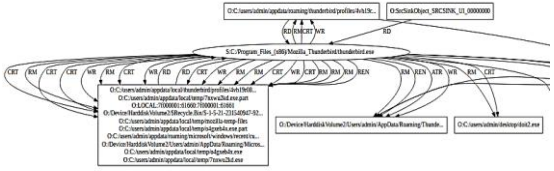
Improving Dependence Explosion by Dynamic Tag Update

Sanaz Sheikhi, Md Nahid Hossain, R.Sekar
Secure System Labs

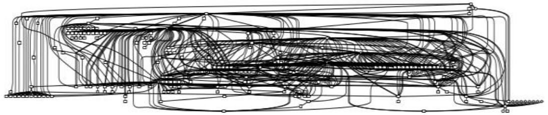
Stony Brook University
Computer Science

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
  manual effort to make model for all attacks
- Analyst-driven search
  manual effort to develop code for all attacks

Our Approach

Tag Decay

Gradually lift data tag d of benign processes to a quiescent value.

$$d = \max(d_0, d_0 * r^t + (1 - r^t) * T_q)$$

Tag Attenuation

Attenuate tags propagating from benign subjects to objects.

$$obj.dtag = sub.dtag + a$$

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: $(proc_i, W_j, alarm_k, count)$
  Number of each alarm, process generates in each time windows
- Object Profile: $(Object_i, W_j, event_k\ count)$
  Number of each event, happening on object in each time window

Dynamic Tag Update

Dynamic attenuation:

$W_{t:}$ ratio of access (read/write) to the object based on the profile

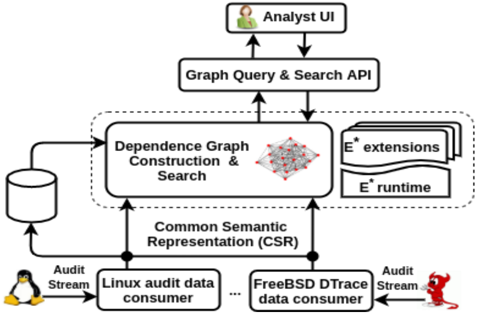
$$obj.datg' = obj.dtag + w_t$$

Dynamic decay:

r_t : ration of process activity in the time window based on the profi

$$Subj.dtag = subj.dtag + r_t * T_q$$

Architecture



Analyst UI
Graph Query & Search API
Dependence Graph Construction & Search
E* extensions
E* runtime
Common Semantic Representation (CSR)
Audit Stream
Linux audit data consumer
…
FreeBSD DTrace data consumer
Audit Stream

Evaluation

Datasets: DARPA TC Engagement 4 Datasets

Dataset	# of Events	Attacks
W_1	45M	SSH/RDP, Phishing Powershell, FireFox Drakon
W_2	49M	Firefox Drakon, Code Injection

Scenario graph from W_2