Exploring Digital Evidence with Graph Theory

Imani Palmer
Department of Computer Science
University of Illinois at Urbana-Champaign

Roy Campbell
Department of Computer Science
University of Illinois at Urbana-Champaign

Boris Gelfand
Advanced Research in Cyber Systems
Los Alamos National Laboratory

whoami?

- Imani Palmer
- 5th-year PhD Student
- Martial Artist Enthusiast

Overview

- Motivation
- Problem Statement
- Graph Theory Applications
- Discussion
- Conclusion

Chat Logs

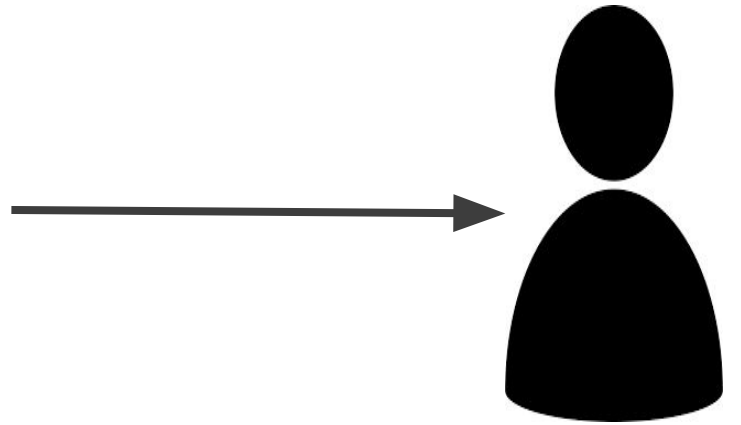
Emails

Photos

System Logs

Metadata

Programs



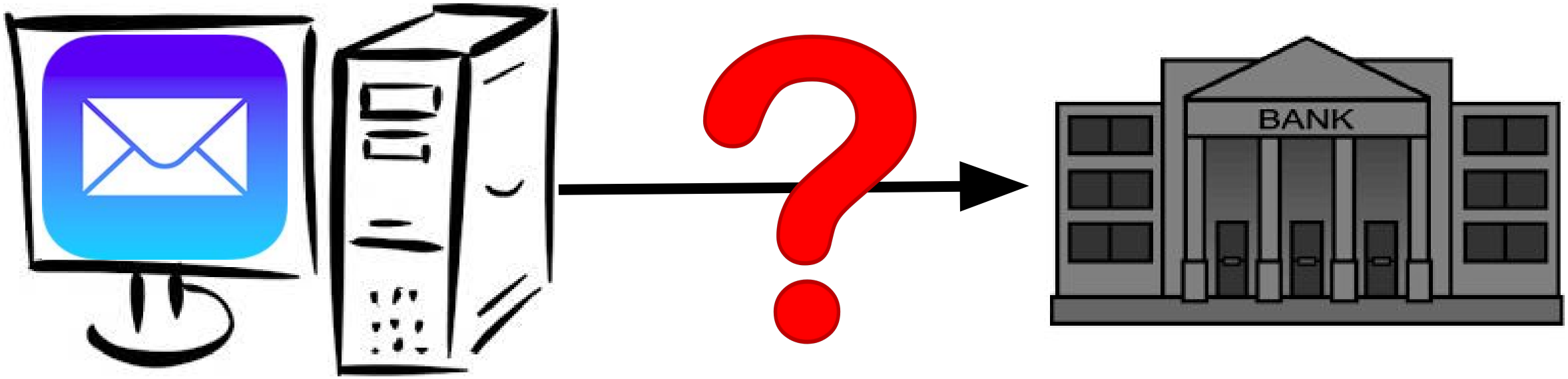
Knowledge Base

Report

Palmer, Imani, et al. "Digital Forensic Education: A Multidisciplinary Curriculum Model." *International Conference on Digital Forensics and Cyber Crime*. Springer International Publishing, 2015.

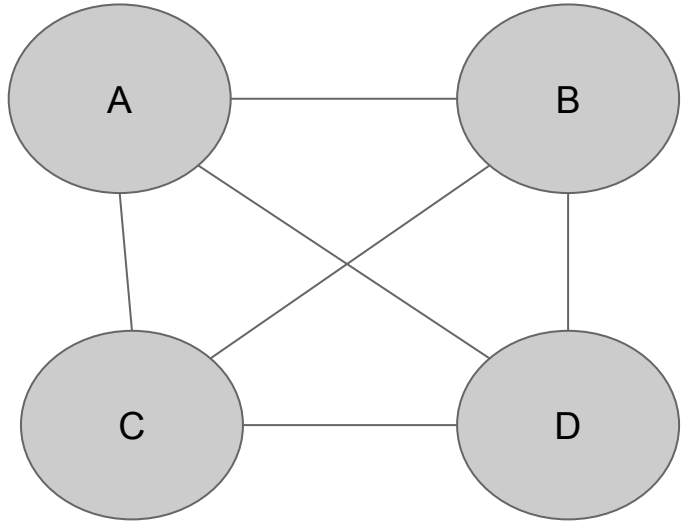
Application of computer science methods towards the digital forensic process

Case Study

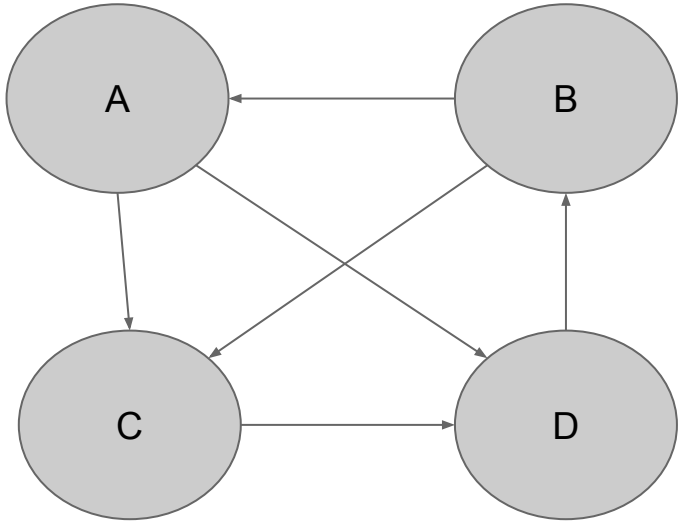


Graph-based Representations

- Different point of view
- Simplify the problem



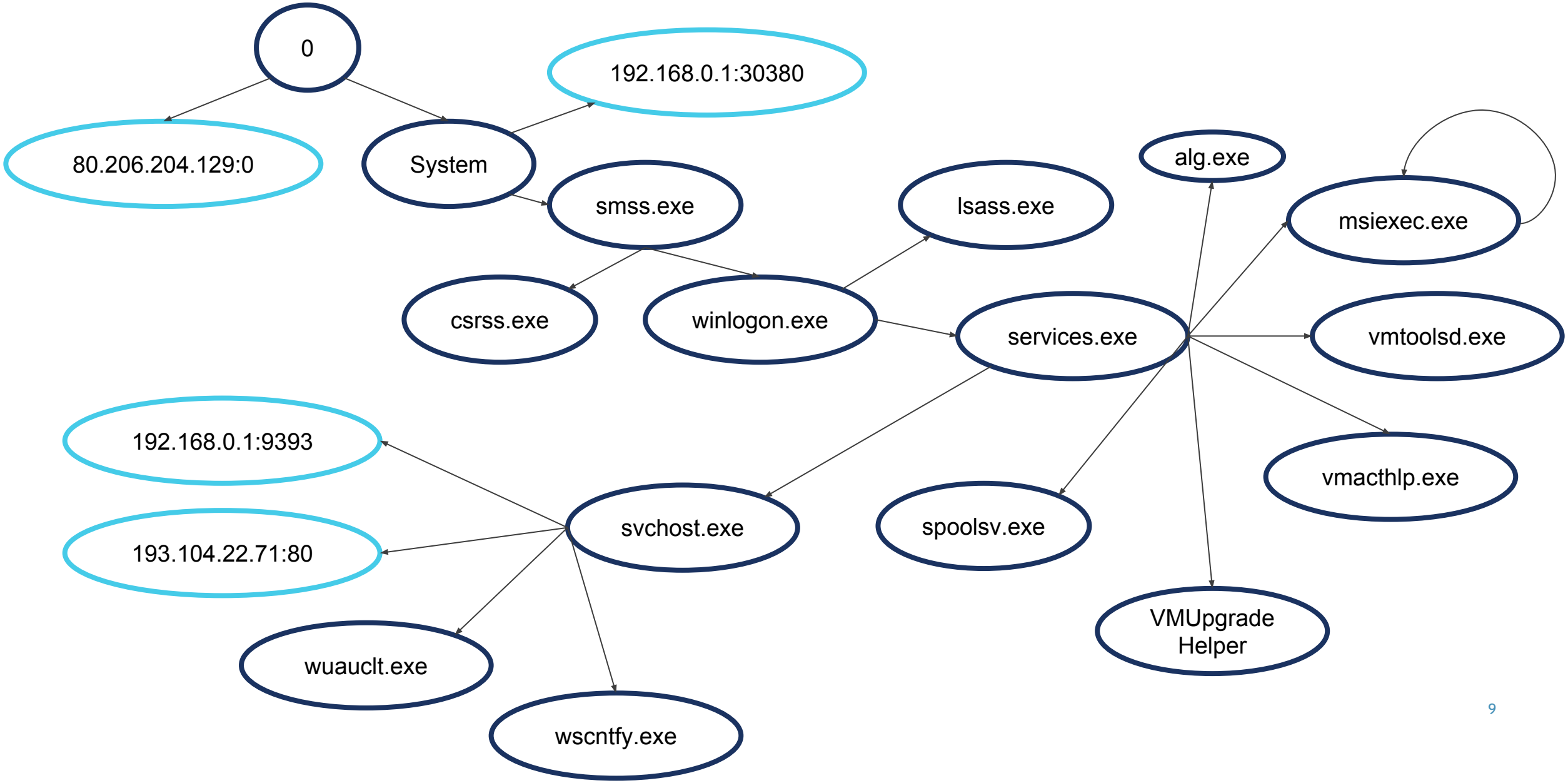
Undirected Graph

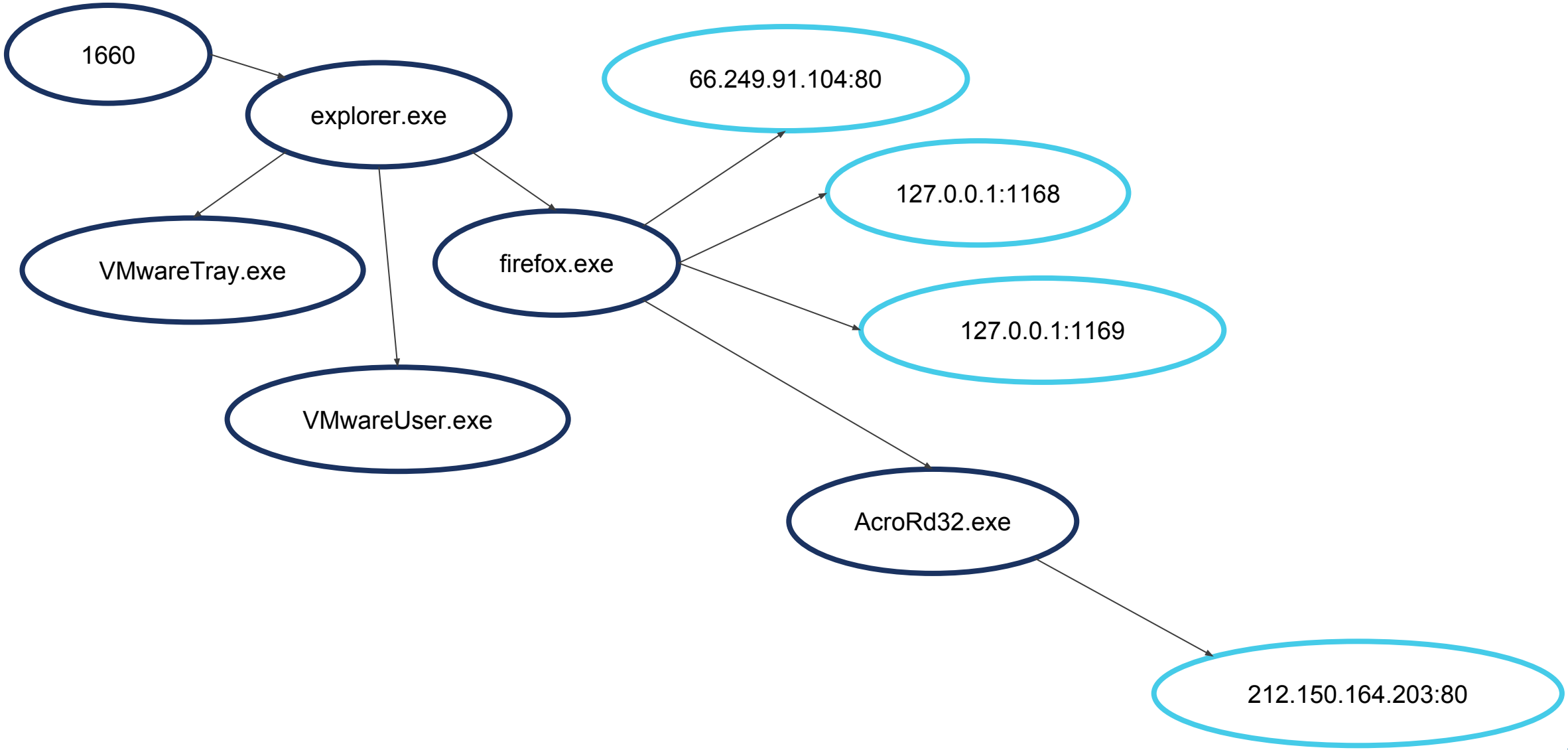


Directed Graph

Observe Evidence

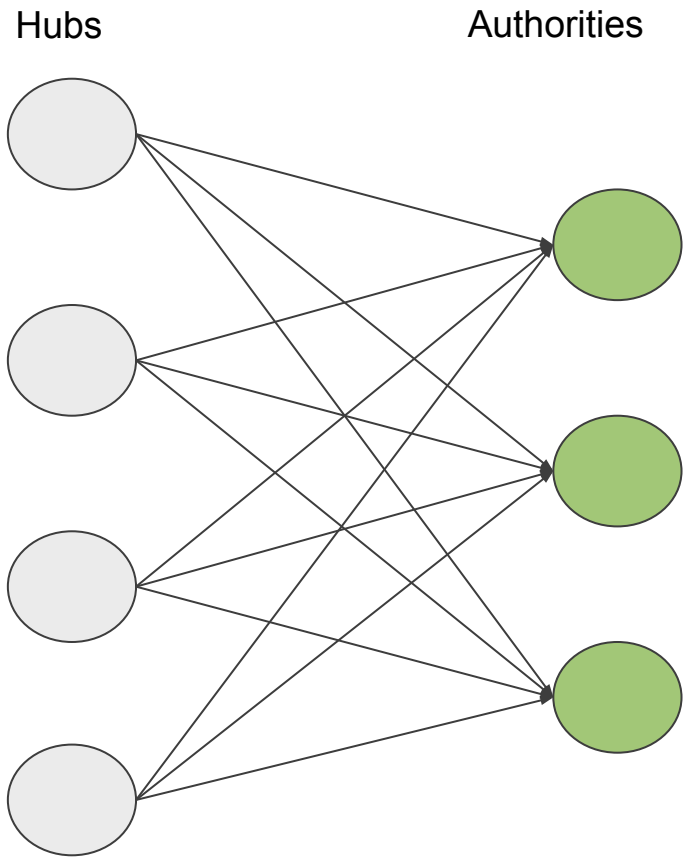
- Graph-based representation of the evidence
  - Nodes
    - Processes and sockets
  - Edges
    - Parent process -> Child Process
    - Process -> Socket

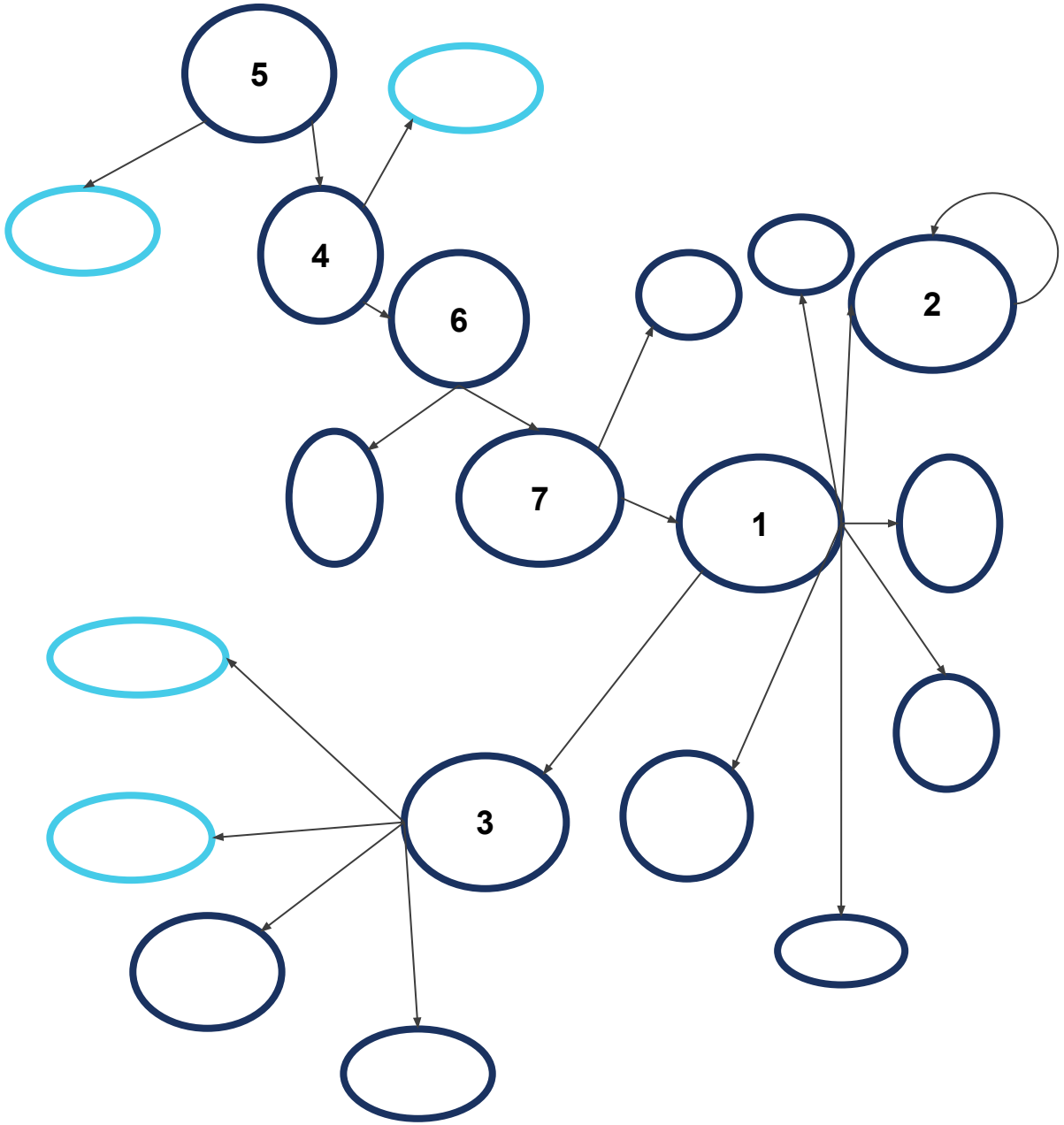

0
192.168.0.1:30380
80.206.204.129:0
System
smss.exe
lsass.exe
alg.exe
msiexec.exe
csrss.exe
winlogon.exe
services.exe
vmtoolsd.exe
192.168.0.1:9393
vmacthlp.exe
svchost.exe
spoolsv.exe
193.104.22.71:80
VMUpgrade
Helper
wuauclt.exe
wscntfy.exe


1660
explorer.exe
66.249.91.104:80
127.0.0.1:1168
VMwareTray.exe
firefox.exe
127.0.0.1:1169
VMwareUser.exe
AcroRd32.exe
212.150.164.203:80

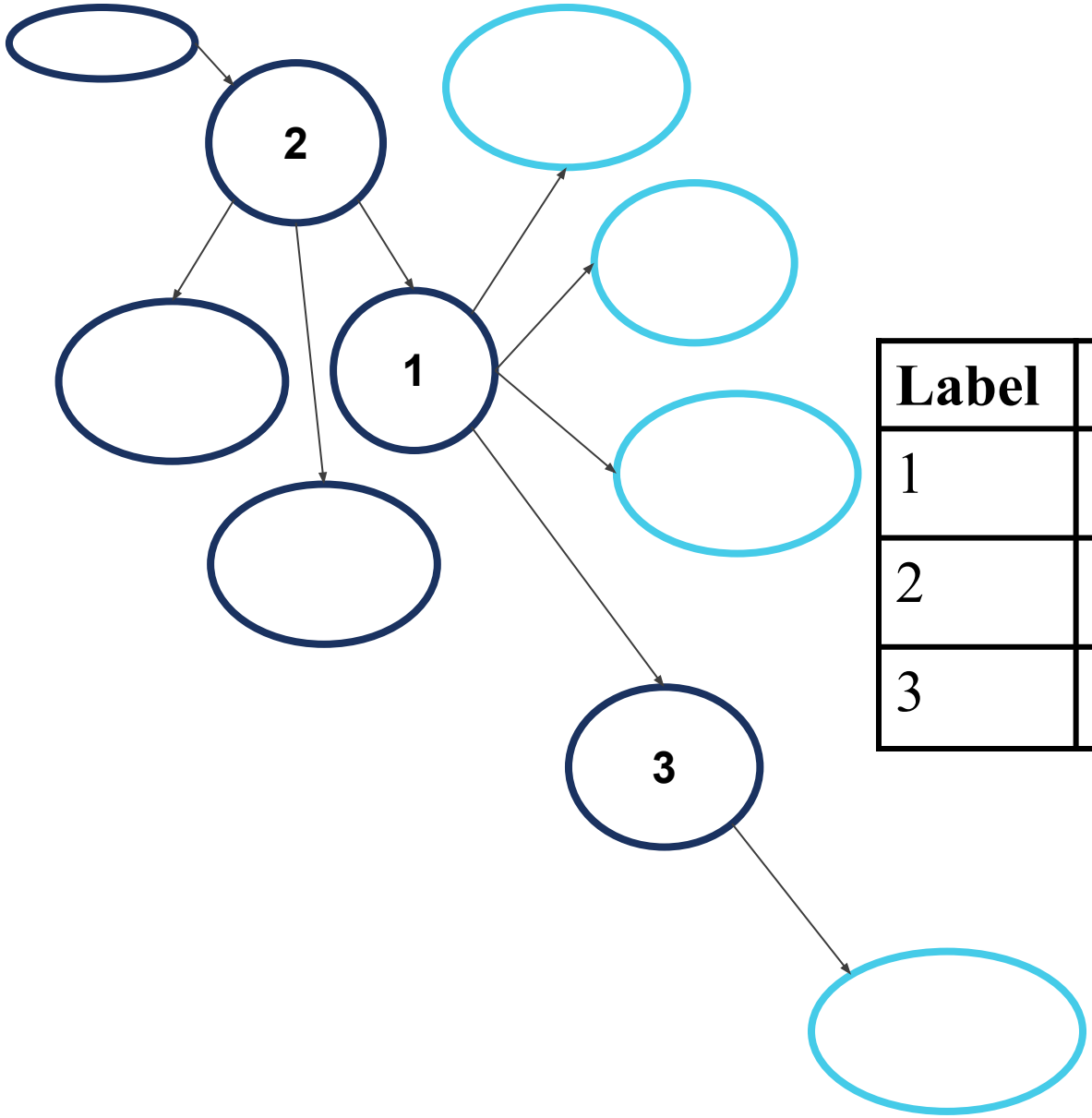
Hyperlink-Induced Topic Search (HITS)

- Identify high-level user actions
- **Authority:** a node that hubs link to
- **Hub::** a node that links to many authorities





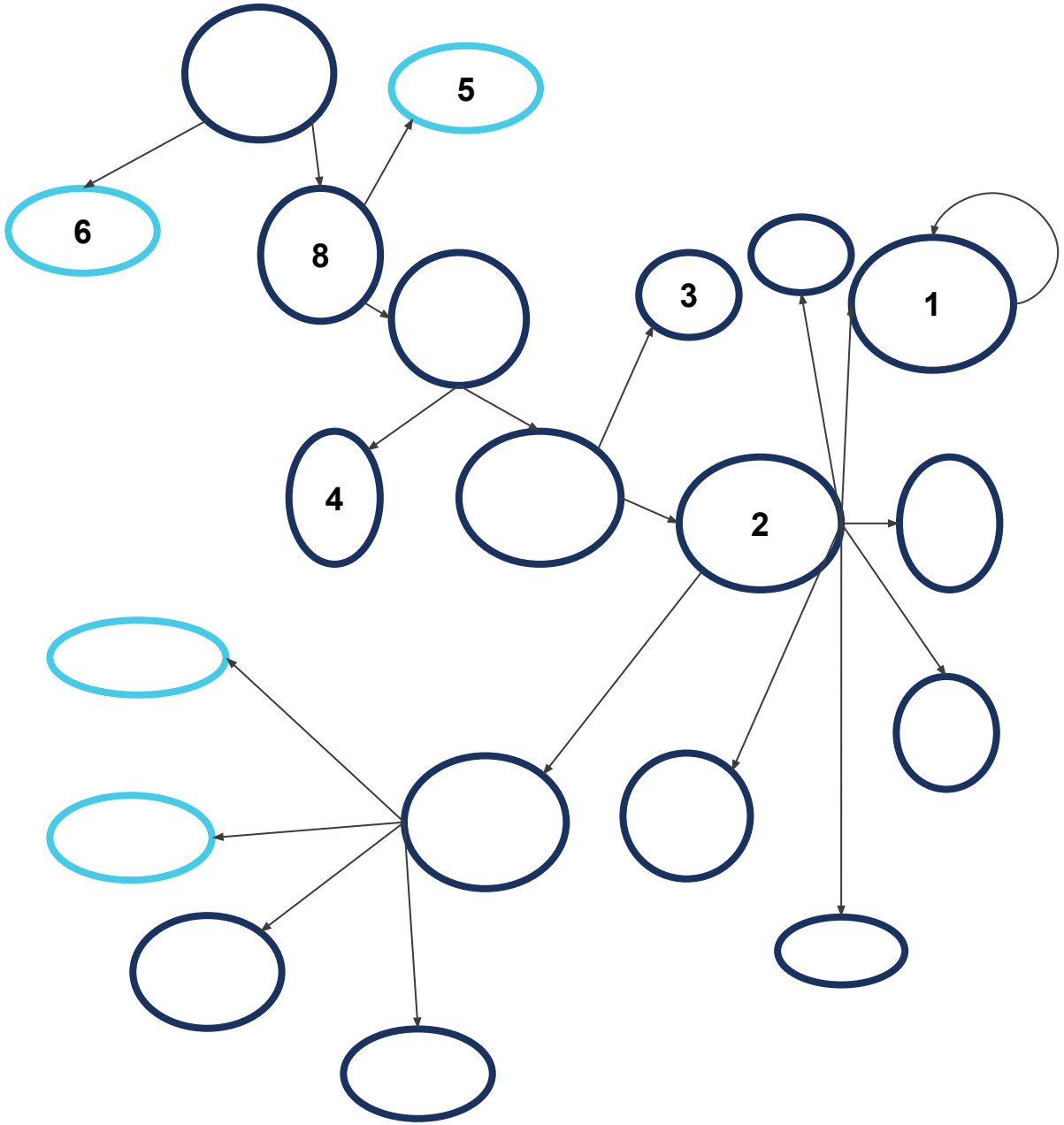
Label	Node	Hub	Authority
1	services.exe	0.8603	8.2465e-19
2	msiexec.exe	0.1396	0.1622
3	svchost.exe	6.0946e-09	0.1396
4	System	1.4190e-18	8.2465e-19
5	0	1.4190e-18	0.0
6	smss.exe	1.4190e-18	8.2465e-19
7	winlogon.exe	1.4190e-18	8.2465e-19



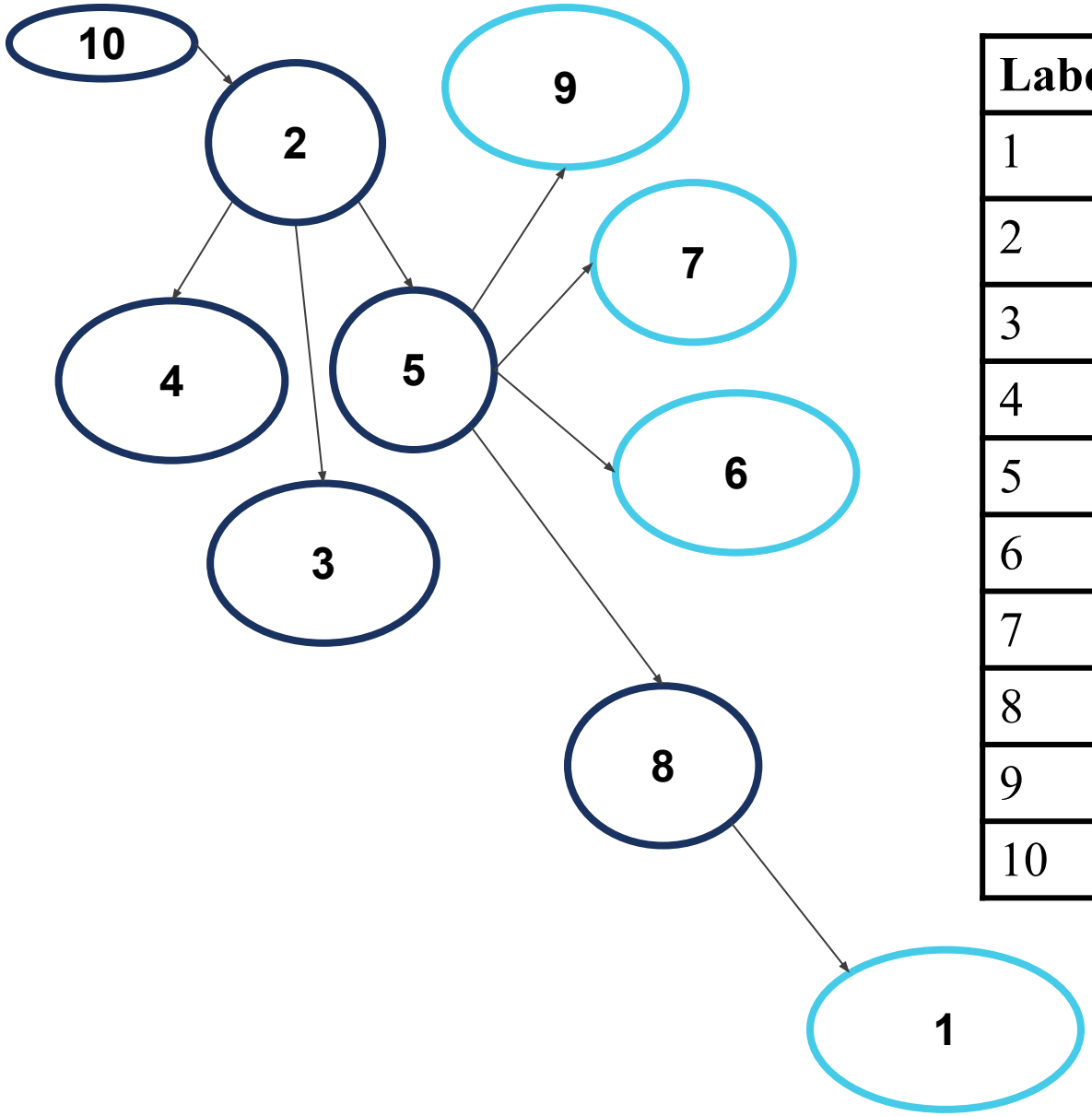
Label	Node	Hub	Authority
1	firefox.exe	0.9999	7.9728e-09
2	explorer.exe	2.3918e-08	1.8807-e37
3	AcroRd32.exe	1.8807e-37	0.2499

PageRank

- Relationship from **Node A** to **Node B** is a vote for **Node B** cast by **Node A**
- Votes cast by nodes that are important weigh more heavily
- Numeric value that represents the importance of a node present on a graph
- Ranking of evidence



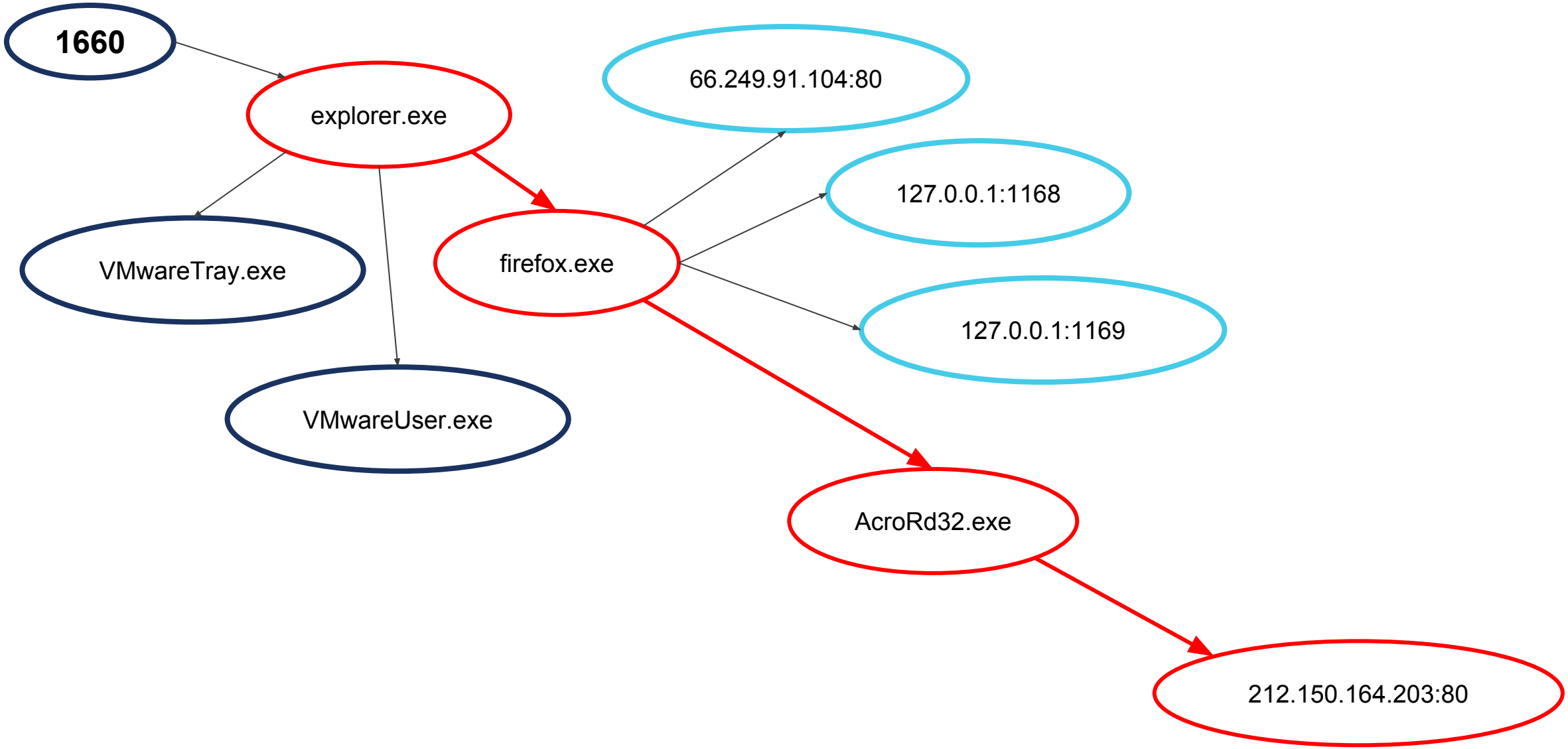
Label	Node	PageRank
1	msiexec.exe	0.2355
2	services.exe	0.0501
3	lsass.exe	0.0501
4	csrss.exe	0.0492
5	192.168.0.1:303080	0.04696
6	smss.exe	0.04696
7	80.206.204.129:0	0.04168
8	System	0.0416



Label	Node	PageRank
1	212.150.164.203:80	0.1431
2	explorer.exe	0.1246
3	VMwareUser.exe	0.1026
4	VMwareTray.exe	0.1026
5	firefox.exe	0.1026
6	127.0.0.1:1169	0.0891
7	127.0.0.1:1168	0.0891
8	AcroRd32.exe	0.0891
9	66.249.91.104:80	0.0891
10	1660	0.0673

Formulate Hypotheses

- Determine that hypothesis ***H*** is supported by a chain of evidence
- Graph traversal
- **Hypothesis**: X downloaded a file that made a network connection


1660
explorer.exe
66.249.91.104:80
127.0.0.1:1168
VMwareTray.exe
firefox.exe
127.0.0.1:1169
VMwareUser.exe
AcroRd32.exe
212.150.164.203:80

Discussion

- Identify user actions
  - Hub
- Evidence ranking
  - PageRank
- Formulate hypotheses
  - Graph traversal

Conclusion & Future Work

- Need for a reliable formalization for digital forensic analysis
- Benefit from structure of mathematics, statistics & probability
- Graph structure establish confidence for semantic interpolation
- Implement a tool