Application Layer Systems
Content inspection

Application layer systems

- Payload inspection in network systems
  - Monitoring
  - Security
  - Content blocking
  - Quality of service
- Our examples:
  - String matching for content filtering firewall
  - Load balancing for web server
Pattern matching

• Given:
  – Alphabet: \{a,c,k,t\}
  – Pattern: attack

• Goal:
  – Match packet payloads against pattern

• Problem variants:
  – Exact string matching
  – Regular expression matching

Exact string matching

• Nondeterministic finite automaton (NFA)
  – Automaton can be in multiple states at the same time
Exact string matching

• States for acaattack:

<table>
<thead>
<tr>
<th>Input</th>
<th>Set of states</th>
<th>Match?</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>{0}</td>
<td>no</td>
</tr>
<tr>
<td>a</td>
<td>{0,1,7,10}</td>
<td>no</td>
</tr>
<tr>
<td>c</td>
<td>{0,8,13}</td>
<td>no</td>
</tr>
<tr>
<td>a</td>
<td>{0,1,7,10,14}</td>
<td>no</td>
</tr>
<tr>
<td>a</td>
<td>{0,1,7,10}</td>
<td>no</td>
</tr>
<tr>
<td>t</td>
<td>{0,2}</td>
<td>no</td>
</tr>
<tr>
<td>t</td>
<td>{0,3}</td>
<td>no</td>
</tr>
<tr>
<td>a</td>
<td>{0,1,4,7,10}</td>
<td>no</td>
</tr>
<tr>
<td>c</td>
<td>{0,5,8,13}</td>
<td>no</td>
</tr>
<tr>
<td>k</td>
<td>{0,6}</td>
<td>yes (“attack”)</td>
</tr>
</tbody>
</table>

Exact string matching

• States for deterministic finite automaton (DFA):
  – What are default transitions?

![DFA Diagram]
Exact string matching

- DFA with default transitions:

![DFA diagram](image)

Regular expression matching

- Richer expressions
  - Alternatives: A | B
  - Zero or more occurrences: A*
  - One or more occurrences: A+

- Example:
  - AB+: AB, ABB, ABBB, ABBBB, ...
  - A(B | C)*: A, AB, AC, ABB, ABC, ACB, ACC, ABBB, ...
Regular expression construction

• NFA with epsilon transitions:

(A|B)

A*

A'

Regular expression construction

• NFA to DFA conversion
  – Closure of all states (to remove epsilon transitions)
  – Subset construction (parallel NFA states in single DFA state)
• Example: ([a,t](ck|t)t+|c(ak)*c)
Content inspection in Internet

- Intrusion detection system (IDS)
  - Check for malicious content in packets
  - Example: known attacks, exploits, malware, etc.
- How to get “signatures” of malicious content?
  - Few tools
  - Mostly created by hand
- Open source tool: snort
  - www.snort.org

Load balancing web server

- Large servers may not keep all pages on one system
  - Need to distribute requests between servers
  - What is the challenge?
Load balancing web server

• Problem is TPC connection setup
  – URL in request only known after connection is established
  – Initial connection setup cannot be routed to actual server
  – Load balancer intercepts and “glues” connections

Network systems

• We have seen all major network systems
  – Interface cards
  – Bridges/switches
  – Routers
  – Transport layer systems
  – Application layer systems

• Next
  – Queuing theory
  – Scheduling
  – Network security