Towards an Operational Semantic Theory of Cyber Defense Against Advanced Persistent Threats

STEVE MECKL, GHEORGHE TECUCI, MIHAI BOICU, DORIN MARCU

GEORGE MASON UNIVERSITY

LEARNING AGENTS CENTER

Overview

- **APT and Modern Network Defense**
- Discovery-based APT Detection Framework
- Representation of APT Detection Models
- Mixed-Initiative Learning of APT Detection Models
- Learning Agent Shell Design
- Conclusions and Future Work

Advanced Persistent Threats

Computer Network Exploitation (CNE) groups

Characterized by superior:

- Resources
- Knowledge
- Tactics

Many are suspected to be state-sponsored.

- Low risk/high reward
- Espionage mission

Attack Lifecycle

Structured, multi-stage approach

Reflects professionalism in its approach

Multiple competing models including:

- Lockeed Martin's Kill Chain
- Mandiant/FireEye Attack Lifecycle

APT1: A Case Study

Attributed to Chinese Military unit 61398.

- Activity traced back as far as 2004.
- Cyber espionage mission.

Multi-phased attack methodology

1. Gain access through spearphishing
2. Use multiple trojan horses to maintain presence
3. Obfuscate attack source through C2 network
4. Privilege escalation
5. Lateral movement
6. Exfiltrate data

Malware Evolution

Malware follows normal software engineering lifecycle.

Evolves slowly over time

Some things are easy to change:

- File hash
- Unique strings
- Network addresses

Some things are hard to change:

- Network protocols
- Persistence mechanism

In some cases, malware lineage can be tracked.

APT1: A Case Study

17 Stage 1 Malware programs:

- Beacon to command-and-control server
- Ask for instructions

27 Stage 2 Malware programs:

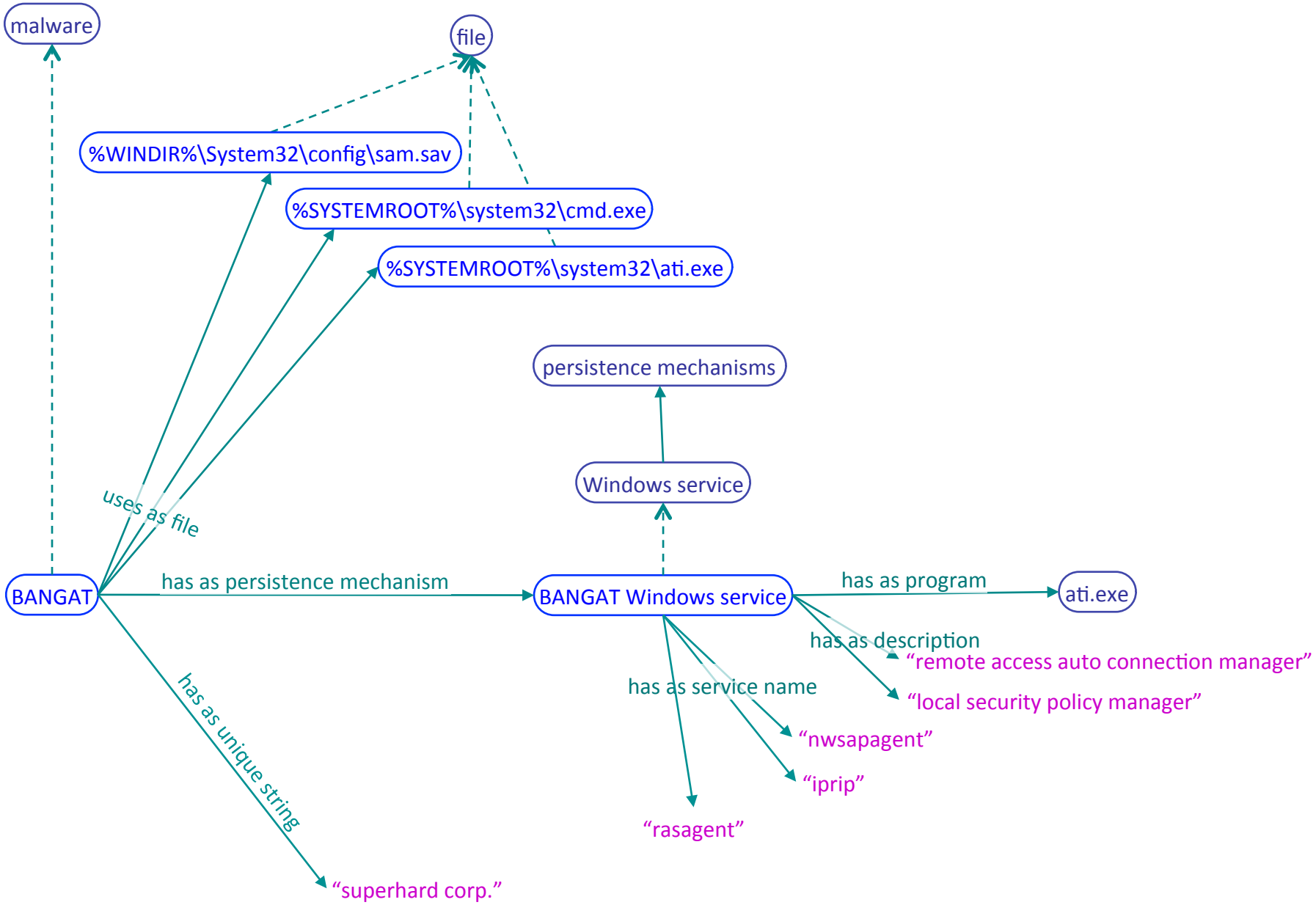
- Provide remote access
- Facilitate lateral movement

4 Data Exfiltration tools:

- Steal data
- Complete the mission

These tools exhibit clusters of behavior.

- We can exploit this!



Tradeoffs in Existing Approaches

	Detection Rate	False Positive Rate	Detect New Attacks?
Misuse Detection	High	Low	No
Anomaly Detection	Medium	High	Yes

Host-based Detection

- Sequences of system calls
- Sets of artifacts
- Memory patterns

Network-based

- 5-tuple (netflow)
- Deep packet inspection (content)
- Protocol validation

State of Network Defense

Cybersecurity Operations Center (CSOC) is the security hub:

People:

- Analysts, forensics experts, system administrators

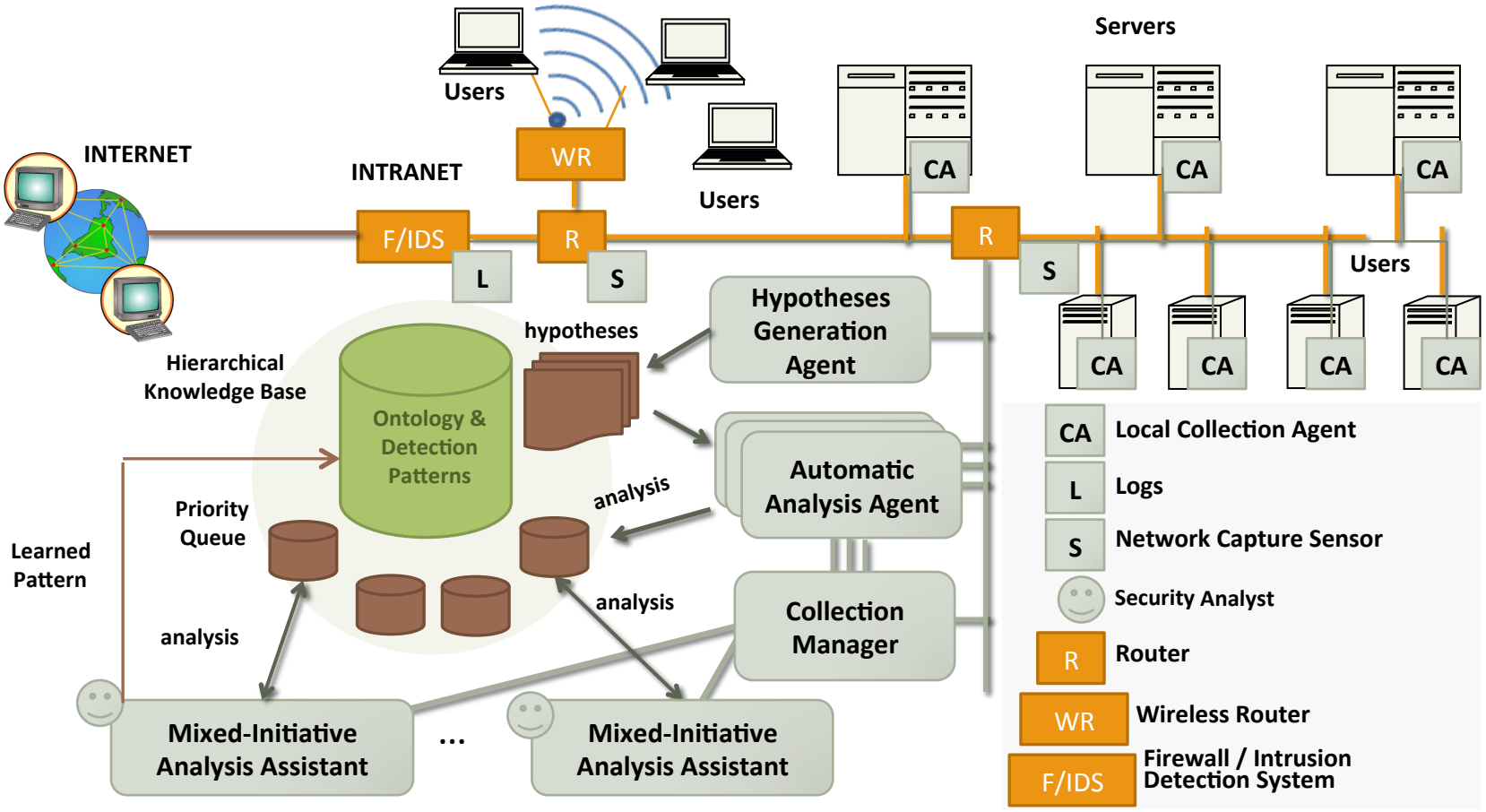
Process:

- Multi-tiered incident handling methodology
  - Tier 1: Incident triage, basic log analysis and intel. gathering
  - Tier 2: In-depth network- and host-based forensics

Technology:

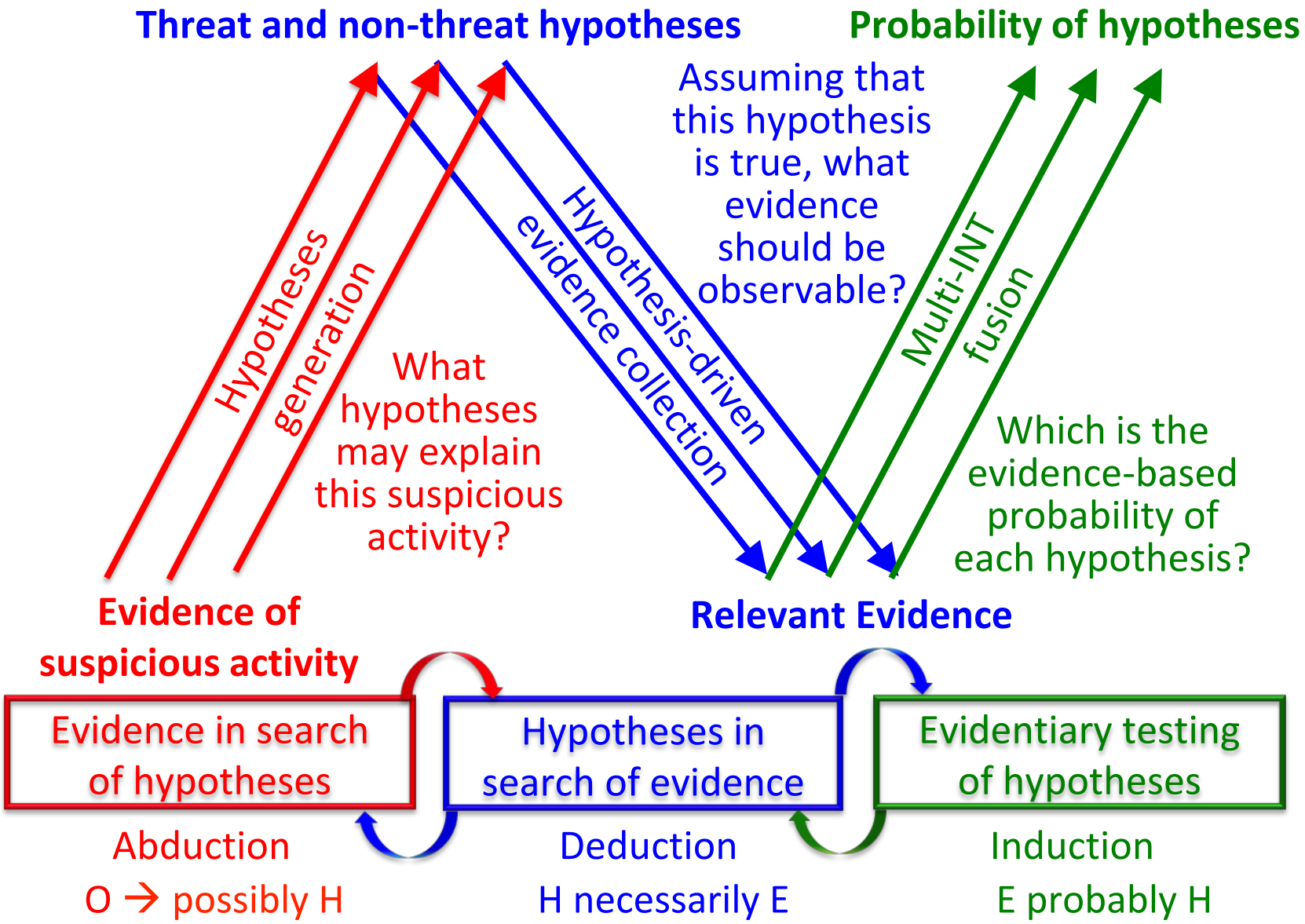
- Network- and host-based sensors
- Intrusion detection systems
- Security Incident/Event Management

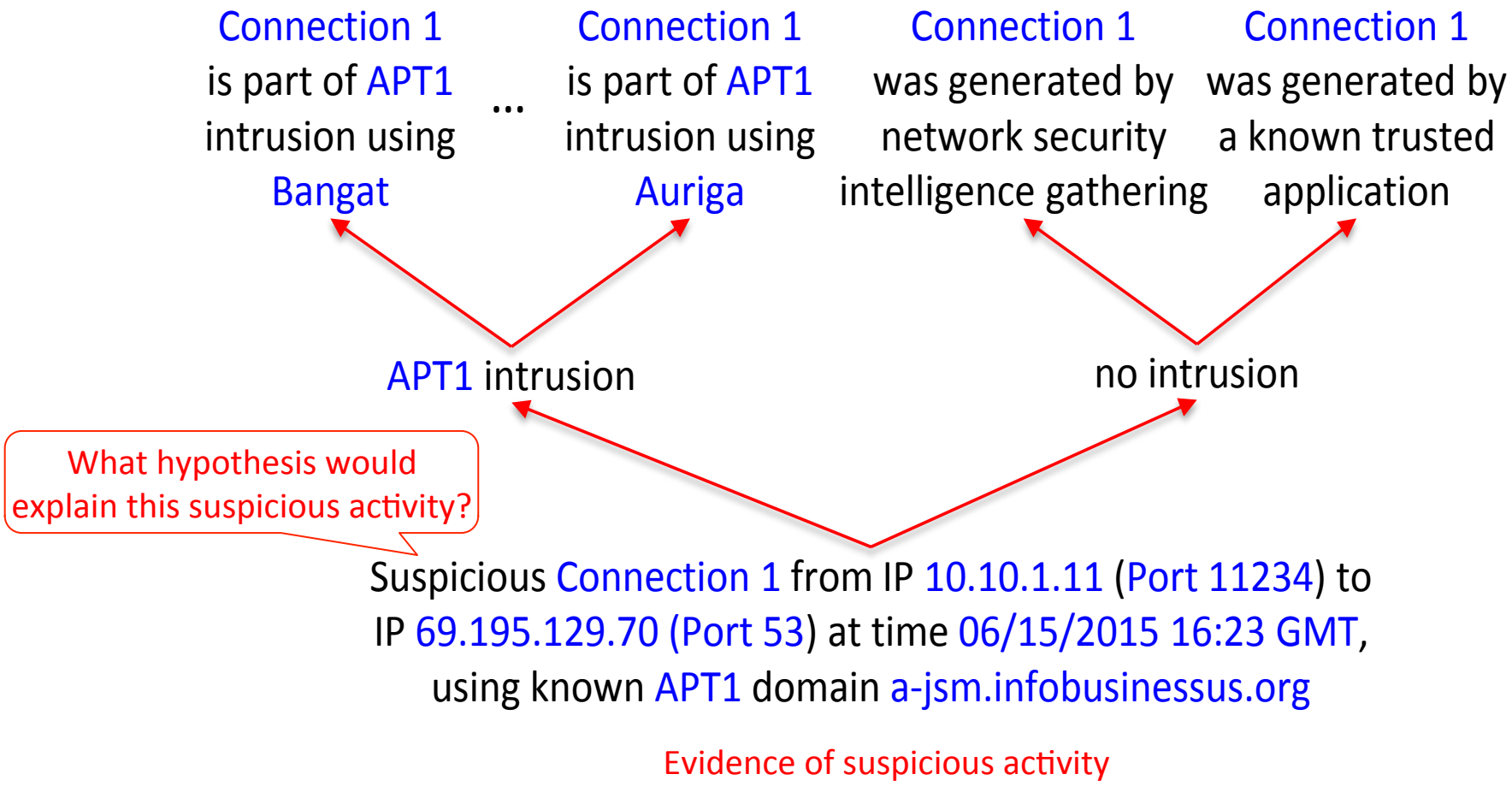
Our Approach

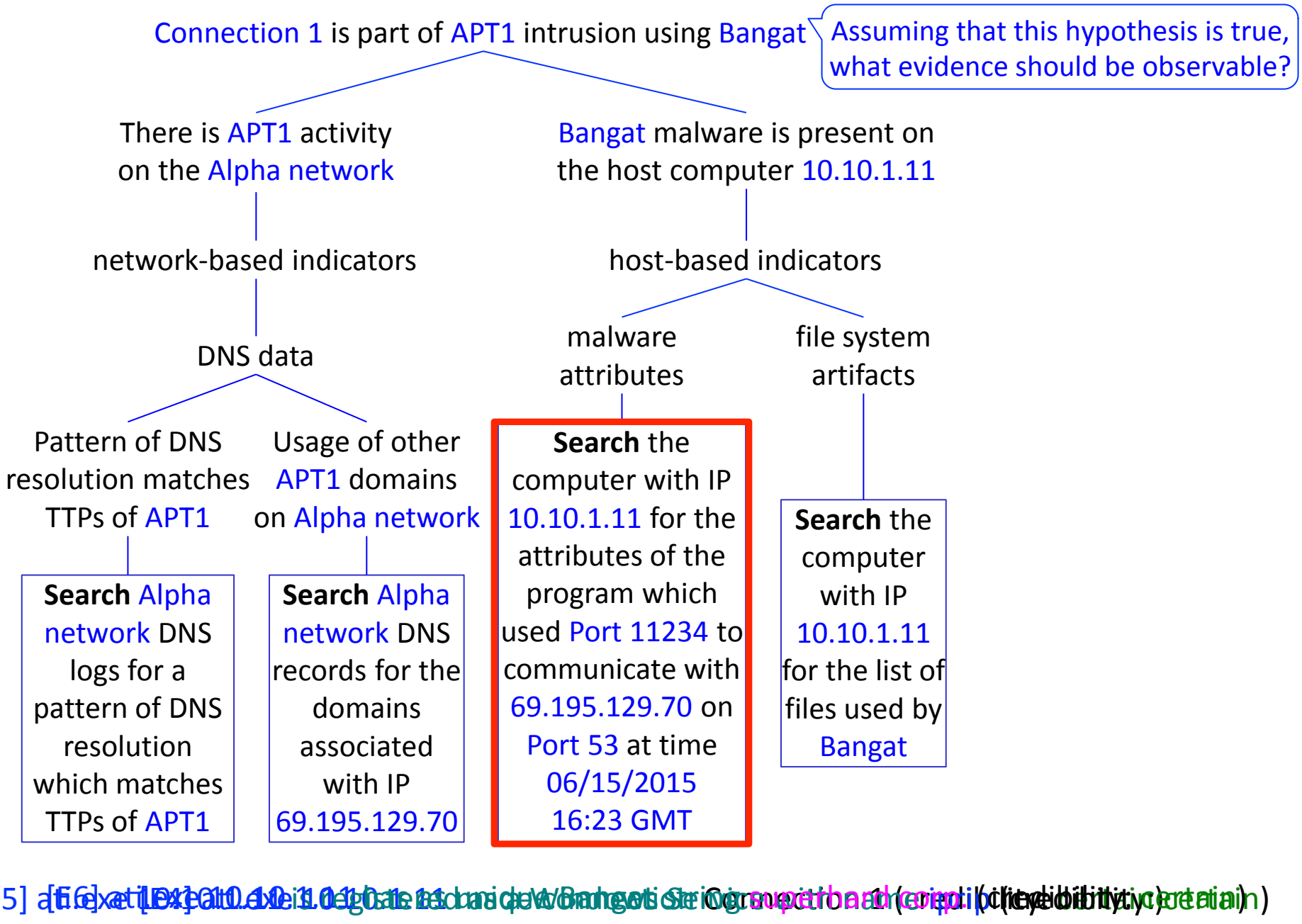


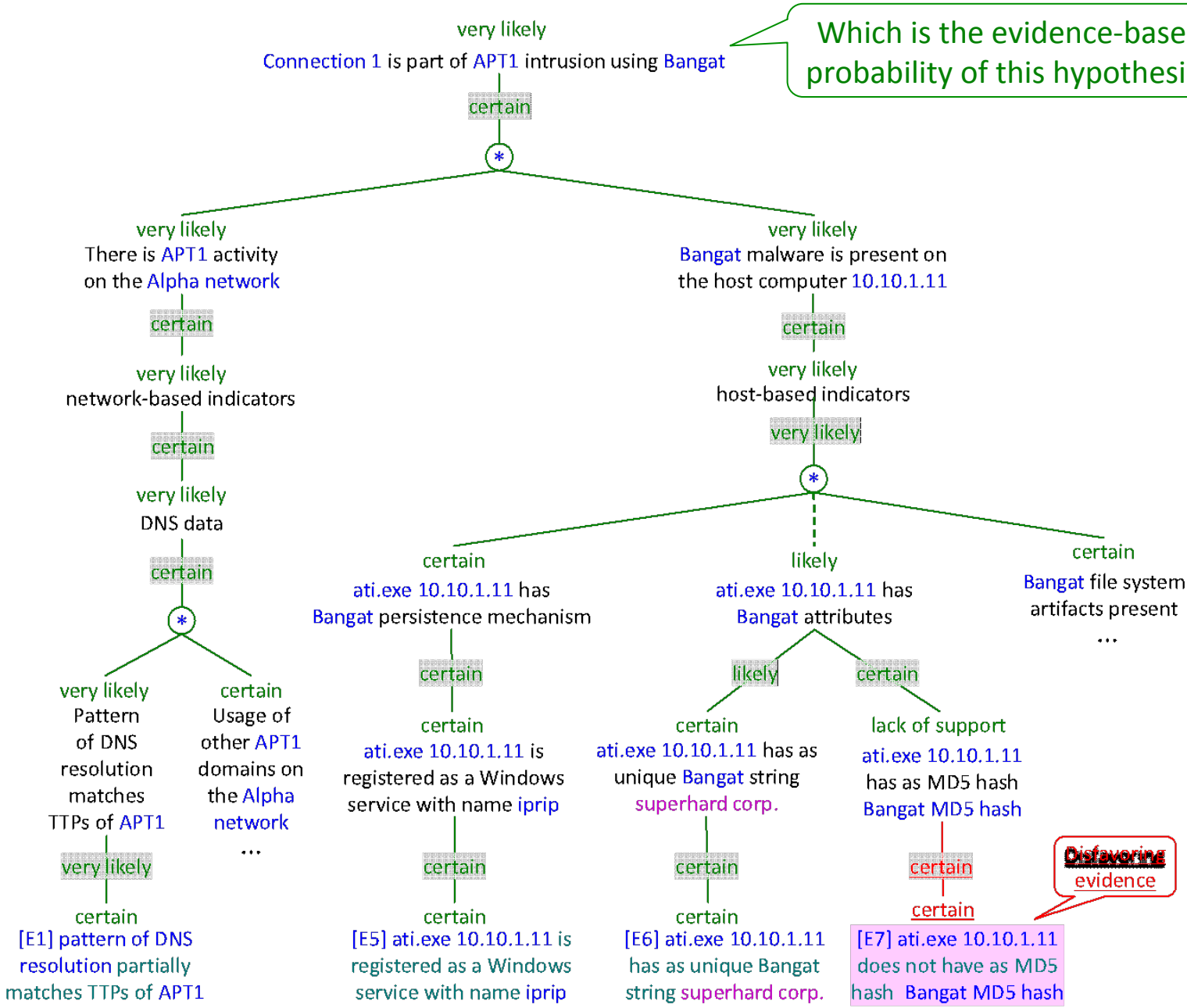
Overview

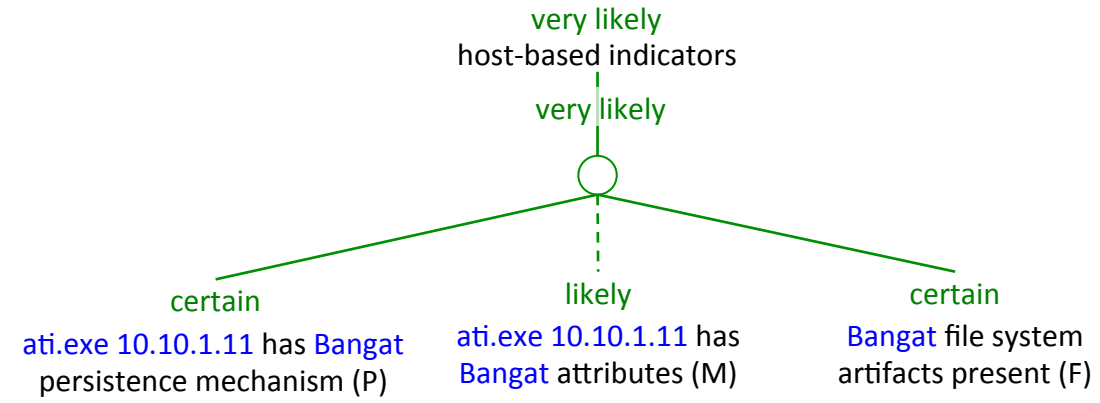
- APT and Modern Network Defense
- **Discovery-based APT Detection Framework**
- Representation of APT Detection Models
- Mixed-Initiative Learning of APT Detection Models
- Learning Agent Shell Design
- Conclusions and Future Work

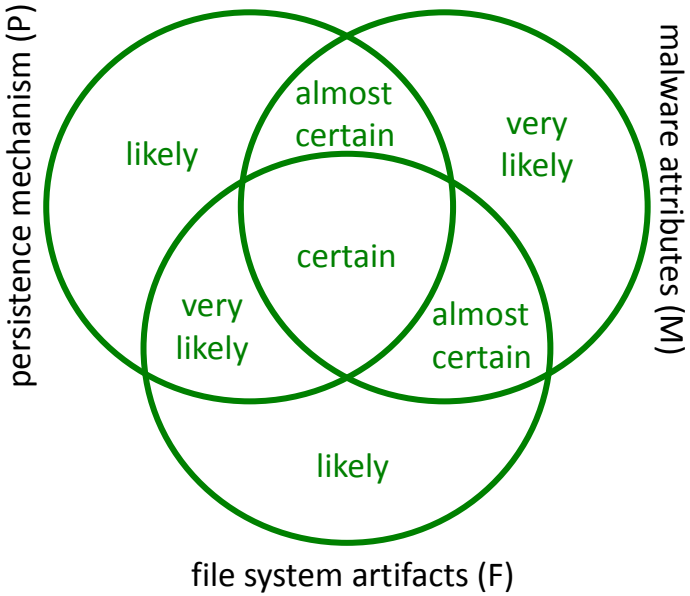


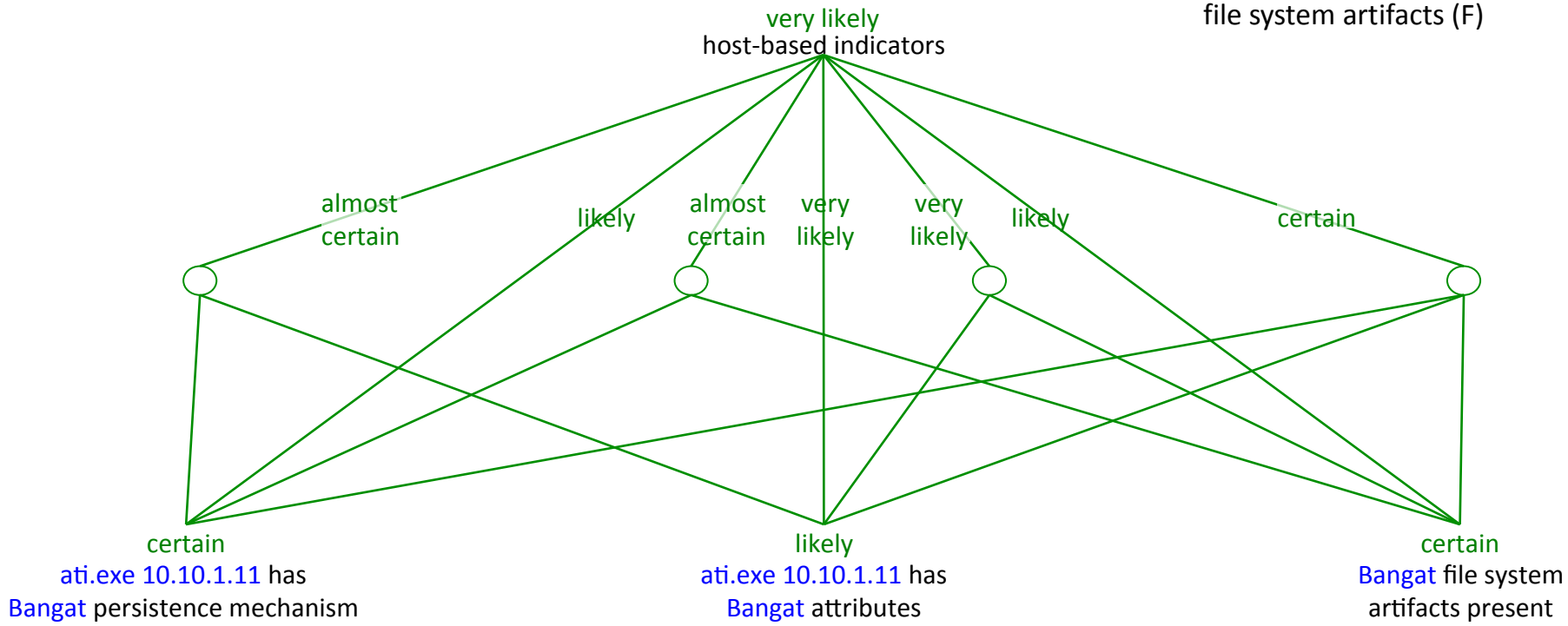








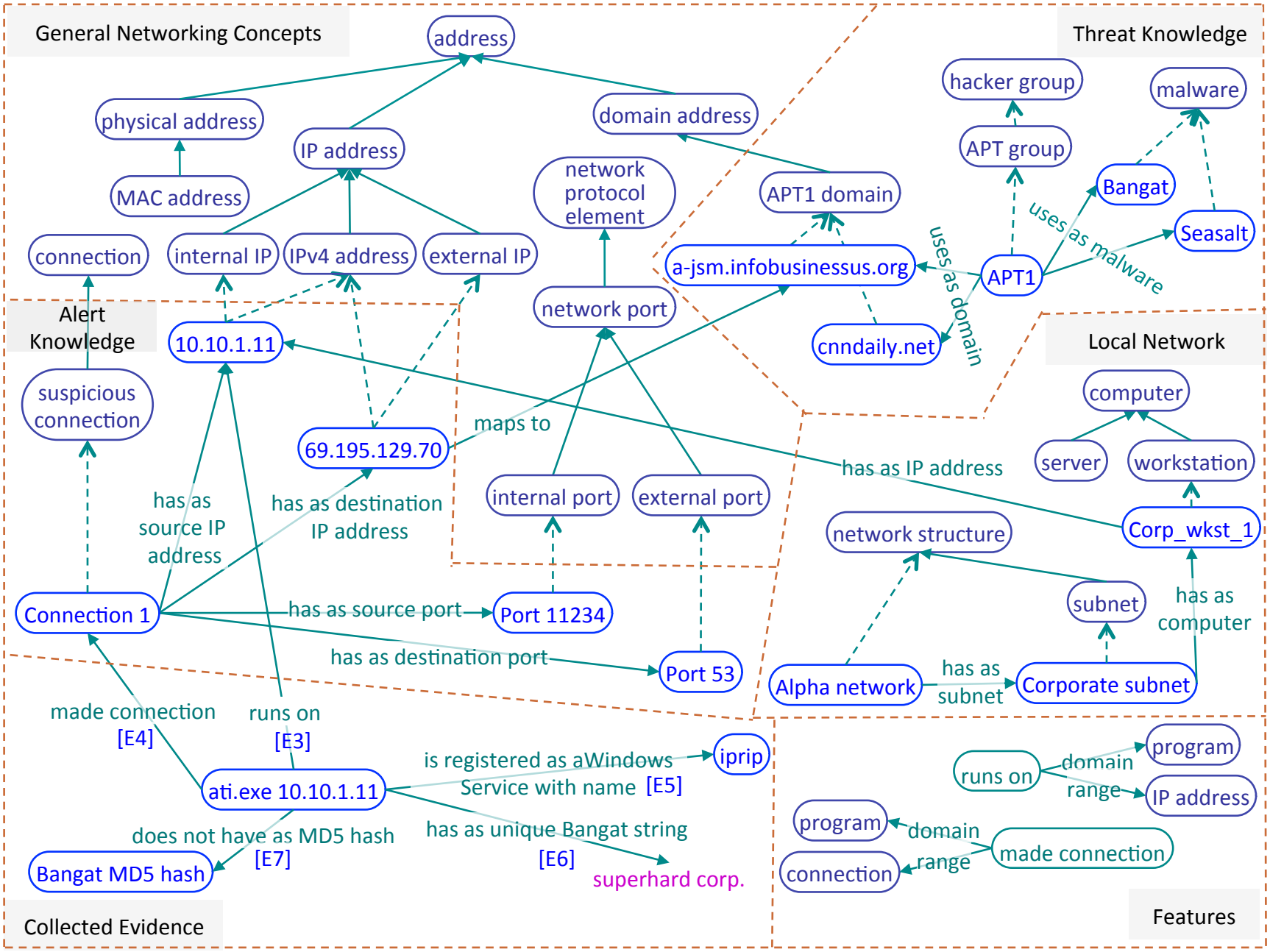




Complex Operator Example

Overview

- APT and Modern Network Defense
- Discovery-based APT Detection Framework
- **Representation of APT Detection Models**
- Mixed-Initiative Learning of APT Detection Models
- Learning Agent Shell Design
- Conclusions and Future Work



Overview

- APT and Modern Network Defense
- Discovery-based APT Detection Framework
- Representation of APT Detection Models
- **Mixed-Initiative Learning of APT Detection Models**
- Learning Agent Shell Design
- Conclusions and Future Work

Condition

V1 is suspicious connection
has as source IP address V5
has as destination IP address V7
has as source port V6
has as destination port V8
has as time stamp V9

V2 is hacker group
uses as domain V10
uses as malware V3

V3 is malware

V4 is network structure
has as subnet V11

V5 is internal IP

V6 is network port

V7 is external IP
maps to V10

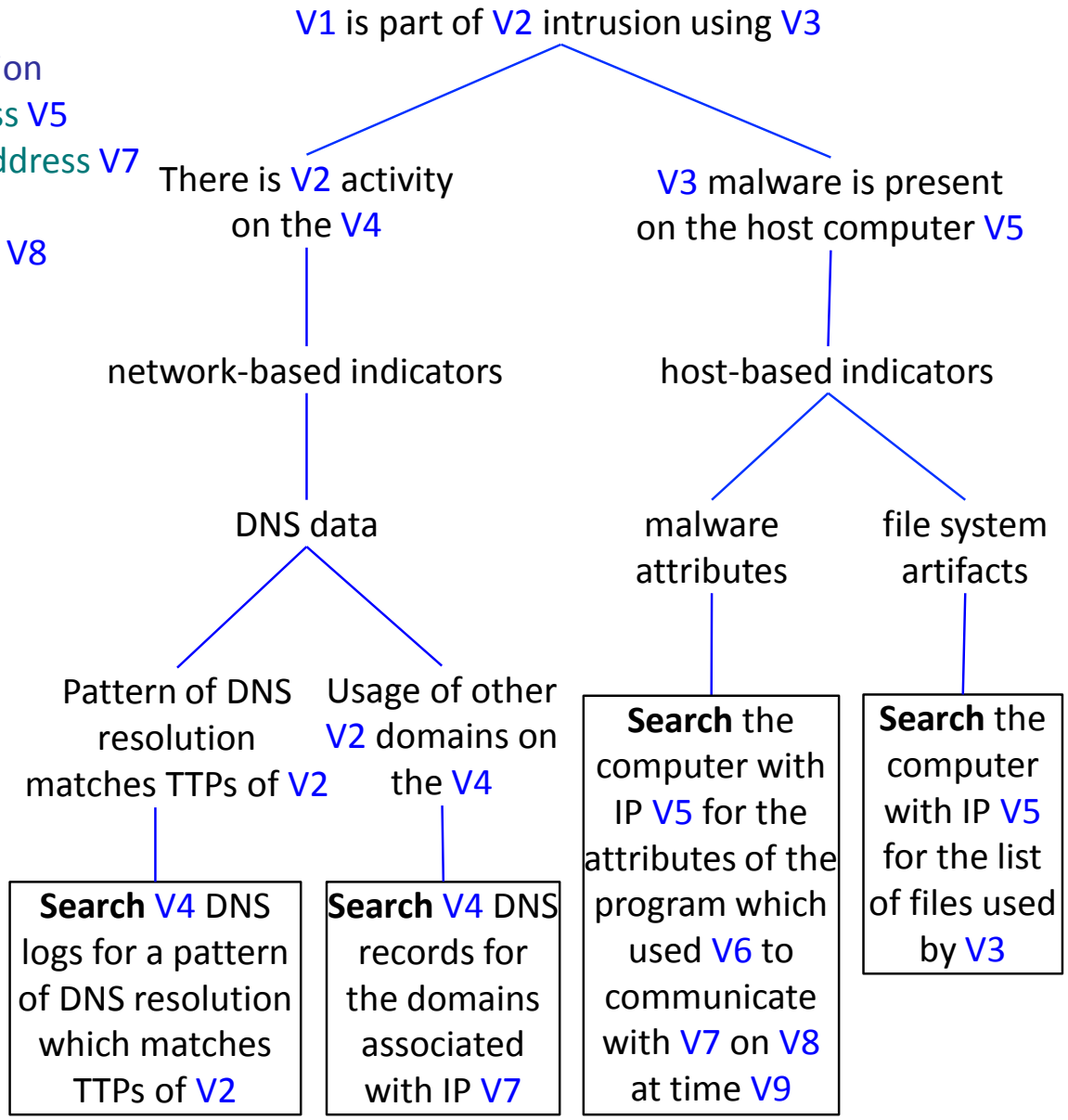
V8 is network port

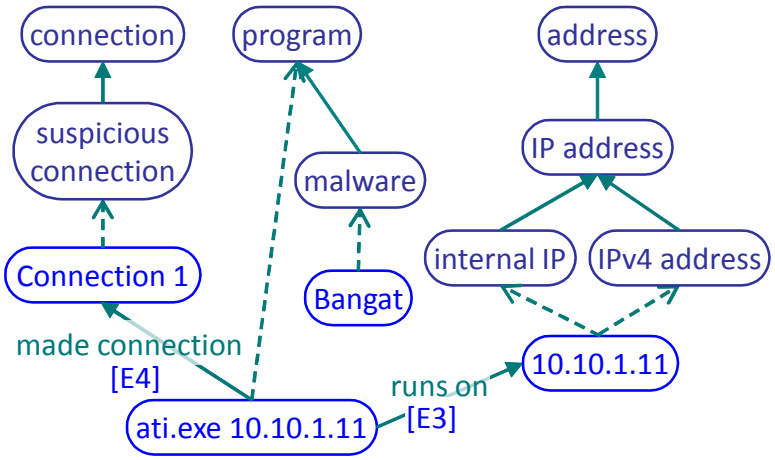
V9 is time stamp

V10 is domain address

V11 is network structure
has as computer V12

V12 is computer
has as IP address V5



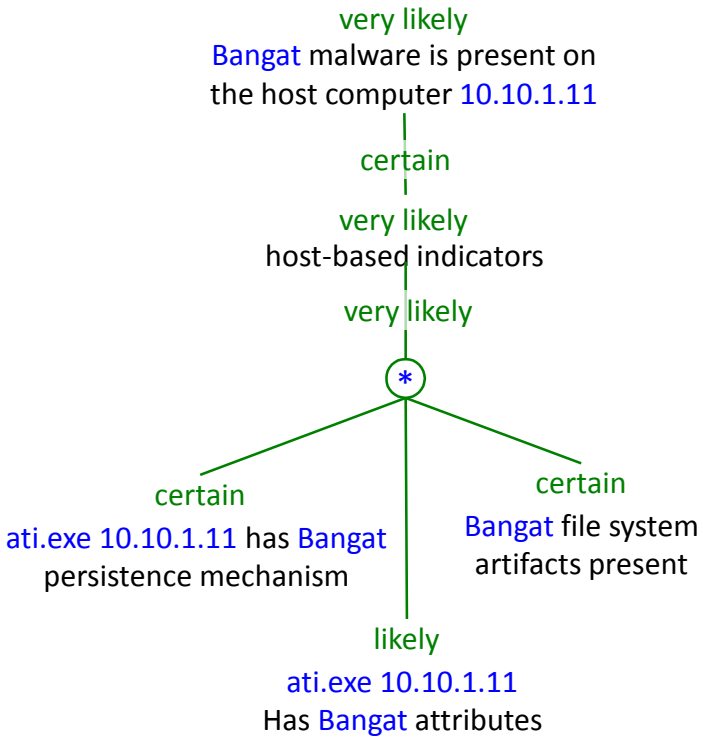


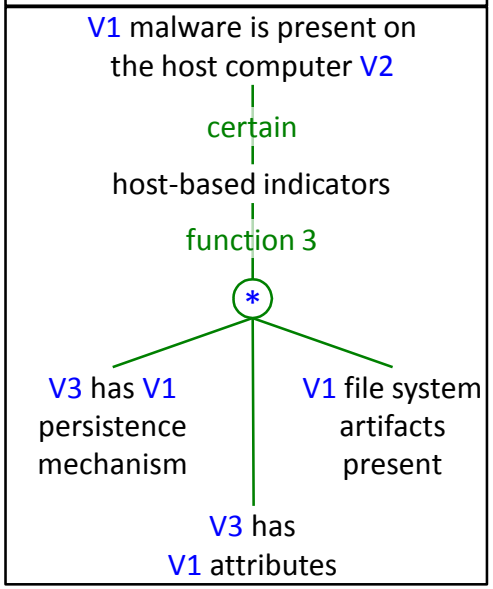
Upper bound condition

V1 is program
V2 is IP address
V3 is program
made connection V4 [V5]
runs on V2 [V6]
V4 is connection
V5 is evidence
V6 is evidence

Lower bound condition

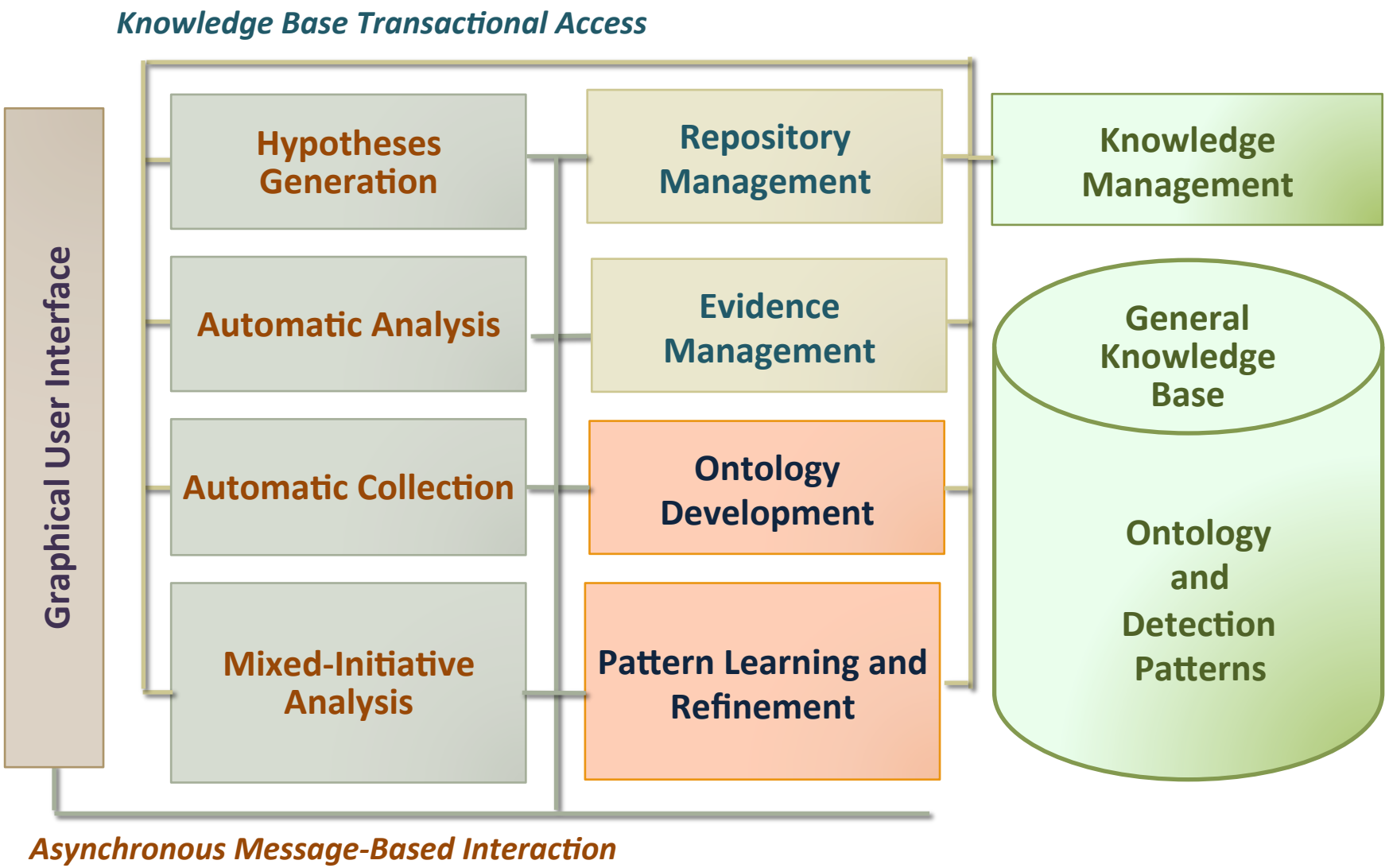
V1 is malware
V2 is (internal IP, IPv4 address)
V3 is program
made connection V4 [V5]
runs on V2 [V6]
V4 is suspicious connection
V5 is evidence
V6 is evidence

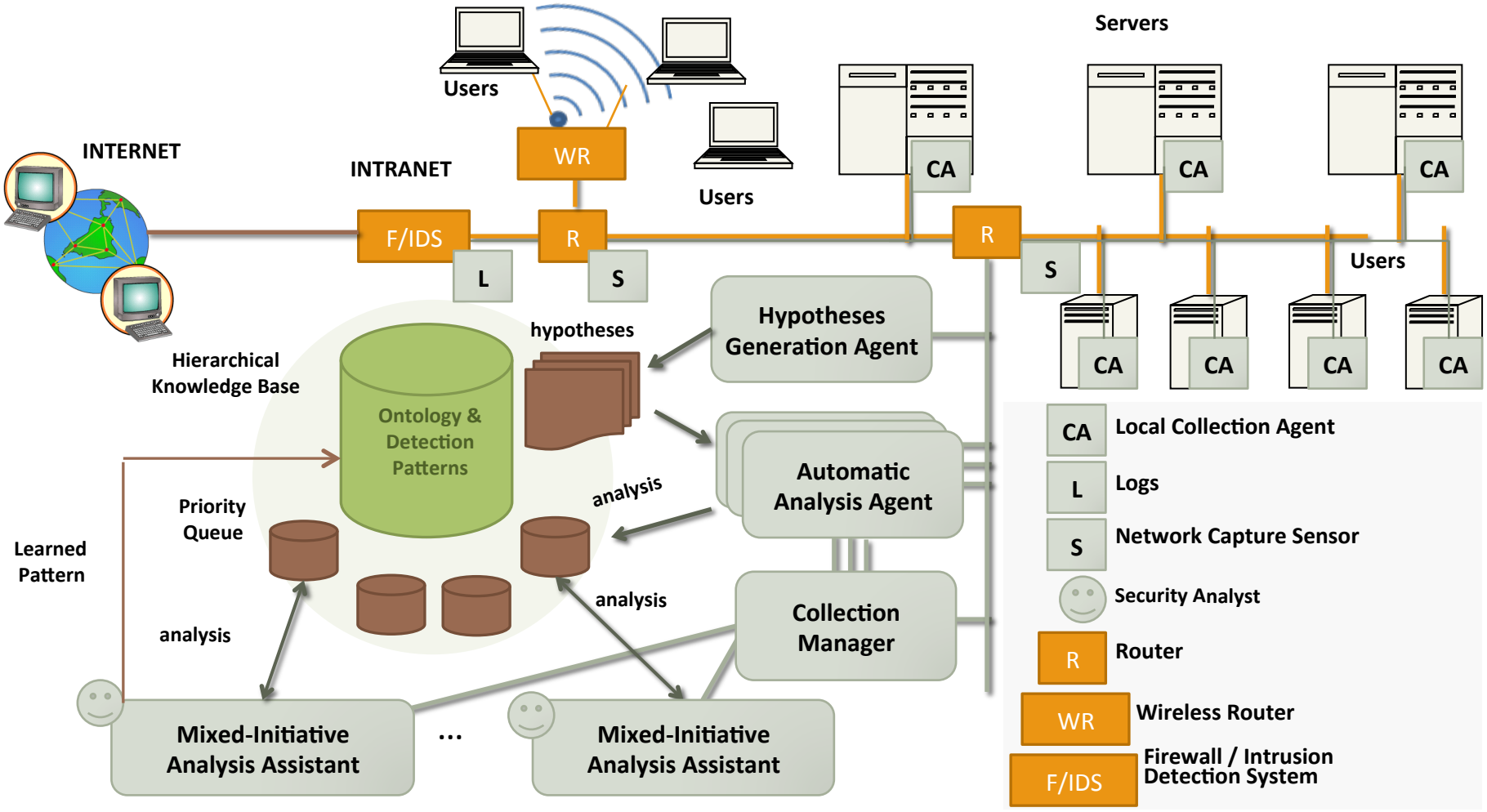




Overview

- APT and Modern Network Defense
- Discovery-based APT Detection Framework
- Representation of APT Detection Models
- Mixed-Initiative Learning of APT Detection Models
- **Learning Agent Shell Design**
- Conclusions and Future Work





Overview

- APT and Modern Network Defense
- Discovery-based APT Detection Framework
- Representation of APT Detection Models
- Mixed-Initiative Learning of APT Detection Models
- Learning Agent Shell Design
- **Conclusions and Future Work**

Conclusions & Future Work

Semantic theory of APT Network Defense

- Basis of cognitive assistance for APT defense.
- Increase CSOC efficiency
- Increase detection accuracy
- Enable continuous learning from SMEs
- Detect new threats

Future Research:

- Further develop reasoning and learning methods
- Develop prototype learning agent shell
- Develop specific agents
- Integrate into real CSOCs
- High defense agility against sophisticated threats

Questions?

Steve Meckl

smeckl@masonlive.gmu.edu

https://github.com/smeckl