CS261 Lecture 2007-10-25

Lecturer: David Wagner Scribe: Bor-Yiing Su

Intrusion Detection System

1. Classification of IDS

1.1. Signature VS. Anomaly Detection:

- **Signature Detection:** Identify the characteristics of existed attacks, record them into a database, and then detect possible attacks by comparing their characteristics with information in the database. Use the blacklist concept to detect attacks.
- **Anomaly Detection:** Learn the normal behaviors of the host or the network that is under protection, record them into a database. If there are some deviation between the current behaviors and the information in the database, alarm the administrator. Use the white list concept to detect attacks.
- **Drawbacks:** Signature detection might have more false negatives, while anomaly detection might have more false positives.
- **Example: Virus Scanner**

Most virus scanners use signature detection to identify virus. The simplest way is to compare some segments of the binaries with some specified features. However, it can be bypassed by injecting some random malicious features in the binaries. Advanced virus scanners will simulate the execution of the binaries, and then see if some malicious features occur.

- **Example: Bro**

  Bro can be used in both schemes as well.

1.2. Single Host VS. Network VS. Application Domain Specific IDS

- **Single Host IDS:** Detect malicious behaviors of processes in a single host. Protection technique can be classified into two modes.
  - Training Mode: Learn about what sequences of system calls could be malicious.
  - Forcing Mode: Monitor the execution of system calls. Compare the execution sequence with the features learned from training mode.
- **Network IDS:** Detect malicious behaviors in network.
  - Bro is a Network IDS.
- **Application Domain Specific IDS:** Detect malicious behaviors in some specific applications.
  - **Example: Credit Card.** If a credit card is debited in CA at 4:00 p.m. and

then debited again in NY at 5:00 p.m., the credit company might identify it as a strange transaction.

  - **Example: Money Transfer.** If a great deal of money is transferred online, the bank might identify it as a malicious behavior.
  - **Example: Router Protection.** We can simply check the number of packets flow in the router and flow out of it. If the two numbers do not match, it is possible that the router is hacked.

1.3. Detect the Exploit VS. Detect the Payload

- **Detect the Exploit:** Detect possible exploits from input streams. For example, detect buffer overrun attack from input stream.
- **Detect the Payload:** Detect the unusual behaviors of the hosts or network that is under protection. For example, detect if the host provide some queer services at some weird ports, say, port 999.
- **Advantages:** Detect the Exploit scheme can detect the attack regardless of attacker's objectives. Detect the Payload scheme can detect the unusual behaviors of the system regardless of how the attacker achieved this.
- **Comment:** Most anomaly detection systems detect the payload.

1.4. Orthogonal VS. Non-Orthogonal

- **Orthogonal:** Protect the system without changing the operating mechanisms of it. The system under protection can not tell whether there is some protection scheme or not.
- **Non-Orthogonal:** Protect the system by changing its operating mechanisms. For example, install specific agents in all the hosts in the network.
- **Advantages:** The orthogonal scheme will not affect the operation of the system under protection. The non-orthogonal scheme can have closer interpretation for the input stream as what the host does. Thus it can provide more secure policies.

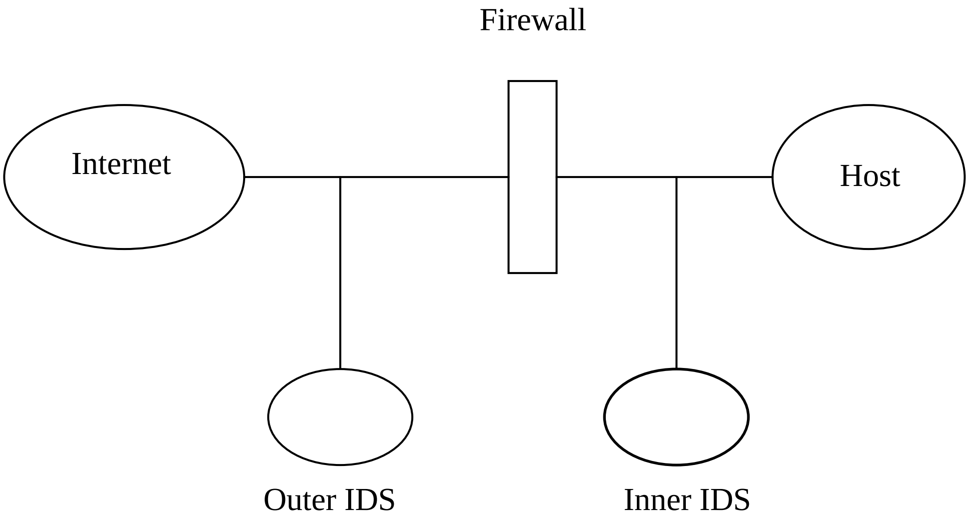
2. Different Objectives Between Firewall and IDS

2.1. Reasons that IDS only alarm and log possible attacks, but not block them.

- Sometimes the data we received is ambiguous, we cannot identify easily whether it comes from an attacker or not.
- It is possible that we will block some innocuous traffic.
- If the IDS block all packets that it classifies as vulnerable, the attacker can send malicious packets to make IDS block all the incoming packets, and results in

DoS condition.

2.2. Capabilities of IDS



- Firewall can block apparent attacks, while IDS can detect delicate attacks.
- The IDS can record information for offline analysis.
- The outer IDS can gather statics about how many attacks we have encountered in some time period.
- The inner IDS can detect insider attacks.
- By comparing logs between outer IDS and inner IDS, we can know how many attacks that have bypassed the firewall.

3. Some special insights from the Bro paper

3.1. False Alarm

- Too many false alarms make user distrust the IDS system.
- **The Base-Rate Fallacy:** We assume that the IDS system is 99% correct. For a malicious packet, the IDS has 99% chance to report that it is malicious. For a usual packet, the IDS has 99% chance to report that it is usual. Moreover, we assume that for 10000 packets, there is only one packet that is malicious. That is, the appearing probability of malicious packets is 1/10000. Under such assumptions, if the IDS reports a malicious packet, the probability that the packet is really malicious is $\dfrac{\frac{1}{10000}\times 0.99}{\frac{1}{10000}\times 0.99+(1-\frac{1}{10000})\times 0.01}=0.00980\approx 1\%$

3.2. Offline Analysis

- The offline analysis can track the attacks that occurred long time ago.

3.3. Evasion / Subterfuge

- It is a hard problem to solve. Some strange behaviors of the network make the problem harder. For example, it is sometimes that the host will receive several TCP packets with the same sequence number, but different contents. However, the packets are not malicious. They might just come from some buggy hosts.
- **Possible Defense:**
  - **Normalization:** We can modify packets in order to avoid ambiguities. That is, when we see some packets that have some deficiencies in the format, we will send to a normalizer to correct the format. For example, if the TTL of the packet is small, the normalizer can enlarge it to make sure that the packet can reach the host. If the checksum of the packet is not correct, the normalizer can correct it.
  - **Active Mapping:** We can map each host with the OS running on it. Therefore, we can apply different policies according to different OS. However, we have to update the mapping frequently to make sure that it is correct. Moreover, the database of the behaviors of all the OS with the policies according to the behaviors will be tremendous.
  - **Agents:** We can install an agent on each host. That is, we can apply Non-Orthogonal IDS on each host according to its OS. However, it is hard to deploy such defense because most hosts do not want an agent to influence its operation.

4. Port Scan

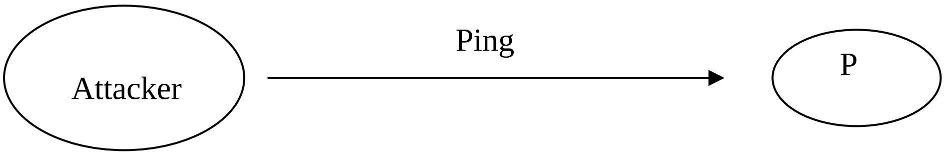
4.1. Advantages of Port Scan

- It is very convenient for an attacker to use port scan to figure out the service provided by the host, and the OS information related to the host.

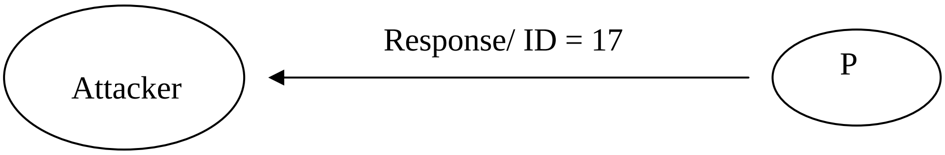
4.2. Port Scan Techniques

- **Multiple Hosts:** We can divide the ports into several segments. Then use different host to scan different segment. If the scanning process is slow, the host is hard to detect such behavior.
- **ID Field:** We can use the continuous character of the ID field in the packet to do port scan.

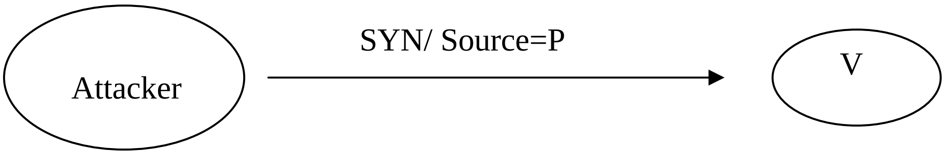
- We assume that the attacker is trying to do port scan on host V by using another host P.
- First of all, the attacker sends some packets to P. For example, a ping request.



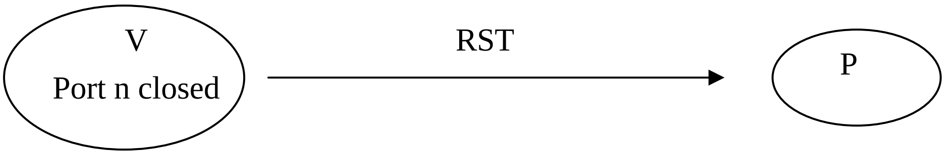
- Then, the attacker can learn the current ID number from P's response.



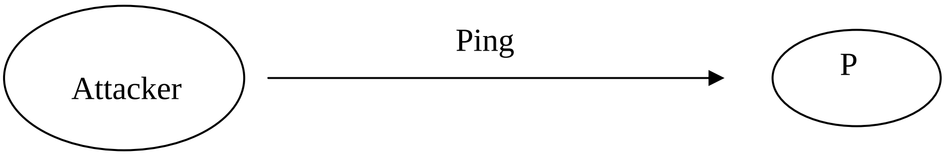
- The attacker sends a SYN packet to host V at a specific port n, with the IP field of the packet filled by host P's IP.



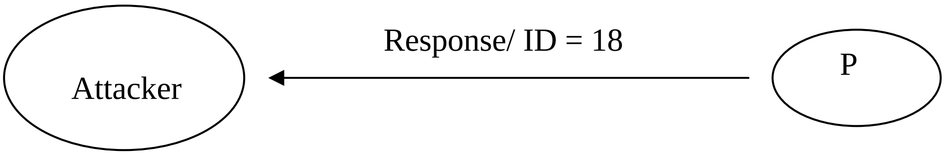
- If host V closes port n, it will send a RST packet to host P.



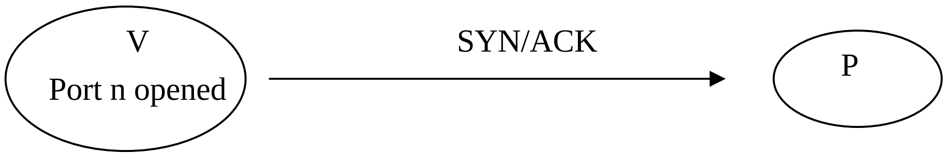
- If attacker pings host P again.



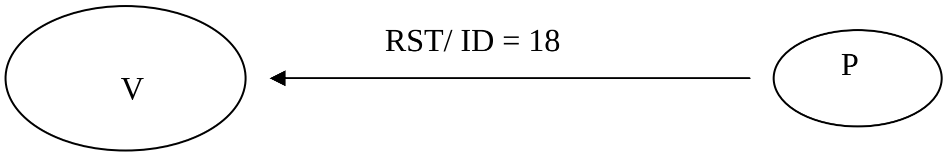
- It will find that the ID number of P is increased by 1.



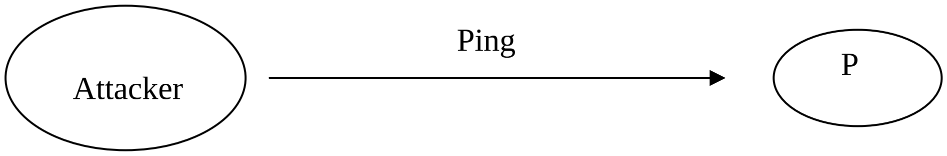
- If host V opens port n, it will send a SYN/ACK packet to host P.



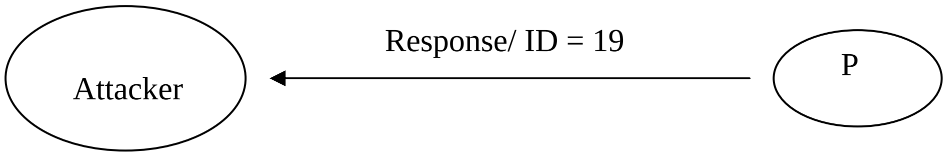
- Host P will then send a RST packet back.



- If attacker pings host P again.



- It will find that the ID number of P is increased by 2.



- Therefore, the attacker can know whether host V opens the port n or not by monitoring the ID field of host P’s response.