CYBER SECURITY DIVISION

2013 PRINCIPAL INVESTIGATORS'

Bio-Inspired Anomaly Detection

Rutgers University

Nina H. Fefferman

9/18/13

Homeland Security

Science and Technology

Team Profile

Rutgers University
- **Dr. Nina H. Fefferman** (PI)
- Dr. Maciej Korczyñski
- Natalie Lemanski

Honeywell
- Dr. Siva Raj Rajagopalan (Co-PI)
- Dr. Haiyang Liu
- Dr. Jun Ho Huh

- Expertise in CyberSecurity/Cryptography (both hardware and software)
- Expertise in Evolutionary Sociobiology of Social Insect Algorithmic Efficiency
- Algorithms Theorists in Academia *with* Hands-on Industrial Partner Creating (Very Cool) Programmable Router Testbed
- Full Complement of Relevant Training: Applied Math, Engineering, Evolutionary Biology, Computer Science (focused on network security)

Customer Need

Computer (and other types of) networks are susceptible to *malicious* attacks and *accidental* anomalies

When computer networks fail, we lose: Money, Information, Time, Security, etc.

Networks are *bigger* than single domains: approach must scale

Attacks have full knowledge of existing sensors and how to evade detection

Early, Sensitive, and Specific threat identification allows us to interrupt propagation

Detecting & Identifying threats to networks faster and more accurately saves money, secures critical infrastructure, and protects industrial, academic, governmental, and public interest

Approach

Leverage bio-inspired distributed detection algorithms to design software to enable 'smart' routers (talk to each other *and* to dumb routers)

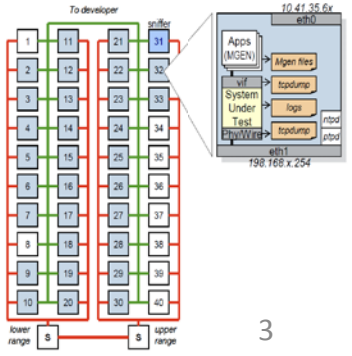
We've started with a Bee-Inspired Anomaly Detection (BIAD)

- Independent analysis of local information
- Shared conclusions (rather than shared data) influence independent analyses (i.e. "I see something interesting in my data, do you in yours?")

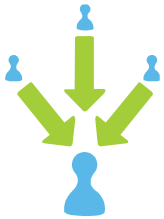
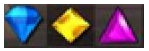
We're developing theory $a_i\left(W_i^{t-w,t}\left(o(b_i)\right), st_{i,t}(b_i)\right)$,

testing it in software simulation,

and testing it in hardware!



Approach: Theory of Anomaly Detection

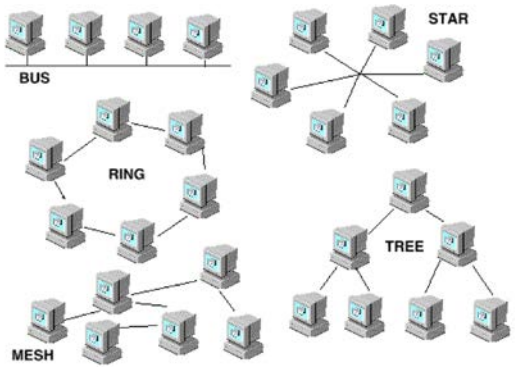
In Nature	vs.	Man-made Solutions
Distributed Knowledge 		Centralized Knowledge 
Known When Seen (like pornography) 		Defined-pattern matching/violation  ? 

Nature Translated into Computer Networks: (basically math)

- Each node analyzes only the packet stream routed through itself, but communicates to other nodes a level of 'concern'
- Each node uses a function of the levels of concern of its neighbor-nodes and its own analysis of packet data to update its sensitivity threshold
- Each node uses its own algorithm for anomaly detection with a dynamic sensitivity threshold for threat definition based on the above function

Approach: Simulation and Testing

Software simulation of nodes:



- Vary topologies
- Vary threats (at least DDOS, Stealth Scans, Worms, Botnets, maybe more)
- Vary the "background" traffic:
  - Rather than just trying to capture single patterns of traffic for particular domains, we are creating a portfolio of types of "normal" traffic patterns.
- Vary individual node anomaly detection algorithms
- Vary sensitivity of the functions of concern

Testing:

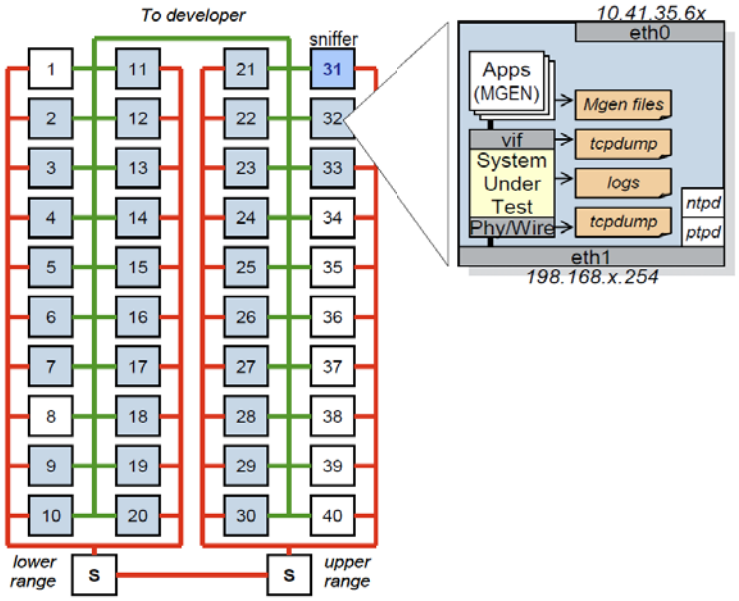
- Benchmarking our performance against existing centralized algorithms for sensitivity, specificity, and rapidity of threat detection

Approach: Hardware Testbed

Real life rarely works as well as software simulation.

In this case, realistic timing of packet transmission vs local decision making vs threat concern consensus communication is the crucial set of things that could behave very differently in real-world switches.

Details for Testbed provided in the Technical/ Demo Poster



Used to snapshot network statuses, control devices in real-time

Used to simulate various network topologies and link qualities

Benefits

- Caveat: this phase is basic research, we will not know if/how we succeed until later in the program (prototype testing/refining is planned for year 3)
- If successful,
  - We will be able to break the scalability problem in network-wide detection (especially true when networks have ***millions*** of devices rather than 100s of thousands of computers)
  - We will be able to detect ***new*** types of malfeasance for which signatures are not yet known.
  - We will be able to detect spreading malware at a speed ***faster*** than propagation rate.
  - All without new hardware in the network, but by extending existing switches and routers.

Current Status

Milestone/Deliverable	Status
Fully documented initial Bee Algorithm Design	Completed
Completed Software Environment for Testing	Completed
Defined Metrics/ Benchmarks of Success	Completed
Integrate Performance Testing of Bee Algorithm on Hardware Testbed	Delayed due to complications in finalizing contracts – new anticipated date 10/13
Completed 1st Pass Software Environment Testing for Bee Algorithm on Specific Anomaly Types	Completed
Completed Performance Evaluation for Revised Bee Algorithm on Specific Anomaly Types	Delayed due to the delay in evaluation on Testbed performance – new anticipated date 11/13

Next Steps: For the coming year

1) Refine BIAD to deliver 20% improvement over centralized detection algorithms in sensitivity, specificity, or rapidity without compromising more than 5% of performance along the other two axes for at least 3 threat types in the software simulation
2) Successful BIAD hardware emulation for at least ¾ of hardware test nodes
3) Successful demonstration of theoretical improvement (as defined in #1) for at least 2 threat types in real-world hardware testbed
4) Complete theoretical analysis of the impact of BIAD smart node density on anomaly detection performance for general networks
   – Needed for practical deployment to understand how much of a network needs to be replaced to achieve benefits from the system

Contact Information

Nina H. Fefferman
feffermn@dimacs.rutgers.edu

14 College Farm Road
Dept of Ecology, Evolution, and Natural Resources
(DIMACS & CCICADA)
Cook Campus, Rutgers University
New Brunswick, NJ
08901