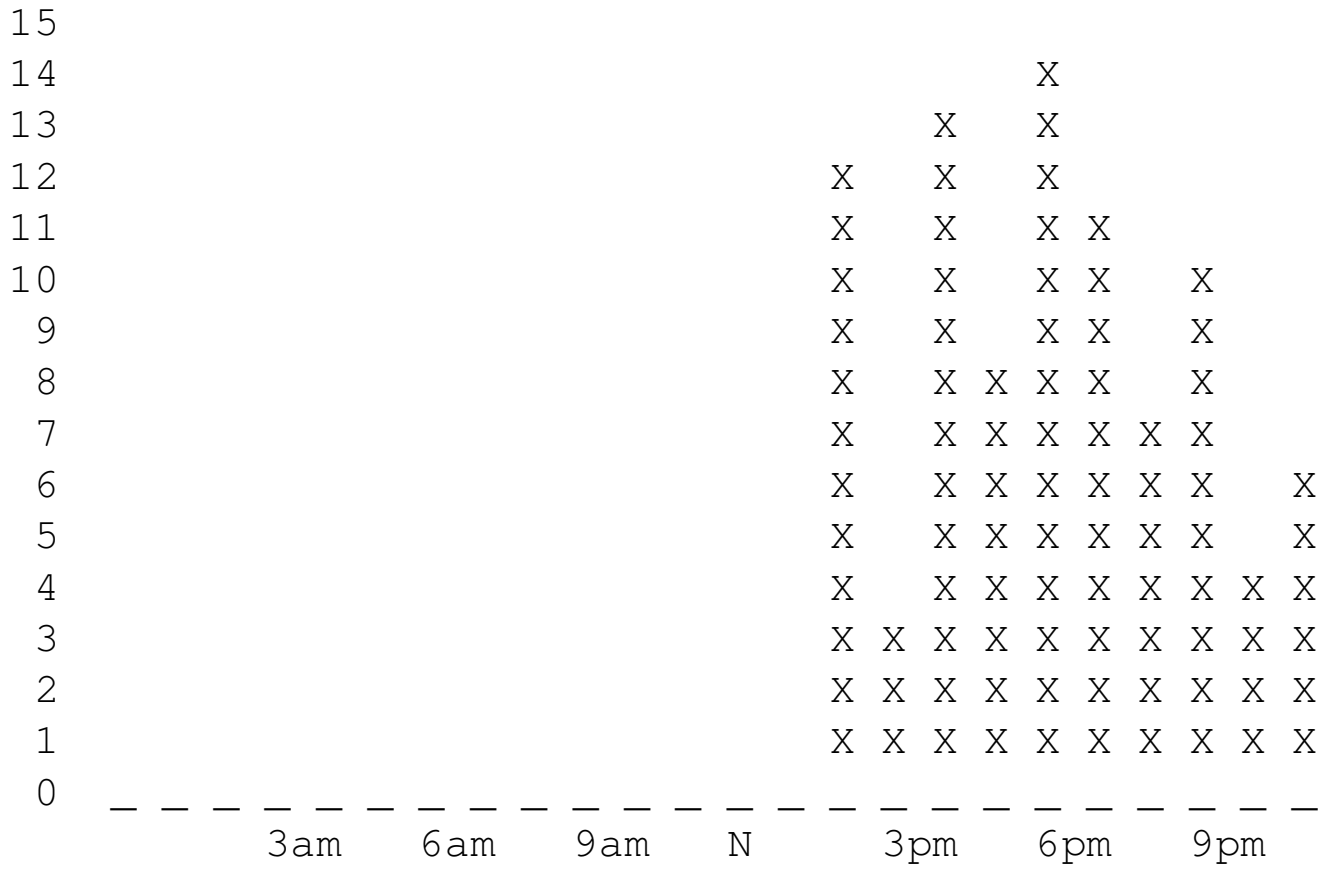
# What does it mean ?

- New detection methods pick up problems that were previously missed
- Spend money on your abuse team & soon reach a steady state (virus outbreaks excepted)
- Email log processing currently picks up 10 customers per day on average
- sFlow data suggests 3400+ more to go ???

# Known “open server”



# Another known “open server”



Work continues...

**Richard Clayton**
**<rnc1@cl.cam.ac.uk>**

**www.spamhints.org**