

Future Directions for Intrusion Detection



Douglas B. Moran

**Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park CA 94025**

SRI AIC: Current & Pending Work



n Practical Computer Security

n Apply AI Technology to Network-based Break-ins

- diagnosing
- repair and recovery: automated security manual
- future integration with intrusion detection

n Tool:

- manages other tools (esp OTS) for data collection
- rule-base to guide/advise user

Procedural Reasoning System



- n **Based on *Acts* (procedures) instead of rules**
- n **Declarative acts and meta-acts**
- n **Based on Beliefs-Desires-Intentions model**
- n **Pre-emptable scheduling**
- n **Supporting tools (eg editor)**

Related Research: Multi-Agent Systems



- n **Community of Agents**
- n **Plug-and-play applications**
- n **Distributed**
- n **User Interface as agents**
 - **Human in the loop**
 - **User as alternative info source**
 - **Access from variety of platforms: desktop, PDA, phone, ...**

Doctrine



- n **Standard Procedures / Training Curriculum**
- n **Largely ineffective in dealing with intrusion**
 - **inadequate training: too much to learn**
 - » **too many threats**
 - » **too many tactics**
 - **rapid evolution**
- n **Automated tools to handle most of doctrine**
- n **Training simulators using statement of doctrine**

Distributed Attacks

Collaborative Defense



For computers under same management:

n Current Tactics: distributed probing & actions

- Distributed Filesystem
 - » probe from multiple hosts within cluster
 - » individual actions below threshold

n Centralized IDS: can it scale up?

n Alternative: local IDS's that query and notify siblings about suspicious events

- parsimonious
- is sibling host still trustworthy?

Encryption of Network Traffic



- n **Widespread use in few years (5 years?)**
- n **Minimizes info available to network monitor**
- n **Each host has to collect raw data on its connections**
 - send to network monitor
 - can these reports be trusted?

Intrusions at Unrelated Sites



- n **Trend: use/depend on data and services from external sources**
 - security policy & capability of other site: unknown, different, none, ...
 - intrusion at any of these other sites can be a virtual intrusion at yours

- n **Issue: how to estimate security status of another**
 - sharing info often not in interest of site
 - » useful to intruder
 - » against organizational self-interest (esp commercial)

Many Sites Many Security Policies



- n **Trend: external sources represent wide range of organizations with wide range of requirements towards security**
- n **Each site must have substantial freedom and capability to specify own policy**
 - all-or-nothing = nothing
 - support tools critical
 - provide baseline policy (augmented by each level of hierarchy)
- n **Dimensions**
 - type of attack
 - category of attacker: capabilities and resources
 - tolerance for loss

Decentralized, Collaborative Intrusion Detection



- n **How much info is enough**
- n **Could this info help intruder more than defender**
- n **Tradeoff: local protection *vs* greater body**
- n **Dimensions:**
 - security policy
 - current events

Clearing Houses



- n **CERT, ASSIST (DISA), CIAC (DOE), ...**
- n **improve quantity and quality of reports**
 - improved reliability
 - support more automated processing and correlation
- n **automated tool for report creation: takes output of tools for detection and analysis**

Cooperating Agents



- n **Info from unsecure or questionable sources**
- n **Negotiation/bidding**
- n **Beliefs, Desires, Intentions (BDI)**
 - Application: robot control
- n **Security of remote system as**
 - cost
 - disadvantage

Security Policy: Repair and Recovery



- n **Tool to advise on risks, prioritize actions, undertake directed actions**
- n **Tradeoff: resuming secure operation *vs* tracking down intruder**
- n **Tradeoff: continued operation *vs* recovery**

Goals of Intruders



- n **Exploration (curiosity or planning damage)**
- n **Resource theft: computation, information**
- n **Disruption**
 - slowdown, shutdown
 - scramble priorities
 - corrupt information
- n **Disinformation**
 - influence down specific path
 - divert into low productive areas (eg responding to slander)

What is Intrusion Detection



- n **Define limits of legitimate use of system**
- n **Phases: detection, diagnosis/analysis, repair**
- n **Discovery**
 - events that are unexplained, anomalous, or against policy
 - data fusion and interpretation
 - data from multiple computers and sites
- n **Repair/Recovery**
 - Identify risks
 - Prioritize repair actions

Capabilities of Intrusion Detection



- n **Detect legitimate user performing improper acts**
- n **Dynamically customize for**
 - site, computer, user
 - enemy, threat, tactics

Categories of Intruders



n Dimensions

- motivation/intentions/goals
- capabilities: expertise and resources
- vulnerability to retaliation and countermeasures

n Example: by size

- individuals and small groups: terrorist, criminals
- middle: major criminal/terrorist orgs, commercial, LDC
- large: countries
- XL: major powers

Types of Attacks



- n **Impersonation of normal user**
- n **Privileged access**
- n **Disruption by starving critical applications using innocuous applications**
- n **Probing attacks that are meant to be detected**
- n **Discourage use of data and services**

Technology Change



- n **New network technologies: ?**
- n **More client-server**
- n **More connection-less services**
- n **Multi-Agent Systems**
 - remote/mobile
 - federated
- n **Encryption of network traffic**

Tripwires



- n **Tripwires and honeypots to simplify monitoring**
- n **Cascading level of tools/expertise activated**
 - **Tradeoff: miss stealthy attack vs reduce consumption of CPU**
 - **Implementation: community of agents**
- n **Can formal specification of policy & doctrine be used to automatically create tripwires??**

Security Policy: Normal vs Crisis



- n **Data collection will occur during normal times**
- n **Simulation of crisis is of limited utility**
- n **ID in normal times: terrorists, criminals, sneak attacks**
- n **ID in crisis:**
 - Not hamper taking initiative
 - redeployment
 - changed job mix: planned, forced, adaptive
 - Alert rates must be at reasonable, but elevated, level

Security Policy: Propagating Changes



- n **During normal operation and during crisis, intruders will develop new tactics**
- n **Propagate rapidly**
- n **Propagate confidentially (if reasonable): aid catching intruders**
- n **Propagate securely: don't become vehicle for intrusions**

Human Interaction



- n **Human interaction critical: too much offline info involved in decisions**
- n **Not in most loops: tools to handle predictable situations**
- n **Human unavailable or overloaded: tools to handle default actions**