
Intrusion Detection

ISA 774
George Mason University
Spring 2006

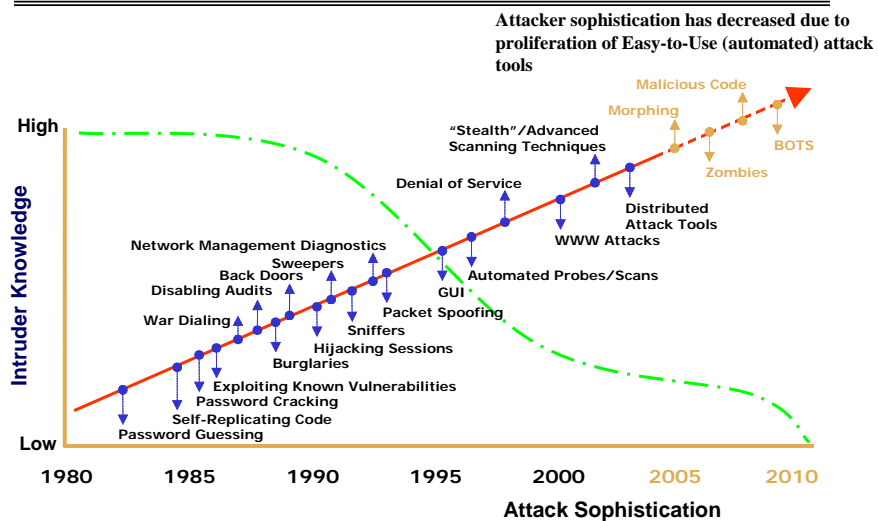
Week 1 Review

Why are Information Systems Vulnerable to Attack?

- The architecture of the Internet makes it susceptible to certain types of attacks
- Interdependent nature of Internet users
- Finite resources
- The National Research Council found in 1999 that:
 - No mechanisms or systematic design methods exist for defending against DoS
 - The ad hoc methods that were useful for defending timesharing systems from DoS are unsuitable for defending network systems

Trust in Cyberspace, Fred B. Schneider, Editor; Committee on Information Systems Trustworthiness, National Research Council (1999), <http://books.nap.edu/books/0309065585/html/index.html>

Threat Trends



What is Information System Security??

- Traditionally, there are three major goals of Information Security:
 - **Confidentiality**
 - The assets are accessible only by authorized parties.
 - **Integrity**
 - The assets are modified only by authorized parties, and only in authorized ways.
 - **Availability**
 - Assets are accessible to authorized parties.

GMU ISA 774 Spring 2006

5

Security Terminology

- Assets – hardware, software, data, people
- Vulnerability
 - a weakness in a system that may make it vulnerable to disruption or attack
- Threats
 - An event that has the potential to cause loss or harm
- Attack
 - A deliberate exploitation of a vulnerability
- Exposure
 - a form of possible loss or harm
- Control – a protective measure

GMU ISA 774 Spring 2006

6

Types of Security Breaches

- Interruption
 - Example: DOS (Denial of Service)
- Interception
 - Peeping eyes
- Modification
 - Change of existing data
- Fabrication
 - Addition of false or spurious data

GMU ISA 774 Spring 2006

7

Computing System Vulnerabilities

- Hardware vulnerabilities
- Software vulnerabilities
- Data vulnerabilities
- Human vulnerabilities ?

GMU ISA 774 Spring 2006

8

People Involved in Computer Crimes

- Amateurs (Script Kiddies?)
- Crackers (Hackers???)
- Career Criminals
- Trusted Insiders
 - Employees
 - Developers
 - Consultants

GMU ISA 774 Spring 2006

9

Methods of Defense

- Encryption
- Software controls
- Hardware controls
- Policies
- Physical controls

GMU ISA 774 Spring 2006

10

Policies

- Guidelines, rules of behavior, standards, etc to provide
- Policy controls can be simple but effective
 - Example: frequent changes of passwords
- Legal and ethical controls
 - Gradually evolving and maturing

GMU ISA 774 Spring 2006

11

Principle of Effectiveness

- Controls must be used to be effective.
 - Efficient
 - Appropriate for the security need
 - Easy to use (ease of *effective* use)
 - Auditable
- The **principle of adequate protection**

GMU ISA 774 Spring 2006

12

Overlapping Controls

- Several different controls may apply to one potential exposure.
- Commonly referred to as **Defense in Depth**

Week 2 Review

Early Intrusion Detection

- System administrator manually monitor user's activity
- Ad hoc and non-scalable
- Clifford Stoll's Cuckoo's Egg
- The Study of Intrusion Detection
 - James P. Anderson's 1980 technical report
 - "Computer Security Threat Monitoring and Surveillance"
- Anderson
 - Introduced the notion of audit trails
 - Suggested that audit trails contain vital information that could be valuable in tracking misuse and understanding user behavior
 - Formed foundation of host-based intrusion and IDS in general

Anderson's Threat Matrix

	Not authorized to use data/program	Authorized to use data/program
Not authorized to use computer	Case A: External	
Authorized to use computer	Case B: Internal	Case C: Misfeasance

Audit Records

- Audit Records
 - Fundamental tool for intrusion detection
 - Native audit records
 - Detection-specific audit records
- Audit Record Analysis
 - Foundation of statistical approaches
 - Analyze records to get metrics over time
 - Use various tests on these to determine if current behavior is acceptable
 - Key advantage is no prior knowledge used

GMU ISA 774 Spring 2006

17

Security Problems

- Causes of Security Problems
 - System design and development
 - System management
 - Trust allocation
- What is an intrusion?
 - Any set of actions that attempt to compromise the confidentiality, integrity, or availability of a computer resource
- Types of Violations
 - Attack
 - Intrusion
 - Misuse

GMU ISA 774 Spring 2006

18

Intruders

- **Classes of Intruders:**
 - masquerader
 - misfeasor
 - clandestine user
- **Basic attack methodology**
 - target acquisition and information gathering
 - initial access
 - privilege escalation
 - covering tracks

GMU ISA 774 Spring 2006

19

Intrusion Detection

- **Must detect intrusions so you can:**
 - Block if detected quickly
 - Act as deterrent
 - Collect info to improve security
- **Differentiating ID Systems**
 - Monitoring strategy
 - Analysis Type
 - Timing
 - Detection Goals
 - Control

GMU ISA 774 Spring 2006

20

Intrusion Detection

- Monitoring Strategy
 - Host
 - Network
 - Application
 - Target-based
- Intrusion Detection Approaches
 - Statistical anomaly detection
 - Signature-based detection
 - Rule-based detection

GMU ISA 774 Spring 2006

21

Intrusion Detection

- Performance Issues
 - False positive rates
 - False negative rates
 - Data volume

GMU ISA 774 Spring 2006

22

Week 3 Review

GMU ISA 774 Spring 2006

23

Data Sources Issues

- Are we collecting the right information?
 - Does it permit identification of violations?
- How much information is enough?
- Where to collect?
 - Host versus network?
- How do you handle the data to use as evidence?
- How can you format for interoperability?
 - IDMEF: XML-based message format

GMU ISA 774 Spring 2006

24

Host Based Information Sources

- Host based IDS was the first method of detecting computers
- Operating System Audit Trails
 - A collection of information about system activities
 - Generated by the O/S Kernel
 - Audit *tokens, records, files*
 - Different vendors developed different audit mechanisms

GMU ISA 774 Spring 2006

25

Audit Trails

- Structure
 - Self-contained
 - Compact file
- Audit System Problems
 - Not designed for IDS
 - May generate too much unnecessary data
 - May not generate enough data
- Pro
 - Protected by the O/S (usually)
 - Fine grained detail
 - Hard to create and insert fake records
- Con
 - High volume of data hard to analyze
 - May mask “set uid” activity

GMU ISA 774 Spring 2006

26

Audit Trails

- Content of Audit Trails
 - Kernel (system) level events
 - User (application) level events
 - Initiating process
 - Used ID
 - Time
 - Event description of code
- Audit Reduction
 - Filtering audit logs to reduce or eliminate redundant or unnecessary information
 - Try to introduce determinism where there is very little!
 - Audit systems may not generate the same log records for 2 identical (?) events
 - May be able to reduce log entries from “trusted” processes

GMU ISA 774 Spring 2006

27

System Logs

- Provided as additional features of systems
- Usually run as an application, rather than as part of the kernel
- Sometimes seen as less trustworthy than audit files
 - Syslog?
- Usually easier to review
- Developed for specific needs, like IDS
- Typical System Log Uses
 - Successful and unsuccessful logins
 - Active users
 - Active processes
 - Attempts to use root privilege

GMU ISA 774 Spring 2006

28

Other Host-Based Sources

- Application logs provide a view of “user space”
 - Criticality of the application may determine the usefulness of the log
- Target-Based Monitoring
 - A static approach
 - Like taking a picture instead of a video
 - Normally take the form of integrity checkers
 - Use cryptographic hash to store an object’s “state” for later comparison

GMU ISA 774 Spring 2006

29

Network-Based Information Sources

- Part of the security perimeter
- Doesn’t impact host performance
- Can be transparent to attacker
- Can detect attacks host based systems may miss
- Based on identifying and interpreting packet data
 - Berkeley Packet Filter
 - Libpcap
 - TCPDump
 - Streams (Solaris)

GMU ISA 774 Spring 2006

30

Other Information Sources

- Network Devices
 - Firewalls
 - Routers/Switches
 - Access Control systems
- Out-of-Band Information Sources
 - Physical security system
 - Caller ID ?
- Other System Components as Data Sources

GMU ISA 774 Spring 2006

31

Host-Based v/s Network-Based IDS

- Host-based IDS
 - Verifies success or failure of an attack
 - Monitors specific system activities
 - Detects attacks that network based systems miss
 - Well-suited for encrypted and switched environments
 - Near-real-time detection and response
 - Requires no additional hardware
 - Lower cost of entry

GMU ISA 774 Spring 2006

32

Host-Based v/s Network-Based IDS ...contd.

- Network-Based IDS
 - Lower cost of ownership
 - Detects attacks that host-based systems miss
 - More difficult for an attacker to remove evidence
 - Real-time detection and response
 - Detects unsuccessful attacks and malicious intent
 - Operating system independence
 - Performance issues

GMU ISA 774 Spring 2006

33

Host-Based v/s Network-Based IDS

- | | |
|--|---|
| <ul style="list-style-type: none">• Host-based IDS<ul style="list-style-type: none">– Verifies success or failure of an attack– Monitors specific system activities– Detects attacks that network based systems miss– Well-suited for encrypted and switched environments– Near-real-time detection and response– Requires no additional hardware– Lower cost of entry | <ul style="list-style-type: none">• Network-Based IDS<ul style="list-style-type: none">– Lower cost of ownership– Detects attacks that host-based systems miss– More difficult for an attacker to remove evidence– Real-time detection and response– Detects unsuccessful attacks and malicious intent– Operating system independence– Performance issues |
|--|---|

GMU ISA 774 Spring 2006

34

Week 4 Review

GMU ISA 774 Spring 2006

35

Thinking About Intrusions

- Defining Intrusion Detection Analysis
 - Organizing and characterizing data about user and system activity to identify activity or events of interest
 - Can be done in real time or after the fact
- Goals of ID Analysis
 - Deterrence
 - Quality control of security
 - Gather information about intrusions

GMU ISA 774 Spring 2006

36

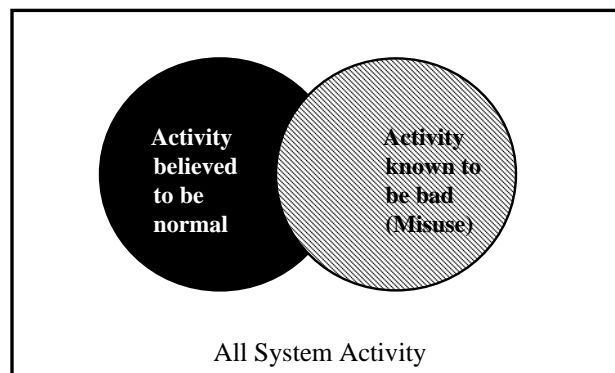
Detecting Intrusions

- Users, administrators and operators
- External Events
- Precursors to Intrusion
- Intrusion Artifacts
- Real time Observations

GMU ISA 774 Spring 2006

37

Misuse vs. Anomaly Detection



GMU ISA 774 Spring 2006

38

A Model for Intrusion Analysis

- Analysis is the core function of intrusion detection (systems)
- Three primary tasks:
 - Constructing the Analyzer
 - Performing Analysis
 - Feedback and Refinement
- The first two tasks have 3 identical “data processing” steps
- Analysis engine forms the core of the IDS
- Tailoring for a specific environment or type of threat expected is required for acceptable performance

GMU ISA 774 Spring 2006

39

A Model for Intrusion Analysis

- Constructing the Analyzer
 - Collect Event Information
 - Preprocess the Information
 - Build a Classification Model
 - Populate the Model with Event Data
 - Store the Populated Model for Use
- Performing Analysis
 - Input new event records
 - Preprocessing
 - Compare the Event to the Knowledge Base
 - Generate a Response
- Feedback and Refinement

GMU ISA 774 Spring 2006

40

Week 5 Review - Intrusion Detection Analysis Techniques

GMU ISA 774 Spring 2006

41

Techniques for Misuse Detection

- Expert Systems
 - Control reasoning is entered as *if-then* rules
 - Inputs are evaluated by the rules
 - Separates the engine from the rules
- Issues with Expert Systems
 - Don't scale well
 - Can only detect known intrusions
 - Can't handle uncertainty

GMU ISA 774 Spring 2006

42

Finite State Machines

- A *model of computation* consisting of a set of *states*, a *start state*, an input *alphabet*, and a *transition function* that maps input symbols and current states to a *next state*.
- Computation begins in the start state with an input string. It changes to new states depending on the transition function.
- Example: *Kosoresow and Hofmeyr*

GMU ISA 774 Spring 2006

43

Petri Nets

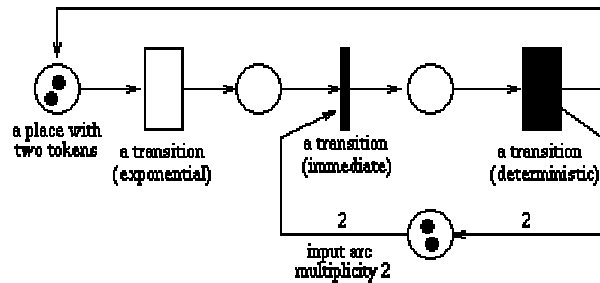
- A Petri net is a graphical and mathematical modeling tool. It consists of *places*, *transitions*, and *arcs* that connect them.
- Colored Petri Nets
 - Colored Petri Nets (CP-nets or CPN) are a graphical oriented language for design, specification, simulation and verification of systems. It is in particular well-suited for systems in which communication, synchronization and resource sharing are important.
 - CP-nets combine the strengths of ordinary Petri nets with the strengths of a high-level programming language.

GMU ISA 774 Spring 2006

44

Petri Nets

Example of a Petri Net



GMU ISA 774 Spring 2006

45

Language Based Approaches

- Optimized languages for developing IDS applications
 - Russell
 - Stalker
 - NFR N-Code
- IDES Model
 - Operational Model
 - Mean and Standard Deviation Model
 - Multivariate Model
 - Markov Process Model

GMU ISA 774 Spring 2006

46

Quantitative Analysis

- Quantitative analysis can be the basis for intrusion signatures and anomaly statistical models
 - Threshold Detection
 - Heuristic Threshold Detection
 - Target Based Integrity Checks
- Quantitative Analysis and Data Reduction

GMU ISA 774 Spring 2006

47

Techniques for Anomaly Detection

- Statistical Measures
 - Strengths
 - Detects intruders masquerading as legitimate users
 - “Might” be able to detect unknown attacks
 - Don’t require constant updating of rules or signatures
 - Drawbacks
 - More adapted to batch mode analysis
 - Selecting the right metrics and thresholds is complex
 - High false alarm rates

GMU ISA 774 Spring 2006

48

Techniques for Anomaly Detection

- Nonparametric Statistical Measures
 - Analysis of data without assuming an underlying distribution
 - Good for data reduction
- Neural Networks
 - Uses ANNs to perform more adaptive learning about the environment and events
 - Still has high false alarm rate

GMU ISA 774 Spring 2006

49

Agent Based Detection

- AAFID
 - Autonomous agents that report to a monitor
 - Multiple agents may reside on one host
 - Transceiver (monitor) performs data reduction
- Emerald
 - Deploys distributed agents to hosts
 - Example: NetBSM

GMU ISA 774 Spring 2006

50

Agent Based Detection

- Advantages
 - Agents can collaborate
 - More resistant to evasion
 - Scalable
- Disadvantages
 - Agent processing may impact host performance
 - Propagation delay from agent to alert
 - Agents become *targets*

GMU ISA 774 Spring 2006

51

Data Mining

- Extracting meaningful data from large amounts of data
 - Classification
 - Link Analysis
 - Sequence Analysis

GMU ISA 774 Spring 2006

52

Week 6 Review - Responses

GMU ISA 774 Spring 2006

53

Responses

- Three Categories of IDS Users
 - Security Managers
 - Systems Administrators
 - Investigators
- Each has their own level of expertise or knowledge domain
- Each has different requirements of the system

GMU ISA 774 Spring 2006

54

Requirements for Responses

- Operational Environment
- System Purpose and Priorities
- Regulatory or Statutory Requirements
- Ease of Use

GMU ISA 774 Spring 2006

55

Types of Responses

- Active Responses
 - Take Action Against the Intruder
 - Amend the Environment
 - Collect Additional Information
- Passive Responses
 - Alarms and Notification
 - SNMP Traps and Plug-Ins
- Control Types for active Response
 - User Driven Responses
 - Automatic Responses

GMU ISA 774 Spring 2006

56

Covering Tracks During Investigation

- IDS benefits from being unobserved by attackers
- Fail-Safe Response Components
- Handling False Alarms
- Archiving and Reporting

GMU ISA 774 Spring 2006

57

Mapping Responses to Policy

- Once again, a robust security policy should dictate what activities are required of the IDS and responses
- Base identifies 4 activity categories based on timeliness and criticality:
 - Immediate
 - Timely
 - Long-Term Actions – Local
 - Long-Term Actions – Global

GMU ISA 774 Spring 2006

58

Mapping Responses to Policy

- Immediate
 - Critical activities performed when discovering an incident
- Timely
 - Activities that may be performed hours or days after an incident
- Long-Term Actions – Local
 - Less critical to handling the incident, but still important to an effective organization
- Long-Term Actions – Global
 - Not critical for incident, but still important from an overall security perspective

GMU ISA 774 Spring 2006

59

Intrusion Detection Errors

- **False negatives:** attack is not detected
 - Big problem in signature-based misuse detection
- **False positives:** harmless behavior is classified as an attack
 - Big problem in statistical anomaly detection
- Both types of IDS suffer from both error types
- Which is a bigger problem?
 - Attacks are fairly rare events
 - IDS often suffer from [base-rate fallacy](#)

GMU ISA 774 Spring 2006

60

Conditional Probability

- Suppose two events A and B occur with probability $\Pr(A)$ and $\Pr(B)$, respectively
- Let $\Pr(AB)$ be probability that both A and B occur
- What is the **conditional probability** that A occurs assuming B has occurred?

$$\Pr(A | B) = \frac{\Pr(AB)}{\Pr(B)}$$

GMU ISA 774 Spring 2006

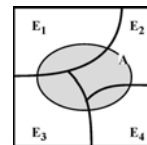
61

Bayes' Theorem

- Suppose mutually exclusive events E_1, \dots, E_n together cover the entire set of possibilities
- Then probability of any event A occurring is

$$\Pr(A) = \sum_{1 \leq i \leq n} \Pr(A | E_i) \cdot \Pr(E_i)$$

- Intuition: since E_1, \dots, E_n cover entire probability space, whenever A occurs, some event E_i must have occurred



- Can rewrite this formula as

$$\Pr(E_i | A) = \frac{\Pr(A | E_i) \cdot \Pr(E_i)}{\Pr(A)}$$

GMU ISA 774 Spring 2006

62

Base-Rate Fallacy

- 1% of traffic is SYN floods; IDS accuracy is 90%
 - IDS classifies a SYN flood as attack with prob. 90%, classifies a valid connection as attack with prob. 10%
- What is the probability that a valid connection is erroneously flagged as a SYN flood by the IDS?

$$\begin{aligned}
 \Pr(\text{valid} \mid \text{alarm}) &= \frac{\Pr(\text{alarm} \mid \text{valid}) \cdot \Pr(\text{valid})}{\Pr(\text{alarm})} \\
 &= \frac{\Pr(\text{alarm} \mid \text{valid}) \cdot \Pr(\text{valid})}{\Pr(\text{alarm} \mid \text{valid}) \cdot \Pr(\text{valid}) + \Pr(\text{alarm} \mid \text{SYN flood}) \cdot \Pr(\text{SYN flood})} \\
 &= \frac{0.10 \cdot 0.99}{0.10 \cdot 0.99 + 0.90 \cdot 0.01} = 92\% \text{ chance raised alarm is false!!!}
 \end{aligned}$$

GMU ISA 774 Spring 2006

63

Base Rate Fallacy in Intrusion Detection

- I: intrusive behavior,
 - I: non-intrusive behavior
 - A: alarm
 - A: no alarm
- Detection rate (true positive rate): $P(A|I)$
- False alarm rate: $P(A|–I)$
- Goal is to maximize both
 - Bayesian detection rate, $P(I|A)$
 - $P(–I|–A)$

GMU ISA 774 Spring 2006

64

Detection Rate vs False Alarm Rate

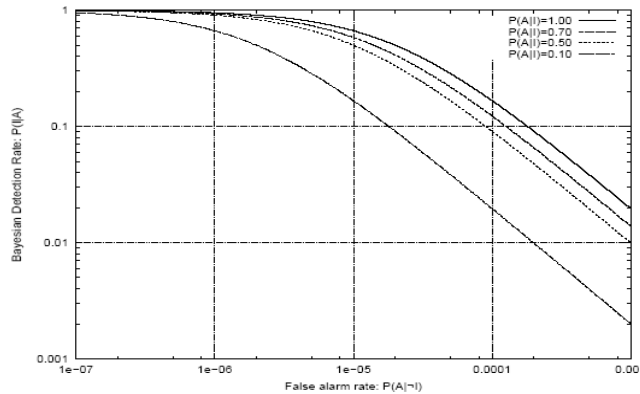
$$P(I|A) = \frac{P(I) \cdot P(A|I)}{P(I) \cdot P(A|I) + P(-I) \cdot P(A|-I)}$$

- Suppose: $P(I) = 1 / \frac{1 \cdot 10^6}{2 \cdot 10} = 2 \cdot 10^{-5};$
 $P(-I) = 1 - P(I) = 0.99998$
- Then: $P(I|A) = \frac{2 \cdot 10^{-5} \cdot P(A|I)}{2 \cdot 10^{-5} \cdot P(A|I) + 0.99998 \cdot P(A|-I)}$
- False alarm rate becomes more dominant if P(I) is very low

GMU ISA 774 Spring 2006

65

Detection Rate vs False Alarm Rate



Axelsson: We need a very low false alarm rate to achieve a reasonable Bayesian detection rate

GMU ISA 774 Spring 2006

66