University of Missouri

VIMAN Lab
VIRTUALIZATION, MULTIMEDIA AND NETWORKING LAB

# CICADA: Cloud-based Intelligent Classification and Active Defense Approach for IoT Security

**Prof. Prasad Calyam (Presenter), Roshan Neupane Kiran Neupane, Trevontae Haughton**
University of Missouri, Department of Electrical Engineering and Computer Science

**Trevor Zobrist, Shaynoah Bedford, Shreyas Prabhudev, Jianli Pan**
Southeast Missouri State Uni., Uni. Of the Virgin Islands, Uni. Of California at San Diego, George Mason Uni.

***10th Annual NSA CAE in Cybersecurity Community Symposium, 2023***

# Outline

- Problem Motivation
- Research Questions
- Solution and Novelty
- CICADA Active Defense Architecture
- Detection Engine
- Active Defense Engine
  - Cost Analysis
  - Risk Analysis
- Evaluation
- Conclusion
- Future Work

# Problem Motivation

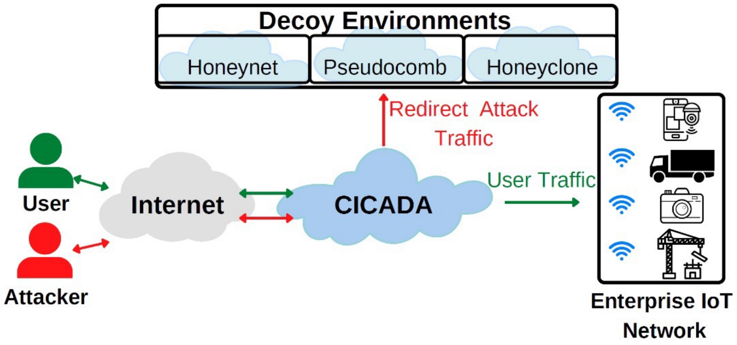
- Internet of Things (IoT) devices are becoming increasingly prevalent in every domain
- These devices capture and process personal, sensitive data such as camera feeds, health data, etc.
- IoT devices are targeted by attacks such as Distributed Denial-of-Service, Command-and-Control, and modern attacks, like Zero-Click, Ransomware

Challenge: To design a system that provides uniform protection to all IoT devices from the Cloud and mitigates risk within an enterprise IoT network using active defense

# Research Questions

How to detect modern and sophisticated threats targeting Enterprise IoT Networks?

How to deploy decoy environments effectively to ensure reduced attack risk and deployment cost in cloud settings?



Overview of an active defense system using decoy environments for securing an IoT-based enterprise network

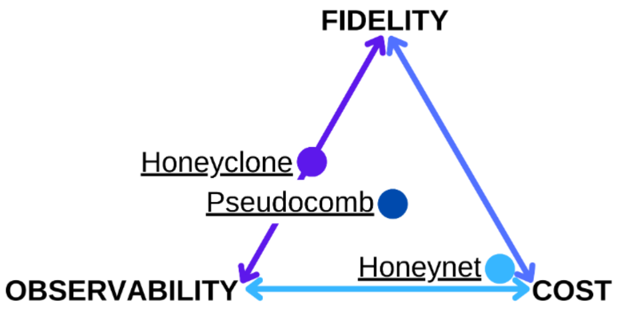
# Solution and Novelty

CFO Triad: Framework to categorize deception environments and analyze them based on tradeoffs, constraints, etc.; created three overarching environments that range according to CFO factors

**Honeynet:** large networks of low-cost and less resource-intensive honeypots that can capture weaker attacks

**Pseudocomb:** balanced architecture that provides fidelity and observability at a reasonable cost

**Honeyclone:** near replica system architecture with production level data flows and systems to deceive attackers into exposing their best techniques



Correlation of cost, observability, and fidelity leading to prescription of pertinent decoy environments in CICADA for different attacks.

# Solution Approach

Implementation of an Active Defense Architecture viz., CICADA for Enterprise IoT Networks

Detection Engine with Multi-model method:

- Neural Network Binary Classifier for traffic (benign and attacker)
- Ensemble model with 8-layer deeply connected NN for further classification of attack traffic

Active Defense Engine with decoy environments for variable observability of attacks

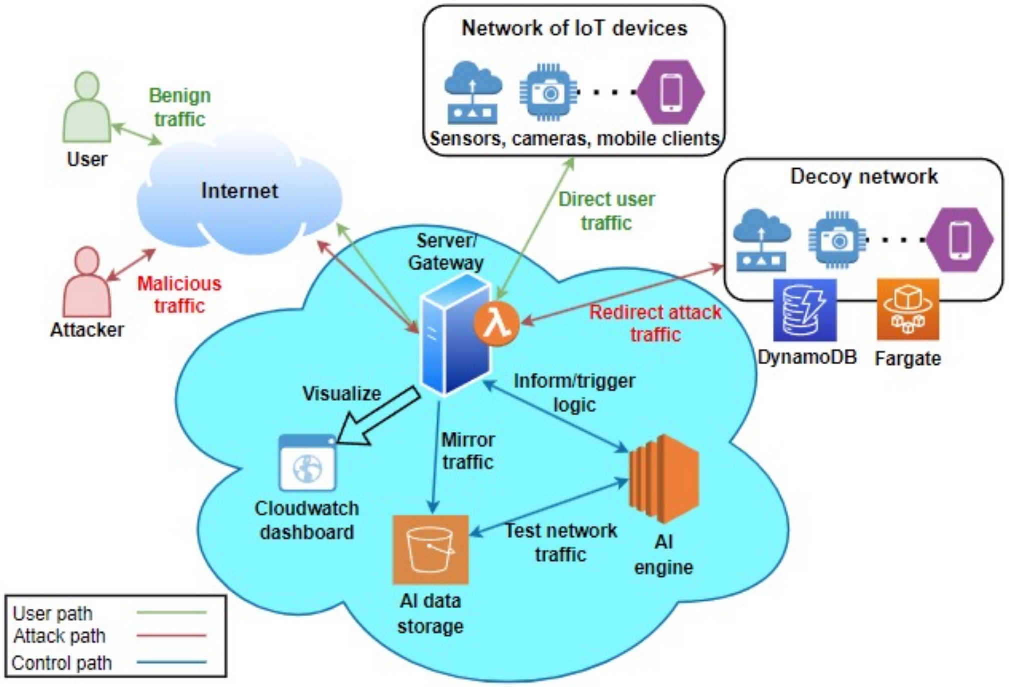
Risk analysis of multiple modern threats and cost analysis of deployment of decoy environments

# CICADA Active Defense Architecture

CICADA operates at the edge for offloading security from resource constrained smart devices

The architecture is equipped with a Detection Engine that uses classifiers and neural networks to detect modern threats

CICADA also has an Active Defense Engine that uses factors such as cost and risk to redirect attack traffic to varying decoy environments viz., Honeynet, Pseudocomb and Honeyclone



# Detection Model Details

**Binary Classifier:**

- Binary classifier classifies the traffic as 'benign' or 'malicious' using 8-layer (6 hidden layers) deeply connected neural network
- Model takes 8 input feature values through MinMax scalar function to feed them through layers with 32,64,24 and consecutive dropout layers
- Output layer has 1 node that outputs a value with sigmoid activation of 0 or 1
- Hyperparameter tuning is used to refine parameters such as epochs and batch sizes

# Detection Model Details

**Ensemble Model:**

- Binary classifier classifies the 'malicious' traffic to specific threats
- Consists of 8 (6 hidden) layer deeply connected NN that uses 16, 64, 24 hidden layers
- Output layer has 5 (consisting 5 classes) node
- When a 'malicious' flow proceeds through the model, the flow is sent to each algorithm in parallel and prediction is gathered
- If collective algorithm's predictions are unanimous or consensual, flow is labeled as concurred classification
- Non-consensual predictions are not labeled as specific threats

# Active Defense Engine

Active Defense Engine is designed with goal of intelligently routing network traffic based on analysis of Detection engine

Cost and Risk factors are considered for assigning different threats to different decoy environments

When a packet is classified as malicious, ADE prevents the flow from reaching the IoT based server

# Cost Analysis

Cost analysis considers the cost of deploying detection engine, analysis service, monitoring service, traffic mirroring, and storage

Total operational cost regardless of the decoy environment deployment is:

$$C_{operation} = C_d + C_a + C_m + C_t + C_s$$

Where,

$$C_d = N * s * c_i$$

$$C_a = q_n * d$$

$$C_m = m_n * \lambda_n + \lambda_r * \lambda_n + a_n * c_a$$

$$C_t = t_n * h * c_{sh}$$

# Cost Analysis (contd.)

Decoy environment costs vary for the different decoy environments Honeynet, Pseudocomb and Honeyclone

Cost of setting up decoy network is given by:

$$C_{decoy} = (p_n * a_d * v_{ch} * v_{cpu}) + (p_n * a_d * m_{ch} * m_{alloc})$$

Deployment of a honeyclone is given by:

$$H_{decoy} = \sum_{i=1}^{n} P^i_{decoy} + C_{load}$$

# Cost analysis (contd.)

As feature utilization of an environment increases, the cost and maintenance demand also increase

The cost incurred is analyzed based on the capabilities of the different decoy/deception environments

We perform this cost calculation based on the resources and services used in the Amazon Web Services cloud platform

| Capability | AWS Service | Cost ($) Per Month | | |
|---|---|---|---|---|
| | | Honeynet | Psuedocomb | Honeyclone |
| Decoy IoT Network | Fargate, EC2 | 72.09 | 1081.20 | 10812.00 |
| Load Generator | VPC, API Gateway | 109.50 | 1642.50 | 16425.00 |
| Network Behavior Analysis | Cloudwatch, VPC | N/A | 1925.55 | 19255.50 |
| Cross-service Load Generator | VPC, API Gateway | N/A | N/A | 8760.00 |
| Data Storage | DynamoDB | N/A | 53.93 | 539.3 |
| Honey Token | SpaceSiren | N/A | 5 | 50 |

# Risk Assessment

We used methodology in NIST Risk Assessment guideline to calculate the potential risk levels for various threats

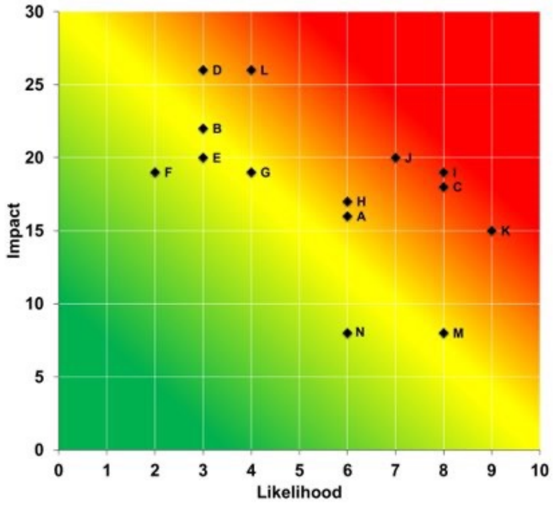
The assessment considers the impact of an attack on the Enterprise IoT network and its likelihood

$$\textit{Risk} = \textit{Impact} * \textit{Likelihood}$$

# Risk Assessment (contd.)

The heat map shows risk of different data actions

- Access control (A, B, C)
- Storage (D, E, F)
- Visualization (G)
- Transfer (H, I)
- Collection (J, K)
- Fulfillment (L, M, N)



Heat map of risk without any decoy environment for Enterprise IoT Network

Example threat events are

A - Modification of access role (escalation of privilege),

B - Update operation on database (data tampering),

D - Unauthorized users having access to the relational database to retrieve private data (information disclosure),

K - Overwhelming network with requests (denial of service),

L - Unlicensed users have access to critical data/system (spoofing identity).

# Evaluation

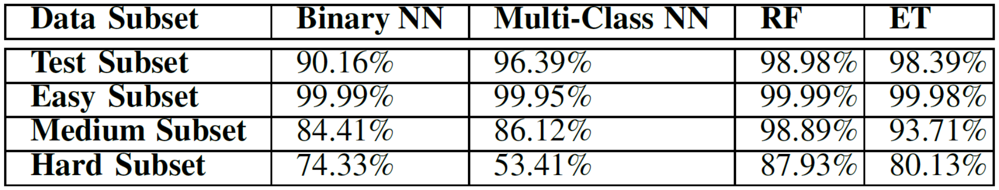
Evaluation of Detection Engine for varying data subset representing observability ranging from Easy, Medium to Hard to detect attacks (classifier accuracy)

Risk Assessment for various deception environments within CICADA-protected Enterprise IoT Network

Risk versus Cost of deploying Deception environments

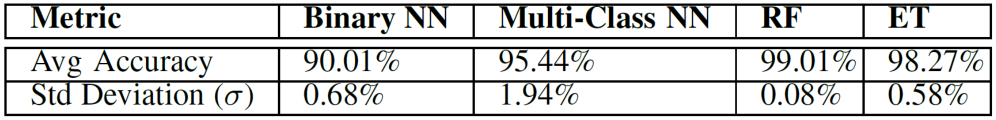
# Performance of Detection Engine

- Used IoT traffic datasets to evaluate the performance of the Detection Engine consisting of attack types DDoS, MiTM, Port Scanning, Malware, C&C
- Threats separated by observability characteristics
- Bluetooth traffic modified to imitate Bleeding Tooth Zero-Click attack



| Data Subset | Binary NN | Multi-Class NN | RF | ET |
|---|---|---|---|---|
| **Test Subset** | 90.16% | 96.39% | 98.98% | 98.39% |
| **Easy Subset** | 99.99% | 99.95% | 99.99% | 99.98% |
| **Medium Subset** | 84.41% | 86.12% | 98.89% | 93.71% |
| **Hard Subset** | 74.33% | 53.41% | 87.93% | 80.13% |

Detection engine performance for varying network data subsets based on malware observability

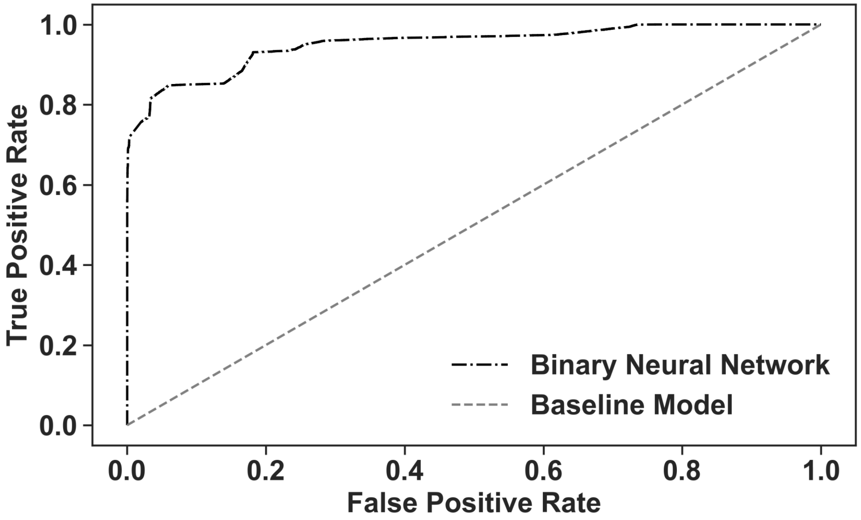


| Metric | Binary NN | Multi-Class NN | RF | ET |
|---|---|---|---|---|
| Avg Accuracy | 90.01% | 95.44% | 99.01% | 98.27% |
| Std Deviation ($\sigma$) | 0.68% | 1.94% | 0.08% | 0.58% |

Detection engine performance for 10-fold cross validation

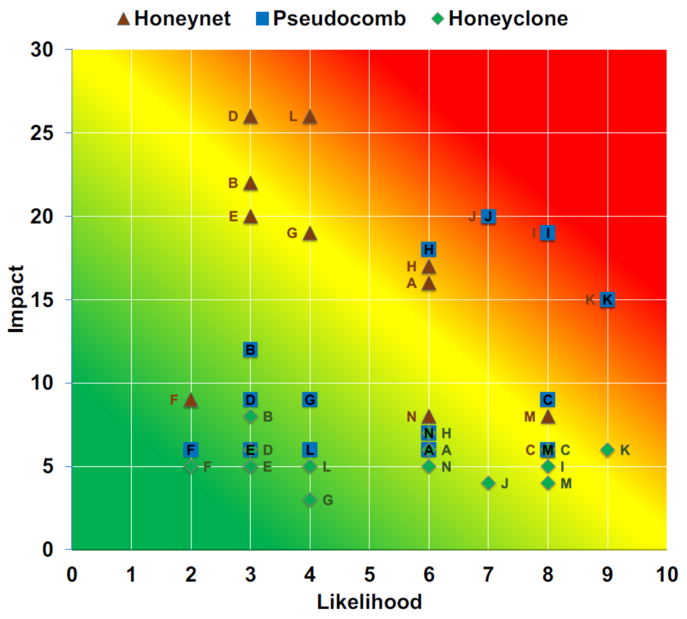
# Performance of Detection Engine

ROC Curve for Binary NN classifying benign and malicious traffic



# Risk Assessment of Decoy Environments

- The risk assessment shows risks of different data actions for different decoy environments
- It can be seen that Honeyclone performs better than the other two environments
- Honeyclone reduces risks by:
    - 78% more when compared to Honeynet
    - 40% more when compared to Pseudocomb



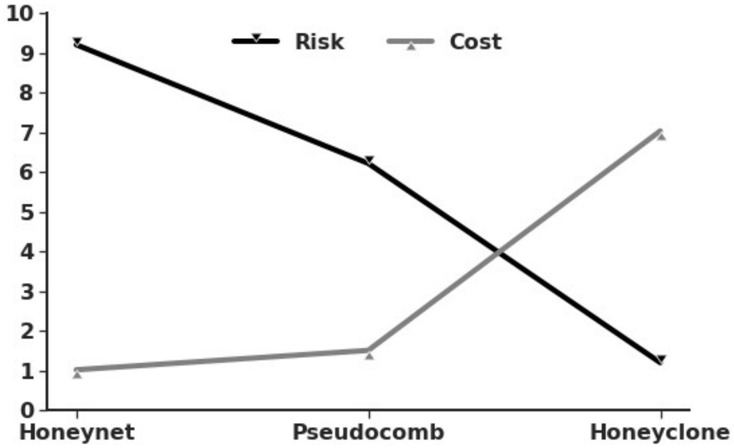
# Risk versus Cost

Graph shows risk versus cost relation

Cost amount and risk values are normalized into scale of [0,10]

It can be seen that Honeyclone is capable reduce risk of threat events significantly but incurs high cost

Honeyclone reduces risk by 88% when compared to network with no defense



# Conclusion

- We presented CICADA, an Active Defense Architecture against modern threats in Enterprise IoT Networks
- CICADA is equipped with different tier decoy environments designed for varying modern threats and their observability
- CICADA can detect attacks with different observability levels using an ensemble of neural networks and classifiers, with up to 73% accuracy for low observability attacks such as Zero Click
- We show the cost analysis for deploying various decoy environments, along with the assessment of risks associated with the different decoy environments, with up to 88% risk reduction via 'Honeyclone' when compared to a defenseless enterprise IoT-based network

# Future Work

- Explore anomaly-based detection instead of binary classification
- Utilize Raspberry Pi to generate traces of Zero-Click
- Extend Active Defense Engine logic to include more sophisticated procedures
- Detection and Defense as a dynamic process by applying a game theoretic approach to intelligently assign deception environments and reduce associated costs for defenders

THANK YOU!

ANY QUESTIONS?