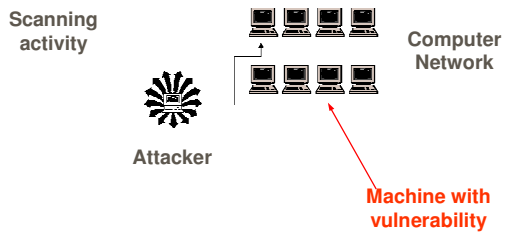
Intrusion Detection Systems

Costantina Caruso
Dipartimento di Informatica
Università degli Studi di Bari
16th January 2004

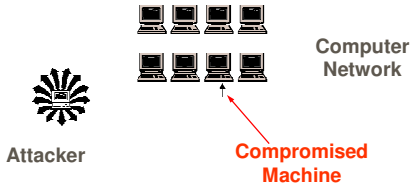
Typical intrusion scenario: phase 1



Introduction: the matter

- Due to the enormous diffusion of Internet, more and more organizations are becoming vulnerable to potential cyber attacks, such as network intrusions
- Both the sophistication and the severity have also notably increased (the most recent and extraordinary case was the SQL slammer worm on Jan 25th, 2003)
- Statistics of the incidents reported by Cert-CC (Computer Emergency Response Team- Coordination Center) and by Mitre Corporation

Typical intrusion scenario: phase 2



Cyber attacks

- Cyber attacks are actions that attempt to bypass security mechanism of computer systems.

  They are caused by:

  - Attackers accessing the system from Internet
  - Insider attackers (authorized users attempting to gain and misuse non-authorized privileges)

Why we need Intrusion Detection?

- Security mechanisms always have inevitable vulnerabilities
- Current firewalls are not sufficient to ensure security in computer networks
  - "Security holes" caused by allowances made to users/programmers/administrators
  - Insider attacks

What is Intrusion Detection?

- **Intrusion Detection**: ID is the process of monitoring the events occurring in a computer system or network and analyzing them for signs of intrusions, defined as the attempts to bypass the security mechanism of a computer or network
- **Intrusion Detection Systems (IDS):**
  - combination of software and hardware that attempts to perform intrusion detection
  - raise the alarm when possible intrusion happens

Types of computer attacks 2/2

- DoS **(Denial of Service) attacks**
  - DoS attacks attempt to shut down a network, computer, or process, or otherwise deny the use of resources or services to the authorized users
  - Distributed DoS attacks
- Probe **(scanning) attacks**
  - Attacker uses network services to collect information about a host (e.g. list of valid IP addresses, what services it offers, what is the operating system)
- Trojan horses / worms **–** attacks that are aggressively replicating on other hosts (worms – self-replicating; Trojan horses are downloaded by users)

Taxonomy of computer attacks

- **Intrusions can be classified according to several categories:**
  - **Attack type (Denial of Service (DoS), Scan, worms/trojan horses, compromises (R2L, U2R), …)**
  - **Number of network connections involved in the attack**
    - single connection cyber attacks
    - multiple connections cyber attacks
  - **Source of the attack**
    - multiple vs. single
    - inside vs. outside
  - **Environment (network, host, P2P, wireless networks, …)**
  - **Automation (manual, automated, semi-automated attacks)**

Number of connections involved in attacks

- **Generally two types of cyber attacks in the computer networks:**
  - attacks that involve multiple network connections (bursts of connections; eg. netscanning)
  - attacks that involve single network connections (eg. to compromise a selected machine with a particular bug)

Types of computer attacks 1/2

- Compromises **-** attackers use known vulnerabilities such as buffer overflows and weak security to gain privileged access to hosts
  - R2L (Remote to Login) attacks - attacker who has the ability to send packets to a machine over a network (but does not have an account on that machine), gains access (either as a user or as a root) to the machine and does harmful operations
  - U2R (User to Root) attacks - attacker who has access to a local account on a computer system is able to elevate the own privileges by exploiting a bug in the operating system or a program that is installed on the system

Source of computer attacks

- **Attacks may be launched from single location or from several different locations**
- **Attacks may be also targeted to single or many different destinations**
- **Need to analyze network data from several sites in order to detect these distributed attacks.**
  - Single source attacks
  - Distributed/Coordinated attacks

Environment of computer attacks

- **Attacks may be categorized according to the environment where they occur**
- **Network intrusions** (intrusions in computer networks)
- **Intrusions on the host machine** (single computers)
- **Intrusions in P2P environment**
  - connected computers act as peers on the Internet, nothing else than clients
  - they are cut off from the DNS system since they do not have fixed IP address, and therefore difficult to trace the attack source
- **Intrusions in wireless networks**
  - Physical layer is less secure than in fixed computer networks
  - Mobile nodes do not have fixed infrastructure
  - There are no traffic concentration points where packets can be monitored

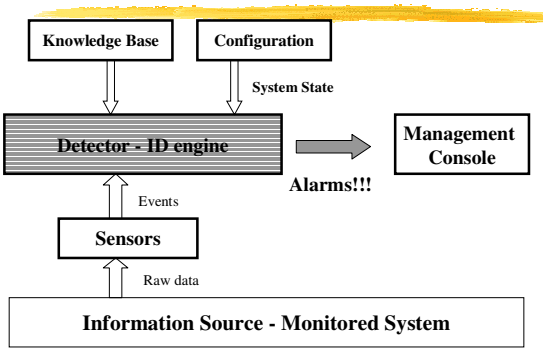
Intrusion detection taxonomy

- **Information source**
  - host-based ID, network-based ID, wireless-network ID, application logs, sensor alerts
- **Time aspects in analysis**
  - Real-time analysis vs. off-line analysis
- **Architecture**
  - Single centralized vs. distributed & heterogeneous
- **Activeness**
  - Active reaction vs. passive reaction
- **Continuality**
  - Continuous analysis vs. periodic analysis
- **Analysis strategy**
  - **Anomaly detection vs. misuse detection**
    - Data mining approach vs. traditional techniques

Automation of computer attacks

- **Wide-spread availability of automated tools, often used by "script kiddies"**
- **These attacker tools are capable of scanning a large part of the Internet in a short time period**
  - Automated attacks use these tools
  - Semi-automated (the attacker deploys automated scripts for scanning and compromise of network machines and installation of attack code). Attacker then uses the handler (master)machines to specify the attack type and victim's address
  - Manual (the attacker scans machines manually, not used often nowadays)

Basic IDS model



Difficulties in detection intrusion

- **Attacks Stealthiness**
  - **Attackers tries to hide their actions from either an individual who is monitoring the system, or an IDS**
    - Cover their tracks by editing system logs
    - Reset a modification date on a file that they replaced modified
  - **Novel Intrusions**
    - Undetectable by signature based IDSs
    - Should be detected as anomalies by observing significant deviations from the normal network behavior
  - **Distributed/coordinated attack**
    - Need for attack correlation

IDS information source

- **Host-based IDS**
  - base the decisions on information obtained from a *single host* (e.g. system log data, system calls data)
- **Network-based IDS**
  - make decisions according to the information and data obtained by *monitoring the traffic in the network* to which the hosts are connected
- **Wireless network IDS**
  - detect intrusions by analyzing traffic between mobile nodes
- **Application Logs**
  - detect intrusions analyzing for example database logs, web logs
- **IDS Sensor Alerts**
  - analysis on low-level sensor alarms
  - Analysis of alarms generated by other IDSs

IDS - Architecture

- **Centralized IDS**
  - Data analysis is performed in a fixed number of locations, independent of how many hosts are being monitored
- **Distributed IDS**
  - Data analysis is performed in a number of locations proportional to the number of hosts that are being monitored
  - Necessary for detection of distributed/coordinated attacks targeted at multiple networks/machines

IDS - Time aspects in analysis

- **Real-time IDS**
  - Analyzes the data while the sessions are in progress (e.g. network sessions for network intrusion detection, login sessions for host based intrusion detection)
  - Raises an alarm immediately when the attack is detected
- **Off-line IDS**
  - Analyzes the data when the information about the sessions are already collected –post-analysis
  - Useful for understanding the attackers' behavior

IDS - Activeness

- **Passive reaction**
  - Merely generates the alarms for the attacks
  - No countermeasure is actively applied to thwart the attack
- **Active response on attack detection**
  - Corrective response (closing security holes, reconfiguring firewalls, routers and switches)
  - Pro-active (logging out attackers, turning off IP addresses, terminating network connections)
    - NetProbe (cut network connections)
    - CISCO Net Ranger (reconfigures routers and switches)
    - Ballista (shutdowns vulnerable services, modifies configuration files, …)

Measures for evaluating IDS

- ***Detection rate*** **- ratio between the number of correctly detected attacks and the total number of attacks**
- ***False alarm (false positive)*** **rate - ratio between the number of normal connections that are incorrectly misclassified as attacks and the total number of normal connections**
- ROC Curve: trade-off between detection rate and false alarm rate; ideal system should have 100% detection rate with 0% false alarm
- Performance (Processing speed + propagation + reaction)
- Fault Tolerance (resistant to attacks, recovery)

IDS - Continuality

- **Continuous Monitoring**
  - IDS performs a continuous, real-time analysis by acquiring information about the actions immediately after they happen
  - Costly process due to transporting the audit data and processing them quickly
- **Periodic Analysis**
  - IDS periodically takes the snapshot of the environment (monitored system), analyzes the data snapshot looking for vulnerable software or spots and their exploits, configuration errors, etc.
  - Widely used by system administrators, but not satisfactory to ensure high security, since the security exposure between two consecutive runs is sufficient for active exploit of a vulnerability

IDS - analysis strategy 1/2

- **Misuse (signature) detection is based on extensive knowledge of patterns associated with known attacks provided by human experts**
- **Existing approaches: pattern (signature) matching, expert systems, state transition analysis, data mining**
- **Major limitations:**
  - Unable to detect novel & unanticipated attacks
  - Signature database has to be revised for each new type of discovered attack

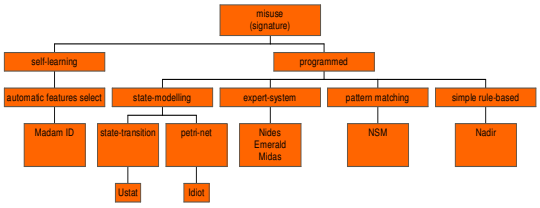
IDS - analysis strategy 2/2

- **Anomaly detection is based on profiles that represent normal behavior of users, hosts, or networks, and detecting attacks as significant deviations from this profile**
- Major benefit - potentially able to recognize unforeseen attacks.
- Major limitation - possible high false alarm rate, since detected deviations do not necessarily represent actual attacks
- Major approaches: statistical methods, clustering, neural networks, support vector machines, outlier detection schemes (*data mining*)

Usual approach of a traditional IDS

- **Traditional intrusion detection system (IDS) tools (e.g. SNORT, open source signature-based Network IDS) are based on signatures of known attacks**
  - Example of SNORT rule (MS-SQL "Slammer" worm) any -> udp port 1434 (content:"|81 F1 03 01 04 9B 81 F1 01|"; content:"sock"; content:"send")
- **Limitations**
  - Signature database has to be manually revised for each new type of discovered intrusion
  - They cannot detect emerging cyber threats
  - Substantial latency in deployment of newly created signatures

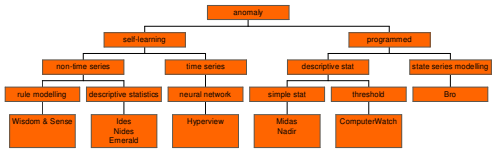
Classification as to the detection technique 1/2



Key technical challenges

- **Large data size**
  - E.g. Millions of network connections are common for commercial network sites, …
- **High dimensionality**
  - Hundreds of dimensions are possible
- **Temporal nature of the data**
  - Data points close in time - highly correlated
- **Data Preprocessing**
  - Converting data from monitored system into data appropriate for analysis

Classification as to the detection technique 2/2



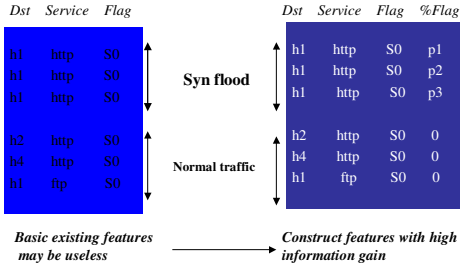
Data mining for Intrusion Detection

- **Misuse detection**
  - Predictive models are built from labeled data sets (instances are labeled as "normal" or "intrusive")
  - These models can be more sophisticated and precise than manually created signatures
  - Unable to detect attacks whose instances have not yet been observed
- **Anomaly detection**
  - Build models of "normal" behavior and detect anomalies as deviations from it
  - Possible high false alarm rate - previously unseen (yet legitimate) system behaviors may be recognized as anomalies

Basic steps in data mining for ID

- **Converting the data from monitored system (computer network, host machine, …) into data (features) that will be used in data mining models**
  - For misuse detection, labeling data examples into normal or intrusive may require enormous time for many human experts
- **Building data mining models**
  - Misuse detection models
  - Anomaly detection models
- **Analysis and summarization of results**

Madam ID: feature extraction from network data



Basic existing features may be useless → *Construct features with high information gain*

Use temporal and statistical patterns e.g. "a lot of S0 connections to same service/host within a short time window"

Projects: Data mining in ID

- **MADAM ID** (Mining Audit Data for Automated Models for Intrusion Detection) – Columbia University, Georgia Tech, Florida Tech
- **MINDS** (University of Minnesota)
- **ADAM** (Audit Data Analysis and Mining) - George Mason University
- **Intelligent Intrusion Detection** – IIDS (Mississippi State University)
- **Data Mining for Network Intrusion Detection** (MITRE corporation)
- **Agent based data mining system** (Iowa State University)
- **IDDM** – Department of Defense, Australia
- …..

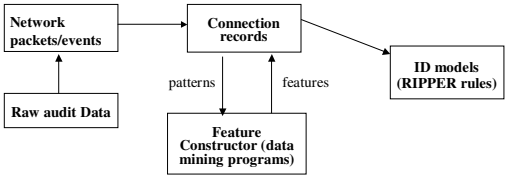
Madam ID: feature construction

- **Three groups of features are constructed:**
  - **"content-based" features within a connection**
    - number of packets, acknowledgments, data bytes from *src* to *dest*
    - intrinsic characteristics of data packets
  - **time-based traffic features** included number of connections or different services from the same source or to the same destination considering recent time interval (e.g.a few seconds)
    - useful for detecting scanning activities
  - **connection based features** included number of connections from same source or to same destination or with the same service considering in last *N* connections
    - useful for detecting *slow* scanning activities

MADAM ID data mining for misuse detection

- **Network traffic data is collected using "sniffers" (e.g. *tcpdump*, *net-flow tools*, …) – DARPA '98 data set**
- **Collected data are in the form of network connections or network packets (a network connection may contain several packets)**
- **Basic information collected for individual network connections include e.g.**
  - start time and duration
  - protocol type
  - source IP address and port,
  - destination IP address and destination port (service)
  - number of bytes, packets in connection
  - …

Madam ID: data mining process to build ID models



- **Association rules and frequent episodes are applied to network connection records to obtain additional features for data mining algorithms**
- **Apply RIPPER classifier to labeled data sets and learn the intrusions**

Data mining techniques for misuse detection

- **Classification techniques:**
  - **Rule based techniques (eg. RIPPER)**
    - Projects: MADAM ID, ADAM, MINDS, …
  - **Tree based approaches (decision trees)**
    - MADAM ID, MITRE, …
  - **Association rules, fuzzy association rules**
    - Projects: MADAM ID, ADAM, MINDS, IIDS
  - **Bayesian classifiers, genetic algorithms, ...**
  - **Neural networks**

Statistics based outlier detection scheme

- **Statistics based approaches –** data points are modeled using stochastic distribution ⇒ points are determined to be outliers depending on their relationship with this model
  - With high dimensions, difficult to estimate distributions
- **Major approaches**
  - probability distribution: measures how likely a point is an outlier
  - information theory measures: entropy measures the uncertainty of data items

Data mining for anomaly detection

- **Build models of "normal" behavior** and detect anomalies as deviations from it
- **Possible high false alarm rate -** previously unseen (yet legitimate) system behaviors may be recognized as anomalies
- **Major approaches**
  - Outlier detection
  - Profiling based techniques
- **Two types of techniques**
  - with access to normal data
  - with no access to normal data (not known what is "normal")

Distance based outlier detection schemes

- **Nearest neighbor based approaches -** Outliers are points that do not have enough neighbors
- **Density based approach finds outliers based on the densities of local neighborhoods**
  - Concept of locality becomes difficult to define due to data sparsity in high dimensional space
- **Clustering based approaches define outliers as points which do not lie in clusters**
  - Implicitly define outliers as background noise or very small clusters

Outlier detection scheme

- **Outlier is defined as a data point which is different from the rest of the data based on some measure**
- **Detect novel attacks/intrusions by identifying them as deviations from "normal", i.e. anomalous behavior**
  - **Identify normal behavior**
  - **Construct useful set of features**
  - **Use outlier detection algorithm**
    - Statistics based approaches
    - Distance based approaches: nearest neighbor, clustering based, density based
  - **Model based scheme**

Model based detection schemes

- **Use a prediction model to learn the normal behavior**
- **Every deviation from learned prediction model can be treated as anomaly or potential intrusion**
- **Recent approaches:**
  - Neural networks: the input variables are the output variables so that the NN forms a compressed model of the data during training
  - Support Vector Machines (SVMs): support vectors and a discriminant linear function; the entire set of training data is separated in classes the support vectors are the boundaries of.

Profiling based anomaly detection

- **Profiling methods are usually applied to host based intrusion detection where users, programs, etc. are profiled**
  - **Profiling sequences of Unix shell command lines**
  - **Profiling users' behavior**

Modeling systems calls data

- **Learn program behavior profiles** from previous execution (short sequences of system calls)
  - …open read mmap mmap open close … unique sequences for a sliding window with size K
- **Learn only traces from system calls from normal data**
  - Detect deviation from this profile
- **Learn traces from both normal and intrusive system calls**
  - Train a RIPPER classifier that will learn classes

Profiling: temporal sequence learning

- **Data –** sequences of Unix shell command lines
- **Set of sequences (user profiles) reduced and filtered** to reduce data set for analysis
- **Build an instance model that stores historic examples of "normal" data**
  - Compares new data stream
  - Distance measure that favors long temporal similar sequences
  - Event sequences are segmented

… a lot of work to do yet!

- **Lack of a structured development methodology**
- **Efficiency**
- **Portability**
- **Upgradability**
- **Maintainability**
- **Benchmarking**

Profiling: anomaly detection using NN

- **Modeling the behavior** of individual users
- **Data –** audit logs for each user for several days
- **Form a distribution vector –** how often user executes each command
- **Train Neural Network** with these vectors as inputs
- **Identify whether the user is regular or illegal** for each new command distribution vector, i.e. for each new login session

To begin…

- **Links :**
  - **www.cs.purdue.edu/homes/clifton/cs590m - security issues in data mining**
  - **www.snort.org**
- **Articles:**
  - **IDSs: A survey and taxonomy, S. Axelsson, 2000**


“...to be continued ...„