Development of a Network Intrusion Detection System

(I): Agent-based Design (FLC1)
(ii): Detection Algorithm (FLC2)

Supervisor: Dr. Korris Chung

Please visit my personal homepage www.comp.polyu.edu.hk/~cskchung/FYP04-05/ for more information.
Why Intrusion Detection?

- Current firewalls are not sufficient to ensure the security in computer networks
- Some intrusions take advantages of vulnerabilities in computer systems or use socially engineered penetration techniques that traditional intrusion prevention techniques are not enough to protect systems
- Intrusion detection System (IDS) will be another wall for protection
Motivations

- Problems of existing IDS:
  - Attack stealthiness
    - Attackers try to hide their actions from either an individual who is monitoring the system or an IDS
  - Novel intrusions
    - Undetectable by signature-based IDSs that they can only be detected as anomalies by observing significant deviations from the normal network behavior
  - Distributed attack
    - Needed for detecting the correlation of attack
Objectives

- Develop an adaptive IDS by using agent-based data mining techniques.
- The data mining approaches are used to capture the actual behavior of network traffic accurately.
- The portfolio mined is useful in differentiating “normal” and “attack” traffics.
- The overall architecture of the proposed IDS is constructed with different types of agent.
Types of Attack (1/2)

- Denial of Service (DoS)
  - SYN Flood
  - Smurf
  - Teardrop
- Probe
  - SYN Stealth Port Scan
  - Connect Mode Port Scan
  - FIN Stealth Mode Port Scan
  - NULL Scan Mode Port Scan
  - Xmas Tree Mode Port Scan
  - ACK Mode Port Scan
  - UDP Scan
Types of Attack (2/2)

- Compromises
  - Buffer Overflow
  - Unauthorized Access from Remote Machine (U2R)
  - Unauthorized Access to Local Superuser Privileges by a Local Unprivileged User (R2L)

- Trojan horses/worms
## Traffic Behaviors (1/2)

### Probe

<table>
<thead>
<tr>
<th></th>
<th>URG</th>
<th>ACK</th>
<th>PSH</th>
<th>RST</th>
<th>SYN</th>
<th>FIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connect scan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Set</td>
<td></td>
</tr>
<tr>
<td>SYN scan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Set</td>
<td></td>
</tr>
<tr>
<td>FIN scan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Set</td>
</tr>
<tr>
<td>NULL scan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xmas tree scan</td>
<td>Set</td>
<td></td>
<td>Set</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACK scan</td>
<td></td>
<td>Set</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Traffic Behaviors (2/2)

## Denial of Service (DoS)

<table>
<thead>
<tr>
<th>Attack type</th>
<th>Port range</th>
<th>No of connection attempt</th>
<th>No of RST packets</th>
<th>Mean packet size</th>
<th>Time covered by 100 packets</th>
<th>No of ICMP packets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connect scan</td>
<td>Large</td>
<td>Large</td>
<td>Don’t mind</td>
<td>Small</td>
<td>Small if not stealthy</td>
<td>Don’t mind</td>
</tr>
<tr>
<td>SYN scan</td>
<td>Large</td>
<td>Large</td>
<td>Don’t mind</td>
<td>Small</td>
<td>Small if not stealthy</td>
<td>Don’t mind</td>
</tr>
<tr>
<td>ACK scan</td>
<td>Large</td>
<td>No</td>
<td>Don’t mind</td>
<td>Small</td>
<td>Small if not stealthy</td>
<td>Don’t mind</td>
</tr>
<tr>
<td>NULL scan</td>
<td>Large</td>
<td>No</td>
<td>Don’t mind</td>
<td>Small</td>
<td>Small if not stealthy</td>
<td>Don’t mind</td>
</tr>
<tr>
<td>Xmas tree scan</td>
<td>Large</td>
<td>Large</td>
<td>Don’t mind</td>
<td>Small</td>
<td>Small if not stealthy</td>
<td>Don’t mind</td>
</tr>
<tr>
<td>UDP scan</td>
<td>Large</td>
<td>No</td>
<td>N/A</td>
<td>Small</td>
<td>Small if not stealthy</td>
<td>Many destination unreachable</td>
</tr>
<tr>
<td>SYN flood</td>
<td>Small</td>
<td>Large</td>
<td>Large after connect pool full</td>
<td>Small</td>
<td>Small</td>
<td>N/A</td>
</tr>
<tr>
<td>PING flood</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Don’t mind</td>
<td>Small</td>
<td>Large</td>
</tr>
<tr>
<td>Victim reply from attack</td>
<td>Large for port scanned</td>
<td>N/A</td>
<td>Large for TCP scan on closed port and SYN flooded</td>
<td>Small</td>
<td>Small</td>
<td>Many ICMP if UDP scanned or ping flooded</td>
</tr>
<tr>
<td>Normal traffic</td>
<td>Small</td>
<td>???</td>
<td>Small</td>
<td>Don’t mind</td>
<td>Larger</td>
<td>Small</td>
</tr>
</tbody>
</table>
Proposed Agent-based Intrusion Detection System
Proposed System

- data mining approaches are introduced
  - Clustering
  - Sequential Association Rules
  - Outlier detection
- Each approach is represented by a number of agents
System Architecture

Network Traffic

Feature Extractor

Trainer

Detection Engine

Update

Real Time

Batch Processing

Proposed IDS

Alarm
Detection Engine
Agent Trainer

Set of Feature

Feature Distributor

Approach 1

Approach 2

Agent $n$

Update Corresponding Agents
Two types of System Models

- **Predictive model**
  - Built from labeled data sets (i.e. instances are labeled as “normal” or “attack”)
  - Misuse detection
  - More sophisticated and precise than manually created signatures
  - Expected to be unable to detect attacks whose instances have not yet been observed

- **Anomaly detection model**
  - Built only based on “normal” behavior (i.e. normal behavior of the network is profiled)
  - Detecting anomalies as deviations
  - Possibly high in false alarm rate as previously unseen
  - Higher adaptive ability
Thank You!