Using Early Results from the 'spamHINTS' Project to Estimate an ISP Abuse Team's Task

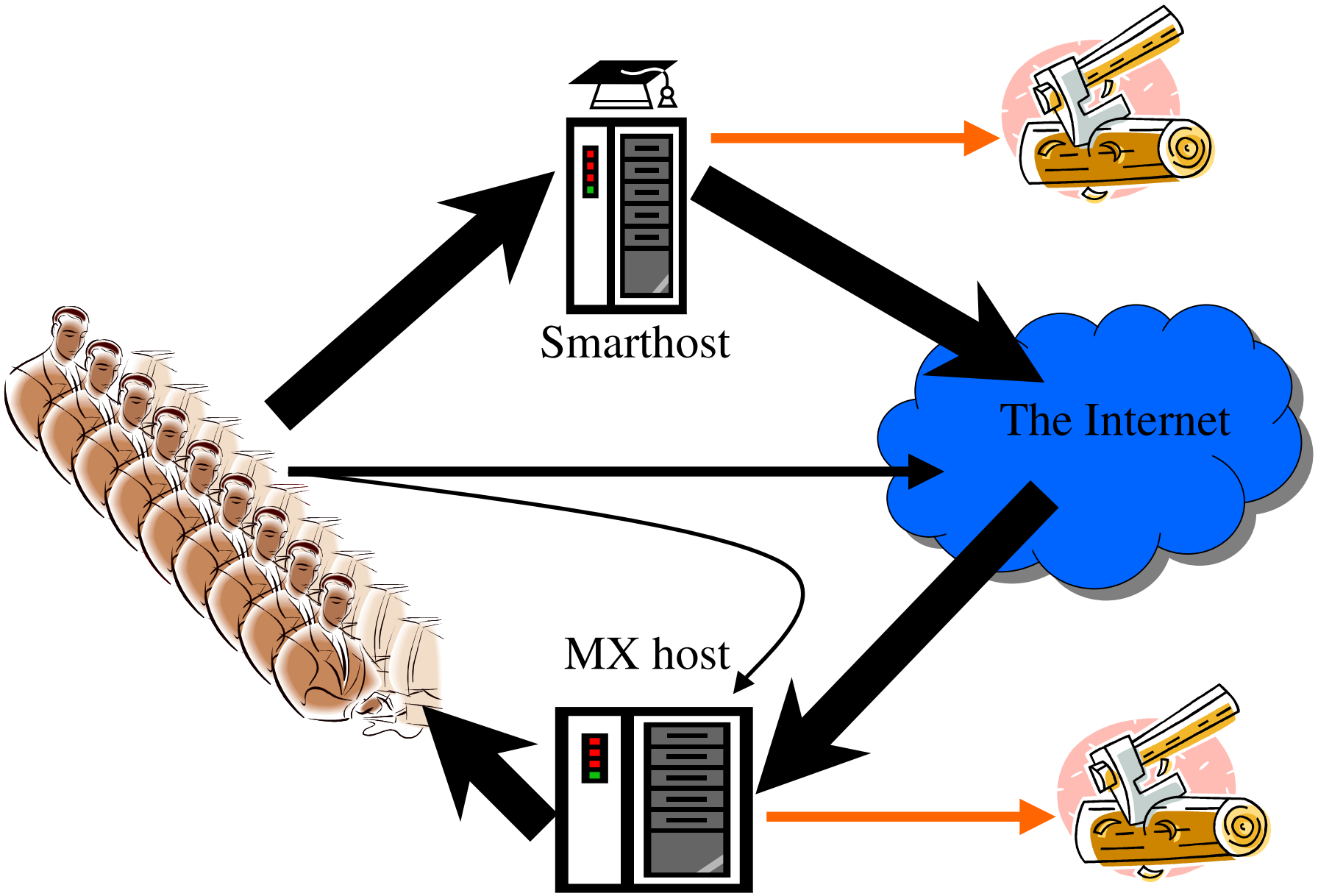
Richard Clayton

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

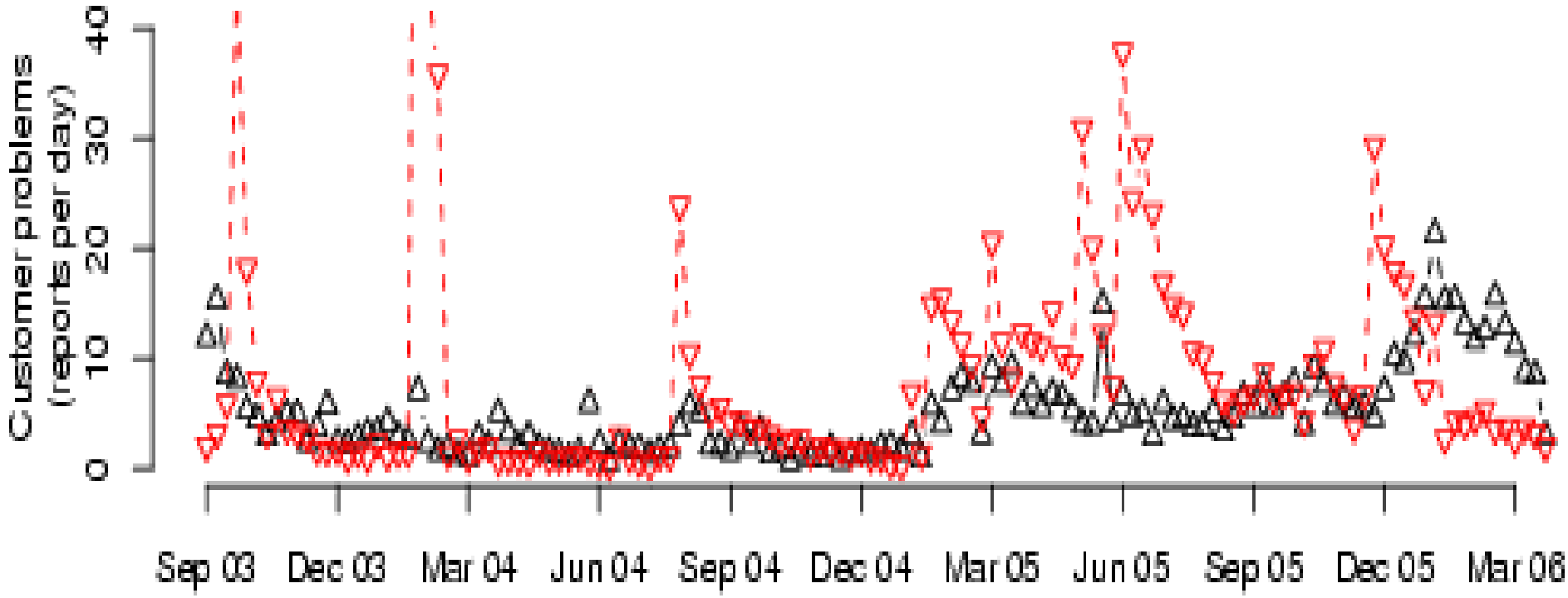
CEAS, Mountain View

28th July 2006

ISP Email Handling



Email Log Processing @ Demon



Detection of spam (black) and viruses (red)

Detection Ratios (spam)

Count	AS	Description	Ratio
80319	AS4134	CHINANET (CN)	55%
75980	AS4766	Korea Telecom (KR)	62%
47578	AS4812	China Telecom (CN)	59%
18683	AS9318	Hanaro Telecom (KR)	53%
12609	AS4837	CNC (CN)	38%
5792	AS12322	Proxad (FR)	38%
4941	AS3786	Dacom Corporation (KR)	67%
4779	AS7738	TeleBahia (BR)	21%
3929	AS9277	Thrunet (KR)	62%
3911	AS3320	Deutsche Telecom (DE)	16%

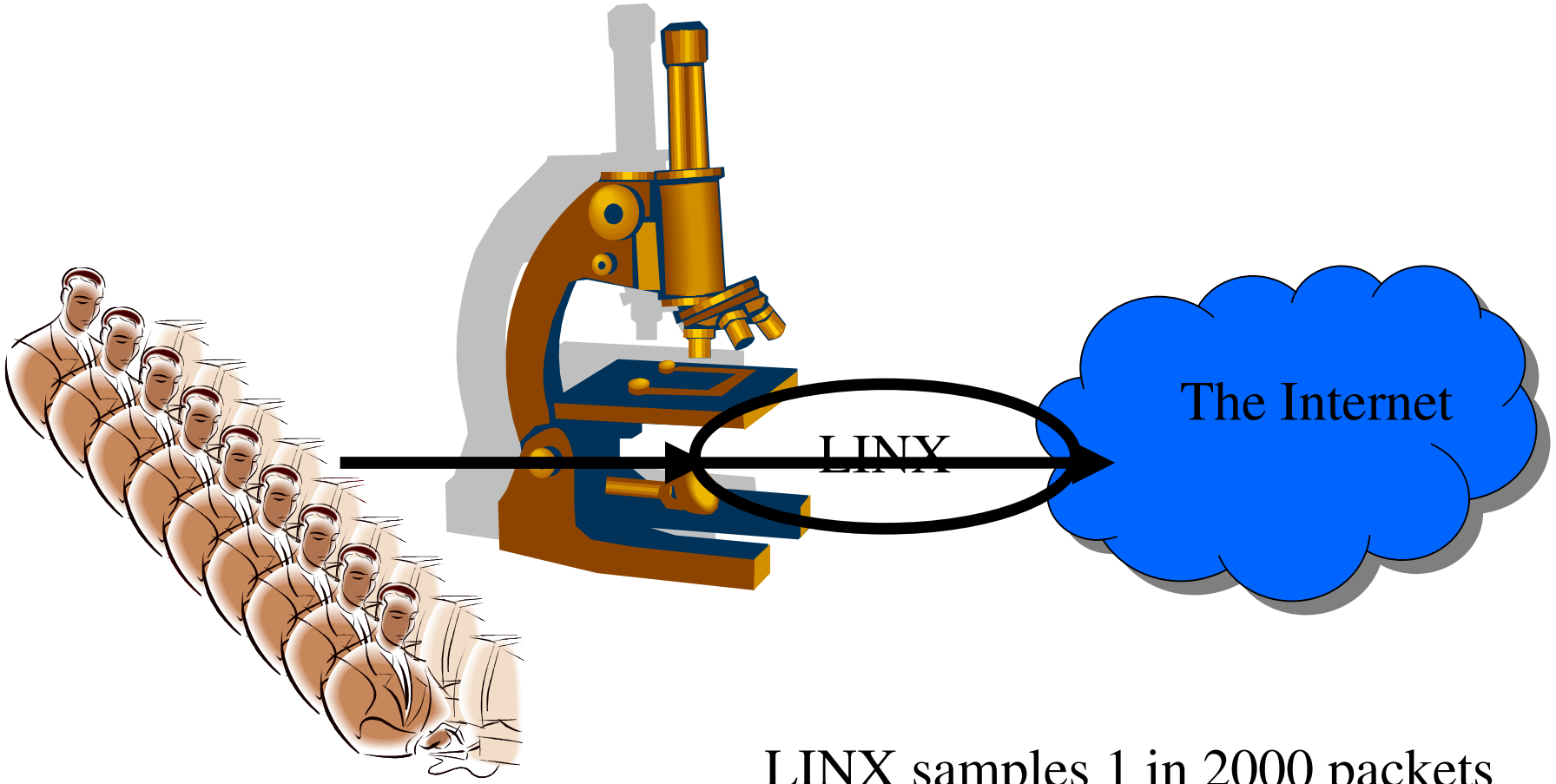
Ratios

- Significantly better than previously reported
  - was circa 1% at time of CEAS 2005
  - now 25% - 50%
  - Main reason is better heuristics (esp HELOs)
- Applying to Demon… we are detecting 20 customers a day, so maybe 40-80 are actually infected at maximum

Suppose All Email is Bad ☹

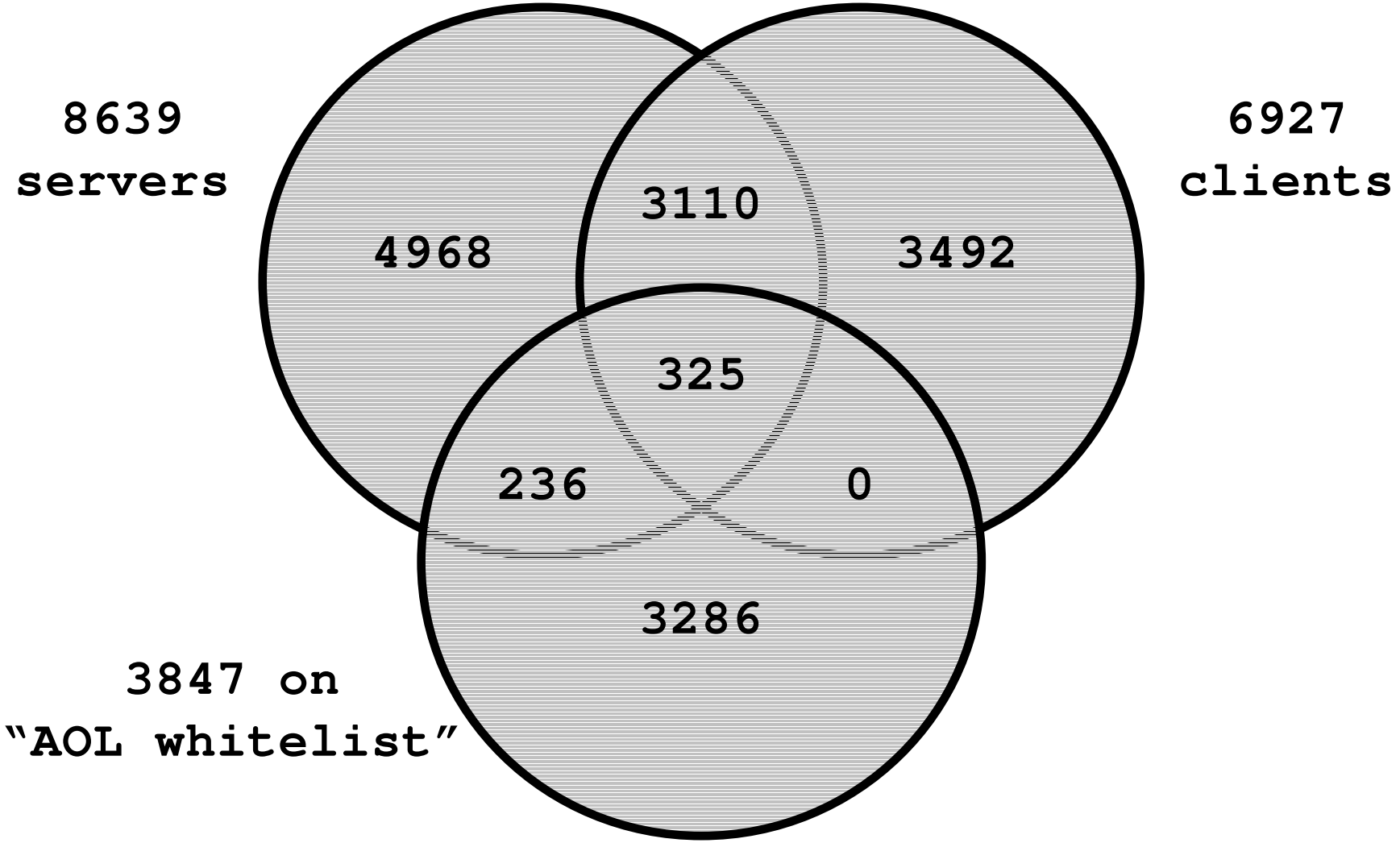
- Typically, remote ASs are sending email from 0.25% … 1.25% of their address space
- Applying this ratio to Demon address space (~200K customers) means 500 to 2500 customers will have a problem and so far we have detected almost none of them ☹

spamHINTS Research Project



LINX samples 1 in 2000 packets (using sFlow) and makes the port 25 traffic available for analysis…

sFlow Results : Demon Customers



Conclusions

- Log processing spots ~20 problems a day
- Remote data suggests 40–80 Demon customers actually have a problem
- Worst case analysis on customer counts suggests 500–2500 with a problem
- sFlow data suggests 3500+ problems
  - but much more work is needed!

Using Early Results from the ‘spamHINTS’ Project to Estimate an ISP Abuse Team’s Task

Richard Clayton

`http://www.cl.cam.ac.uk/~rnc1/earlyResult.pdf`

UNIVERSITY OF CAMBRIDGE
Computer Laboratory