Intrusion Detection/Response

- Denning: Systems under attack fail to meet one or more of the following characteristics:
  1. Actions of users/processes conform to statistically predictable patterns
  2. Actions of users/processes do not include sequences of commands to subvert security policy
  3. Actions of processes conform to specifications describing allowable actions
Intrusion Detection

- Idea: Attack can be discovered by one of the above being violated
  - Problem: Definitions hard to make precise
- Practical goal of intrusion detection systems:
  - Detect a wide variety of intrusions
  - Detect in a timely fashion
  - Present in a useful manner
  - Be (sufficiently) accurate

IDS Types: Anomaly Detection

- Compare characteristics of system with expected values
  - report when statistics do not match
- Threshold metric: when statistics deviate from normal by threshold, sound alarm
- Statistical moments: based on mean/standard deviation of observations
- Markov model: based on state, expected likelihood of transition to new states
Anomaly Detection: How do we determine normal?

- Capture average over time
  - But system behavior isn’t always average
- Correlated events
- Machine learning approaches

IDS Types: Misuse Modeling

- Does sequence of instructions violate security policy?
  - Problem: How do we know all violating sequences?
- Solution: capture *known* violating sequences
  - But won’t the attacker just do something different?
- Often, no
  - *kiddie scripts* Rootkit, …
- Alternate solution: State modeling
  - Known “bad” state transition from attack
  - Capture when transition has occurred (user → root)
Specification Modeling

- Does sequence of instructions violate system specification?
  - What is the system specification?
- Need to formally specify operations of potentially critical code
  - trusted code
- Verify postconditions met

IDS Architecture

- Similar to Audit system
  - Log events
  - Analyze log
- Difference: happens real-time
  - timely fashion
- (Distributed) IDS idea:
  - Agent generates log
  - Director analyzes logs
  - Notifier decides how to handle result
Where is the Agent?

- Host based IDS
  - watches events on the host
  - Often uses existing audit logs
- Network-based IDS
  - Packet sniffing
  - Firewall logs

IDS Problem: Base Rate Fallacy

- IDS useless unless accurate
  - Significant fraction of intrusions detected
  - Significant number of alarms correspond to intrusions
- Assume 99% of intrusions detected
  - 1% of non-intrusions generate alarm
  - 1 in 10,000 events really an intrusion
- 1% of alarms “real”
Intrusion Response

- Incident Prevention
  - Stop attack before it succeeds
  - Measures to detect attacker
    - Honeypot / Jailing
- Intrusion handling
  - Contain attack
  - Eradicate attack
  - Recover to secure state
  - Punish attacker

Containment

- Passive monitoring
  - Track intruder actions
  - Eases recovery and punishment
- Constraining access
  - Downgrade attacker privileges
  - Protect sensitive information
  - Why not just pull the plug?
Eradication

- Terminate network connection
- Terminate processes
- Block future attacks
  - Close ports
  - Disallow specific IP addresses
  - Wrappers around attacked applications

Follow-Up

- Legal action
  - Trace through network
- Cut off resources
  - Notify ISP of action
- Counterattack
  - Is this a good idea?