Improving Dependence Explosion by Dynamic Tag Update
Sanaz Sheikhi, Md Nahid Hossain, R.Sekar
Secure System Labs

**Dependence Explosion Problem**

- **Dependency graph** captures casual relations between system entities (processes, files, sockets, ...)
- **Used for** attack detection and scenario reconstruction

- **Dependence explosion**: every output of a process becomes dependent on every earlier input operation.
- **Long running processes** cause dependence explosion and make the graph so huge.

**Existing Approaches Drawbacks**

- Fine-grained dependence tracking: instrumentation of applications and/or OS code
- Model-assisted search
- Analyst-driven search

**Our Approach**

- **Tag Decay**: Gradually lift data tag $d$ of benign processes to a quiescent value.
  \[ d = \max(d_0, d_0 + r^i + (1 - r^i) \cdot T_q) \]
- **Tag Attenuation**: Attenuate tags propagating from benign subjects to objects.
  \[ \text{obj.
dtag} = \text{sub.
dtag} + a \]

**Improved Attenuation and Decay**

- Attenuation/Decay are not affective on Windows audit data
- Observing broken data or specific behavior of processes in Windows.
- **Solution**: learning benign behavior of the system and update subject and object tags accordingly.
- **Attenuation/Decay rates** are dynamic regarding the training results.

**Learning System Behavior**

- **Process profile**: $(\text{proc}, W_j, \text{alarm}_k, \text{count})$
  Number of each alarm, process generates in each time windows
- **Object Profile**: $(\text{Object}, W_j, \text{event}_k, \text{count})$
  Number of each event, happening on object in each time window

**Dynamic Tag Update**

- **Dynamic attenuation**: $W_j$: ratio of access (read/write) to the object based on the profile
  \[ \text{obj.
dtag}' = \text{obj.
dtag} + w_j \]
- **Dynamic decay**: $r_i$: ratio of process activity in the time window based on the profile
  \[ \text{Subj.
dtag} = \text{sub.
dtag} + r_i \cdot T_q \]

**Evaluation**

Datasets: DARPA TC Engagement 4 Datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th># of Events</th>
<th>Attacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_1$</td>
<td>45M</td>
<td>SSH/RDP, Phishing, PowerShell, Firefox Draxon</td>
</tr>
<tr>
<td>$W_2$</td>
<td>49M</td>
<td>Firefox Draxon, Code Injection</td>
</tr>
</tbody>
</table>

Scenario graph from $W_2$