Malware Research at SMU

Tom Chen
SMU
tchen@engr.smu.edu
www.engr.smu.edu/~tchen

Outline

- About SMU and Me
- Virus Research Lab
- Early Worm Detection
- Epidemic Modeling
- New Research Interests

About SMU

- Small private university with 6 schools - engineering, sciences, arts, business, law, theology
- 6,300 undergrads; 3,600 grads; 1,200 professional (law, theology) students
- School of Engineering: 51 faculty in 5 departments
- Dept of EE: specialization in signal processing, communications, networking, optics

About Me

- BS and MS in EE from MIT, PhD in EE from U. California, Berkeley
- GTE (Verizon) Labs: research in ATM switching, traffic modeling/control, network operations
- 1997 joined EE Dept at SMU: traffic control, network security

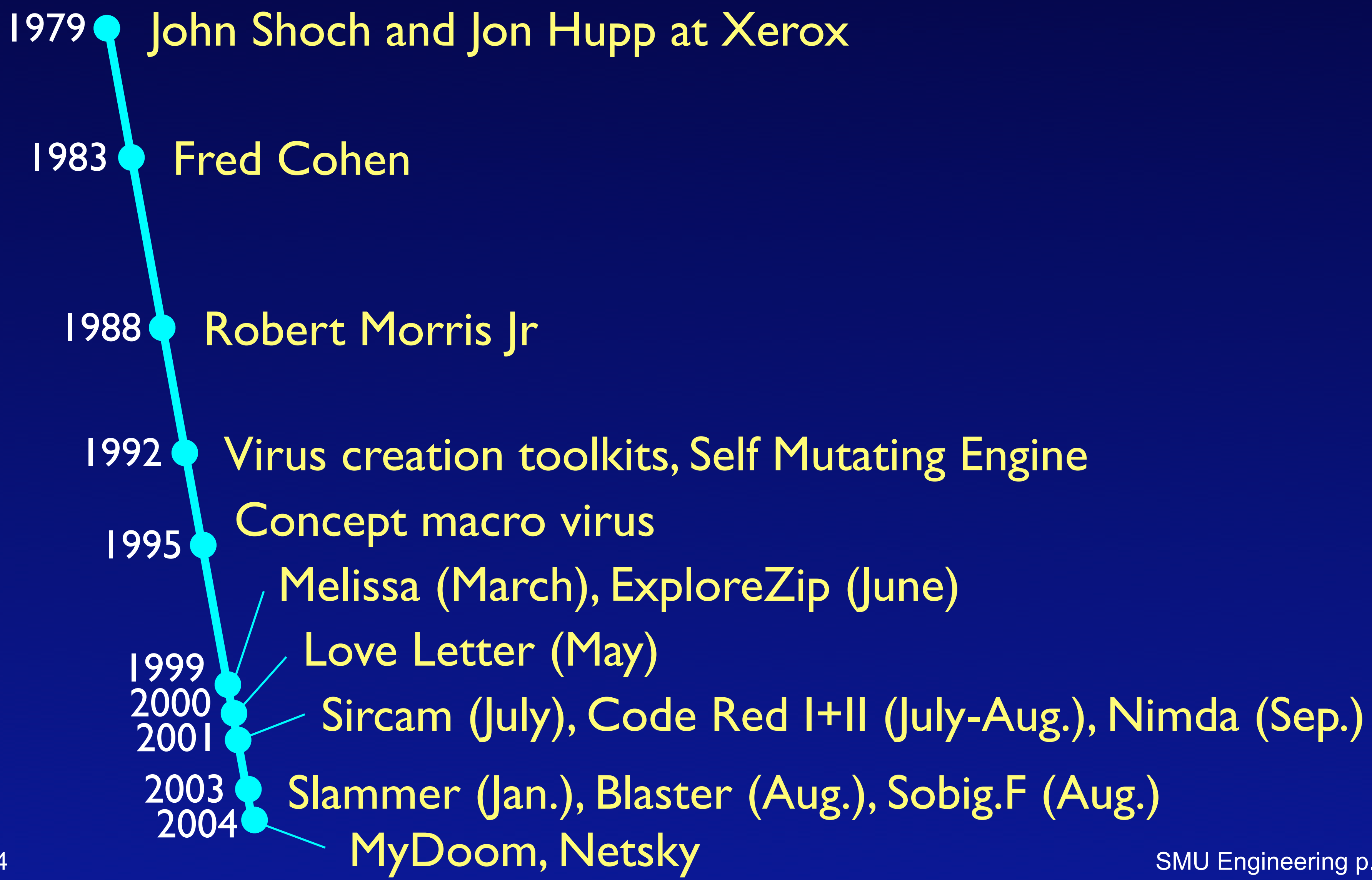
Research Interests

- Convergence of traffic control and Internet threats
  - Large-scale traffic effects of worm epidemics
  - Traffic control (packet classification, filtering/ throttling) for detection and defenses
- Deception-based attacks and defenses
  - Social engineering, honeypots

Motivations

- Worms and social engineering attacks (phishing, spam) have widespread effects in Internet
  - Top worms (Loveletter, Code Red, Slammer,...) causes billions in damages
  - 78% organizations hit by virus/worm, $200k average damage per organization [2004 FBI/ CSI survey]
  - 40% Fortune 100 companies hit [Symantec report]

- 25 years- problem continues to get worse
- We want to apply theories (traffic control, epidemiology) towards detection and control



1979 John Shoch and Jon Hupp at Xerox

1983 Fred Cohen

1988 Robert Morris Jr

1992 Virus creation toolkits, Self Mutating Engine

1995 Concept macro virus

1999 Melissa (March), ExploreZip (June)

2000 Love Letter (May)

2001 Sircam (July), Code Red I+II (July-Aug.), Nimda (Sep.)

2003 Slammer (Jan.), Blaster (Aug.), Sobig.F (Aug.)

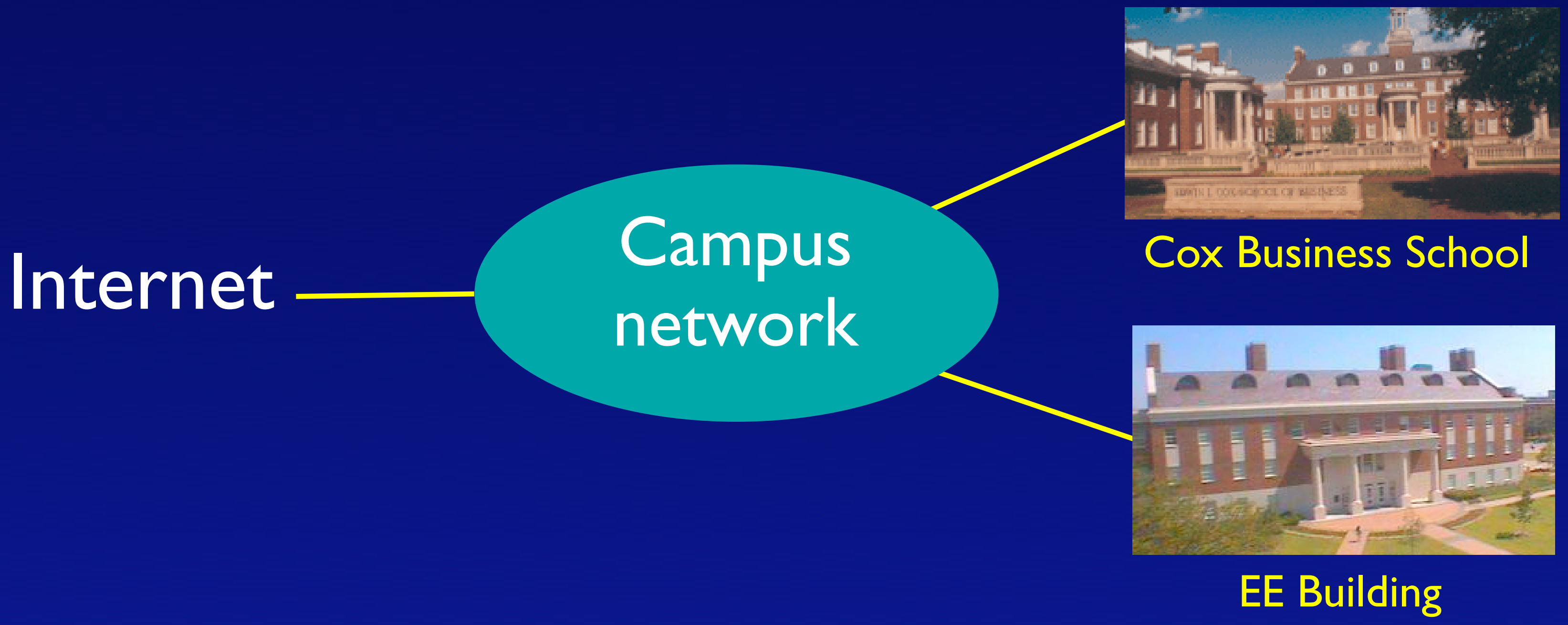
2004 MyDoom, Netsky

Research Activities

- Virus research lab
- Early worm detection
- Epidemic modeling

Virus Research Lab

- Distributed computers in EE building and Business School



Internet

Campus network

Cox Business School

EE Building

Virus Research Lab (cont)

- Intrusion detection systems to monitor live traffic
  - Snort (network IDS), Prelude (event correlation), Samhain (host-based IDS), Nagios (network manager)
- Honeypots for worm detection/capture
  - Honeyd (honeypot), Logwatch (log monitoring)

Virus Research Lab (cont)

- Network/worm simulator (Java)
  - To simulate different worm behaviors in different network topologies
  - To find worm-resistant network topologies

Early Detection of Worms

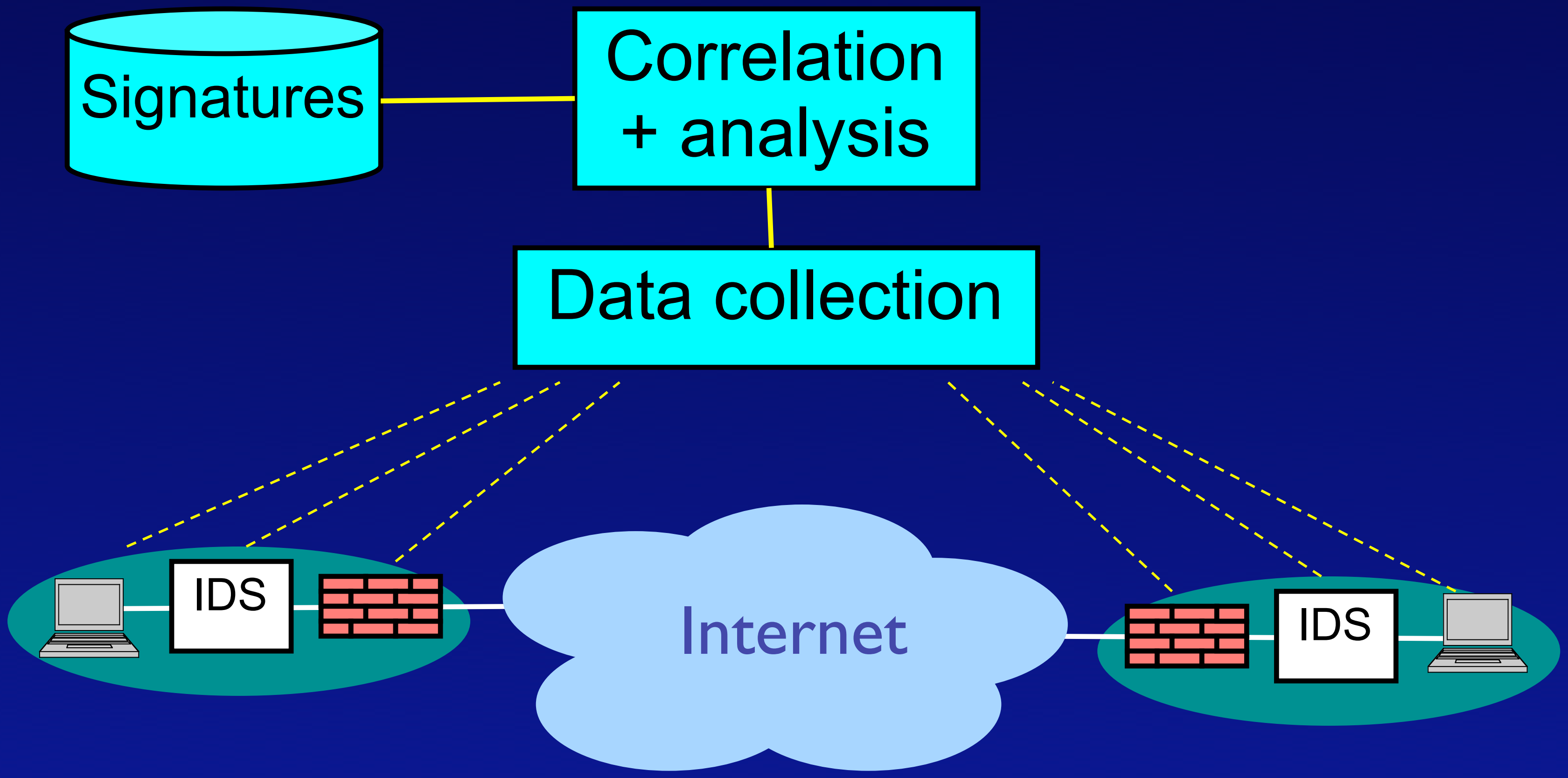
- Goal is global system including honeypots for early warning of new worm outbreaks
- Honeypots are traditionally used for post-attack forensics
- For early warning, honeypots need augmentation with real-time analysis

Early Detection (cont)

- Jointly with Symantec to enhance their DeepSight Threat Management System
  - DeepSight collects log data from hosts, firewalls, IDSs from 20,000 organizations in 180 countries
  - Symantec correlates and analyzes traffic data to track attacks by type, source, time, targets

Early Detection (cont)

- Architecture of DeepSight



Early Detection (cont)

- We want to add honeypots to DeepSight
- Honeypot sensors have advantage of low false positives (a problem with IDSs)
- DeepSight has correlation/analysis engine to make honeypots useful for real-time detection
  - Modifications to correlation engine needed

Epidemic Modeling

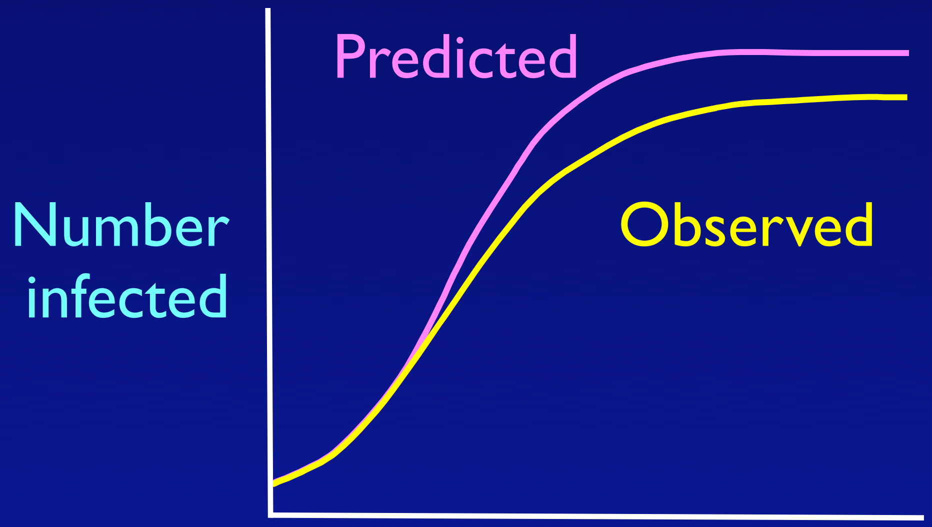
- Epidemic models predict spreading of diseases through populations
  - Deterministic and stochastic models developed over 250 years
  - Helped devise vaccination strategies, eg, smallpox
- Our goal is to adapt epidemic models to computer viruses and worms
  - Take into account network congestion

Basic Epidemic Model

- Assumes all hosts are initially Susceptible, can become Infected after contact with an Infected
  - Assumes fixed population and random contacts
- Then basic epidemic model predicts number of Infected hosts has logistic growth

Basic Epidemic (cont)

- Logistic equation predicts “S” growth
- Observed worm outbreaks (eg, Code Red) tend to slow down more quickly than predicted



Basic Epidemic (cont)

- Initial rate is exponential: random scanning is efficient when susceptible hosts are many
- Later rate slow downs: random scanning is inefficient when susceptible hosts are few
- Spreading rate also slows due to network congestion caused by heavy worm traffic

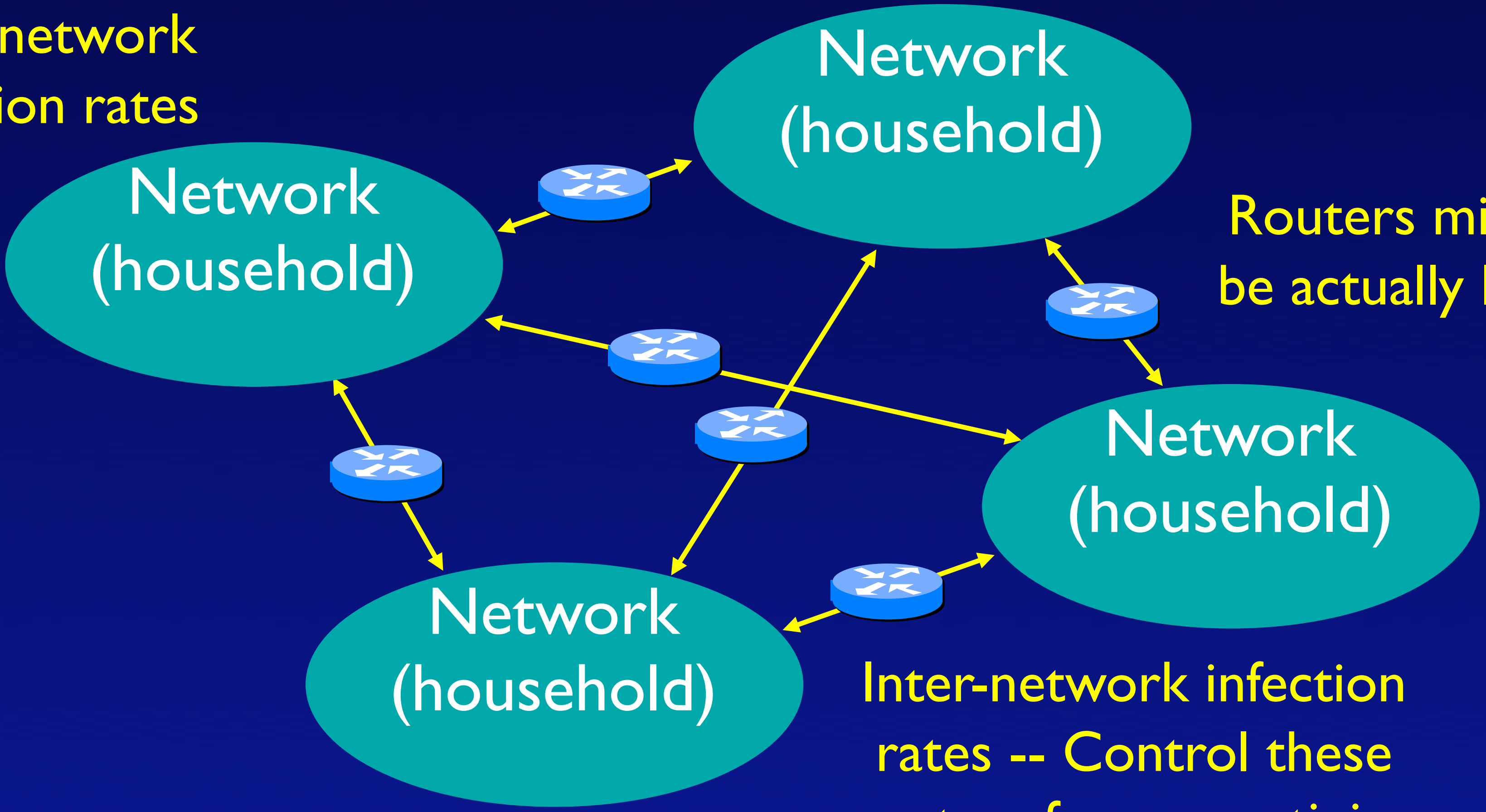
Dynamic Quarantine

- Recent worms spread too quickly for manual response
- Dynamic quarantine tries to isolate worm outbreak from spreading to other parts of Internet
  - Cisco and Microsoft proposals
  - Rate throttling proposals
- Epidemic modeling can evaluate effectiveness

Quarantining (cont)

- “Community of households” epidemic model assumes
  - Population is divided into households
  - Infection rates within households can be different than between households
- Similar to structure of Internet as “network of networks”
  - Household = organization’s network

Quarantining (cont)



Intra-network infection rates

Routers might be actually ISPs

Inter-network infection rates -- Control these routers for quarantining

Quarantining

- As outbreak spreads, congestion causes inter-network infection rates to slow down outbreak naturally (seen empirically)
- Dynamic quarantining: quickly shutting down or throttling inter-network rates should slow down outbreak faster
  - Reaction time is critical
  - In practice, rate throttling may be preferred as gentler than blocking

New Research Interests

- Phishing
  - Damages: $1.2 billion to US financial organizations; 1.8 million consumer victims [Symantec]
  - 1,974 new unique phishing attacks in July 2004; 50% monthly growth rate in attacks [Anti-Phishing Working Group]

Phishing (cont)

- Our approach: email honeypots (spamtraps) are honeypots modified to receive and monitor email at fake addresses
  - Reliably capture spam
- Modify spam filters to detect phishing emails
- Analyze contents and links to fake Web sites, generate new email filter rules

New Research (cont)

- Bot nets
  - Symantec tracking 30,000+ compromised hosts; around 1,000 variants each of Gaobot, Randex, Spybot
  - Used for remote control, information theft, DDoS
  - Potentially useful for fast launching worms
  - Perhaps used by organized crime

Bot Nets (cont)

- Bots typically use IRC (Internet relay chat) channels for command and control
- We are seeking signs of bot nets on IRC channels

Conclusions

- Interests in traffic control and modeling applied to network security
    - Early detection, dynamic quarantining, epidemic modeling
- Interests in deception-based threats and defenses
    - Phishing, honeypots