Responses to Cyber Attacks in Distributed Systems

Saurabh Bagchi
The Center for Education and Research in Information Assurance and Security (CERIAS)
School of Electrical and Computer Engineering
Purdue University

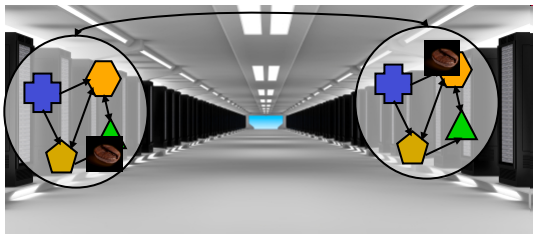
Supported by: NSF, Lockheed Martin, NEHRP

Joint work with: Eugene H. Spafford, Guy Lebanon

Outline

- **Problem Statement**
- Solution Directions
- Some Promising Solutions
- Ongoing Challenges

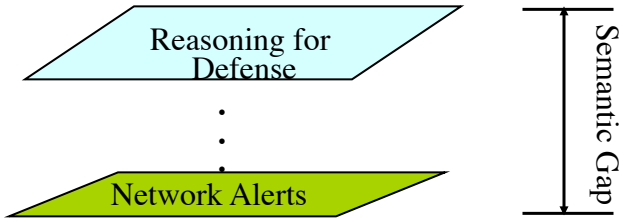
Defending Distributed Systems



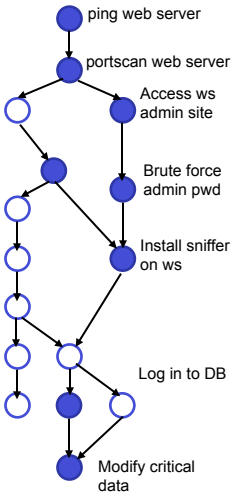
- Large-scale distributed systems to defend
  - Heterogeneous third-party services
- Lots of points for attacks
  - Lots of points to introduce cybersecurity mechanisms
- Interactions between the services allow for attack escalation

Drowning in a Sea of Alerts

- Large distributed systems get tons of alerts
  - Up to 20,000 per day
- Many of these are false alarms



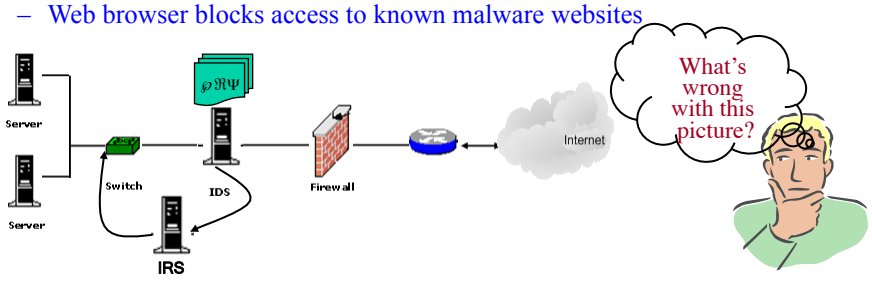
Fast Moving Attacks



- Multi-stage attacks
  - Compromise outward facing services
  - Use transitive trust and privilege escalation
  - Compromise internal services
  - Access crown jewels
- Attack progresses in machine time, rather than human time
- Examples: Worms and other self-propagating malware

Signature-based Responses

- Intrusion Response Systems (IRS) take reports from IDS and carry out actions to counter the intrusion
- Many examples of IRS
  - Anti-virus software disables access to worm executables or files infected with virus
  - Iptables which terminates a session on matching a malware signature
  - Web browser blocks access to known malware websites



Outline

- Problem Statement
- **Solution Directions**
- Some Promising Solutions
- Ongoing Challenges

Solution Directions

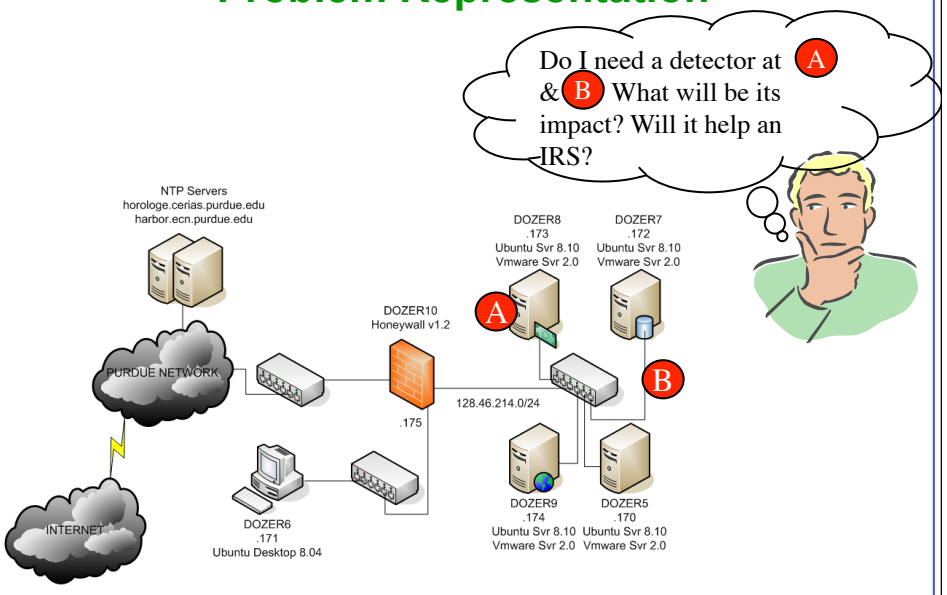
- We want to perform secure configuration and intrusion response in the face of threats that are fast-changing and therefore unknown

1. We want to learn from past behavior
   - But not overlearn
2. We want to grow our knowledge structures with runtime information
   - But not learn untruths
3. We want to perform the learning at runtime
   - This implies expensive batch mode processing is out
4. We do not want to rely only on signature-based security
   - Abstractions of attack steps are useful

Outline

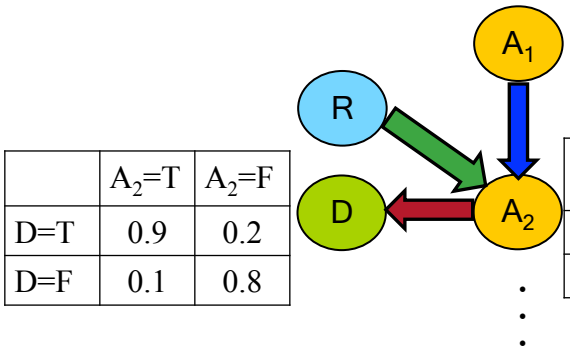
- Problem Statement
- Solution Directions
- **Some Promising Solutions**
- Ongoing Challenges

Problem Representation



Our Solution Approach: Detector Placement (SMARTS)

- Bayesian network used to model the causality in the network



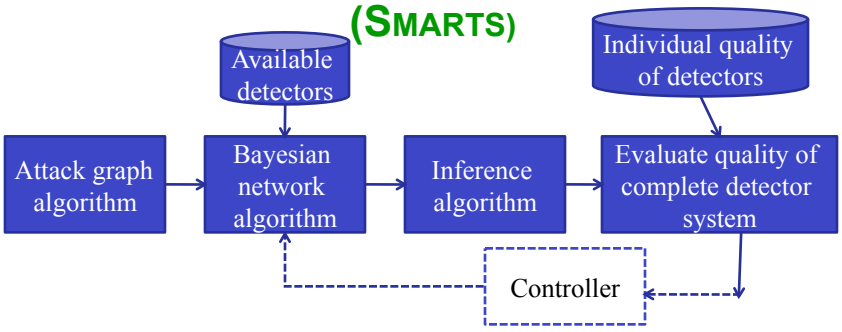
	A_2=T	A_2=F
D=T	0.9	0.2
D=F	0.1	0.8

	R=F, A_1=T	R=T, A_1=T	R=F, A_1=F	R=T, A_1=F
A_2=T	0.90	0.15	0.25	0.05
A_2=F	0.10	0.85	0.75	0.95

A Attack Step D Detector R Response

PURDUE UNIVERSITY

Our Solution Approach: Detector Placement (SMARTS)



- Inference on the Bayesian network performed through different choice and placements of detectors
- Heuristic-driven choice of one detector and its placement at a time
- Heuristic depends on individual detector quality and overlap with previously chosen detectors
- Controller to adjust detector setting when network changes

PURDUE UNIVERSITY

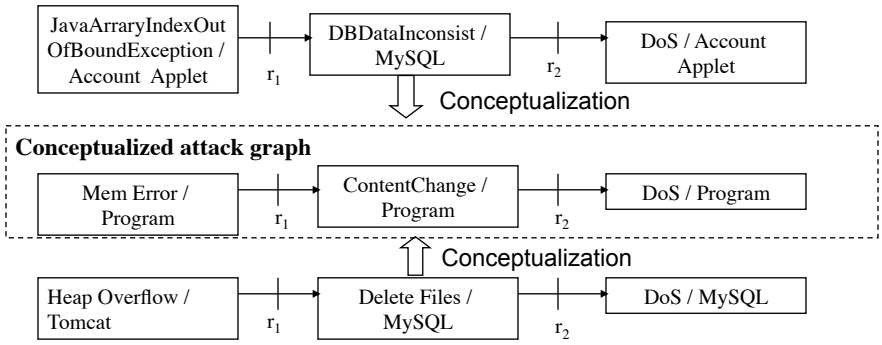
Our Solution Approach: Intrusion Response (ADEPTS)

- Short-term as well as long-term goals
  - Contain the current attack
  - Recover affected services to a functional state
  - Proactive defenses for future attacks
- Leverage distributed system's characteristics
  - Determine if the alert is false
  - Determine if the impact is worth responding to
- Learn from thy observations and mistakes
  - Calibrate prior responses
  - Learn characteristics of interactions in the system through past attacks
  - Quick customized responses to polymorphs of prior attacks

Responding to the Unknown

- Zero-day attack
  - Knowledge of the steps in the attack does not exist in the IRS
- Current solution: Take a drastic response, such as disconnecting the service
- Problem:
  - May be reacting to spurious alarms
  - Cannot learn from the spread of the attack
- Our solution approach:
  - Abstract the specifics of the attack
  - At a higher level of abstraction, map the attack to a previously seen attack
  - Use the learning on the previous attack to guide the responses for the current zero-day attack

Responding to the Unknown: Example



Responses: r_1: Disable connection from tomcat/applet to MySQL; r_2: Rollback to last data files checkpoint

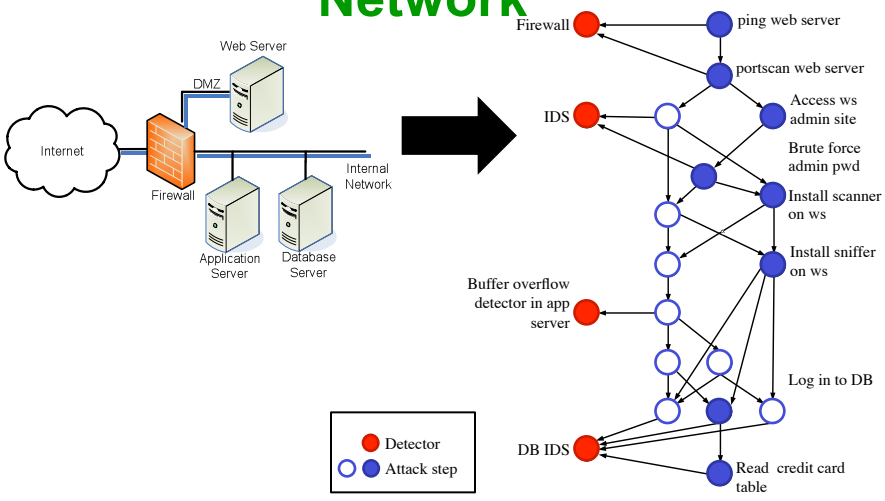
- Challenges: (1) High similarity does not necessarily give you the best response; (2) To what level should each node be conceptualized

Goals of Ongoing Work

- Secure Configuration Management
  - Detector placement is a specific example of security configuration
  - Tool should detect (when insecure configuration is introduced) and diagnose (which component has been mis-configured)
  - Tradeoffs exist between security of configuration and usability
  - Tool must not make arbitrary decisions on this spectrum
- Automated Intrusion Response
  - Resilience to zero-day attacks through more effective responses (i.e., less drastic than rebooting the servers)
  - Correlation of multiple detectors to increase confidence that an attack is underway before responding

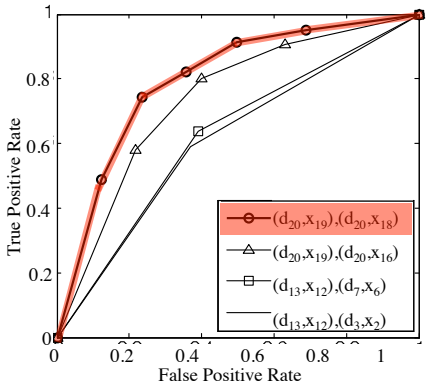
Backup Slides

Sample Network and Corresponding Bayesian Network



Impact on Choice and Placement of Detectors

- System: Three-tier web-based online service
- Objective: determine impact of selecting detectors and corresponding locations
- Performance of detector pair (selected from algorithm) is compared against randomly selected pairs



Selection made by our algorithm