accenture

Strategy | Consulting | Digital | Technology | Operations

Automating Incident Investigations: Sit Back and Relax, Bots are Taking Over…..

Elvis Hovor
Mohamed El-Sharkawi

November 3, 2016

Automating Incident Investigations: Sit Back and Relax, Bots are Taking Over…..

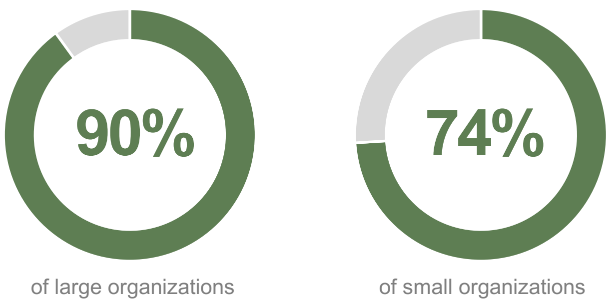
Content **Slide #**

Current Challenges With Incident Investigation

The Hard Facts – Get ready to be “Breached”

Suffering a breach is almost certainly inevitable

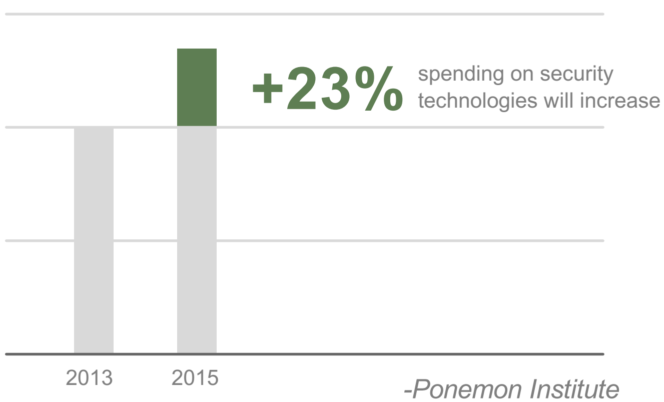
In 2015 Information Security Breaches survey show that **90% of large organizations** and **74% of small organizations** suffered a data breach in 2014



-Department for Business, Innovation and Skills (BIS)

Breaches are getting more expensive for organizations

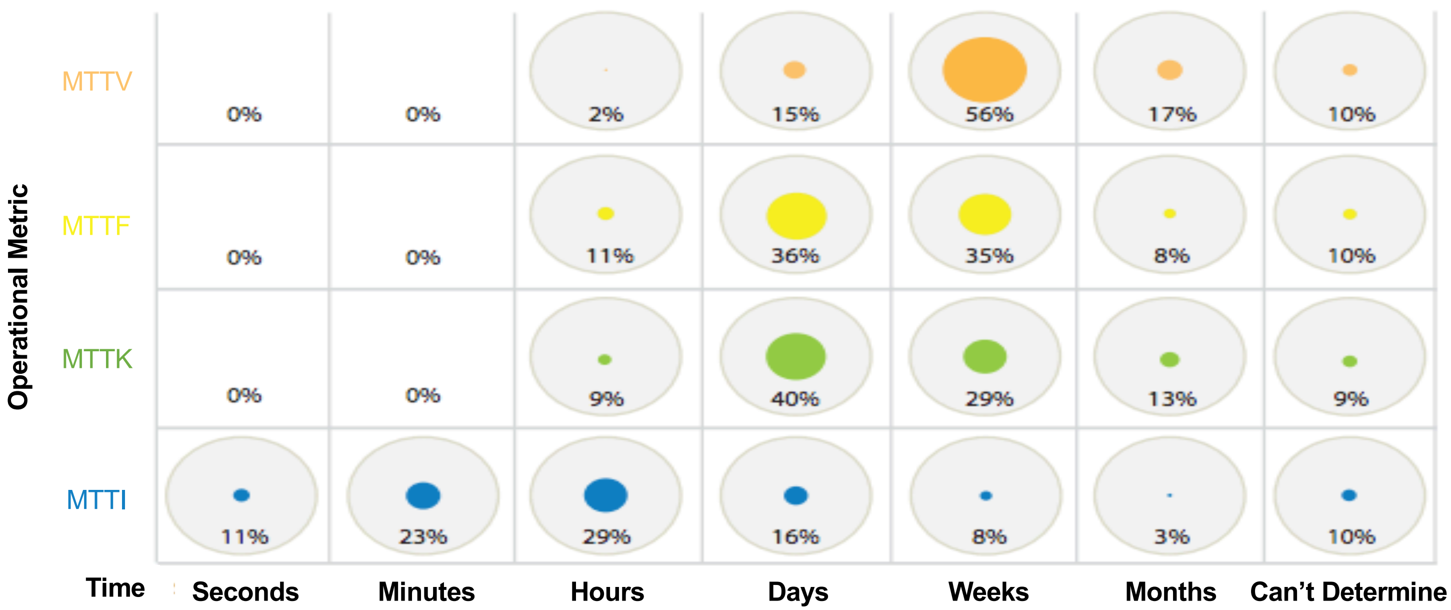
The average total cost of a data breach for the participating companies **increased 23 percent** over the past two years **to $3.79 million**



-Ponemon Institute

The Problem: Incident Investigation & Response

Incident response is **time consuming**, **resource intensive,** and a **costly process**. Ineffective triaging allows early warnings of compromises to **slip through the cracks**.



MTTV	Meantime to verify
MTTF	Meantime to Fix
MMK	Meantime to Know
MTTI	Meantime to Identify

-Ponemon Institute

The Problem: Incident Investigation & Response

Time	Seconds	Minutes	Hours	Days	Weeks	Months	Can't Determine
MTTK	0%	0%	9%	40%	29%	13%	9%

-Ponemon Institute

Detect, Understand, & Respond

Organizations are focused on **detecting compromises quicker** and **understanding the scope of the compromise** in order to adequately remediate the breach.

WHERE the attacker entered and **HOW** the network was compromised → the attackers TTPs

WHEN and how long has the adversary compromised my network

Systems that were compromised and WHAT damage was done by adversary

WHY you were compromised and what the motivations of the adversary is

Attribution, WHO compromised your network and what type of adversary it is

Asking the RIGHT Questions

The Solution: Process Orchestration & Task Automation in IR

Key Benefits

Accelerated Response Times

Automate the incident response processes to speed reaction and scale human skill.

Respond to threats in real time and without human intervention

Efficiency

Streamlines repeated processes.
Reduces the risk of human error in the investigation process.

Saves time, money and resources.

Risk Reduction

Reduces risk of threats slipping through the cracks by providing the ability to investigate more alerts

Simplified IR Process

Reduces Complexity, and Disruption of the Incident Response process.

Empower Users

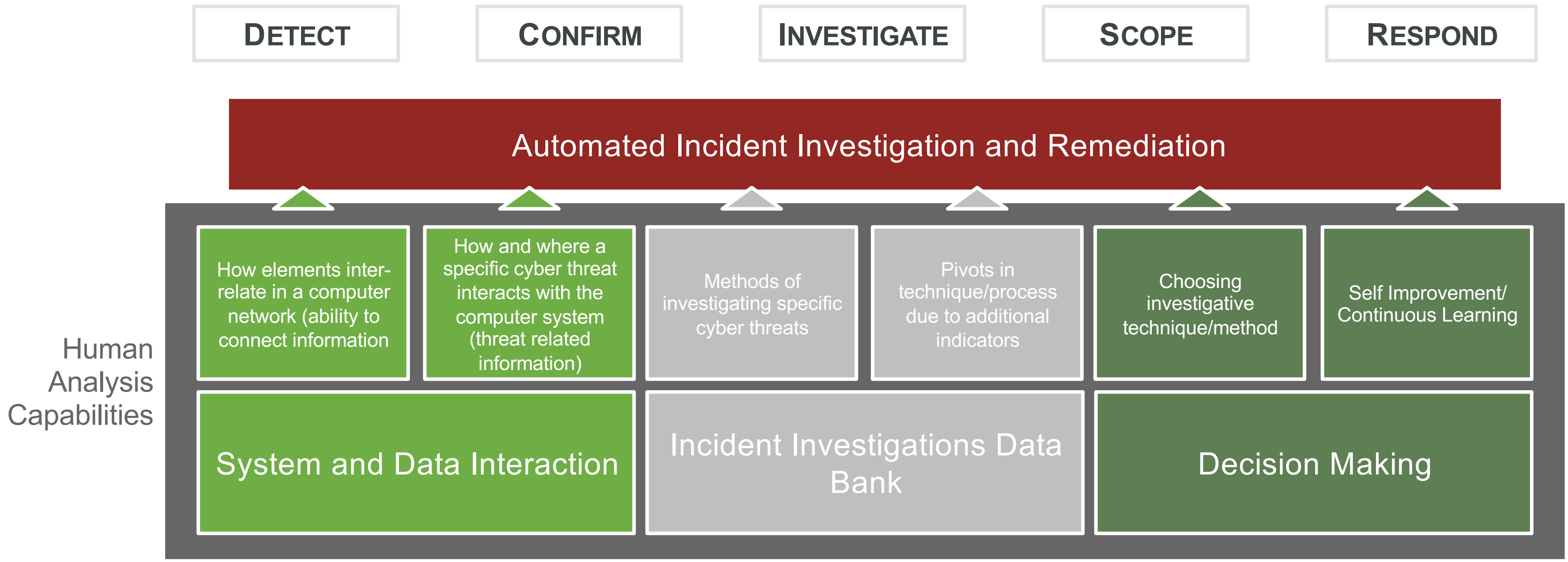
Bridge the skills gap to enable less skilled analysts tackle more complex threat investigations.

Workforce Satisfaction

Employee attrition from long hours of work will be reduced.

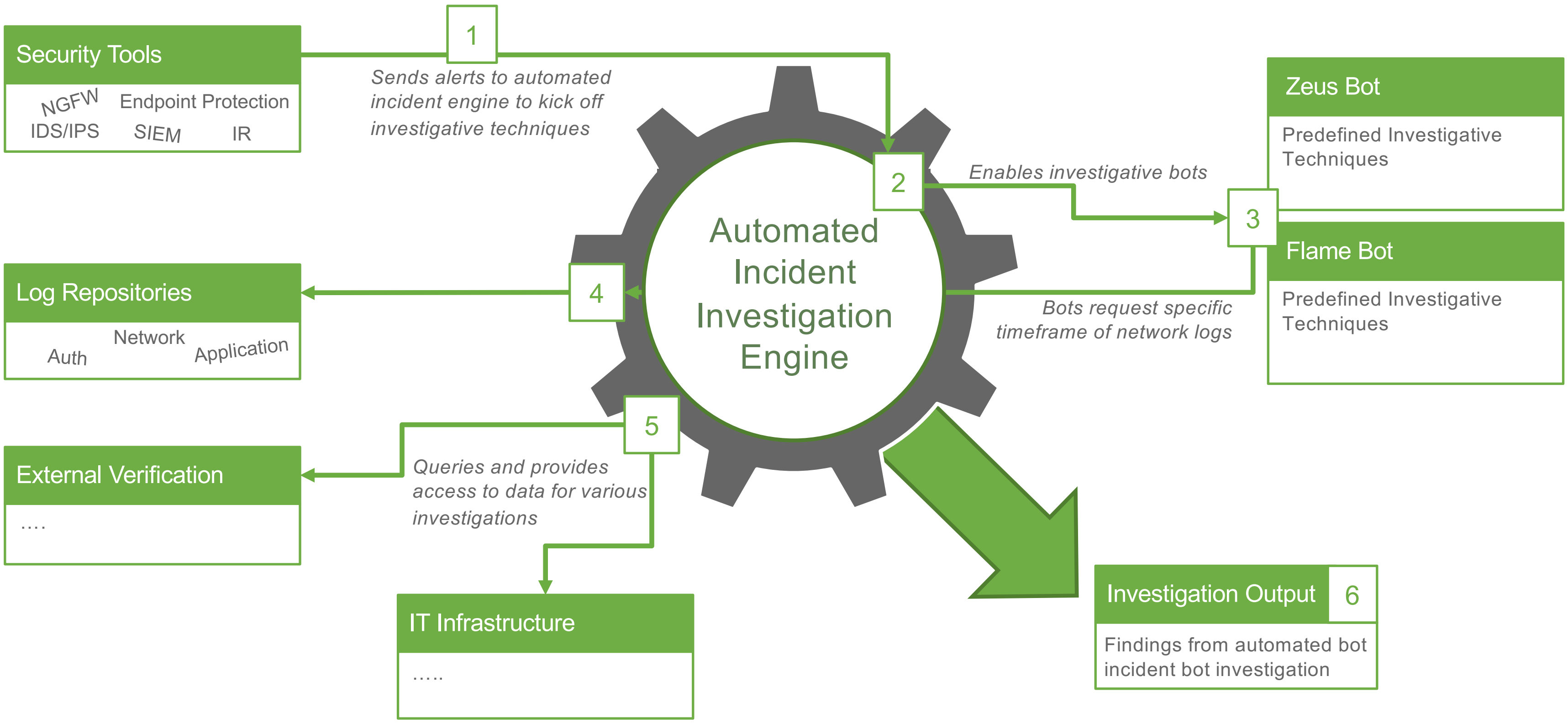
Working on mundane uninteresting work will be eliminated

Capability Framework



Our Beginning - Unleashing the Bots

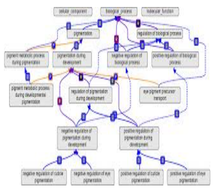
Our initial research started with one objective in mind, to build a platform agnostic engine to moderate various investigative bots that can carry out specific investigative tasks.



Enabling an Incident Investigation Capability

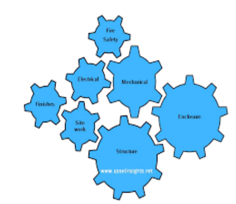
Effective Automation of Incident Investigations involves leveraging key capabilities

Comprehensive Cyber Ontology



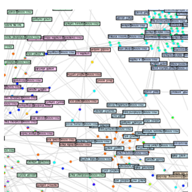
A combination of multiple multi-focused cyber security ontologies that transforms the investigative task by connecting information

System Interconnectivity and Data Sharing



Data drives all incident response investigations. Access to the right type of data and having the ability to utilize data-on-demand from any system to enrich the investigation is critical

Global Incident Investigation Corpus



A dictionary of investigative techniques used for detecting and investigating threats and the data needed for investigation

Cognitive Decision Making & Learning

Automation in complex processes when a machine has ability to pivot and address unaccounted specified workflows

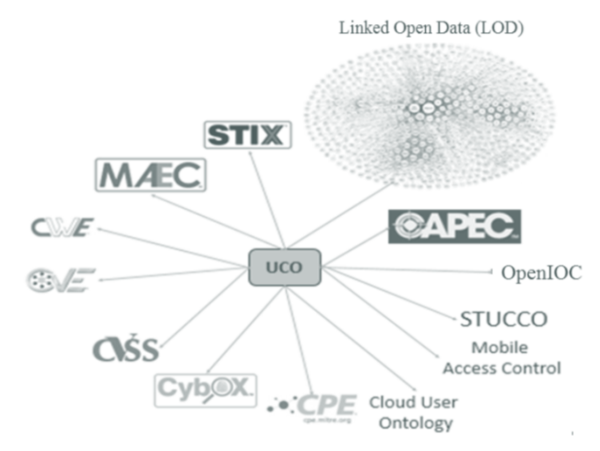
Modeling Manual Investigations

A framework to describe an investigation at different levels of detail. This includes independent vocabulary can be used to describe the researching process in more detail and the ability for different technologies to communicate and exchange data based on well-defined widely adopted standards.

Layer 1 Ontology

- Unified Cyber Ontology (UCO)
- Integrated Cyber Analysis System (ICAS)
- CMU Insider Threat Ontology
- Malware Classes Ontology

Layer 2 Interoperability

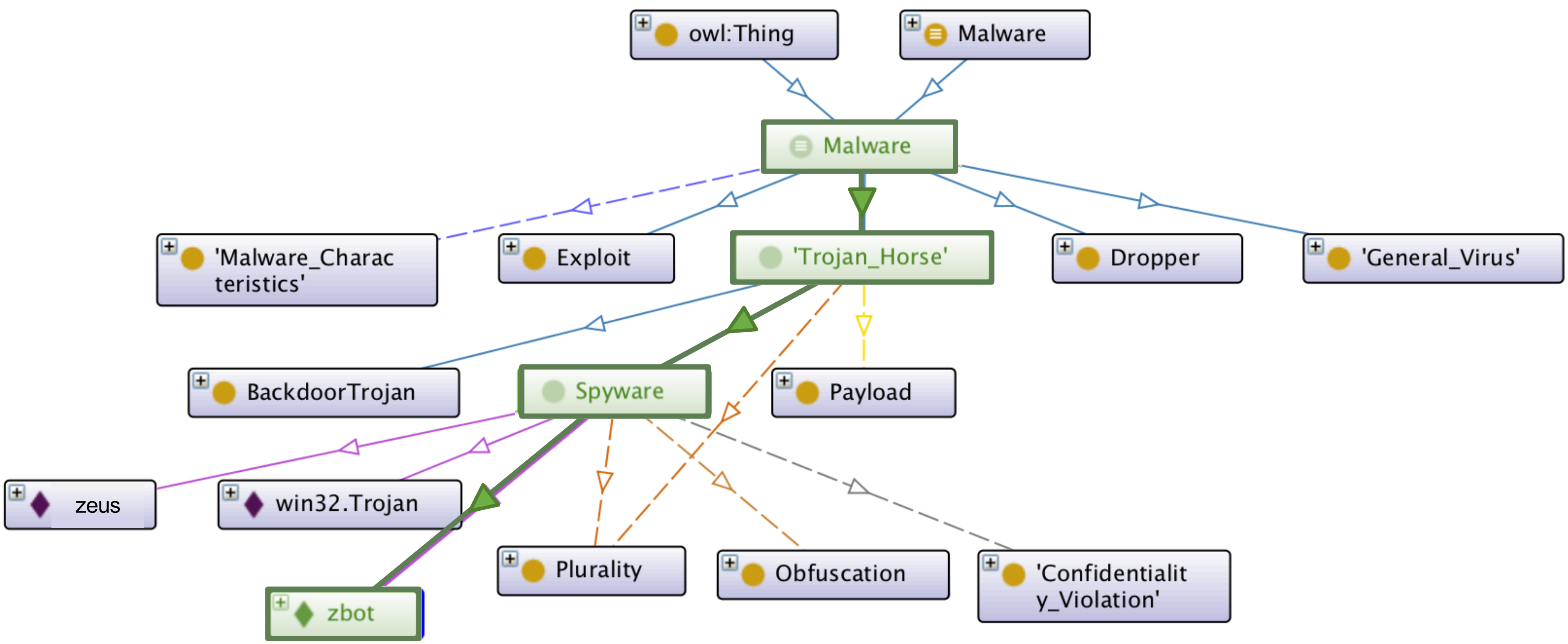


Layer 3 Investigation Techniques

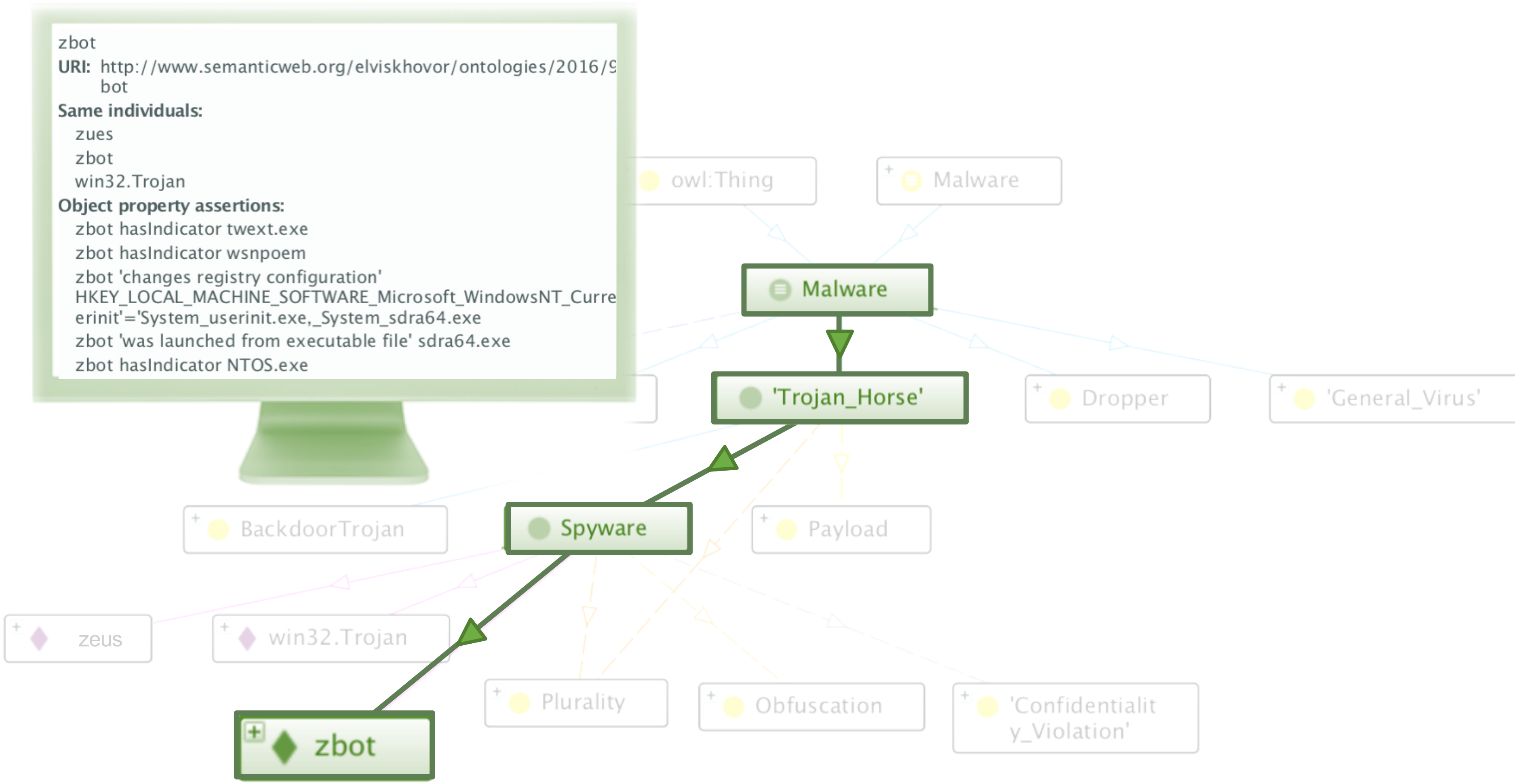
- Incident Investigation Reports
- Incident Investigations Corpus
- Digital Forensics Analysis Expression (DFAX)
- Common Remediation Enumeration (CRE)

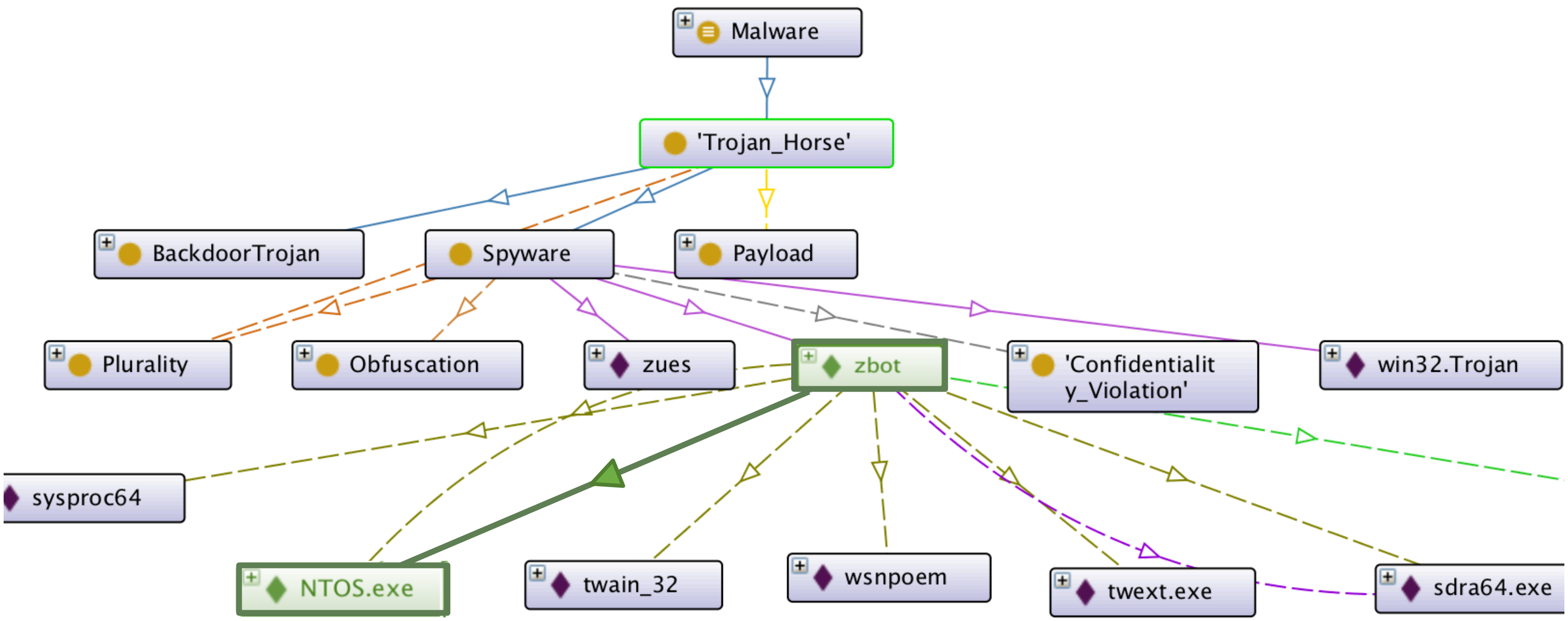
- Cognitive Decision Making

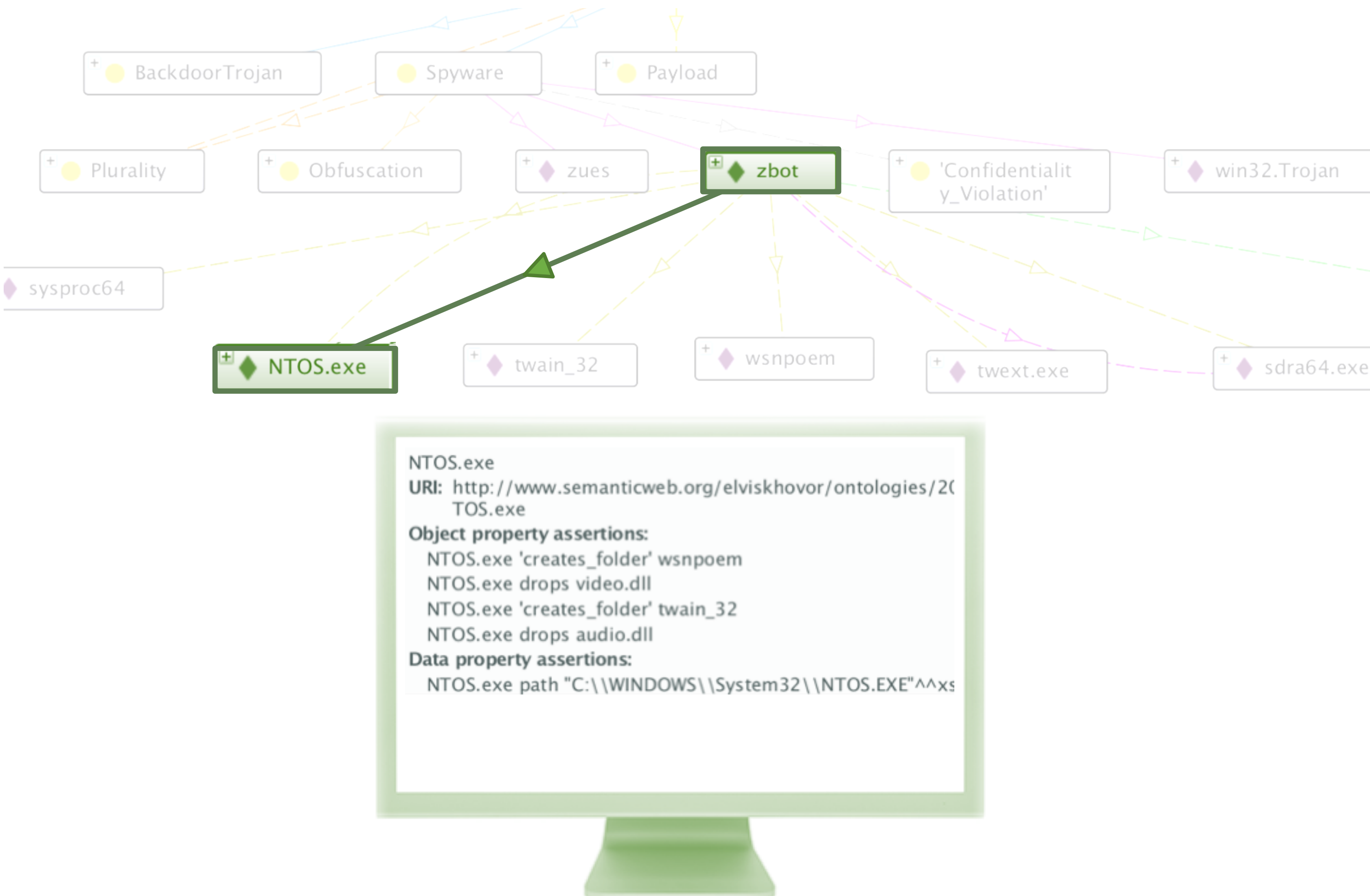
COMPREHENSIVE CYBER ONTOLOGY



owl:Thing
Malware
Malware
'Malware_Characteristics'
Exploit
'Trojan_Horse'
Dropper
'General_Virus'
BackdoorTrojan
Spyware
Payload
zeus
win32.Trojan
Plurality
Obfuscation
'Confidentiality_Violation'
zbot


zbot
URI: http://www.semanticweb.org/elviskhovor/ontologies/2016/9
bot
Same individuals:
zues
zbot
win32.Trojan
Object property assertions:
zbot hasIndicator twext.exe
zbot hasIndicator wsnpoem
zbot 'changes registry configuration'
HKEY_LOCAL_MACHINE_SOFTWARE_Microsoft_WindowsNT_Curre
erinit'='System_userinit.exe,_System_sdra64.exe
zbot 'was launched from executable file' sdra64.exe
zbot hasIndicator NTOS.exe
owl:Thing
Malware
Malware
'Trojan_Horse'
Dropper
'General_Virus'
BackdoorTrojan
Spyware
Payload
zeus
win32.Trojan
Plurality
Obfuscation
'Confidentialit
y_Violation'
zbot


Malware
'Trojan_Horse'
BackdoorTrojan
Spyware
Payload
Plurality
Obfuscation
zues
zbot
'Confidentiality_Violation'
win32.Trojan
sysproc64
NTOS.exe
twain_32
wsnpoem
twext.exe
sdra64.exe


BackdoorTrojan
Spyware
Payload
Plurality
Obfuscation
zues
zbot
'Confidentiality_Violation'
win32.Trojan
sysproc64
NTOS.exe
twain_32
wsnpoem
twext.exe
sdra64.exe
NTOS.exe
URI: http://www.semanticweb.org/elviskhovor/ontologies/2(
TOS.exe
Object property assertions:
NTOS.exe 'creates_folder' wsnpoem
NTOS.exe drops video.dll
NTOS.exe 'creates_folder' twain_32
NTOS.exe drops audio.dll
Data property assertions:
NTOS.exe path "C:\\WINDOWS\\System32\\NTOS.EXE"^^xs

Human-Like Decision Making

Cognitive Decision Making in Incident Investigations

Making Complex Decisions

For automation in incident investigations to be as effective as human incident responders it to be able to:

- Make decision on how investigative an alert. Which path/ technique from the investigative corpus is most suitable for each alert and the reasons for taking that path
- Pivot and provide alternative investigative techniques if the codified path/technique for investigating an alert is not sufficient

Learning & Continuous Improvement

To keep automated incident investigations current and updated it needs to:

- Have the ability to capture human decision making and model it into new automated investigative path/techniques to be used in automated decision making.
- Learn new ways to investigate threats that are not currently captured in the incident investigation corpus and existing ontologies

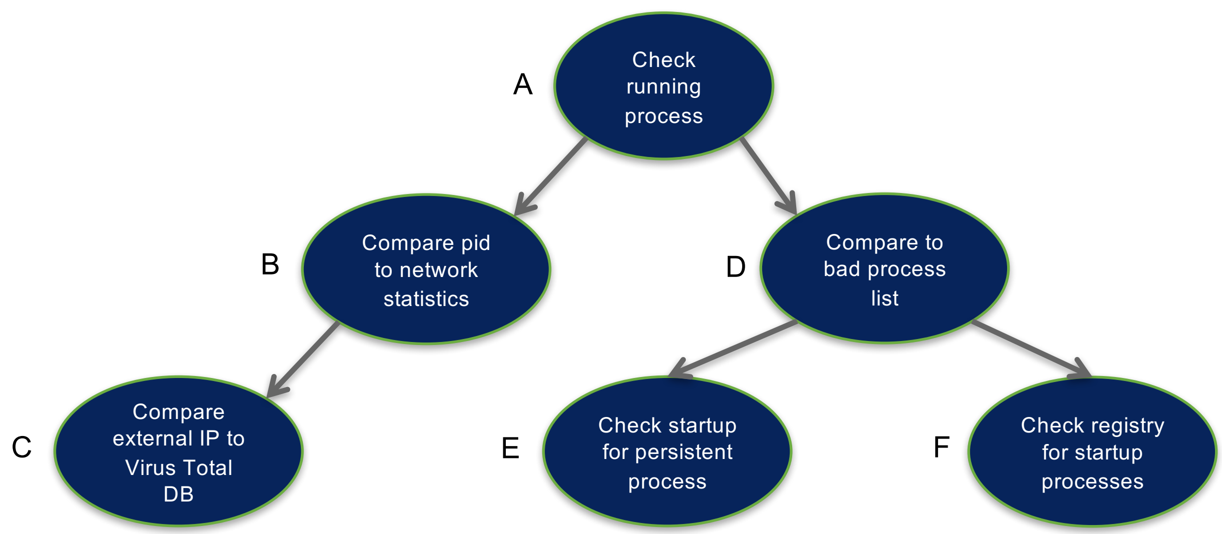
Cognitive Modeling – Core and Alternate Path Selection

Initial/core path for an Incident Investigation

- We use the **Binary Tree Traversal algorithm (BTT)** – Preorder traversal to determine the most efficient path for each type of investigation. This decision making technique is used every time a new manual investigative path is added to the list of paths in the incident investigation engine.
- Currently we are manually scoring the relevance of a path
- With BTT we plot out all possible paths of an incident investigation, computer generated workers go through every path possible while relaying the efficiency and quality of the path. This allows future iterations to go down the more efficient paths.

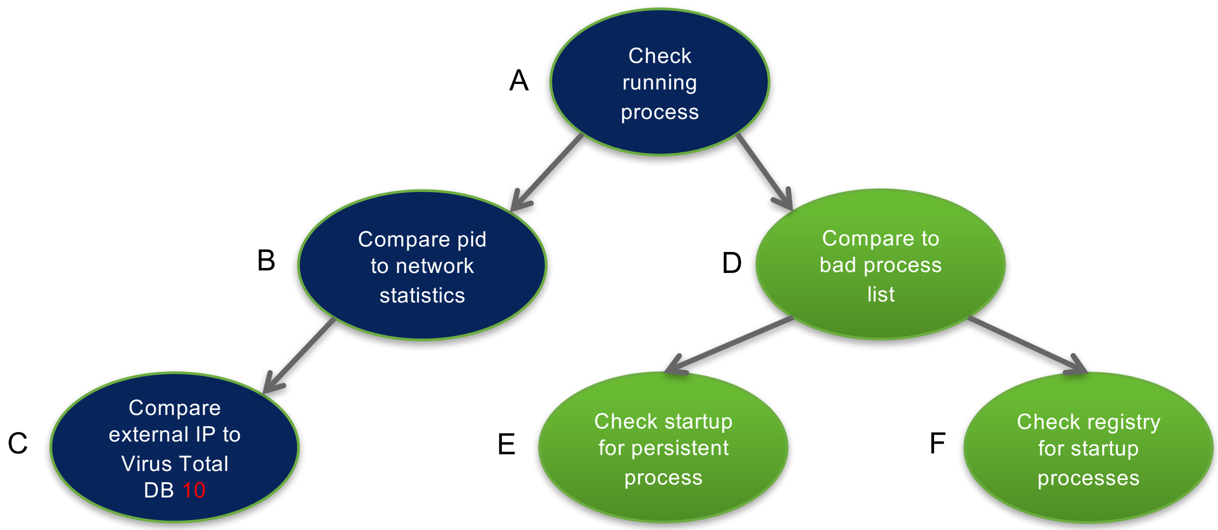
Cognitive Modeling – Binary Tree Traversal

Sample paths checking for malicious process



Cognitive Modeling – Binary Tree Traversal

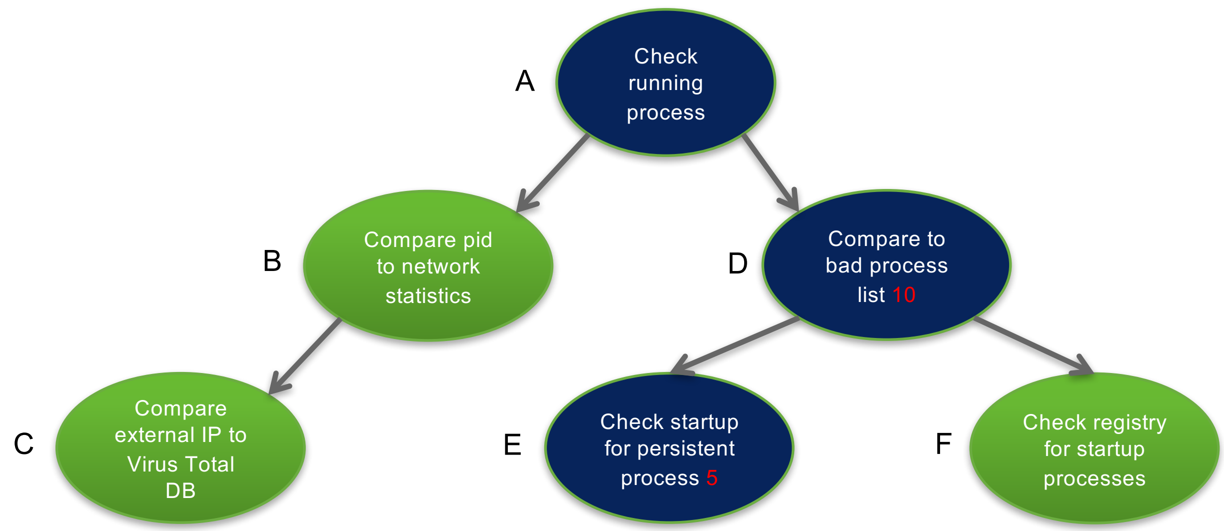
Binary tree goes down the left most path for each node and takes the score of that path



Path 1: A, B, C = 10

Cognitive Modeling – Binary Tree Traversal

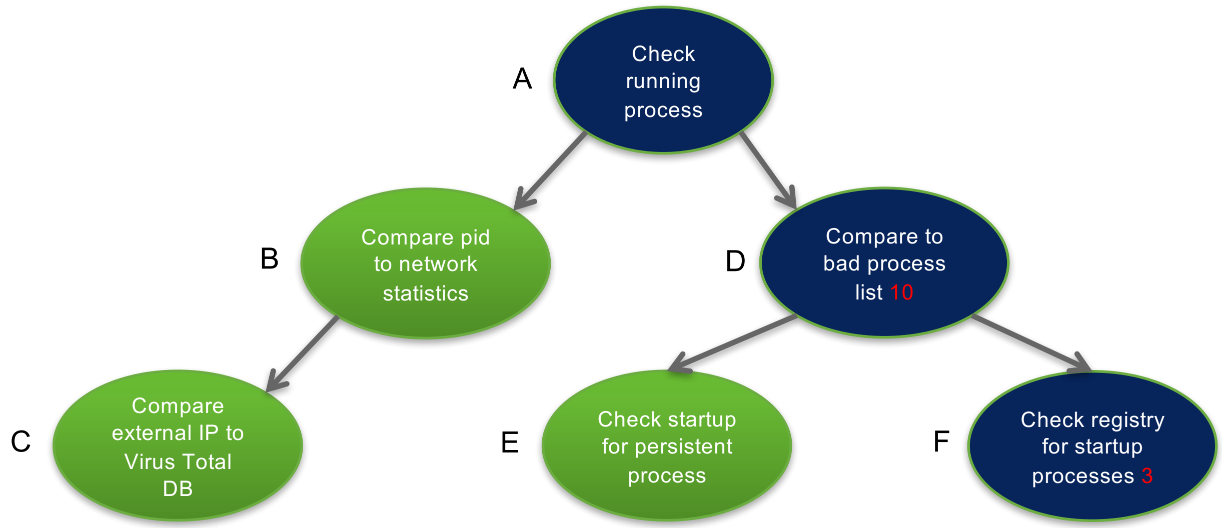
When going left is no longer an option the tree starts traversing down the right path, takes the score of that path



Path 2: A, D, E = 15

Cognitive Modeling – Binary Tree Traversal

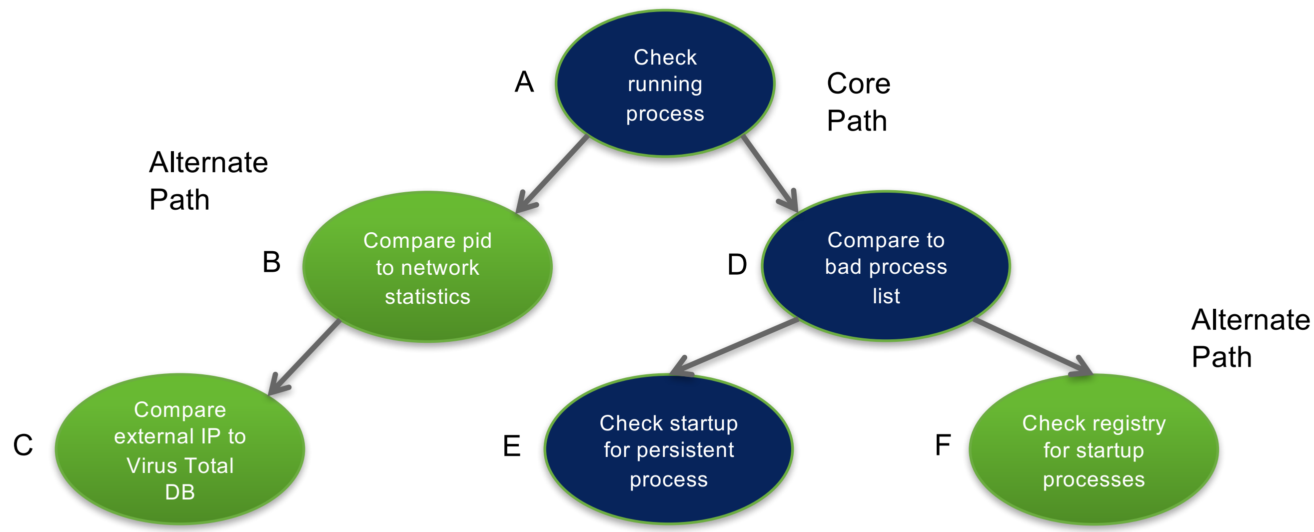
Binary tree goes down the final path and obtains the score of that path



Path 3: A, D, F = 13

Cognitive Modeling – Binary Tree Traversal

Path A, D, E becomes the core path since it retrieved the best score, the rest become alternate paths



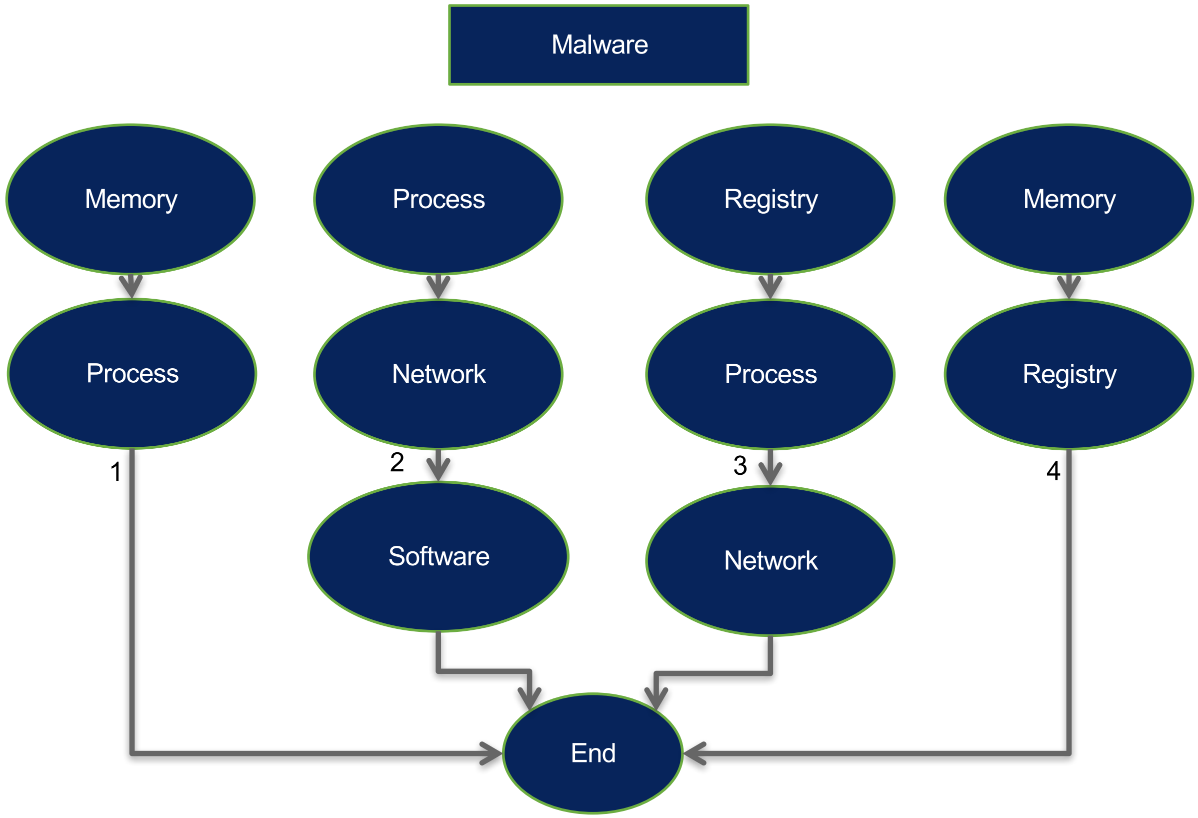
Machine Learning for Continuous Workflow Improvement

Modelling Manual Investigation & Updating Workflows

- This process records manual incident investigations by analyst & replicates in automated workflows that can be used to update workflows in the automated investigation engine to provide new investigative paths and its associated scores.
- The algorithm that helps achieve this is a modified version of **Critical Path Analysis**
- After each step of a manual investigation by an incident responder is recorded we see which paths have been taken the most by these incident responders. That path is added as an additional branch to the workflow to later be analyzed by our binary tree traversal model. The other paths are stored for later, and used when an investigation wasn't resolved by the specified paths then we go back to these alternate paths for further investigation.

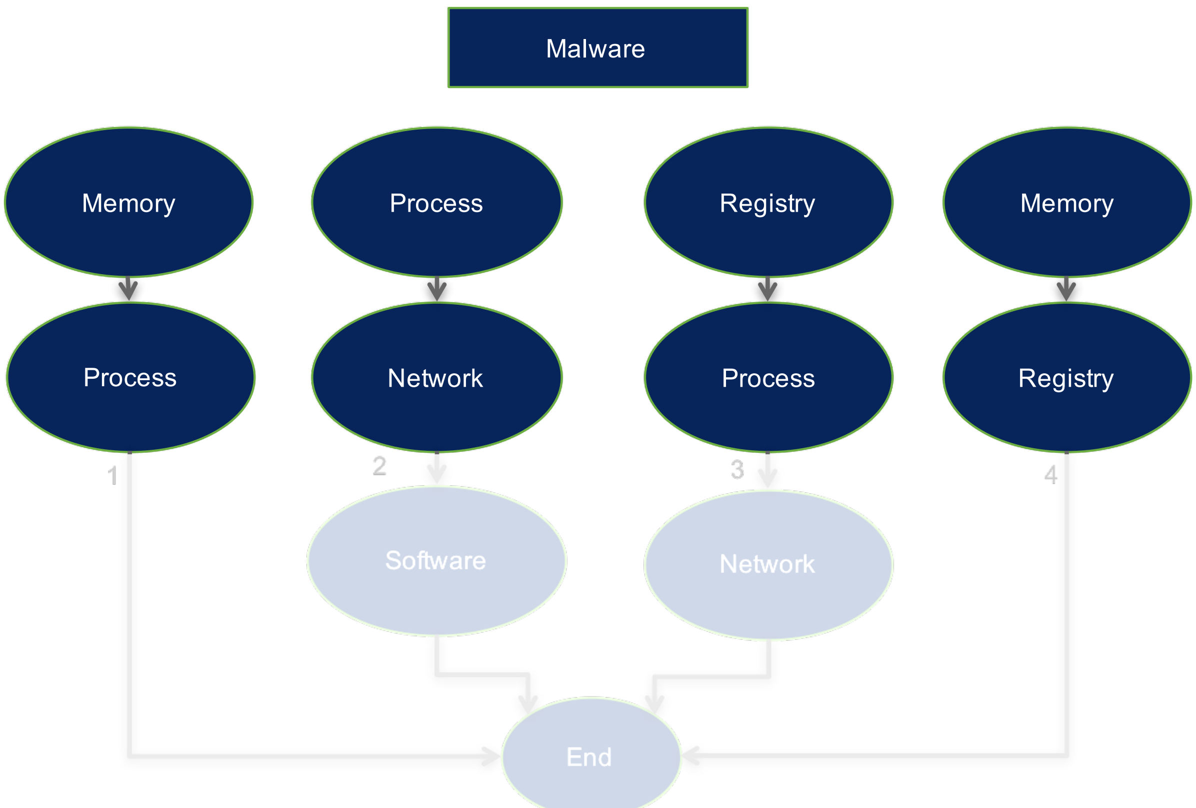
Cognitive Modeling – Modified Critical Path Analysis

Manual Scenario: Diagram shows an example of an incident investigation for the same event by multiple incident responders. Each path recorded shows the investigative path that the incident responder took to arrive at their conclusions (end).



Cognitive Modeling – Critical Path Analysis

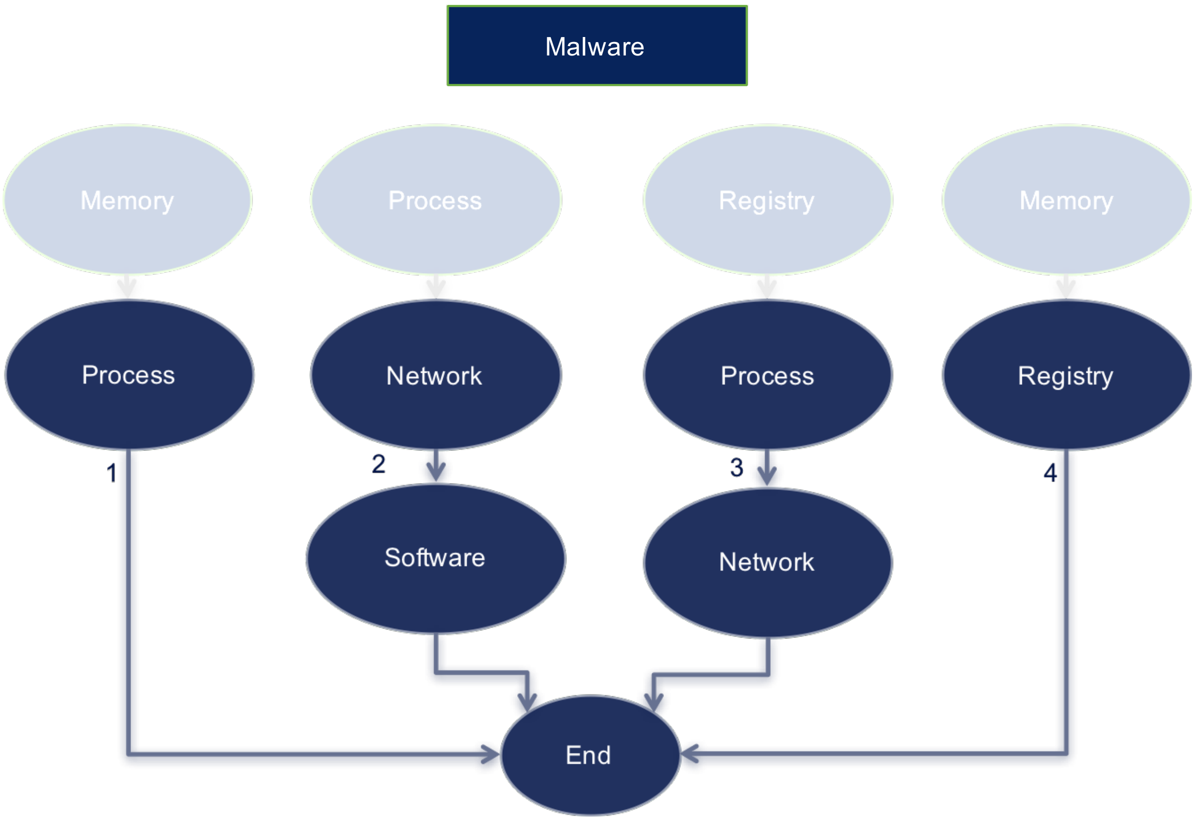
Part 2: Check the first node of every path and it's following node for similarities



First Node	Second Node	Count
Memory	Process	1
Process	Network	1
Registry	Process	1
Memory	Registry	1

Cognitive Modeling – Critical Path Analysis

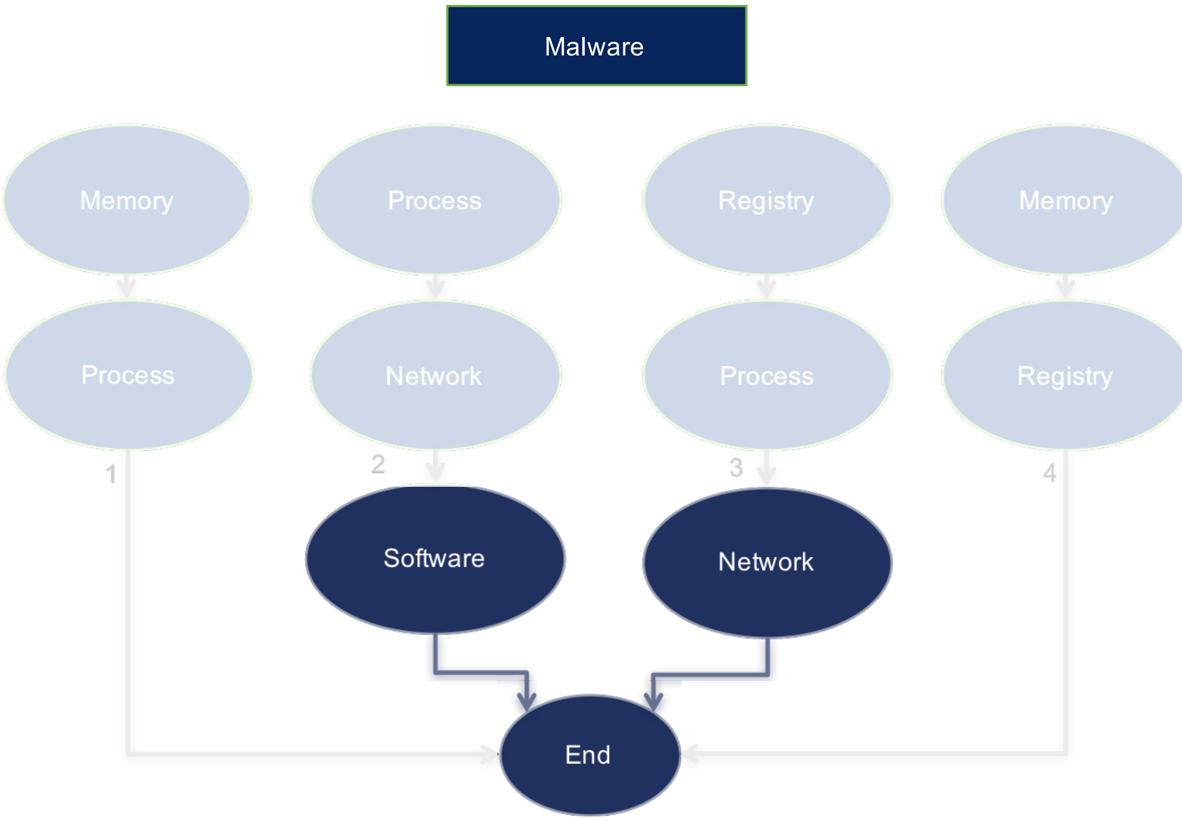
Part 3: Check the second node of every path and it's following node for similarities



First Node	Second Node	Count
Memory	Process	1
Process	Network	2
Registry	Process	1
Memory	Registry	1
Network	Software	1
Registry	End	1

Cognitive Modeling – Critical Path Analysis

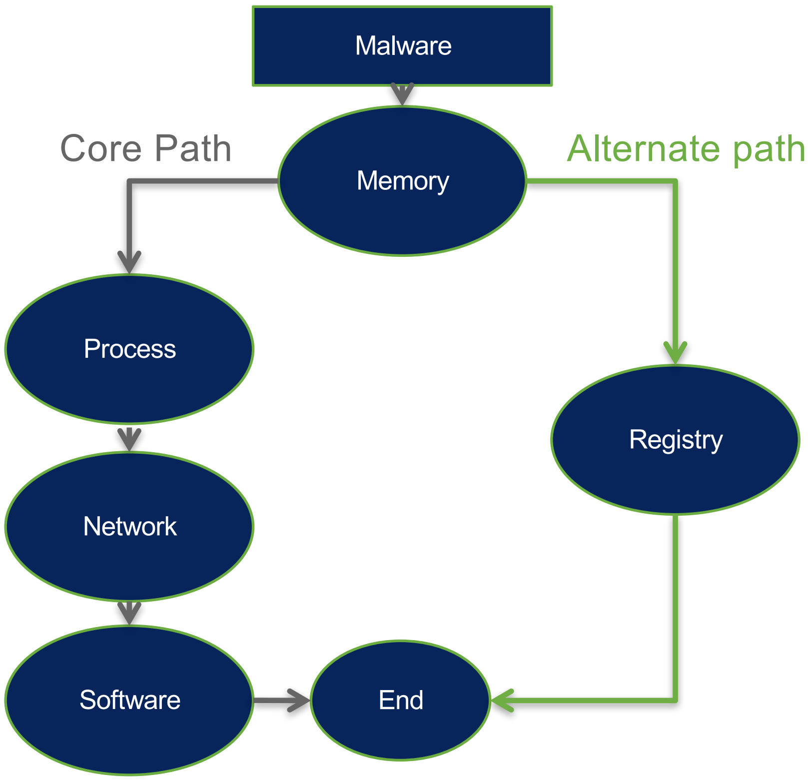
Part 4: Check the third node of every path and it's following node for similarities



First Node	Second Node	Count
Memory	Process	1
Process	Network	2
Registry	Process	1
Memory	Registry	1
Network	Software	1
Registry	End	1
Software	End	1
Network	End	1

Cognitive Modeling – Critical Path Analysis

Part 5: Create new path based on previous nodes investigations



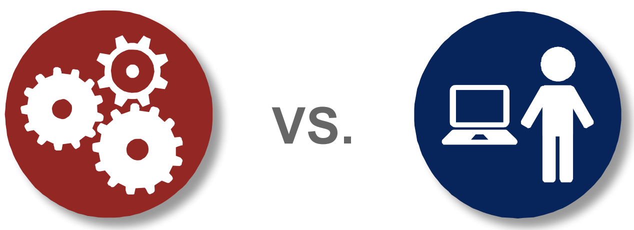
First Node	Second Node	Count
Memory	Process	1
Process	Network	2
Registry	Process	1
Memory	Registry	1
Network	Software	1
Registry	End	1
Software	End	1
Network	End	1

Sample Investigation Engine

Future Work & Improvements

Benchmarking

To compare the the speed and accuracy of the automated system to the speed with which a human will perform the same investigation.

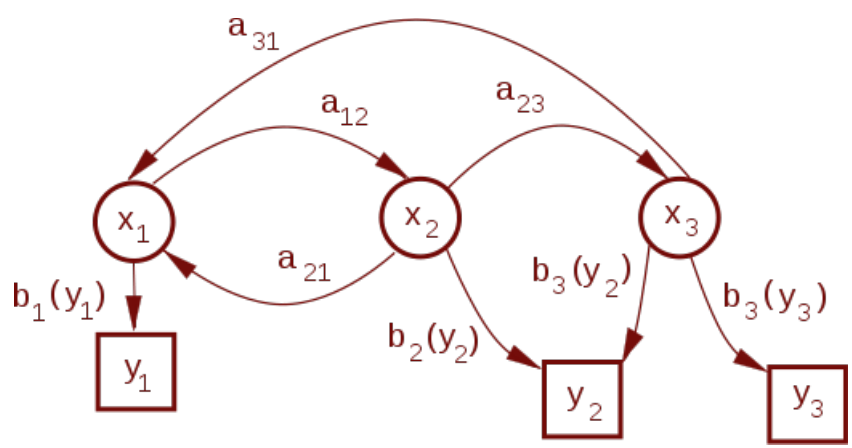


Improvements

- Incorporate the BBT algorithm for picking core and alternate paths as part of one automated system
- Build investigative bots directly from SPARQL queries from ontologies and automatically pick core and alternate paths without any scoring

Improved Cognitive Learning Techniques

– Markov Chain
  - With more investigative data we can start using the Markov chain to improve the advanced decision making model



Automating Incident Investigations

Blockers

Comprehensive Ontology

- Ontologies available currently address piecemeal incident investigations concepts and not the incident investigation life cycle as a whole.

Investigative Corpus

- Most incident responders are not willing to freely submit their knowledge on incident investigations into a global database.

Automated Investigations

- We do not have enough data to build full cognitive models to drive automated investigations. It is also very reliant on organizations being able to connect to all systems and query data on demand

Existing Work

ICAS

- The DARPA funded Integrated Cyber Analysis System (ICAS) ontologies had a primary focus of incident response.

Source: https://github.com/invincealabs/icas-ontology/tree/master/ontology
http://stids.c4i.gmu.edu/papers/STIDS_2015_T06_BenSalem_Wacek.pdf

CMU Insider Threat Ontology

- Funded by DARPA and FBI and developed by CMU. Has an insider threat focus.

Source: http://resources.sei.cmu.edu/asset_files/TechnicalReport/2016_005_112_465537.owl

Unified Cyber Ontology

- Attempt at combining knowledge schemas from different cybersecurity systems and most commonly used cybersecurity standards

Source: https://github.com/Ebiquity/Unified-Cybersecurity-Ontology
http://ebiquity.umbc.edu/_file_directory_/papers/781.pdf

Cyber Security Ontologies (user: Spriley) http://cps-vo.org/node/26324

Black Hat Sound Bites (3)

High performance. Delivered.

Q&A

Elvis Hovor
@KofiBaron

Mohamed El-Sharkawi
@HackerShark