Intrusion Detection System

(1) Intrusion Detection Basics

- What is intrusion detection
  - Process of monitoring the events occurring in a computer system or network and analyzing them for signs of *intrusion*.

- Types of Intrusion Detection Systems
  - Information Sources: the different sources of event information used to determine whether an intrusion has taken place.
    - Network-based IDS
    - Host-based IDS
    - Application-Based IDS
  - Analysis: the most common analysis approaches are
    - Misuse Detection
    - Anomaly Detection
  - Response: the set of actions that the system takes once it detects intrusions.
    - Passive measure: reporting IDS findings to humans, who are then expected to take action based on those reports.
    - Active measure: involving some automated intervention on the part of the system.

- Misuse Detection (signature-based ID)
  - Looking for events or sets of events that match a predefined pattern of events that describe a known attack. The patterns are called *signatures*.
  - Rule-based systems: encoding intrusion scenarios as a set of rules.
  - State-based intrusion scenario representations.
  - Advantages:
    - Very effective at detecting attacks without generating an overwhelming number of false alarms.
  - Disadvantages
    - Can only detect those attacks they know about—therefore they must be constanly updated with signatures of new attacks.
    - Many misuse detectors are designed to use tighly defined signatures that prevent them from detecting variants of common attacks.

- Anomaly Detection
  - Identify abnormal unusual behavior (anomalies) on a host or network. They function on the assumption that attacks are different from “normal” (legitimate) activity and can therefore be detected by systems that identify these differences.
  - Static and dynamic:
    - Static: Static means a portion of the system remain constant, e.g. data integrity, tripwire, virus checkers.
    - Dynamic: profile. A profile consists of a set of observed measures of behavior for each of a set of dimensions. Frequently used dimensions include:
      - Preferred choices, e.g., log-in time, log-in location, and favorite editor.
      - Resources consumed cumulatively or per unit time.

        - Representative sequences of actions.
        - Program profiles: system call sequence.
- Methods
    - Threshold detection: certain attributes of user and system behavior are expressed in terms of counts, with some level established as permissible. Such behavior attributes can include the number of files accessed by a user in a given period of time, the number of failed attempts to login to the system, the amount of CPU utilized by a process, etc.
    - Statistical measures
        - Parametric: The distribution of the profiled attributes is assumed to fit a particular pattern
        - Non-parametric: The distribution of the profiled attributes is "learned' from a set of historical values, observed over time.

    - Rule-based measures: similar to non-parametric statistical measures in that ooberved data defines acceptable usage patterns, but differs in that those patterns are specified as rules, not numeric quantities.

    - Other methods:
        - Machine learning
        - Data mining
        - Neural networks, genetic algorithms, etc.
- Advantages
    - Can detect unusual behavior and thus have the ability to detect symptoms of attacks without specific knowledge of details.
    - Can produce information that can in turn be used to define signatures for misuse detectors.
- Disadvantages
    - Usually produce a large number of false alarms due to the unpredictable behaviors of users and networks.
    - Often require extensive "training sets" of system event records in order to characterize normal behavior patterns.

- Host-based IDS
    - Using OS auditing mechanisms: e.g. BSM in Solaris logs all direct and indirect events generated by a user; `strace` monitors system calls made by a program.
    - Monitoring user activities: analyzing shell commands.
    - Monitoring executions of system programs, e.g. `sendmail`'s system calls.
    - Advantages
        - Can detect attacks that cannot be seen by NIDS
        - Can operate in an environment in which network traffic is encrypted
        - Unaffected by switched networks
        - Can help detect Trojan horse or other attacks that involve software integrity breaches
    - Disadvantages
        - Since at least the information sources reside on the host targeted by attacks, the IDS may be attacked and disabled as port of the attack
        - Are not well suited by detecting network scans or other such surveillance that targets an entire network
        - Since they use the computing resources of the hosts they are monitoring, therefore inflicting a performance cost on the monitored systems.

- Network Intrusion Detection Systems (NIDS)
  - Using packet sniffing.
  - Looking at IP header as well as data parts.
  - Disadvantages of Network-Based IDSs:
    - NIDS may have difficult processing all packets in a large or busy network and therefore, may fail to recognize an attack launched during periods of high traffic.
    - Modern switch-based networks make NIDS more difficult: Switches subdivide networks into many small segments and provide dedicated links between hosts serviced by the same switch. Most switches do not provide universal monitoring ports
    - NIDS cannot analyze encrypted information.
    - Most NIDS cannot tell whether or not an attack was successful.

- Evaluating an IDS
  - False positive
  - False negative
  - ROC curve: Receive Operating Characteristic

- IDS strengths and limitations
  - Up side:
    - Detect an ever-growing number of serious problems
    - New signatures are added.
    - New methods are being developed.
  - Down side:
    - IDs look for known weaknesses (patterns or normal behavior)
    - False positive

(2) Eluding Network Intrusion Detection

- Insertion: Defeating signature analysis
  - Conceptual Example

End System sees: A T T A C K

Network Monitor: A T X T A C K

Attacker's data stream: T X T C A A K

-
  - Real example: "Get /cgi-bin/phf?"
  - Solution: make the IDS as strict as possible in processing packets read off the wire.

- Evasion
  - Conceptual Example

End System sees: A T T A C K

Network Monitor: A T T C K

Attacker's data stream: T T C A A K

- How to achieve Insertion/Evasion Attacks based on IP
  - Checksum (easy to solve)
  - TTL: large enough for IDS monitor, but not enough for the end system.
  - Don't fragment
  - IP Options:
    - Many OS automatically reject source routed packets.
    - Timestamp: discard packets with illegal formats
  - MAC address: address the faked packet to IDS's Mac address, so the end system will not receive it.
  - IP Reassembly Problem
  - IDS also needs to reassembly packets.
  - Subject to DOS attacks.
  - IDS must drop incomplete fragments (or late fragments) the same manner as the end system does. Otherwise inconsistence exists.
  - Overlapping fragments: must process them in the same manner as the end system.
    - Windows NT 4.0: always favors old data
    - Solaris 2.6: always favors old data
    - 4.4BSD: Favors New data for forward overlap
    - Linux: Favors New data for forward overlap

- How to achieve Insertion/Evasion Attacks based on TCP?
  - TCP Code: packets with illegal code will be discarded.
  - SYN packet may carry data, and some implementation may not process these data.
  - TCP Window size: inconsistence between end system and IDS can cause problems.
  - TCP Overlapping: NT 4.0 favors old data; others favor new data.
  - Establishing TCP Connections: consistency between IDS and end systems.
  - Tearing Down TCP Connections: consistency ...

- Denial of Service Attacks on IDS
  - CPU, memory, bandwidth