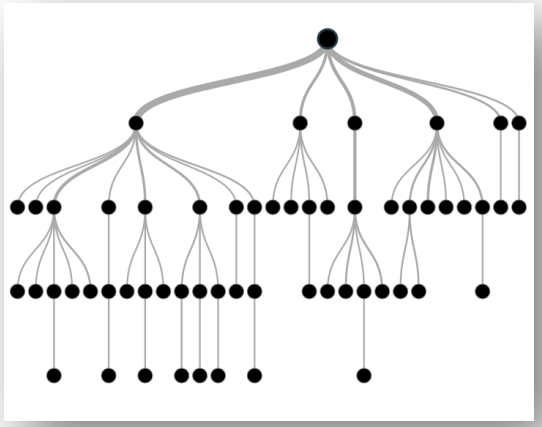
ACCELERATING ANALYSIS WITH DECISION TREES

SEPTEMBER 12, 2017

Agenda



- **Introduction**
- Original Research
- Limitations
- Where are we now?
- Case Studies
- Future

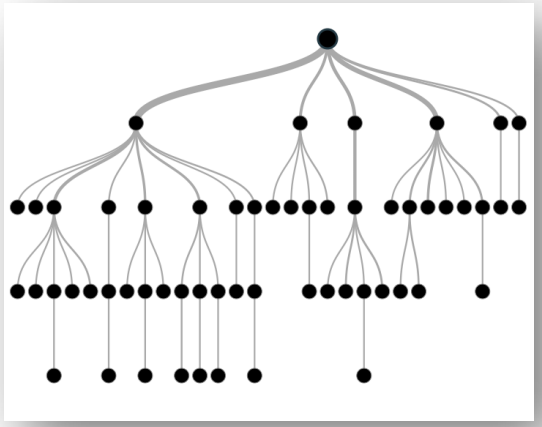
Introduction

- NIC (www.egov.com)
- Provide services to state and local governments in 29 states and a few federal contracts
- Me (www.linkedin.com/in/rodney-caudle-a48700)

Introduction

- Me (www.linkedin.com/in/rodney-caudle-a48700)
- BS in Electrical Engineering
- MS in Information Security Engineering
- Fancy myself as a bits-and-bytes guy who ends up speaking to management a lot

Agenda



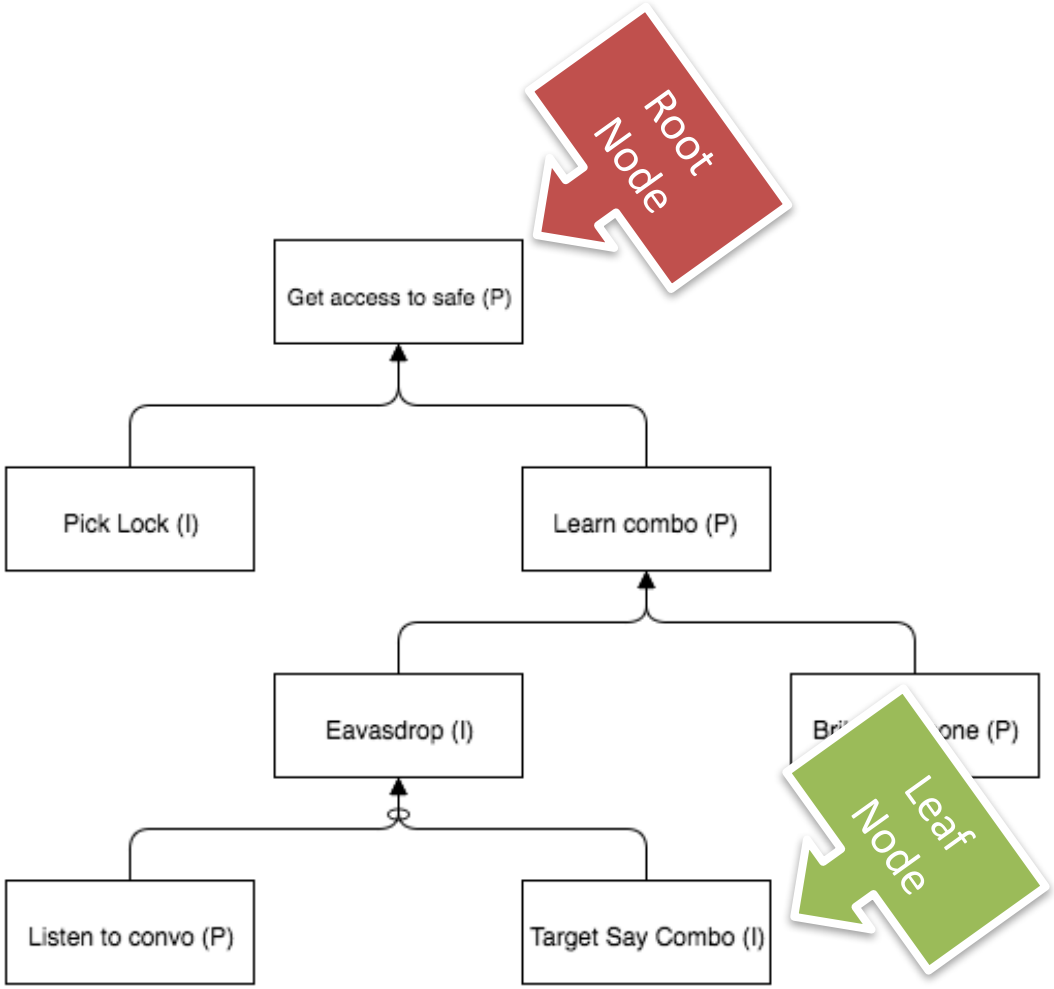
- ~~Introduction~~
- **Original Research**
- Limitations
- Where are we now?
- Case Studies
- Future

Original Research

- SANS Gold Paper in 2009
- https://www.sans.org/reading-room/whitepapers/incident/investigative-tree-models-33183
- Structured approach provides for definable, reproducible structures to be created facilitating controlled cost exposure during an incident response cycle.

Attack Trees

- Bruce Schneier (2000)
- A methodical way of describing threats against, and countermeasures protecting, a system



Multi-Parameter Attack Trees

- Ahto Buldas, Peter Laud, Jaan Priisalu, Mart Saarepera, and Jan Willemson
- CRITIS ’06
- Expands on attack trees to determine whether an IT infrastructure is protected (a) sufficiently, (b) reasonable, or (c) not protected adequately.
- ”Rational Attack” to achieve “Optimal Outcome”

Optimal Security Hardening

- "Optimal Security Hardening Using Multi-Objective Optimization on Attack Tree Models of Network" - 2007
- Rinku Dewri, Nayot Poolsappasit, Indrajit Ray and Darrel Whitley
- How to select the optimal subset of security hardening measures while minimizing residual damage

http://sarazervos.com/wp-content/uploads/2016/01/plug-dam.png

Defining "Attack"

Let S be a set of attributes of a system.

We define Att to be a mapping Att: SxS → {true,false} and Att(S_c,S_p) = value of S_p

a = Att(S_c,S_p) is an attack if $S_c \neq S_p$ ^ a

EXAMPLE → whiteboard

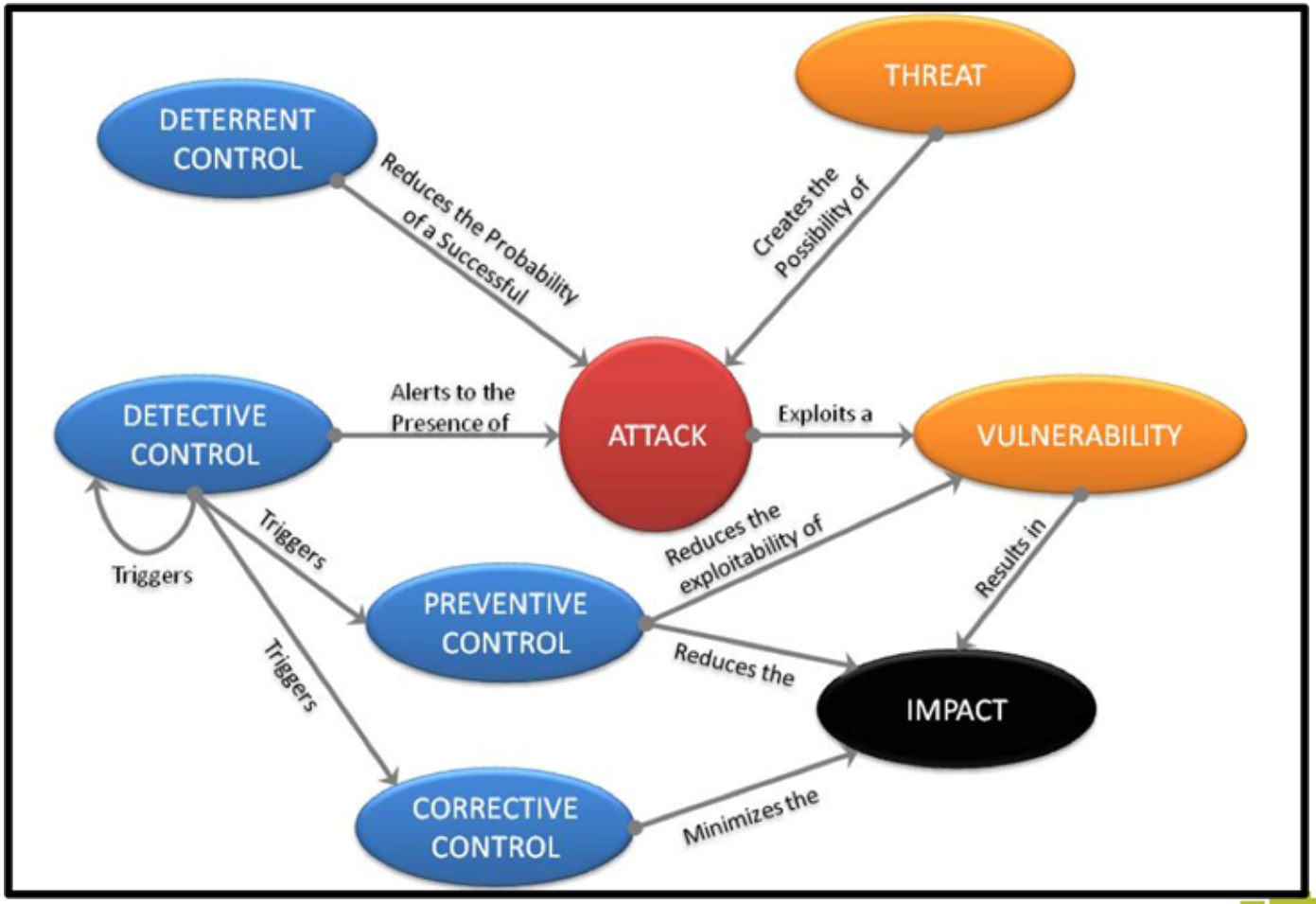
Protection Trees

- Given an attack tree shows the weaknesses in a system, a protection tree shows a methodical means of mitigating these weaknesses
- Kenneth Edge, Richard Raines, Michael Grimaila and Rusty Baldwin (2007)

Defining "Security Control"

- Given an attack tree, a security control is effective in defending against an attack "a" if the post condition (S_p) in the presence of the security control (sc) is the same as the pre-condition (S_c).
- But every control comes at a cost …

Security Controls



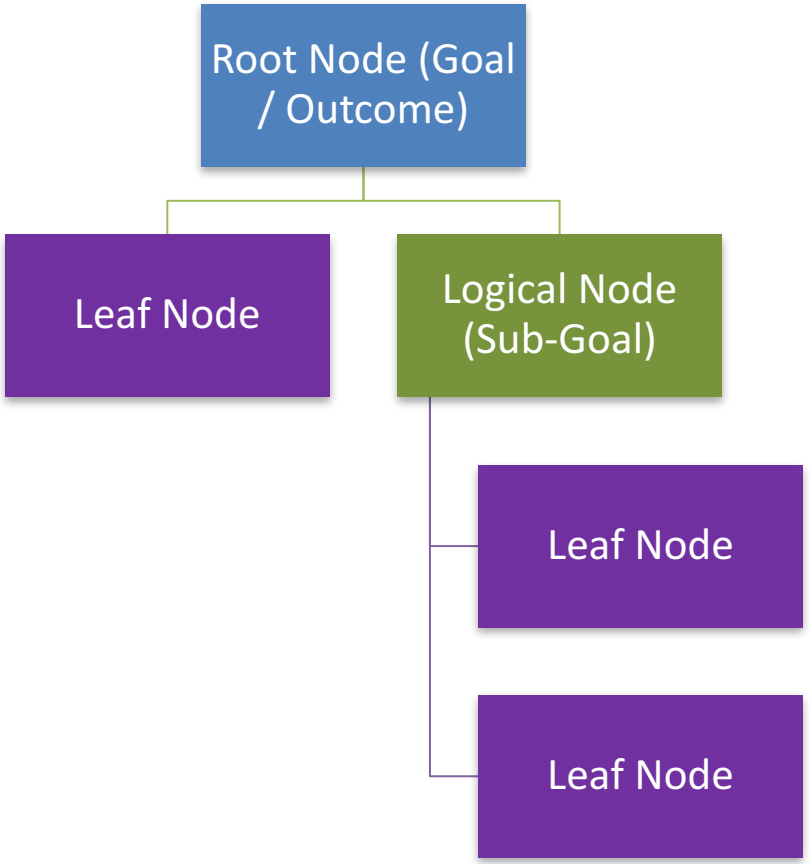
Attacks leave residue

- Since an attack can only be considered an attack if the value of an attribute changed as a result of the attack
- You can derive that an attack leave residual evidence that can be collected
- That's the focus of the investigation tree used to accelerate analysis

Investigative Trees

- Tree model focused on the response to an attack
- Starts with a root node or objective
- Builds layers of the tree until you reach atomic leaf nodes (doesn't make sense to be more specific)
- Focuses on the questions an investigation needs to answer

Components of the Tree Model



Email Investigation

- **Outcome / Goal**
  Can we verify email was sent from CEO's mailbox? (Both must be true)
- **Logical Nodes:**
  - Can we verify the vector for delivering the e-mail?
  - Can we verify the access to the CEO's account?

Vectors for Email Delivery

- Exchange (5 ways for clients to connect)
  - Outlook Anywhere
  - ActiveSync (mobile devices)
  - POP3/IMAP4
  - Exchange Web Services
  - Mapi/RPC
- OWA / web-mail
- SMTP gateway

Expanding the Tree

- Can we verify the vector for delivering the e-mail? (any of these will work)
  - Were emails sent via MAPI/RPC?
  - Were emails sent via ActiveSync?
  - Were emails sent via SMTP?
  - Were emails sent via OutlookAnywhere?
  - Were emails sent via EWS?

One More Level

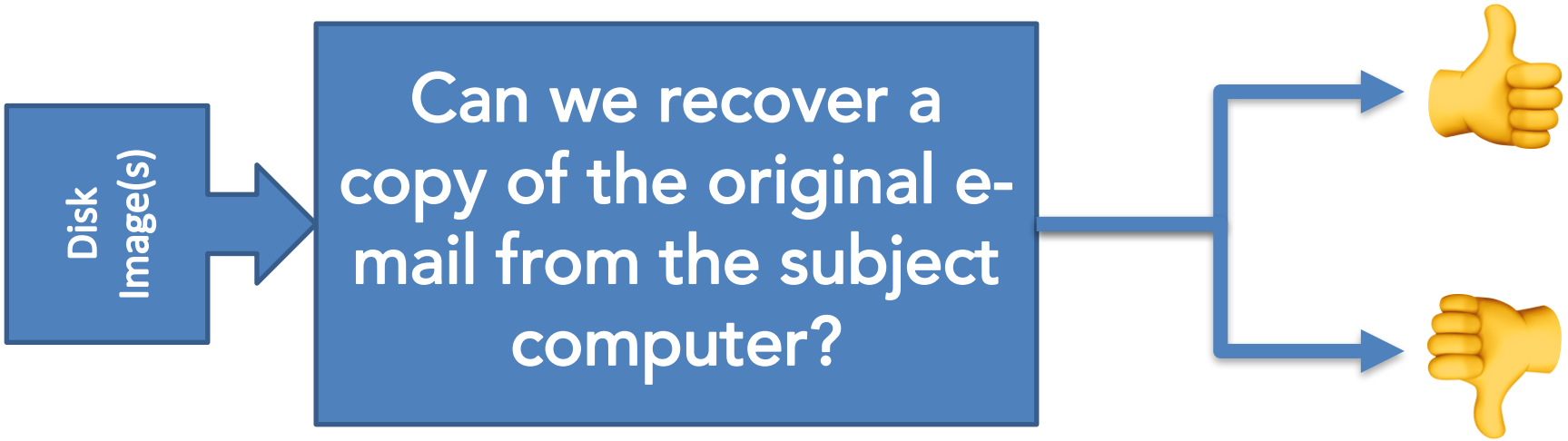
- Were the emails sent via SMTP?
  - Can we recover the presence of SMTP headers in the original email?
  - OR
  - Can we verify the presence of the emails in the log events from the MTA (SMTP Gateway) server?

And Again… Now we are good

- Can we verify the presence of SMTP he[illegible] in the original email?
  - Can we recover a copy of the original e-mail from the subject computer?
  - Does the email contain SMTP headers 821)?

Atomic Operation View

- To break the question down further would not show value over answering the question



Switch to procedural steps

- Can we recover a copy of the original e-mail from the subject computer?
  - Acquire a forensic image of the subject computer
  - Extract PST/OST files
  - If OST, convert to PST
  - Process PST to recover deleted items from the file structure
  - Extract original emails into EML format
  - Hash and store original emails securely as new evidence
  - Return “True” else return “False”

Node Parameters (Core)

- Confidence – un-impacted confidence level in the Boolean value represented in this node
- $Confidence_i$ – impacted confidence level (something outside of this node caused the confidence level to change)
- Impacted – Boolean value
- Weight – Weighted value of node compared to peers
- Category – Category of action

Impacted Confidence

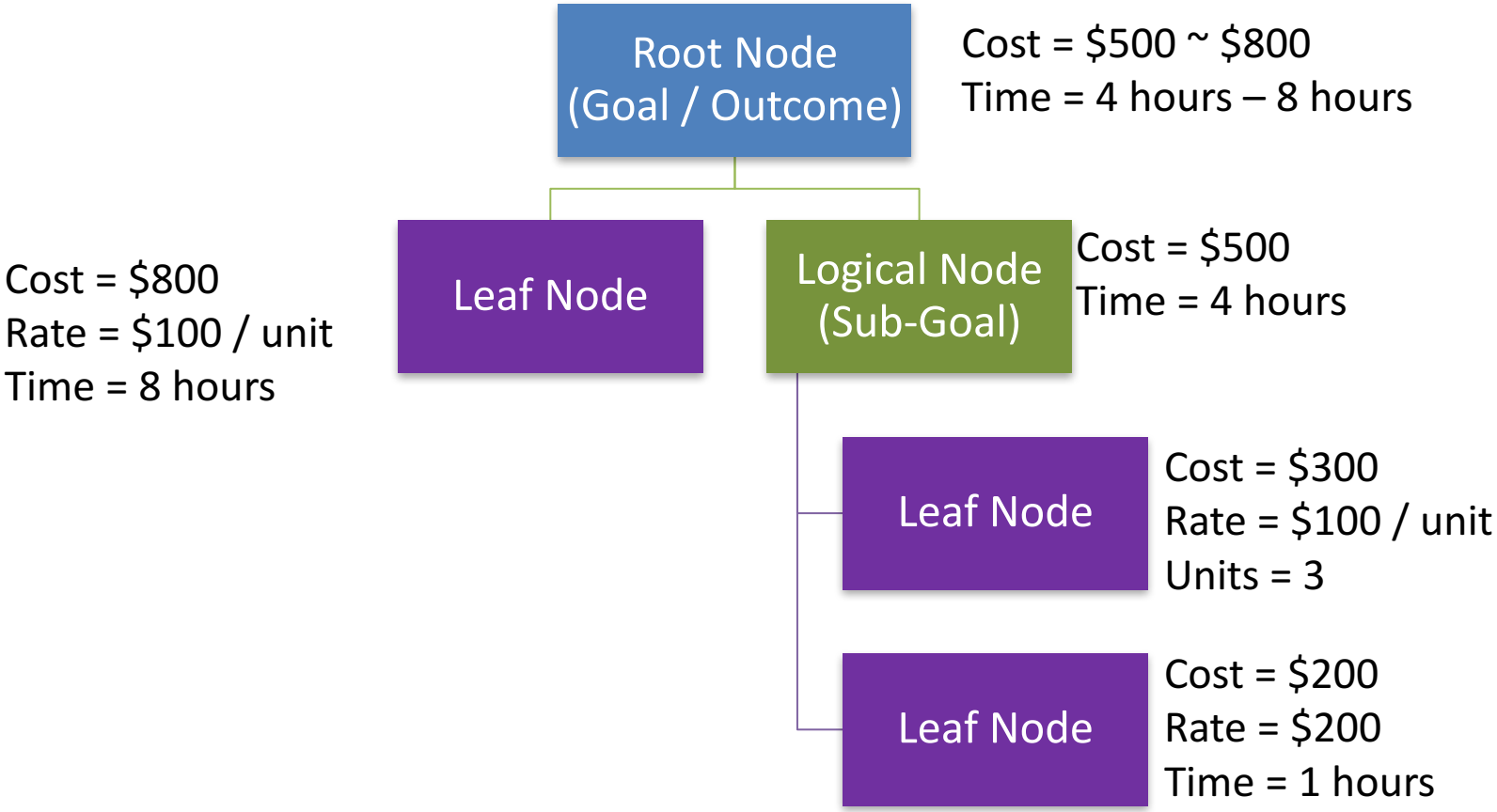
- Some artifact that turned up in another branch of the investigation impacted the confidence in the results of this node
- Example: Time stomping tools present on the system would impact confidence in logs residing on the subject system

Financial Parameters

- Cost – calculated or manually set
- Time – elapsed time to reach this node
- *Rate – amount to bill per unit
- *Units – number of units represented

* - only applies to leaf nodes (atomic operations)

Financial Parameters



Investigation Trees

- By adding parameters to a tree model and focusing that tree model on answering Boolean questions you can build structure to investigations that are repeatable and control costs by focusing on the outcome
- Show a financial perspective, confidence perspective, etc.

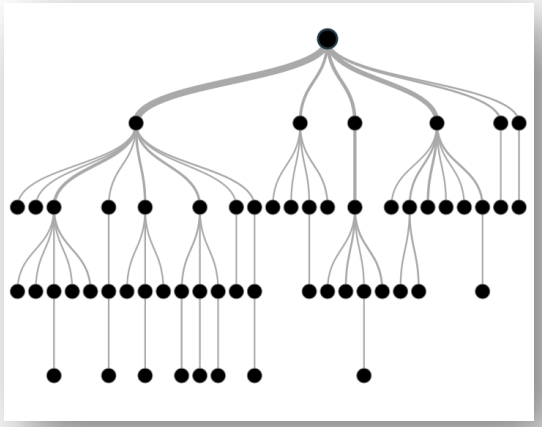
Historical Perspective

- Can handle re-opening of an investigation
- Add a data source to an investigation and suddenly new answers become available

Example

- Exploring an example
  - E-mail Investigation

Agenda



- ~~Introduction~~
- ~~Original Research~~
- **Limitations**
- Where are we now?
- Case Studies
- Future

Boolean != true

- Sometimes you need more than two options
- What if your test does not return conclusive results?
  - Discovery/Containment: Which systems did this IP address interact with?

Parallel Execution

- A lot of tasks can be executed in parallel if
  - Different data sources involved
  - Not dependent on each other

- **Need to factor this into the model**

Limited Set of Parameters

- Need to consider additional parameters tied to a perspective or layer
- Tie these perspectives to the calculations
- Visualize the “best route” to a conclusion based on a set of parameters

Dependencies should be more

- Dependency only shows relation or execution order
- Could be linked to confidence or another parameter
- Could be assigned a scope so only certain types of nodes are dependent

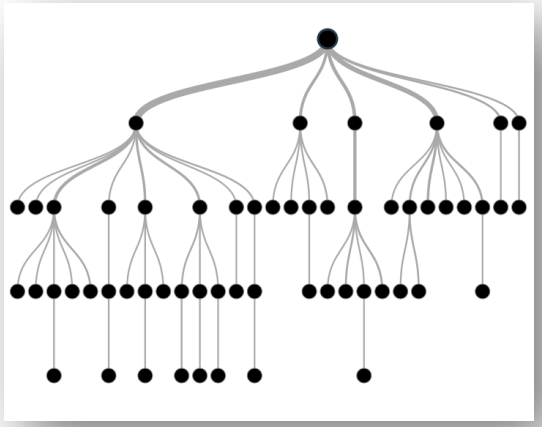
Automation

- No concept of workflow
- No concept of feedback or control loops

Confidence

- Confidence should be a calculation or equation based on the parameters of the nodes beneath

Agenda



- ~~Introduction~~
- ~~Original Research~~
- ~~Limitations~~
- **Where are we now?**
- Case Studies
- Future

Playbooks

- A series of steps to accomplish an objective
- Usually focused on technology specific solutions
- Essentially “leaf nodes” or atomic operations

Playbooks - Cisco

- http://blogs.cisco.com/security/using-a-playbook-model-to-organize-your-information-security-monitoring-strategy (November, 2013)
- Collection of plays
- "Plays are self-contained, fully documented prescriptive procedures for finding some sort of undesired activity"

Sections of a Play

- Report ID
- Report Type with Name
- Objective Statement
- Result Analysis
- Data Query / Code
- Analyst Comments / Notes

Objective Statement

- English statement of the "what and why" of a play.

Why

"Our HIPS logs contains suspicious network connections which allow for the detection of Bitcoin P2P activity on hosts. This report looks for the processes that appear to be participating in the Bitcoin network that don't obviously announce they are Bitcoin miners."

What

Result Analysis

- Explanation of the query/code section to better explain how the query results in an answer
- Expect this section to be very tied to the specific technology
- This can result in a large amount of variance from technology to technology

Open Source Playbooks

- https://www.isecom.org/Open-Source-Cybersecurity-Playbook.pdf

Security Automation and Orchestration

- Technology products that provide automated, coordinated, and policy-based action of security processes across multiple technologies, making security operations faster, less error-prone, and more efficient.

  - Forrester, April 2017

Playbooks - Challenges

- While this approach will get you answers, they may not be to the questions you need answered
- A lot of extra narrative that, at this level, may not be informative or related to the overall investigation
- Not everything can be reduced to a query or piece of code

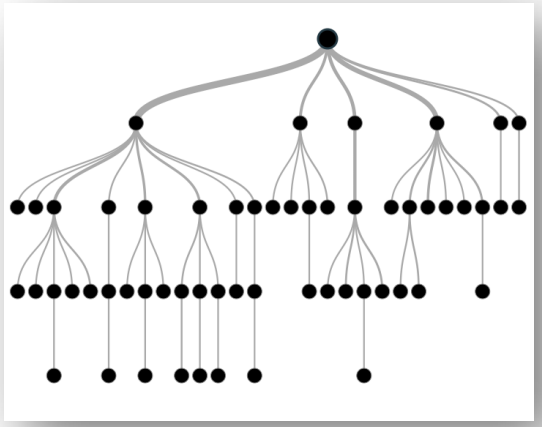
Playbooks = Answers

- By executing a playbook you are answering questions
- Starts with a single event, maybe from a SIEM
- Information is parsed out, enriched and maybe a decision is reached

Playbooks = more questions

- What additional questions need to be answered to complete the investigation?
- What additional questions can I answer now that I have completed this playbook?

Agenda



- ~~Introduction~~
- ~~Original Research~~
- ~~Limitations~~
- ~~Where are we now?~~
- **Case Studies**
- Future

Case Studies

- DDOS Attack Investigation
- Malicious IP Investigation
- Average Security Event Investigation
- Website Defacement Investigation

DDOS Attack Investigation

- **Objective:** Is this attack originated from spoofed IPs or real hosts?

- We feel there are indicators in the raw pcap data that provide a high level of confidence in an answer to this question.

- Demo - DDOS

Malicious IP Investigation

- Objective: Should we request a block for this IP address?

- Whack-a-mole

- Demo - BlockIP

Security Event Investigation

- Objective: Do we need to escalate this event for further action?

- We've built decision trees for each of our monitoring technologies for initial investigations.

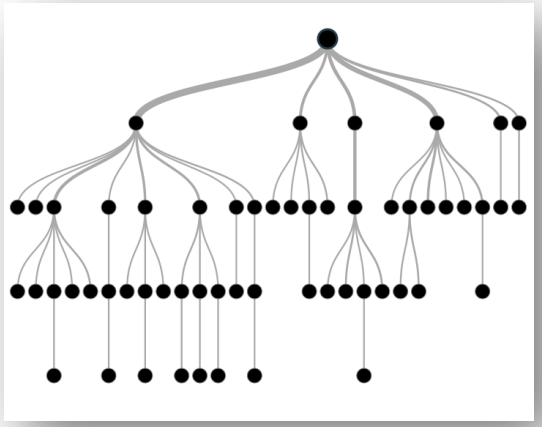
- Demo - Security Events

Website Defacement Investigation

- Objective: Has this website been defaced?

- Has content been changed?
- Was it changed by unauthorized access?
- Was it changed by authorized access?

Agenda



- ~~Introduction~~
- ~~Original Research~~
- ~~Limitations~~
- ~~Where are we now?~~
- ~~Case Studies~~
- Future

Open Source Tool

- Whiskered Cat - OSS
- Create an investigation tree
- Explore scenarios based on parameters (cost, confidence, etc.)
- Provide guidance for the forensic investigator through the phases of investigation
  - Preservation of evidence ➔ Reporting

QUESTIONS?