Quantitative Evaluation of the Target Selection of Havex ICS Malware Plugin

Julian Rrushi
Western Washington University
Department of Computer Science
Bellingham, WA
julian.rrushi@wwu.edu

Outline

- Research problem investigated
- Target selection features measured
- Decoy OPC tag deployment
- Trials
- Target selection measures
- Quantitative analysis
- Conclusions
- Questions

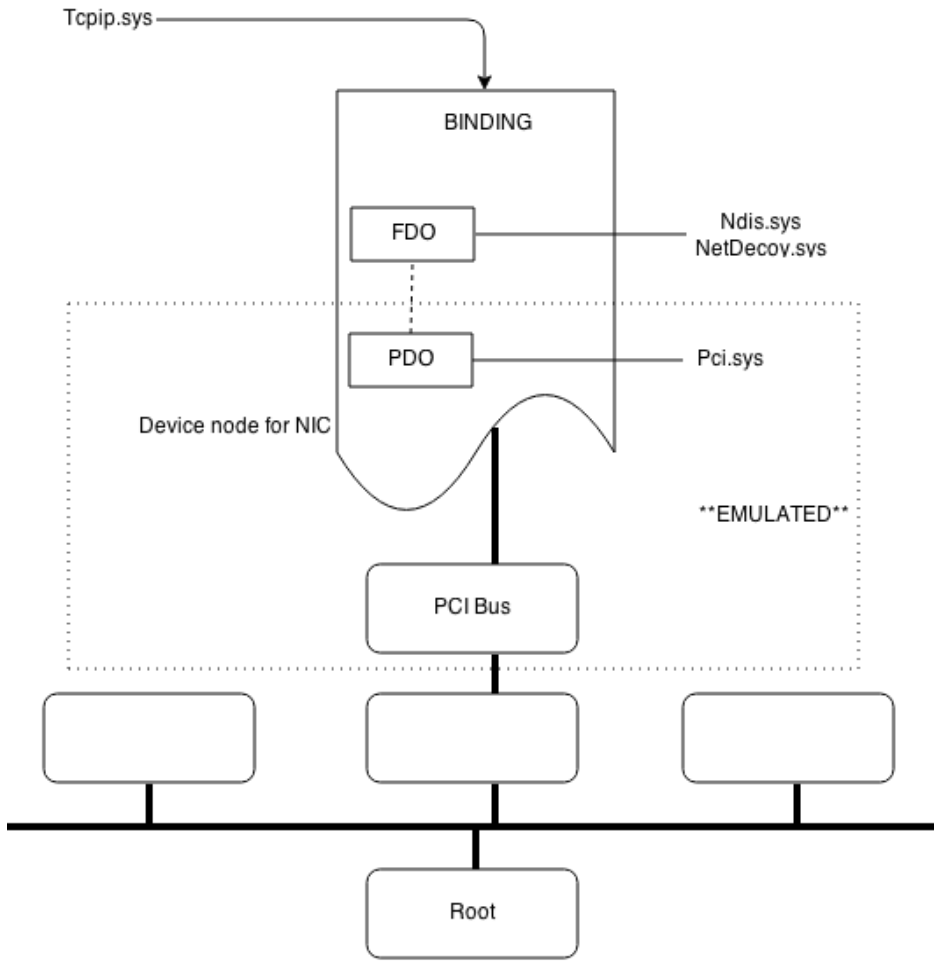
Target Selection Features Measured

- Ability to discover true servers over the network from the compromised machine
    - Ability to ignore or discard nonexistent or absent servers on the network
- Ability to determine whether or not a network server hosts COM objects and interfaces
- Ability to find true OPC server
    - …and dismiss COM objects that are not OPC

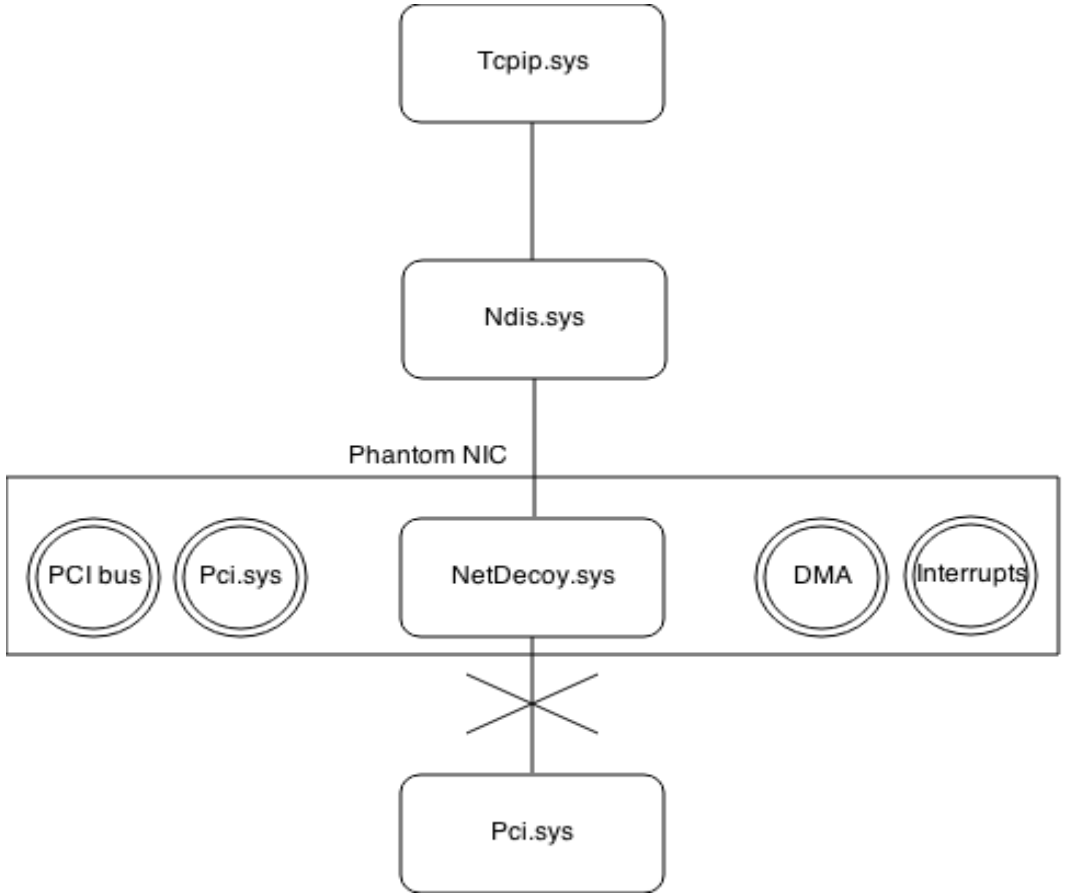
Other Important Feature

- Ability to differentiate between valid and invalid OPC tags
  - Honeytoken OPC tags
  - OPC tags that are no longer mapped to a location in the memory of a controller
- Not implemented due to safety reasons
  - Requires an IED configured to monitor and control the passage of electrical power from one circuit to another
  - OPC tags updated based on the IED scans
  - Those would be the target tags

Decoy OPC Tag Display



Deceptive Emulation



Trials

- Signal trials
  - Consist of true targets, i.e., server machines, COM objects, OPC server objects
  - Targets exposed to Havex
  - Empirically observed whether Havex recognizes those targets as valid
- Noise trials
  - Consist of fake or nonexistent targets
  - Fake targets exposed to Havex as well
  - Empirically observed whether Havex pursues those targets

Factors of Interest

- Response bias
    - A general tendency to deem a target to be valid or invalid, i.e., signal or noise, respectively
- Sensitivity
    - The degree of overlap between the valid-target and invalid-target probability distributions
    - Involves the internal reasons that cause Havex to pursue a target
- Both factors are affected by the hit rate and the false-alarm rate

Measures of Sensitivity (I)

- d’ measures the distance between the mean values of those probability distributions in standard deviation units
- d’ close to 0 indicates inability to distinguish between valid and invalid targets

$$d' = \Phi^{-1}(H) - \Phi^{-1}(F)$$

Measures of Sensitivity (II)

- A’ is a measure that ranges between 0.5 and 1.0
- 0.5 indicates inability to distinguish between valid and invalid targets
- 1.0 indicates full ability to distinguish valid targets from invalid targets

$$A' = \begin{cases} 0.5 + \frac{(H-F)(1+H-F)}{4H(1-F)}, & \text{if } H \geq F \\ 0.5 - \frac{(F-H)(1+F-H)}{4F(1-H)}, & \text{if } H < F \end{cases}$$

Measures of Response Bias

- β measure
- When β<1, there is bias towards accepting a target as being valid
- When β>1, there is bias towards discarding a target as invalid

$$\beta = \frac{e^{-0.5[\Phi^{-1}(H)]^2}}{\sqrt{2p}} \div \frac{e^{-0.5[\Phi^{-1}(F)]^2}}{\sqrt{2p}}$$

Server Trials

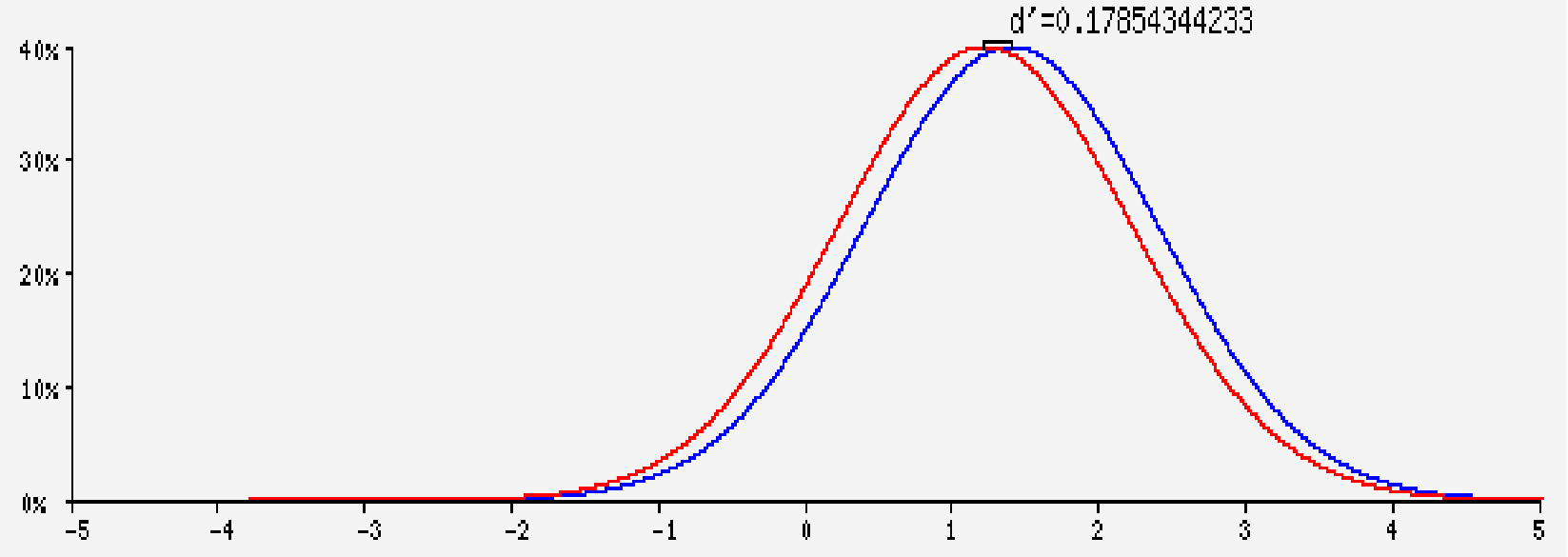
- Windows machine infected by Havex
- Signal trials
    - The machine had access to real servers over the network
    - Havex recognized most existing servers as valid targets
- Noise trials
    - No real servers, only server displays
    - Havex pursued most of the phantom servers as valid targets

Measurements

- d’=0.179, and thus close to 0
- A’=0.576, and thus close to 0.5

- β=0.791 and thus <1
    - Havex has the tendency to recognize as a valid server any software component that can respond to network queries

Probability Distributions



COM Object Trials

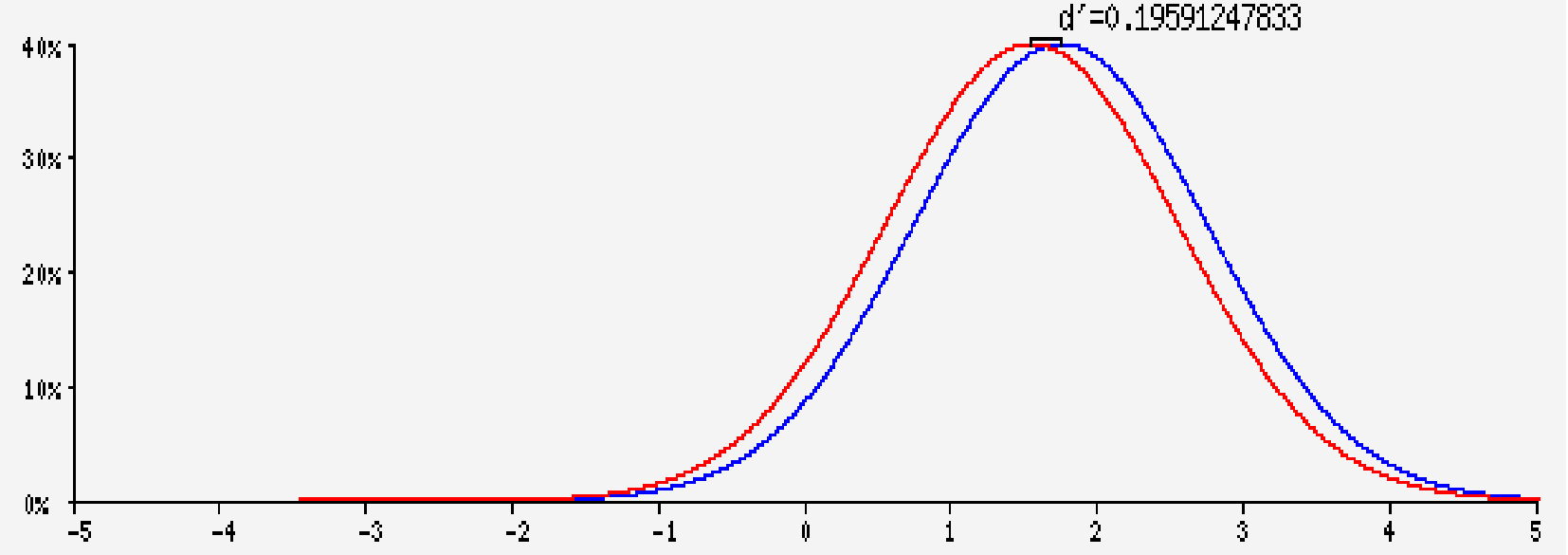
- A real server was reachable by Havex over the network
- Signal trials
    - The server hosted true COM objects and interfaces
    - Havex recognized most of the existing COM objects as valid targets
- Noise trials
    - The server generated a fake response when queried for COM objects and interfaces
    - No true COM objects and interfaces
    - Havex accepted most of those nonexistent objects as valid targets

Measurements

- d’=0.196, and thus relatively close to 0
- A’=0.589, and thus close to 0.5

- β=0.723 and thus <1
    - Havex is biased towards accepting as a valid target any server that claims to host COM objects and interfaces

Probability Distributions



OPC Server Object Trials

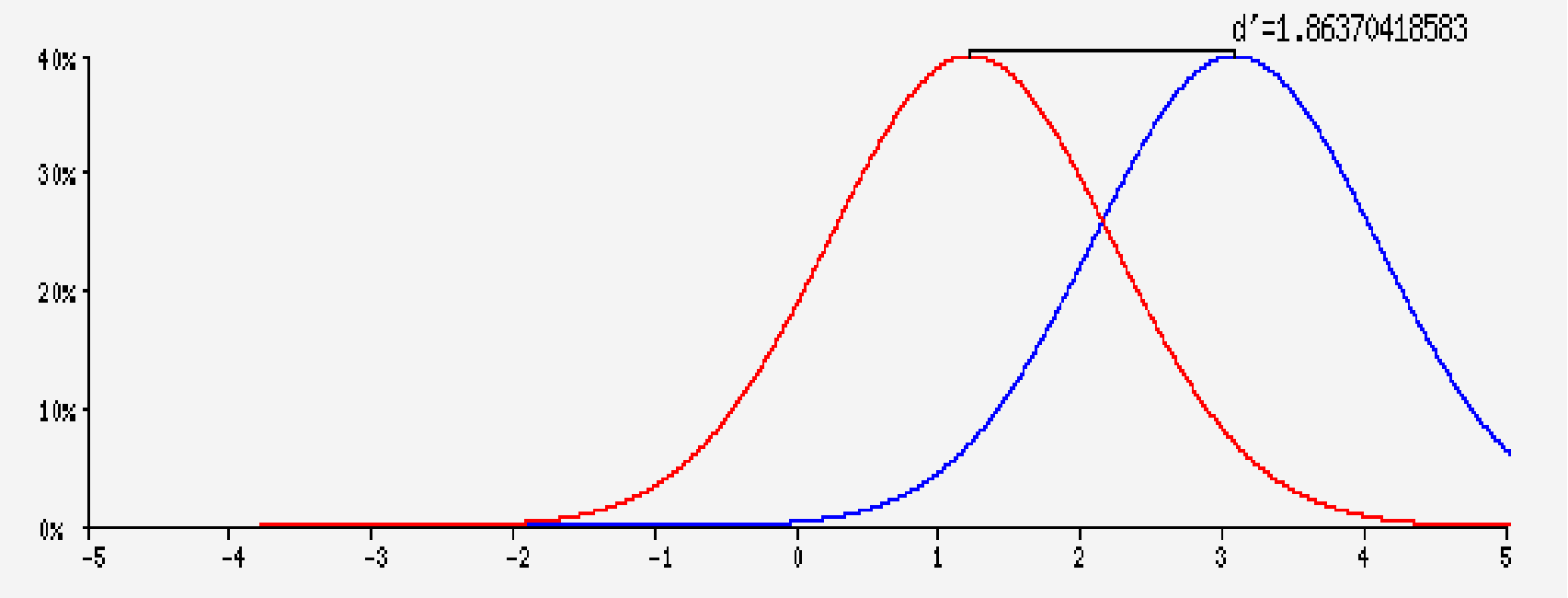
- A real server with support for COM was reachable by Havex over the network
- Signal trials
    - The server hosted true OPC server objects
    - Havex recognized most of the existing OPC server objects as valid targets
- Noise trials
    - The server returned lists of OPC server objects that did not exist
    - No true OPC server objects were involved
    - Havex accepted most of those nonexistent OPC server objects as valid targets

Measurements

- d’=0.1864, and thus relatively close to 0
- A’=0.775, and thus still relatively close to 0.5

- β=0.018 and thus <1
    - Havex is biased towards accepting any claim of OPC server object as valid

Probability Distributions



All questions and feedback are welcome!