Evaluating Machine Learning-Based IoT Device Identification Models for Security Applications

Eman Maali, Omar Alrawi, and Julie McCann

IMPERIAL

Georgia Institute of Technology

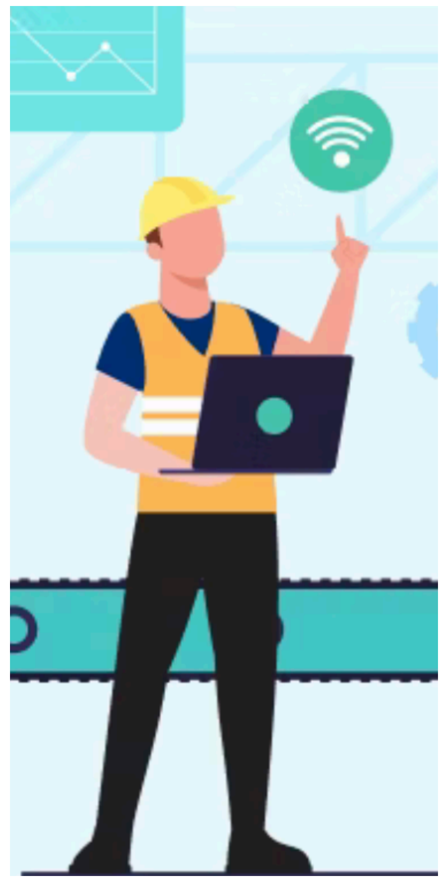
Feb, 2025

Images source: https://www.intuz.com/blog/iot-applications-in-smart-warehouse-management

What is the motivation and the problem statement?

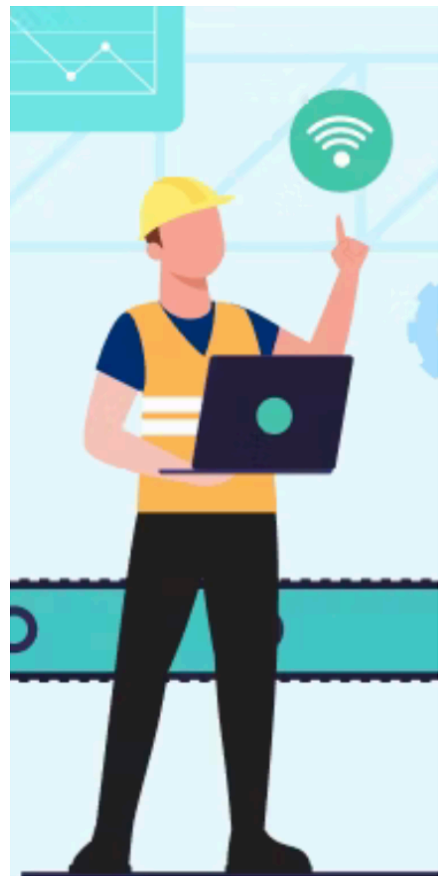
Images source: https://www.intuz.com/blog/iot-applications-in-smart-warehouse-management

What is the motivation and the problem statement?



Images source: https://www.intuz.com/blog/iot-applications-in-smart-warehouse-management

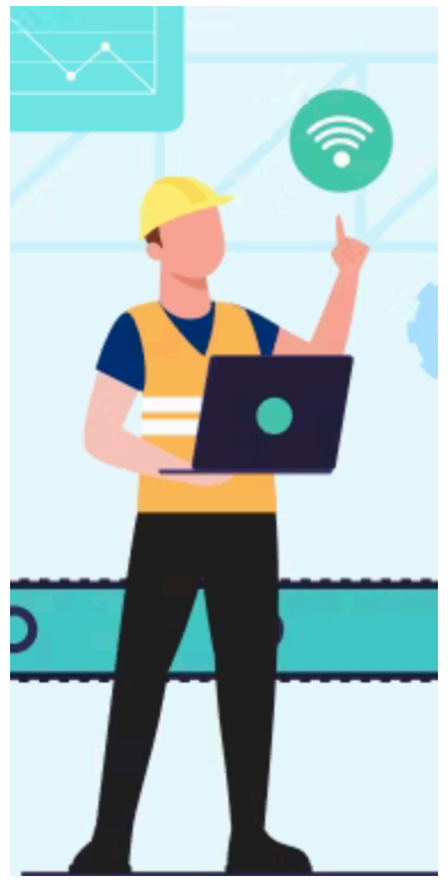
What is the motivation and the problem statement?



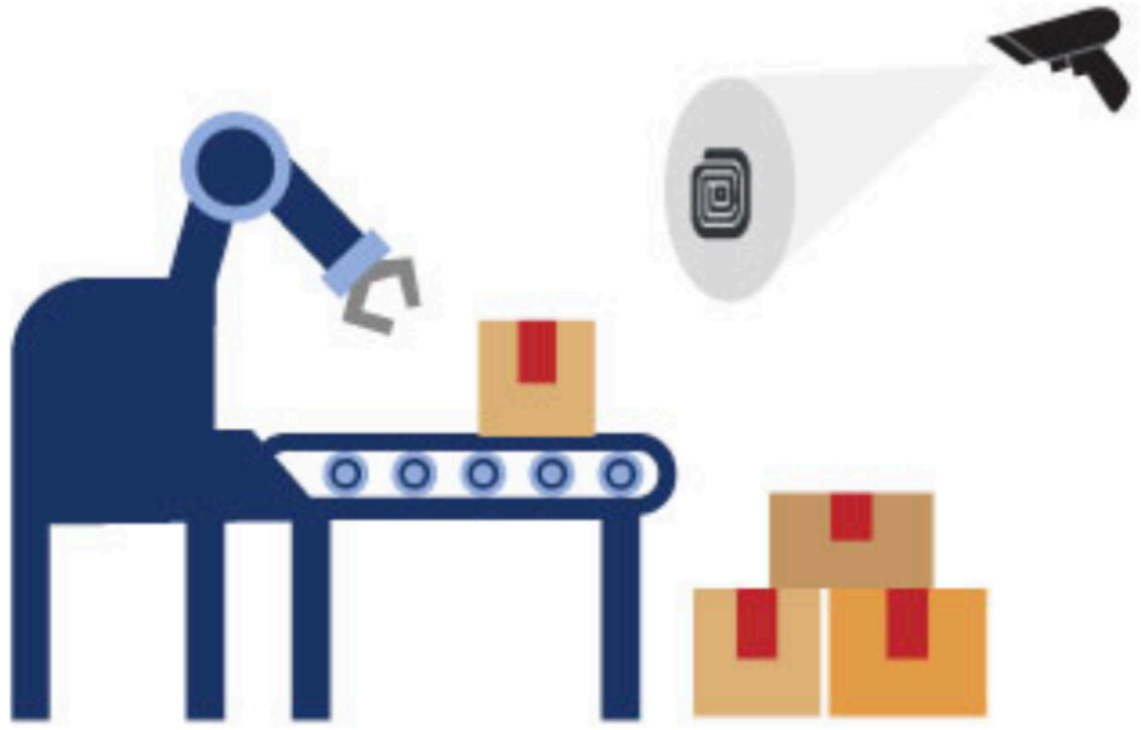
Network Operator

Images source: https://www.intuz.com/blog/iot-applications-in-smart-warehouse-management

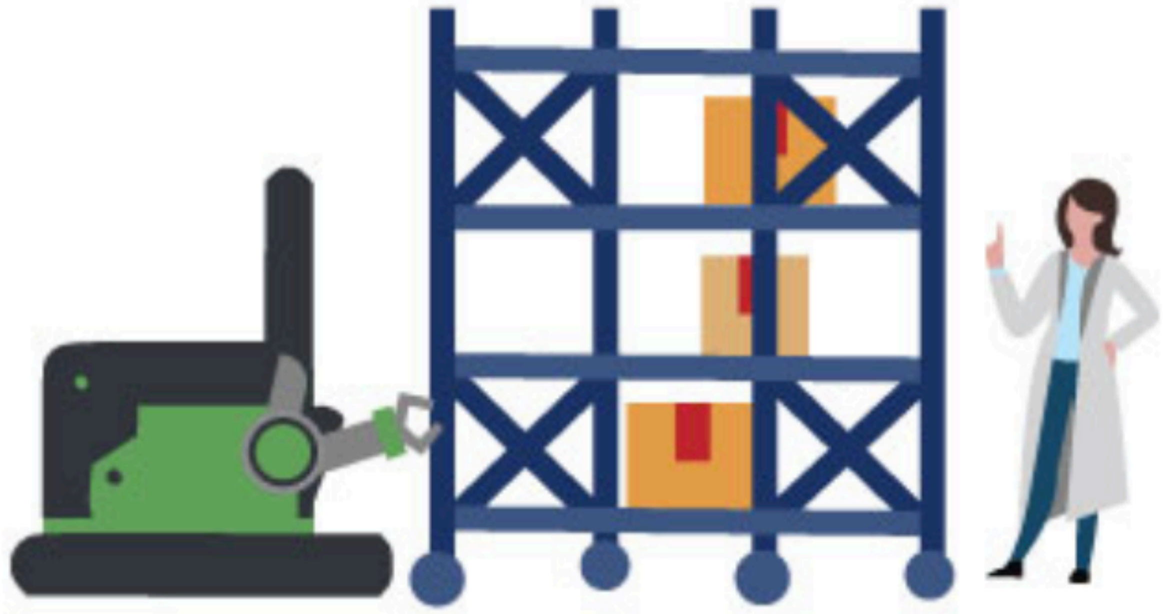
What is the motivation and the problem statement?



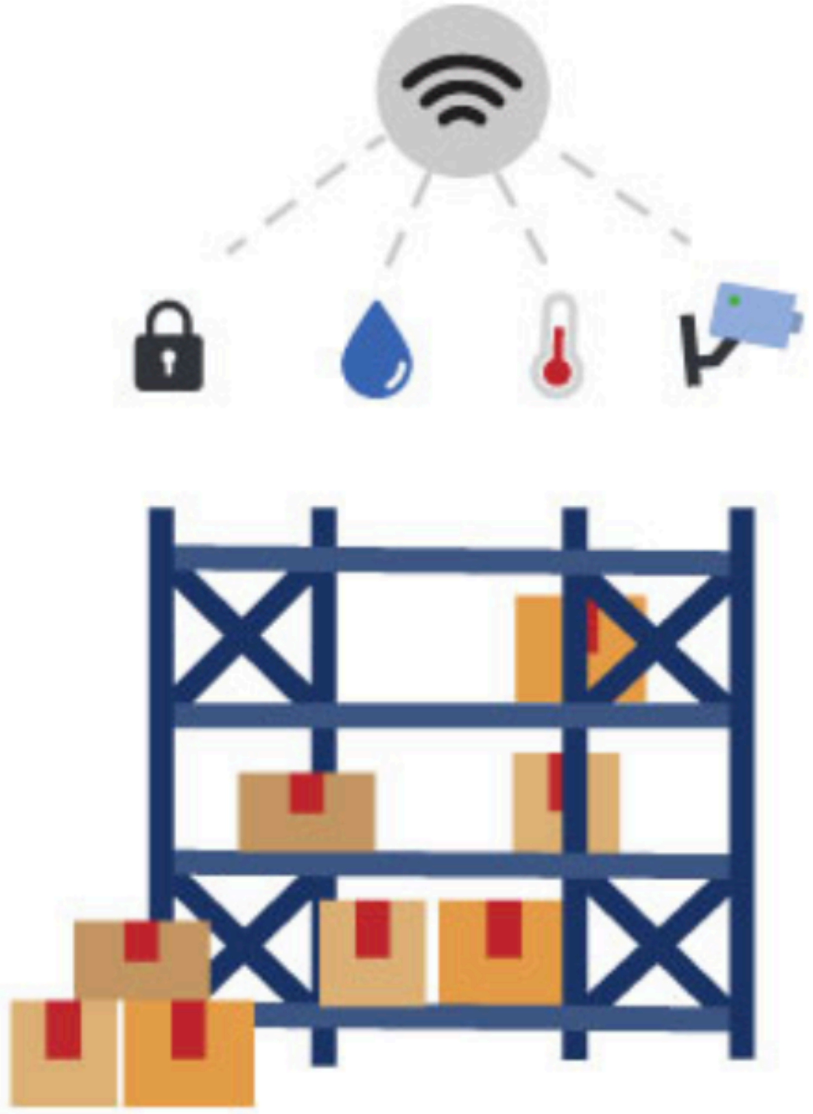
Network Operator



Automated items tracking



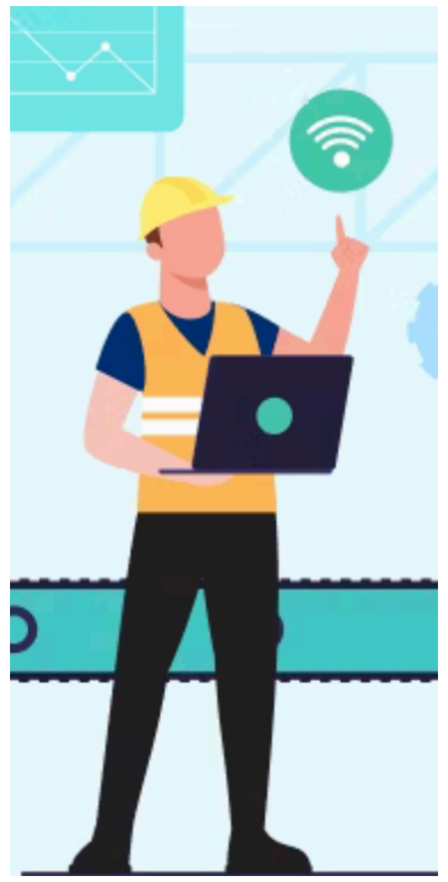
Storage racks moved by robots



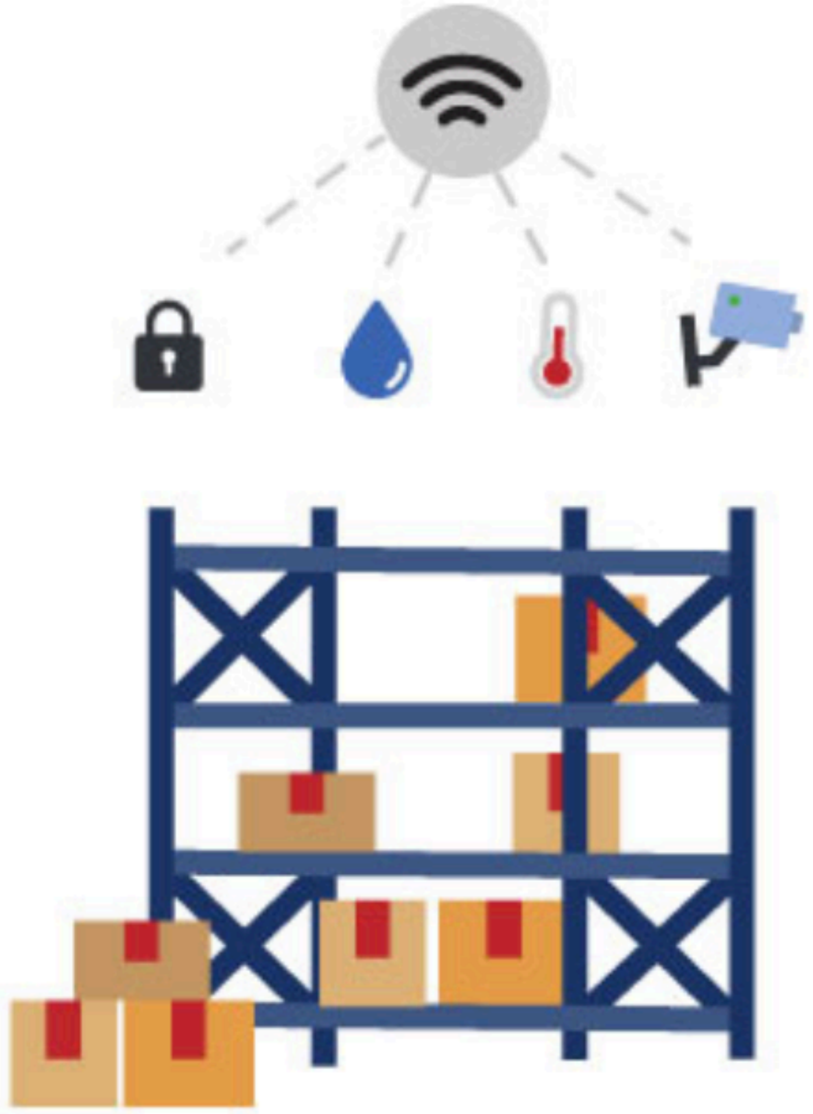
Smart warehouse maintenance

Images source: https://www.intuz.com/blog/iot-applications-in-smart-warehouse-management

What is the motivation and the problem statement?



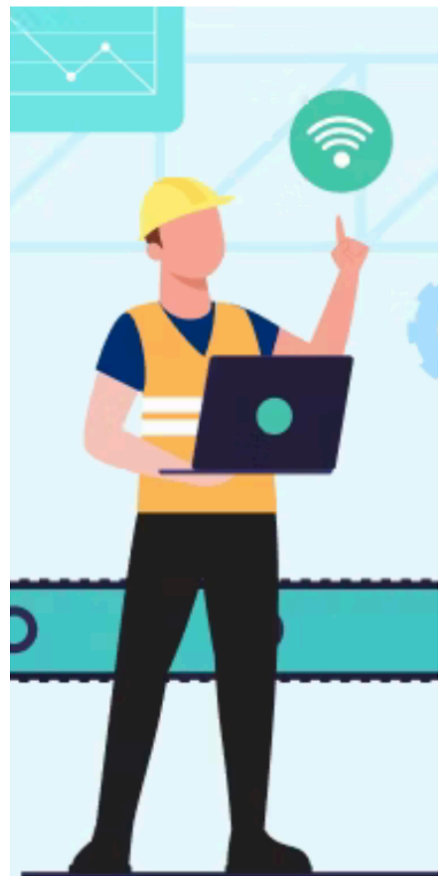
Network
Operator



Smart warehouse
maintenance

Images source: https://www.intuz.com/blog/iot-applications-in-smart-warehouse-management

What is the motivation and the problem statement?



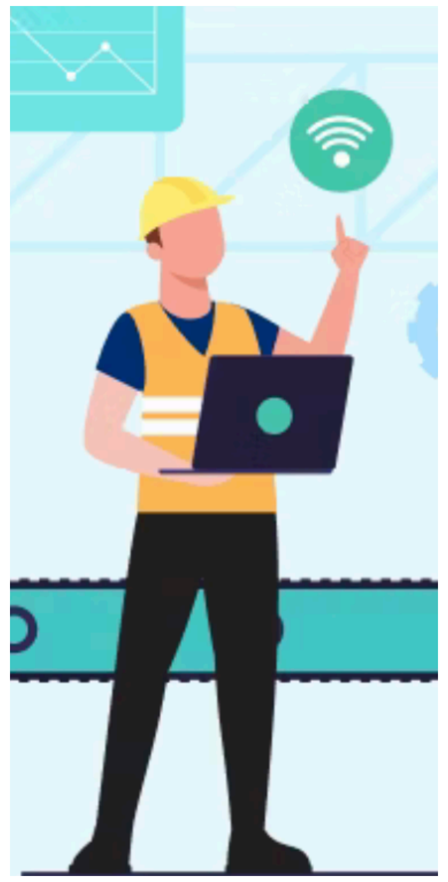
Network Operator

IoT Identification Solution

Smart warehouse maintenance

Images source: https://www.intuz.com/blog/iot-applications-in-smart-warehouse-management

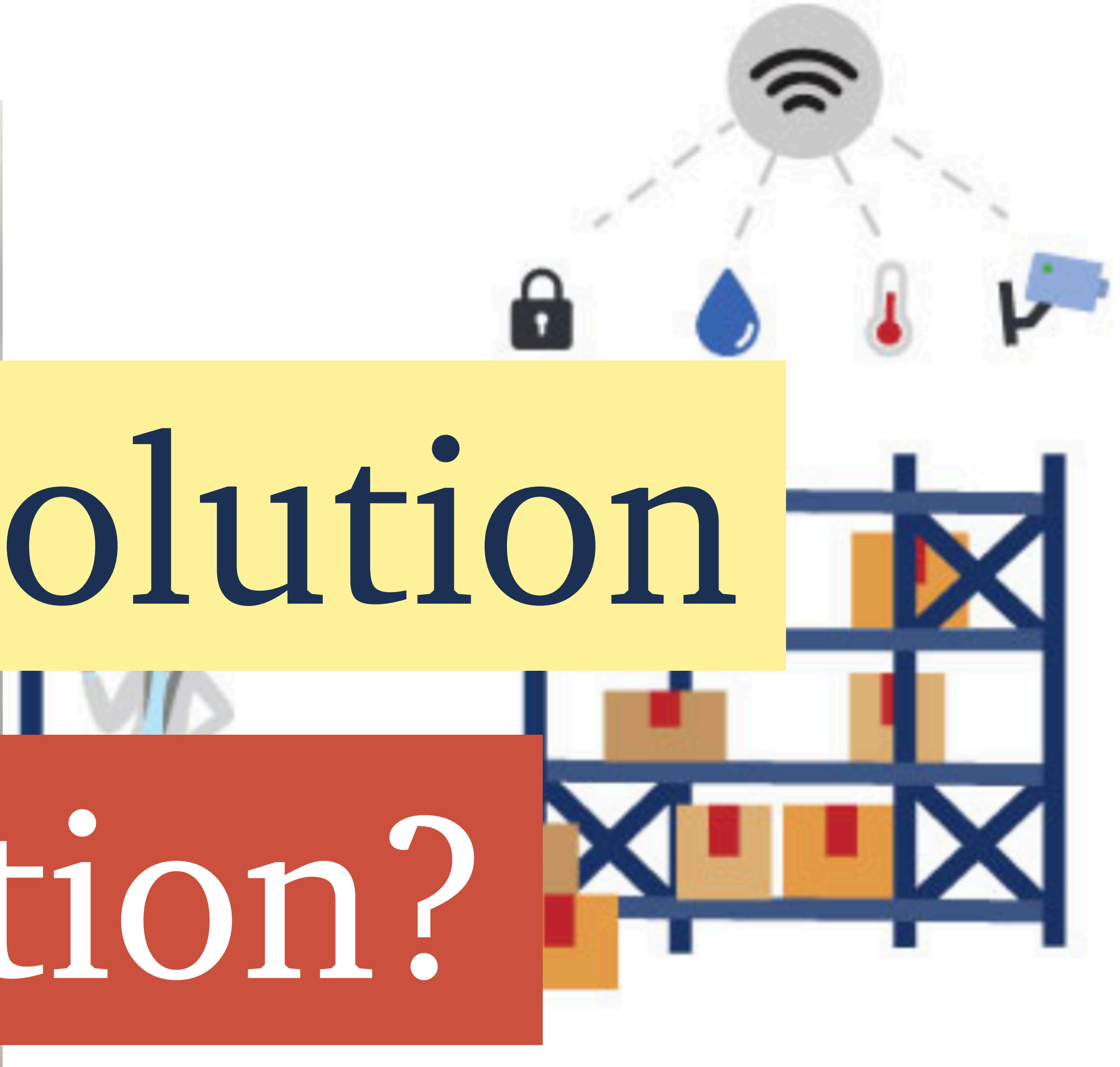
What is the motivation and the problem statement?



Network Operator

IoT Identification Solution

Incorrect Identification?



Smart warehouse maintenance

Images source: https://www.intuz.com/blog/iot-applications-in-smart-warehouse-management

What is the motivation and the problem statement?

What is the motivation and the problem statement?

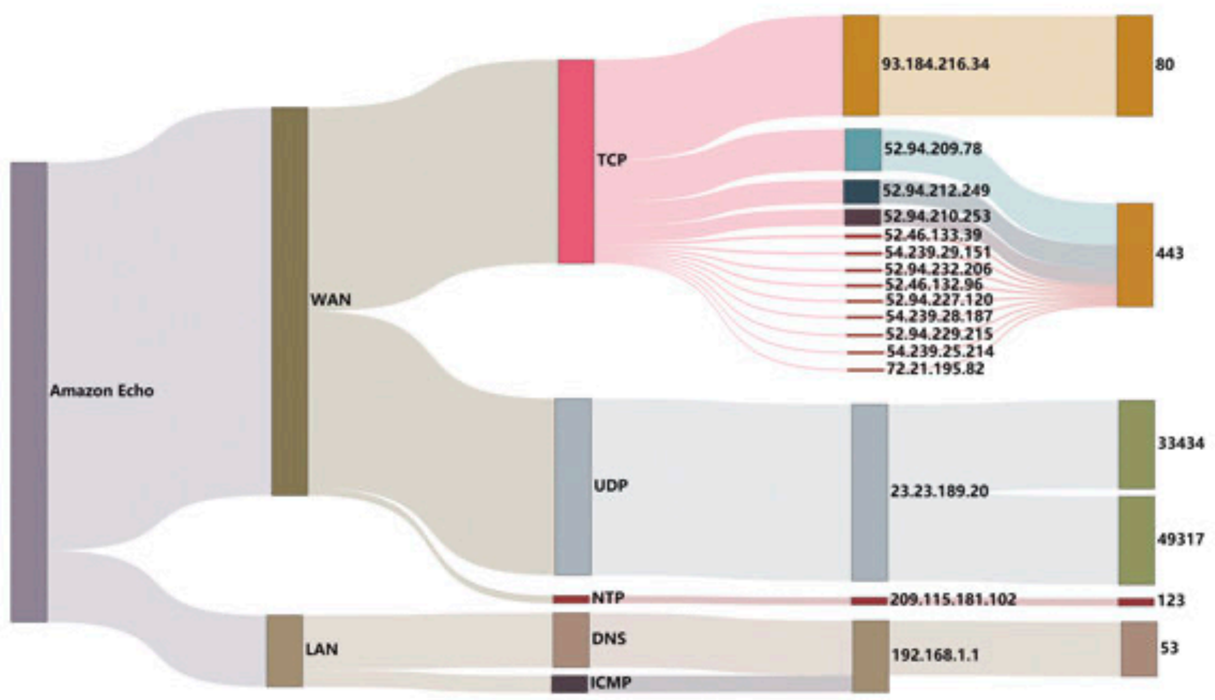
What IoT device identification models are currently available?

What is the motivation and the problem statement?

What IoT device identification models are currently available?

Static/Rule Based

Manufacturer Usage Description (MUD)



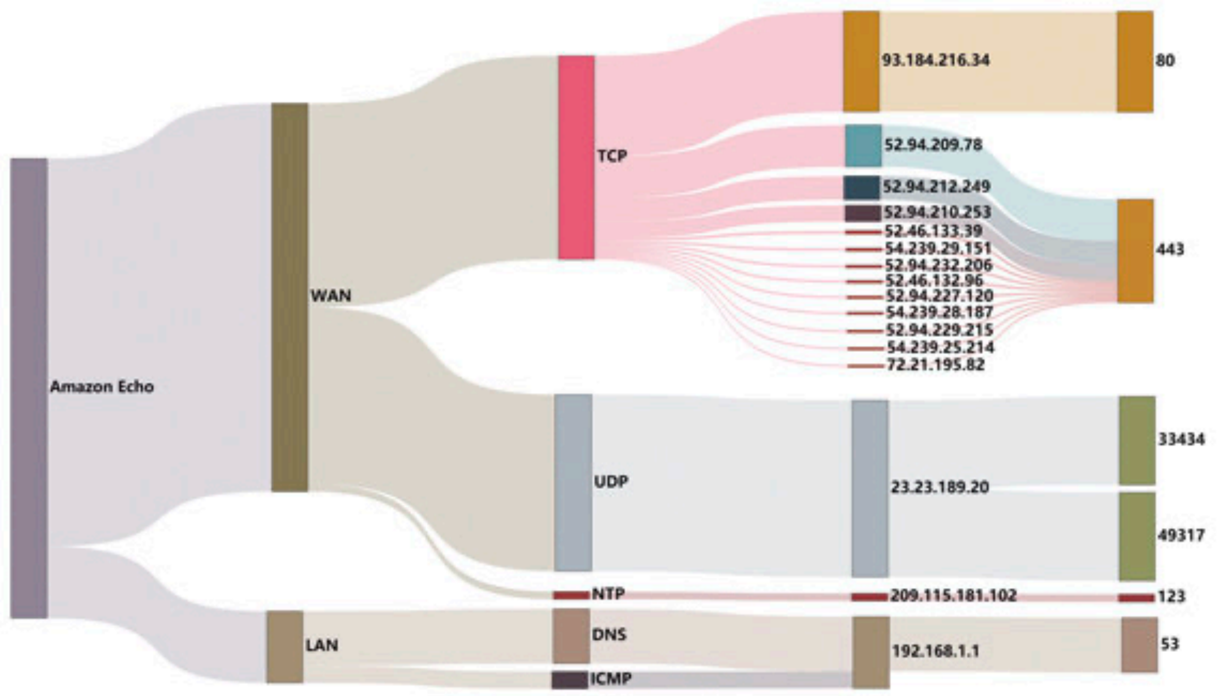
What is the motivation and the problem statement?

What IoT device identification models are currently available?

Static/Rule Based

OR

Manufacturer Usage Description (MUD)



What is the motivation and the problem statement?

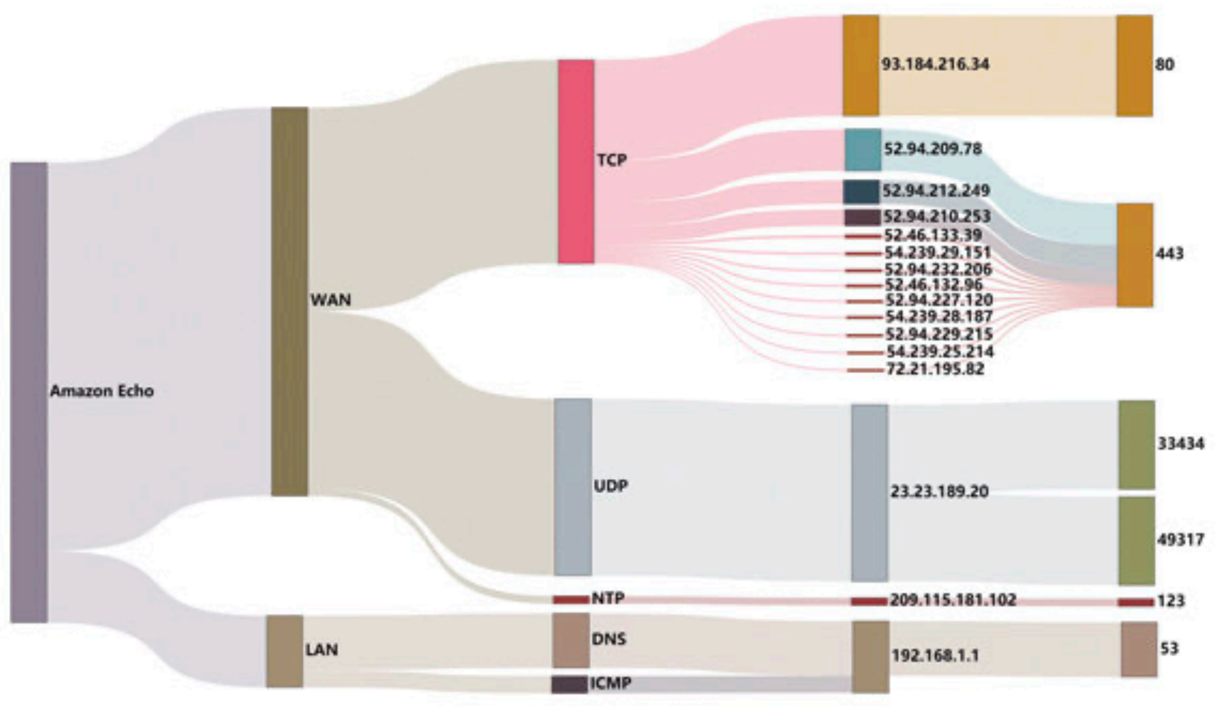
What IoT device identification models are currently available?

Static/Rule Based

OR

Machine Learning Based

Manufacturer Usage Description (MUD)



- Supervised Learning.
- Unsupervised Learning.
- Semi-Supervised Learning.

Algorithms: Random Forest, LSTM, CNN, LightGBM, DBSCAN.

What is the motivation and the problem statement?

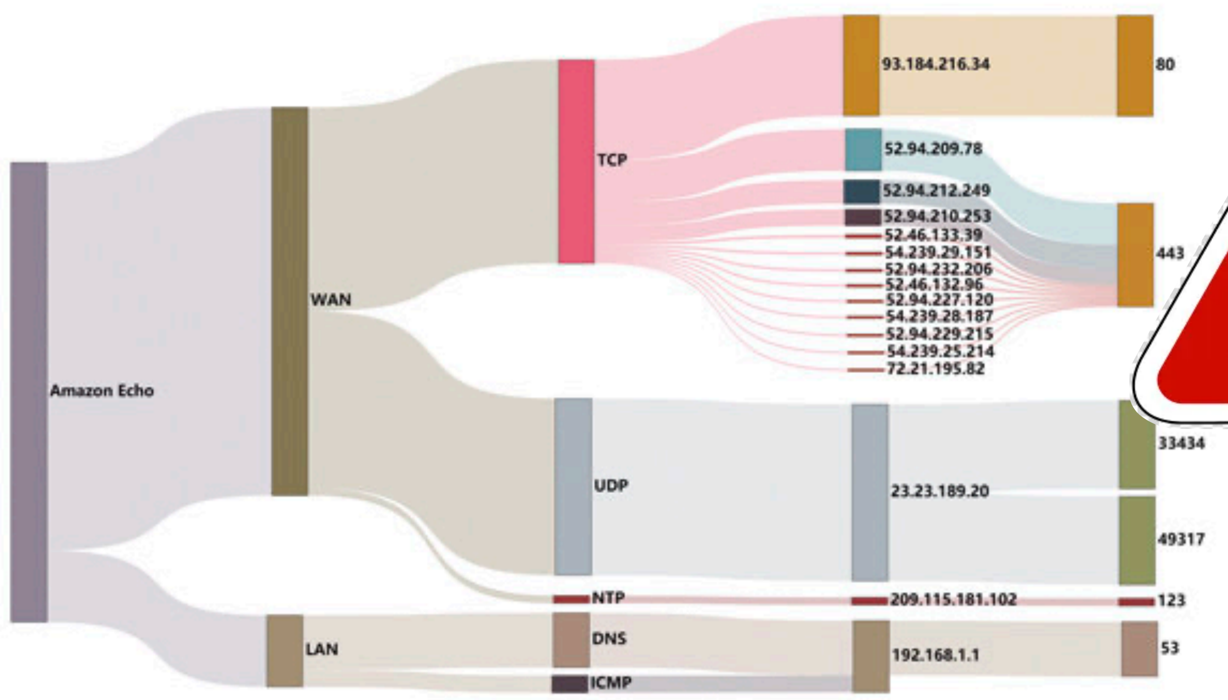
What IoT device identification models are currently available?

Static/Rule Based

OR

Machine Learning Based

Manufacturer Usage Description (MUD)



- Supervised Learning.
- Unsupervised Learning.
- Semi-Supervised Learning.

Algorithms: Random Forest, LSTM, CNN, LightGBM, DBSCAN.

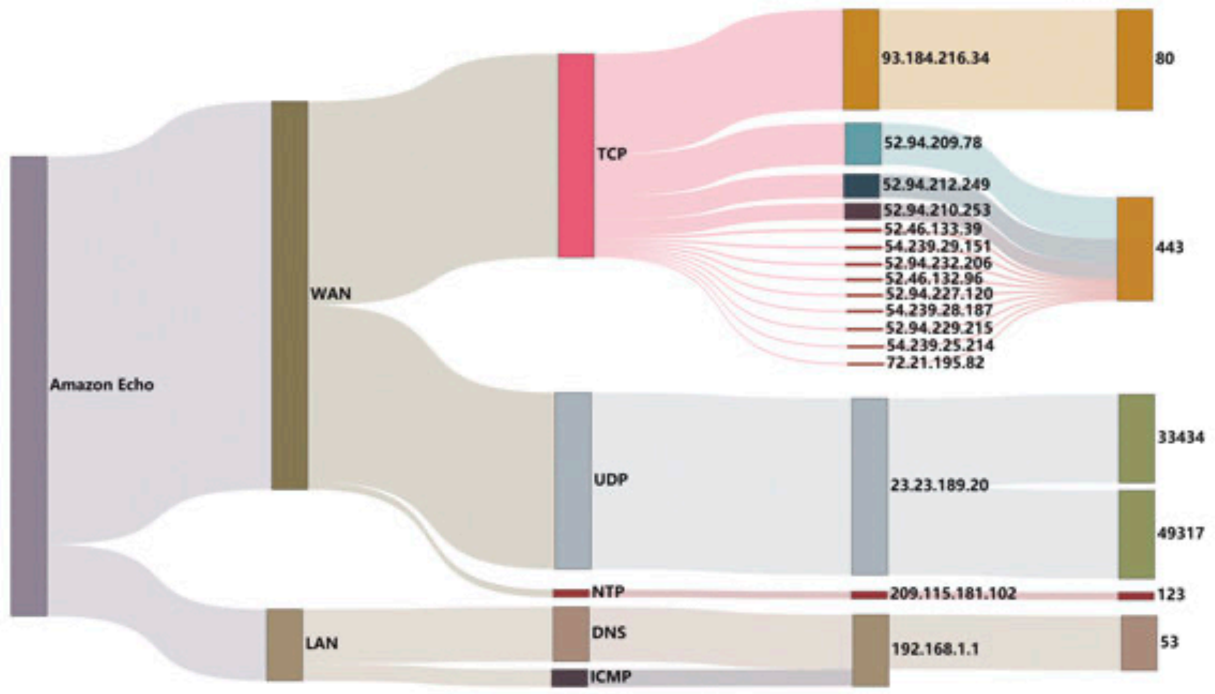
What is the motivation and the problem statement?

What IoT device identification models are currently available?

Static/Rule Based

OR

Manufacturer Usage Description (MUD)



Machine Learning Based

- Supervised Learning.
- Unsupervised Learning.
- Semi-Supervised Learning.

Algorithms: Random Forest, LSTM, CNN, LightGBM, DBSCAN.

What is the motivation and the problem statement?

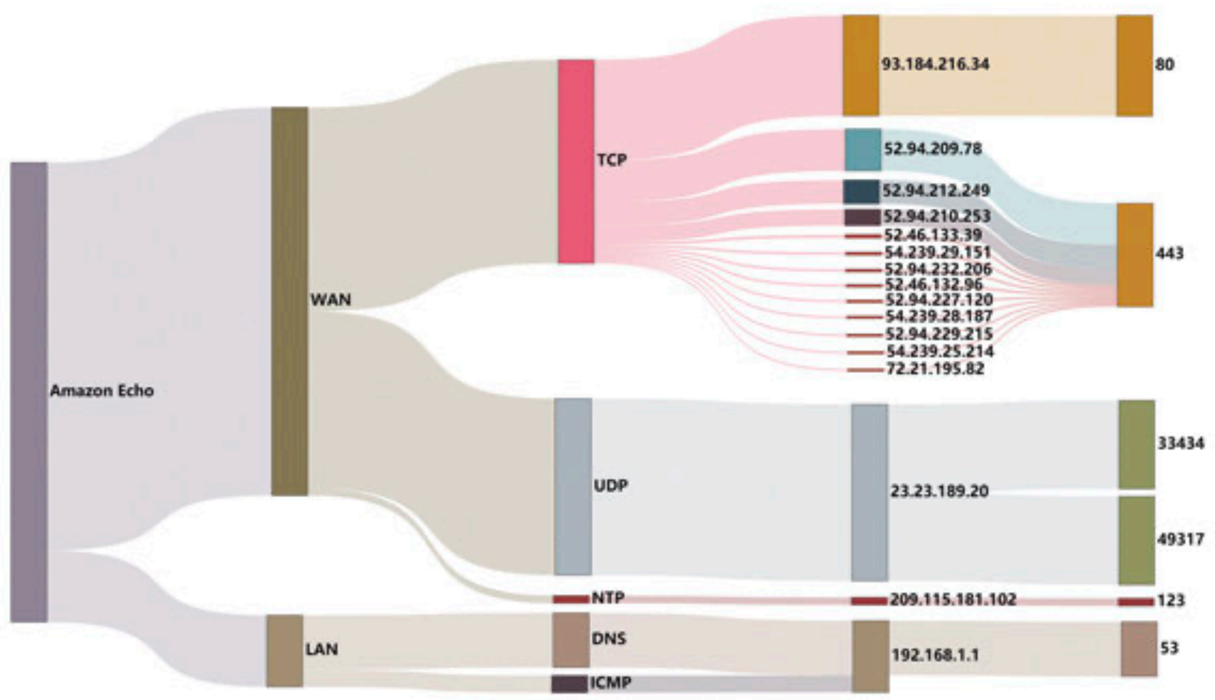
What IoT device identification models are currently available?

Static/Rule Based

OR

Machine Learning Based

Manufacturer Usage Description (MUD)



- Supervised Learning.
- Unsupervised Learning.
- Semi-Supervised Learning.

Algorithms: Random Forest, LSTM, CNN, LightGBM, DBSCAN.

70% of current SoA

What is the motivation and the problem statement?

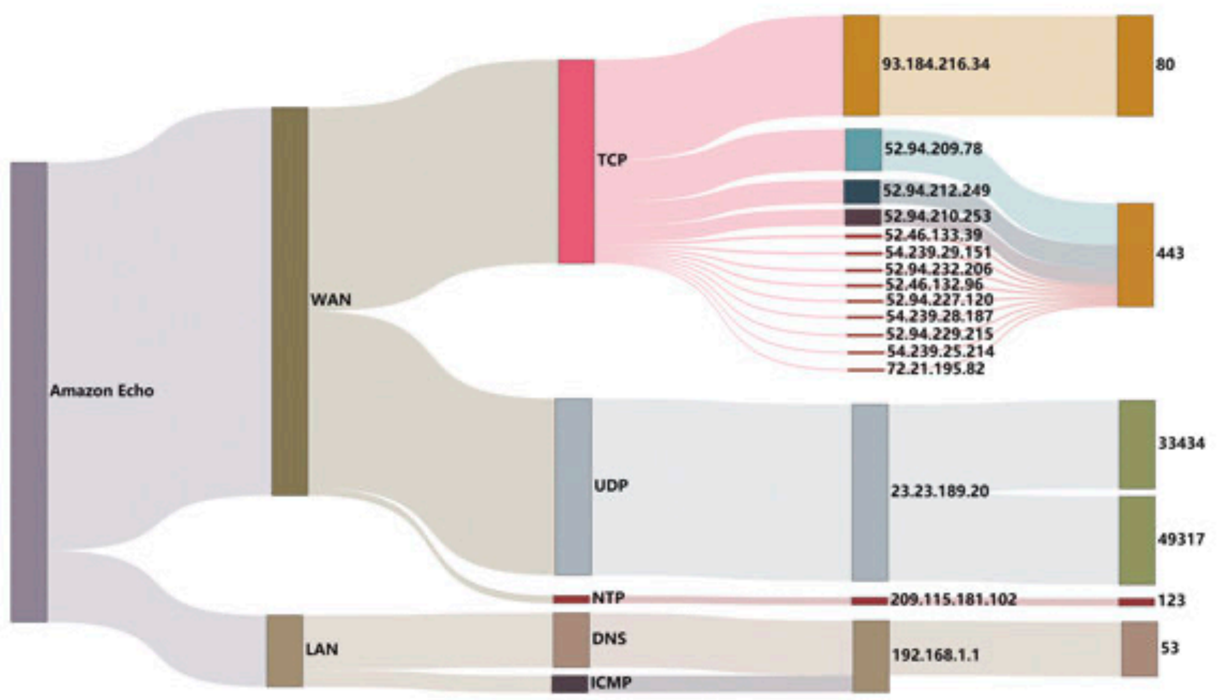
What IoT device identification models are currently available?

Static/Rule Based

OR

Machine Learning Based

Manufacturer Usage Description (MUD)



- Supervised Learning.
- Unsupervised Learning.
- Semi-Supervised Learning.

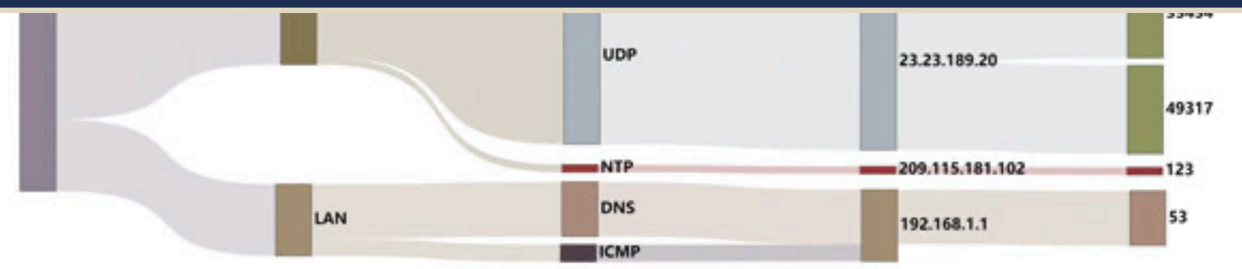
Algorithms: Random Forest, LSTM, CNN, LightGBM, DBSCAN.

70% of current SoA

What is the motivation and the problem statement?

What IoT device identification models are currently available?

Can IoT device identification models be deployed?



Machine Learning Based

- Supervised Learning.
- Unsupervised Learning.
- Semi-Supervised Learning.

Algorithms: Random Forest, LSTM, CNN, LightGBM, DBSCAN.

70% of current SoA

What defines a practical ML-based device identification model?

What defines a practical ML-based device identification model?

Practicality Definition and Attributes

What defines a practical ML-based device identification model?

Practicality Definition and Attributes

The model's capability to ensure robust and reliable IoT device identification across different operational modes, deployment environments, and network conditions.

What defines a practical ML-based device identification model?

Practicality Definition and Attributes

The model's capability to ensure robust and reliable IoT device identification across different operational modes, deployment environments, and network conditions.

What defines a practical ML-based device identification model?

Practicality Definition and Attributes

The model's capability to ensure robust and reliable IoT device identification across different operational modes, deployment environments, and network conditions.

What defines a practical ML-based device identification model?

Practicality Definition and Attributes

The model's capability to ensure robust and reliable IoT device identification across different operational modes, deployment environments, and network conditions.

What defines a practical ML-based device identification model?

Practicality Definition and Attributes

The model's capability to ensure robust and reliable IoT device identification across different operational modes, deployment environments, and network conditions.

What defines a practical ML-based device identification model?

Practicality Definition and Attributes

The model's capability to ensure robust and reliable IoT device identification across different operational modes, deployment environments, and network conditions.

What defines a practical ML-based device identification model?

Practicality Definition and Attributes

The model's capability to ensure robust and reliable IoT device identification across different operational modes, deployment environments, and network conditions.

What defines a practical ML-based device identification model?

Practicality Definition and Attributes

The model's capability to ensure robust and reliable IoT device identification across different operational modes, deployment environments, and network conditions.

1. What defines a robust and reliable solution?

What defines a practical ML-based device identification model?

Practicality Definition and Attributes

The model's capability to ensure robust and reliable IoT device identification across different operational modes, deployment environments, and network conditions.

1. What defines a robust and reliable solution?
2. What do varying modes and deployment environments mean?

What defines a practical ML-based device identification model?

Practicality Definition and Attributes

The model's capability to ensure robust and reliable IoT device identification across different operational modes, deployment environments, and network conditions.

1. What defines a robust and reliable solution?
2. What do varying modes and deployment environments mean?
3. What network conditions must a solution consider?

What attributes define practicality in ML-based models?

What attributes define practicality in ML-based models?

- Generalisation and Robustness
- Stability Over Time
- Model Scalability
- Data Efficiency
- Deployment Compatibility
- Cost Metric
- Ethics and Societal Impact
- Fairness and Accountability

What attributes define practicality in ML-based models?

- Generalisation and Robustness
- Stability Over Time
- Model Scalability
- Data Efficiency
- Deployment Compatibility
- Cost Metric
- Ethics and Societal Impact
- Fairness and Accountability

What attributes define practicality in ML-based models?

- Generalisation and Robustness
- Stability Over Time
- Model Scalability
- Data Efficiency
- Deployment Compatibility
- Cost Metric
- Ethics and Societal Impact
- Fairness and Accountability

Which are relevant to practicality in the context of IoT identification?

What attributes define practicality in ML-based models?

- Generalisation and Robustness
- Stability Over Time
- Model Scalability
- Data Efficiency
- Deployment Compatibility
- Cost Metric
- Ethics and Societal Impact
- Fairness and Accountability

Which are relevant to practicality in the context of IoT identification?

Deploying a solution that can be generalised across different environments/configurations simultaneously with robustness in performance and stability over time.

What generalise, robustness and stability over time mean in IoT environments?

What generalise, robustness and stability over time mean in IoT environments?

Practicality Definition and Attributes

What generalise, robustness and stability over time mean in IoT environments?

Practicality Definition and Attributes

Defintion

What generalise, robustness and stability over time mean in IoT environments?

Practicality Definition and Attributes

Defintion

Attributes

What generalise, robustness and stability over time mean in IoT environments?

Practicality Definition and Attributes

Defintion

The model's ability to **generalise** across variability introduced by the IoT device life cycle (active vs. idle).

Attributes

What generalise, robustness and stability over time mean in IoT environments?

Practicality Definition and Attributes

Defintion

The model's ability to **generalise** across variability introduced by the IoT device life cycle (active vs. idle).

Attributes

What generalise, robustness and stability over time mean in IoT environments?

Practicality Definition and Attributes

Defintion

The model's ability to **generalise** across variability introduced by the IoT device life cycle (active vs. idle).

Attributes

Mode of Operation

What generalise, robustness and stability over time mean in IoT environments?

Practicality Definition and Attributes

Defintion

The model's ability to **generalise** across variability introduced by the IoT device life cycle (active vs. idle).

The model's ability to **generalise** across diverse environments and remains **robust** to temporal changes.

Attributes

Mode of Operation

What generalise, robustness and stability over time mean in IoT environments?

Practicality Definition and Attributes

Defintion

The model's ability to **generalise** across variability introduced by the IoT device life cycle (active vs. idle).

The model's ability to **generalise** across diverse environments and remains **robust** to temporal changes.

Attributes

Mode of Operation

What generalise, robustness and stability over time mean in IoT environments?

Practicality Definition and Attributes

Defintion	The model's ability to **generalise** across variability introduced by the IoT device life cycle (active vs. idle).	The model's ability to **generalise** across diverse environments and remains **robust** to temporal changes.
Attributes	Mode of Operation	Transferability

What generalise, robustness and stability over time mean in IoT environments?

Practicality Definition and Attributes

Defintion	The model's ability to **generalise** across variability introduced by the IoT device life cycle (active vs. idle).	The model's ability to **generalise** across diverse environments and remains **robust** to temporal changes.	The model's ability to maintain **robust** performance across various network conditions and sampling rates.
Attributes	Mode of Operation	Transferability	

What generalise, robustness and stability over time mean in IoT environments?

Practicality Definition and Attributes

Defintion	The model's ability to **generalise** across variability introduced by the IoT device life cycle (active vs. idle).	The model's ability to **generalise** across diverse environments and remains **robust** to temporal changes.	The model's ability to maintain **robust** performance across various network conditions and sampling rates.
Attributes	Mode of Operation	Transferability	

What generalise, robustness and stability over time mean in IoT environments?

Practicality Definition and Attributes

Defintion	The model's ability to **generalise** across variability introduced by the IoT device life cycle (active vs. idle).	The model's ability to **generalise** across diverse environments and remains **robust** to temporal changes.	The model's ability to maintain **robust** performance across various network conditions and sampling rates.
Attributes	Mode of Operation	Transferability	Observability

What generalise, robustness and stability over time mean in IoT environments?

Practicality Definition and Attributes

Defintion

The model's ability to **generalise** across variability introduced by the IoT device life cycle (active vs. idle).

The model's ability to **generalise** across diverse environments and remains **robust** to temporal changes.

The model's ability to maintain **robust** performance across various network conditions and sampling rates.

Attributes

Mode of Operation

Transferability

Observability

How should current solutions be evaluated against the three attributes?

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

1

IoT Identification Problem

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1 Dataset

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1 **Dataset**

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1 **Dataset** — What data should be collected, and how should the dataset be gathered for evaluation?

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1 Dataset — What data should be collected, and how should the dataset be gathered for evaluation?

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1 **Dataset** What data should be collected, and how should the dataset be gathered for evaluation?

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1 **Dataset** — What data should be collected, and how should the dataset be gathered for evaluation?

2

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1. **Dataset** — What data should be collected, and how should the dataset be gathered for evaluation?
2. **Model**

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1 **Dataset** — What data should be collected, and how should the dataset be gathered for evaluation?

2 **Model**

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1. **Dataset** — What data should be collected, and how should the dataset be gathered for evaluation?
2. **Model** — Which model should be used, and how complex should it be?

IoT Identification Problem

2

Model

Which model should be used, and how complex should it be?

IoT Identification Problem

2

Model

Which model should be used, and how complex should it be?

200

Papers

IoT Identification Problem

2

Model

Which model should be used, and how complex should it be?

200

Papers

IoT identification.

IoT Identification Problem

2

Model

Which model should be used, and how complex should it be?

200
Papers

96
Papers

IoT identification.

IoT Identification Problem

2

Model

Which model should be used, and how complex should it be?

200
Papers

IoT identification.

96
Papers

Representative work.

IoT Identification Problem

2

Model

Which model should be used, and how complex should it be?

200 Papers

96 Papers

14 Papers

IoT identification.

Representative work.

IoT Identification Problem

2

Model

Which model should be used, and how complex should it be?

200 Papers

96 Papers

14 Papers

IoT identification.

Representative work.

IoT Identification Problem

2

Model

Which model should be used, and how complex should it be?

200 Papers

IoT identification.

96 Papers

14 Papers

Representative work.

2

Original code public.

IoT Identification Problem

2 **Model** Which model should be used, and how complex should it be?

200 Papers

IoT identification.

96 Papers

14 Papers

Representative work.

2 Original code public.

1 Original code with few modifications.

IoT Identification Problem — 2 — **Model** — Which model should be used, and how complex should it be?

200 Papers

IoT identification.

96 Papers

14 Papers

Representative work.

- 2 — Original code public.
- 1 — Original code with few modifications.
- 7 — Implemented using original paper description.

IoT Identification Problem

2 — **Model** — Which model should be used, and how complex should it be?

200 Papers

IoT identification.

96 Papers

14 Papers

Representative work.

- 2 — Original code public.
- 1 — Original code with few modifications.
- 7 — Implemented using original paper description.
- 4 — No sufficient information about feature extraction.

IoT Identification Problem

2 **Model** Which model should be used, and how complex should it be?

200 Papers

IoT identification.

96 Papers

14 Papers

Representative work.

- 2 Original code public.
- 1 Original code with few modifications.
- 7 Implemented using original paper description.
- 4 No sufficient information about feature extraction.

IoT Identification Problem

2

Model

Which model should be used, and how complex should it be?

200 Papers

IoT identification.

96 Papers

14 Papers

Representative work.

2 Original code public.

1 Original code with few modifications.

7 Implemented using original paper description.

10 Papers

4 No sufficient information about feature extraction.

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1. **Dataset** — What data should be collected, and how should the dataset be gathered for evaluation?
2. **Model** — Which model should be used, and how complex should it be?

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1. **Dataset** — What data should be collected, and how should the dataset be gathered for evaluation?
2. **Model** — Which model should be used, and how complex should it be?
3.

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1. **Dataset** — What data should be collected, and how should the dataset be gathered for evaluation?
2. **Model** — Which model should be used, and how complex should it be?
3. **Features**

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1. **Dataset** — What data should be collected, and how should the dataset be gathered for evaluation?
2. **Model** — Which model should be used, and how complex should it be?
3. **Features**

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1. **Dataset** — What data should be collected, and how should the dataset be gathered for evaluation?
2. **Model** — Which model should be used, and how complex should it be?
3. **Features** — What features should be extracted, and how should they be represented?

How should current solutions be evaluated against the three attributes?

Components of ML-based Model

IoT Identification Problem

1. **Dataset** — What data should be collected, and how should the dataset be gathered for evaluation?
2. **Model** — Which model should be used, and how complex should it be?
3. **Features** — What features should be extracted, and how should they be represented?

What is the experimental setup for practicality evaluation and attributes?

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

Mode of Operation

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

Mode of Operation

Scenarios

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

Mode of Operation

Scenarios

1 *Mix (idle & active) vs idle*

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

Mode of Operation

Scenarios

1. *Mix (idle & active) vs idle*
2. *Mix vs active*

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

Mode of Operation

Scenarios

1. *Mix (idle & active) vs idle*
2. *Mix vs active*
3. *Mix vs Mix*

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

Mode of Operation

Transferability

Scenarios

1. *Mix (idle & active) vs idle*
2. *Mix vs active*
3. *Mix vs Mix*

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

Mode of Operation	Transferability
	Spatial

Scenarios

1. *Mix (idle & active) vs idle*
2. *Mix vs active*
3. *Mix vs Mix*

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

Scenarios

Mode of Operation	Transferability
1 *Mix (idle & active) vs idle*	Spatial 1 *UK vs USA*
2 *Mix vs active*	
3 *Mix vs Mix*	

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

Scenarios

Mode of Operation	Transferability
1 *Mix (idle & active) vs idle*	Spatial 1 *UK vs USA*
2 *Mix vs active*	2 *USA vs UK*
3 *Mix vs Mix*	

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

Scenarios

Mode of Operation

1. *Mix (idle & active) vs idle*
2. *Mix vs active*
3. *Mix vs Mix*

Transferability

Spatial

1. *UK vs USA*
2. *USA vs UK*

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

Scenarios

Mode of Operation	Transferability	
1 *Mix (idle & active) vs idle*	Spatial	1 *UK vs USA*
2 *Mix vs active*		2 *USA vs UK*
3 *Mix vs Mix*	Temporal	

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

Scenarios

Mode of Operation

1 *Mix (idle & active) vs idle*

2 *Mix vs active*

3 *Mix vs Mix*

Transferability

Spatial

1 *UK vs USA*

2 *USA vs UK*

Temporal

7 *1 week- 52 weeks Gaps*

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

Mode of Operation	Transferability		Observability
1 *Mix (idle & active) vs idle*	Spatial	1 *UK vs USA*	
2 *Mix vs active*		2 *USA vs UK*	
3 *Mix vs Mix*	Temporal	7 *1 week- 52 weeks Gaps*	

Scenarios

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes	Mode of Operation	Transferability		Observability
Scenarios	1 *Mix (idle & active) vs idle*	Spatial	1 *UK vs USA*	1 *1:100s*
	2 *Mix vs active*		2 *USA vs UK*	
	3 *Mix vs Mix*	Temporal	7 *1 week- 52 weeks Gaps*	

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes	Mode of Operation	Transferability		Observability
Scenarios	1 *Mix (idle & active) vs idle*	Spatial	1 *UK vs USA*	1 *1:100s*
	2 *Mix vs active*		2 *USA vs UK*	2 *1:1000s*
	3 *Mix vs Mix*	Temporal	7 *1 week- 52 weeks Gaps*	

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes

Scenarios

Mode of Operation		Transferability		Observability
1 *Mix (idle & active) vs idle*	Spatial	1 *UK vs USA*		1 *1:100s*
2 *Mix vs active*		2 *USA vs UK*		2 *1:1000s*
3 *Mix vs Mix*	Temporal	7 *1 week- 52 weeks Gaps*		3 *1:5000s*

What is the experimental setup for practicality evaluation and attributes?

For each of 10 papers, we have a baseline, and perform the following experimental scenarios:

Attributes	Mode of Operation		Transferability			Observability	
Scenarios	1	*Mix (idle & active) vs idle*	Spatial	1	*UK vs USA*	1	*1:100s*
	2	*Mix vs active*		2	*USA vs UK*	2	*1:1000s*
	3	*Mix vs Mix*	Temporal	7	*1 week- 52 weeks Gaps*	3	*1:5000s*

In total, we performed 140 practicality evaluation across three attributes.

What are the key findings of the practicality evaluation?

What are the key findings of the practicality evaluation?

Attributes

What are the key findings of the practicality evaluation?

Attributes

What are the key findings of the practicality evaluation?

Attributes **Key Findings**

What are the key findings of the practicality evaluation?

Attributes	Key Findings
Mode of Operation	

What are the key findings of the practicality evaluation?

Attributes	Key Findings
Mode of Operation	

What are the key findings of the practicality evaluation?

Attributes	Key Findings
Mode of Operation	Idle and active modes introduce behavioural shifts that reduce performance.

What are the key findings of the practicality evaluation?

Attributes	Key Findings
Mode of Operation	Idle and active modes introduce behavioural shifts that reduce performance.
Transferability	

What are the key findings of the practicality evaluation?

Attributes	Key Findings
Mode of Operation	Idle and active modes introduce behavioural shifts that reduce performance.
Transferability	

What are the key findings of the practicality evaluation?

Attributes	Key Findings
Mode of Operation	Idle and active modes introduce behavioural shifts that reduce performance.
Transferability	- Spatial degradation drop of 7.5%–74% . - Temporal degradation begins after 1 week (19.32%) and worsens to 85.90% after a year.

What are the key findings of the practicality evaluation?

Attributes	Key Findings
Mode of Operation	Idle and active modes introduce behavioural shifts that reduce performance.
Transferability	- Spatial degradation drop of 7.5%–74% . - Temporal degradation begins after 1 week (19.32%) and worsens to 85.90% after a year.
Observability	

What are the key findings of the practicality evaluation?

Attributes	Key Findings
Mode of Operation	Idle and active modes introduce behavioural shifts that reduce performance.
Transferability	- Spatial degradation drop of 7.5%–74% . - Temporal degradation begins after 1 week (19.32%) and worsens to 85.90% after a year.
Observability	

What are the key findings of the practicality evaluation?

Attributes	Key Findings
Mode of Operation	Idle and active modes introduce behavioural shifts that reduce performance.
Transferability	- Spatial degradation drop of 7.5%–74% . - Temporal degradation begins after 1 week (19.32%) and worsens to 85.90% after a year.
Observability	Sampled traffic (e.g., sFlow) reduces performance by an average of 70.09%.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

Paper

Meid20*

*Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai, "A novel approach for detecting vulnerable IoT devices connected behind a home NAT," Computers & Security, 2020.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

Paper ⟶

Meid20*

*Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai, "A novel approach for detecting vulnerable IoT devices connected behind a home NAT," Computers & Security, 2020.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

Paper → ***Setup***

Meid20*

Model: LightGBM

Features: ibytes,

ipackets, TCP flags, Port,
IPv4 address

*Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai, "A novel approach for detecting vulnerable IoT devices connected behind a home NAT," Computers & Security, 2020.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

Paper	*Setup*
Meid20*	**Model:** LightGBM **Features:** ibytes, ipackets, TCP flags, Port, IPv4 address

*Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai, "A novel approach for detecting vulnerable IoT devices connected behind a home NAT," Computers & Security, 2020.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

Paper	*Setup*	**Training**
Meid20*	**Model:** LightGBM **Features:** ibytes, ipackets, TCP flags, Port, IPv4 address	***Mode of Operation:*** *MIX* ***Temporal:*** *12-2022, 4-2024* ***Spatial:*** *UK, USA* ***Sample Rate:*** *full sample 1:1*

*Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai, "A novel approach for detecting vulnerable IoT devices connected behind a home NAT," Computers & Security, 2020.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

Paper	*Setup*	**Training**
Meid20*	**Model:** LightGBM **Features:** ibytes, ipackets, TCP flags, Port, IPv4 address	***Mode of Operation:*** *MIX* ***Temporal:*** *12-2022, 4-2024* ***Spatial:*** *UK, USA* ***Sample Rate:*** *full sample 1:1*

*Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai, "A novel approach for detecting vulnerable IoT devices connected behind a home NAT," Computers & Security, 2020.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

Paper	*Setup*	**Training**	*Testing*
Meid20*	**Model:** LightGBM **Features:** ibytes, ipackets, TCP flags, Port, IPv4 address	***Mode of Operation:*** *MIX* ***Temporal:*** *12-2022, 4-2024* ***Spatial:*** *UK, USA* ***Sample Rate:*** *full sample 1:1*	**Scenario1:** *Mix vs idle,* **Scenario2:** *Mix vs active* **Scenario3:** *Mix vs Mix*

*Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai, "A novel approach for detecting vulnerable IoT devices connected behind a home NAT," Computers & Security, 2020.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

Paper	*Setup*	**Training**	*Testing*
Meid20*	**Model:** LightGBM **Features:** ibytes, ipackets, TCP flags, Port, IPv4 address	***Mode of Operation:*** *MIX* ***Temporal:*** *12-2022, 4-2024* ***Spatial:*** *UK, USA* ***Sample Rate:*** *full sample 1:1*	**Scenario1:** *Mix vs idle,* **Scenario2:** *Mix vs active* **Scenario3:** *Mix vs Mix*

*Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai, "A novel approach for detecting vulnerable IoT devices connected behind a home NAT," Computers & Security, 2020.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

Paper	*Setup*	**Training**	*Testing*	*Results (AUCPR)*
Meid20*	**Model:** LightGBM **Features:** ibytes, ipackets, TCP flags, Port, IPv4 address	***Mode of Operation:*** *MIX* ***Temporal:*** *12-2022, 4-2024* ***Spatial:*** *UK, USA* ***Sample Rate:*** *full sample 1:1*	**Scenario1:** *Mix vs idle,* **Scenario2:** *Mix vs active* **Scenario3:** *Mix vs Mix*	**Scenario1:** *0.54* **Scenario2:** *0.57* **Scenario3:** 0.78

*Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai, "A novel approach for detecting vulnerable IoT devices connected behind a home NAT," Computers & Security, 2020.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

Paper	*Setup*	**Training**	*Testing*	*Results (AUCPR)*
Meid20*	**Model:** LightGBM **Features:** ibytes, ipackets, TCP flags, Port, IPv4 address	***Mode of Operation:*** *MIX* ***Temporal:*** *12-2022, 4-2024* ***Spatial:*** *UK, USA* ***Sample Rate:*** *full sample 1:1*	**Scenario1:** *Mix vs idle,* **Scenario2:** *Mix vs active* **Scenario3:** *Mix vs Mix*	**Scenario1:** *0.54* **Scenario2:** *0.57* **Scenario3:** 0.78

*Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai, "A novel approach for detecting vulnerable IoT devices connected behind a home NAT," Computers & Security, 2020.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

Paper	*Setup*	Training	*Testing*	*Results (AUCPR)*
Meid20*	**Model:** LightGBM **Features:** ibytes, ipackets, TCP flags, Port, IPv4 address	***Mode of Operation:*** *MIX* ***Temporal:*** *12-2022, 4-2024* ***Spatial:*** *UK, USA* ***Sample Rate:*** *full sample 1:1*	**Scenario1:** *Mix vs idle,* **Scenario2:** *Mix vs active* **Scenario3:** *Mix vs Mix*	**Scenario1:** *0.54* **Scenario2:** *0.57* **Scenario3:** 0.78

Empirical Observation

*Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai, "A novel approach for detecting vulnerable IoT devices connected behind a home NAT," Computers & Security, 2020.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

Paper	*Setup*	Training	*Testing*	*Results (AUCPR)*
Meid20*	**Model:** LightGBM **Features:** ibytes, ipackets, TCP flags, Port, IPv4 address	***Mode of Operation:*** *MIX* ***Temporal:*** *12-2022, 4-2024* ***Spatial:*** *UK, USA* ***Sample Rate:*** *full sample 1:1*	**Scenario1:** *Mix vs idle,* **Scenario2:** *Mix vs active* **Scenario3:** *Mix vs Mix*	**Scenario1:** *0.54* **Scenario2:** *0.57* **Scenario3:** 0.78

Empirical Observation

30% active	70% Idle

*Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai, "A novel approach for detecting vulnerable IoT devices connected behind a home NAT," Computers & Security, 2020.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

Paper	*Setup*	Training	*Testing*	*Results (AUCPR)*
Meid20*	**Model:** LightGBM **Features:** ibytes, ipackets, TCP flags, Port, IPv4 address	***Mode of Operation:*** *MIX* ***Temporal:*** *12-2022, 4-2024* ***Spatial:*** *UK, USA* ***Sample Rate:*** *full sample 1:1*	**Scenario1:** *Mix vs idle,* **Scenario2:** *Mix vs active* **Scenario3:** *Mix vs Mix*	**Scenario1:** *0.54* **Scenario2:** *0.57* **Scenario3:** 0.78

Empirical Observation

30% active | 70% Idle

Recommendation

*Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai, "A novel approach for detecting vulnerable IoT devices connected behind a home NAT," Computers & Security, 2020.

What are the key findings of the practicality evaluation?

Mode of Operation. Idle and active modes introduce behavioural shifts that reduce performance.

Paper	*Setup*	**Training**	*Testing*	*Results (AUCPR)*
Meid20*	**Model:** LightGBM **Features:** ibytes, ipackets, TCP flags, Port, IPv4 address	***Mode of Operation:*** *MIX* ***Temporal:*** *12-2022, 4-2024* ***Spatial:*** *UK, USA* ***Sample Rate:*** *full sample 1:1*	**Scenario1:** *Mix vs idle,* **Scenario2:** *Mix vs active* **Scenario3:** *Mix vs Mix*	**Scenario1:** *0.54* **Scenario2:** *0.57* **Scenario3:** 0.78

Empirical Observation 30% active | 70% Idle

Recommendation Training the model in idle mode and then conducting predictions for different periods.

*Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, and A. Shabtai, "A novel approach for detecting vulnerable IoT devices connected behind a home NAT," Computers & Security, 2020.

How can the practicality of ML-based IoT device identification be improved?

Components of ML-based Model

IoT Identification Problem

1. **Dataset** — What data should be collected, and how should the dataset be gathered for evaluation?
2. **Model** — Which model should be used, and how complex should it be?
3. **Features** — What features should be extracted, and how should they be represented?

How can the practicality of ML-based IoT device identification be improved?

Components of ML-based Model

IoT Identification Problem

1. **Dataset** — Privacy and real-world challenges
2. **Model** — Which model should be used, and how complex should it be?
3. **Features** — What features should be extracted, and how should they be represented?

How can the practicality of ML-based IoT device identification be improved?

Components of ML-based Model

IoT Identification Problem

1. **Dataset** — Privacy and real-world challenges
2. **Model** — Mode complexities , and how complex
3. **Features** — What features should be extracted, and how should they be represented?

How can the practicality of ML-based IoT device identification be improved?

Components of ML-based Model

IoT Identification Problem

1 Dataset — Privacy and real-world challenges

2 Model — Mode complexities , and how complex

3 Features — Explainable AI and feature importance

How can the practicality of ML-based IoT device identification be improved?

How can the practicality of ML-based IoT device identification be improved?

Figure.1: Meid20, LightGBM, incoming bytes, incoming packets, TCP flags, Port, IPv4 add , idle

How can the practicality of ML-based IoT device identification be improved?

t-SNE

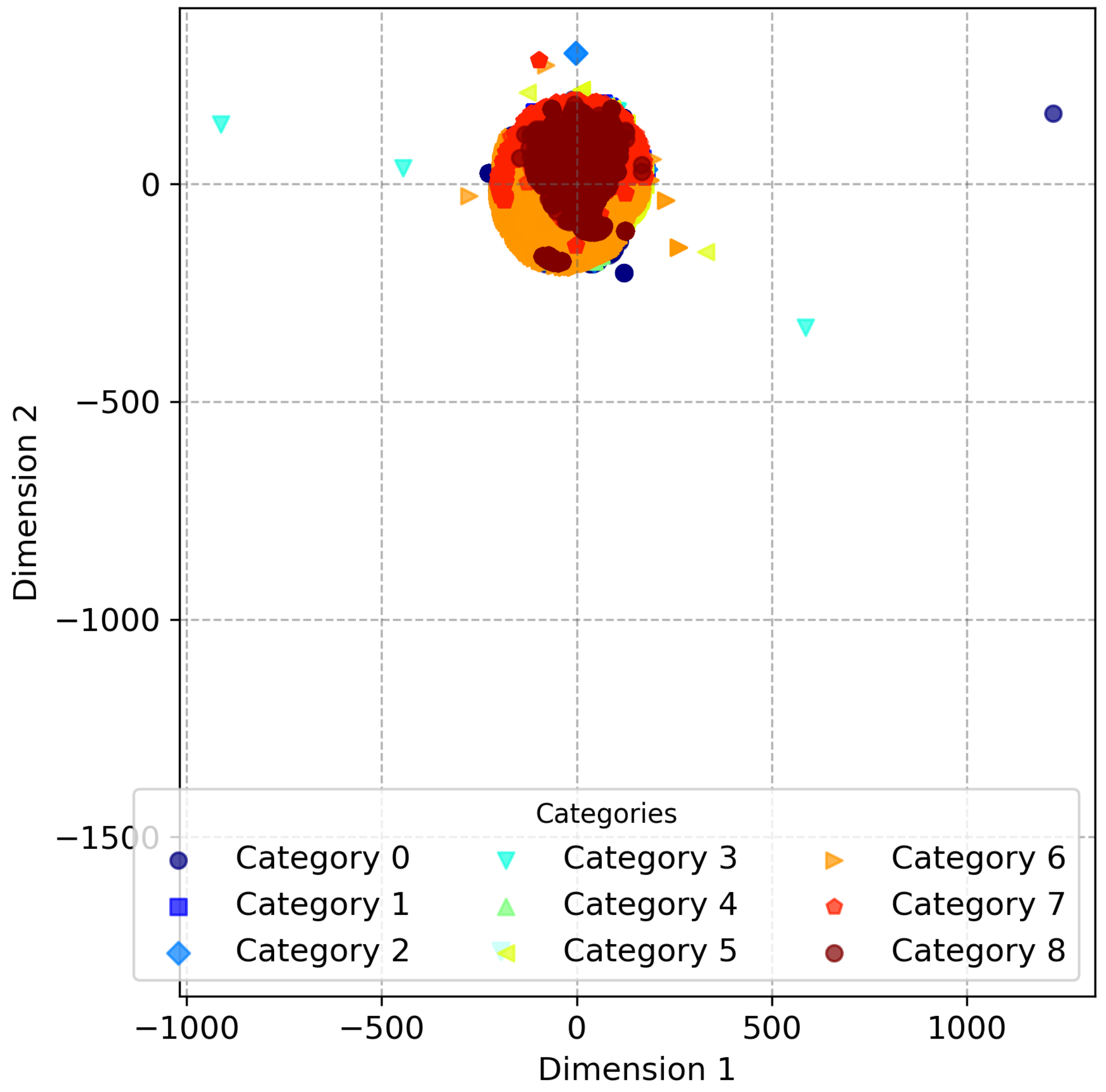
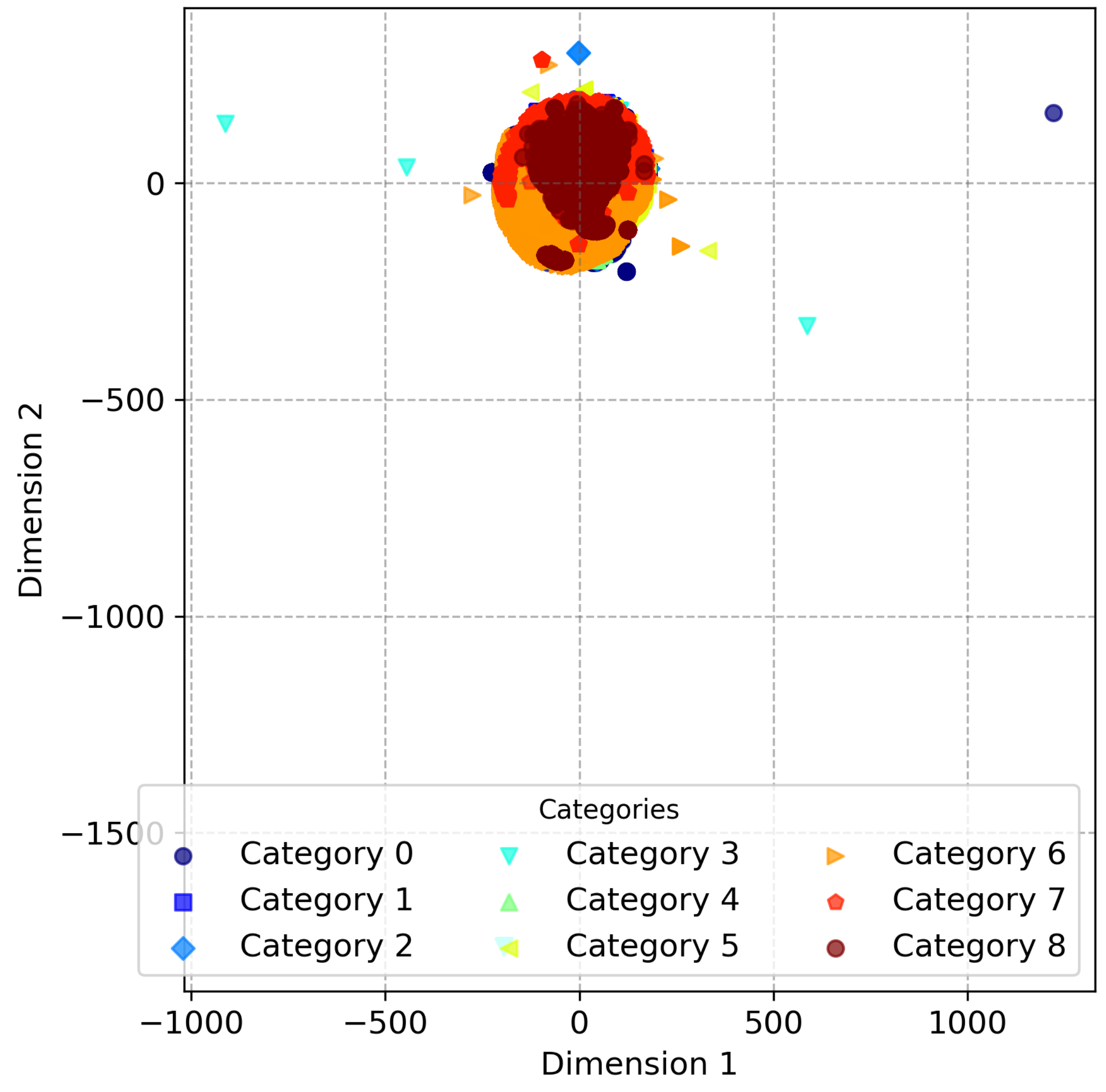


Figure.1: Meid20, LightGBM, incoming bytes, incoming packets, TCP flags, Port, IPv4 add , idle

How can the practicality of ML-based IoT device identification be improved?

t-SNE



SHAP

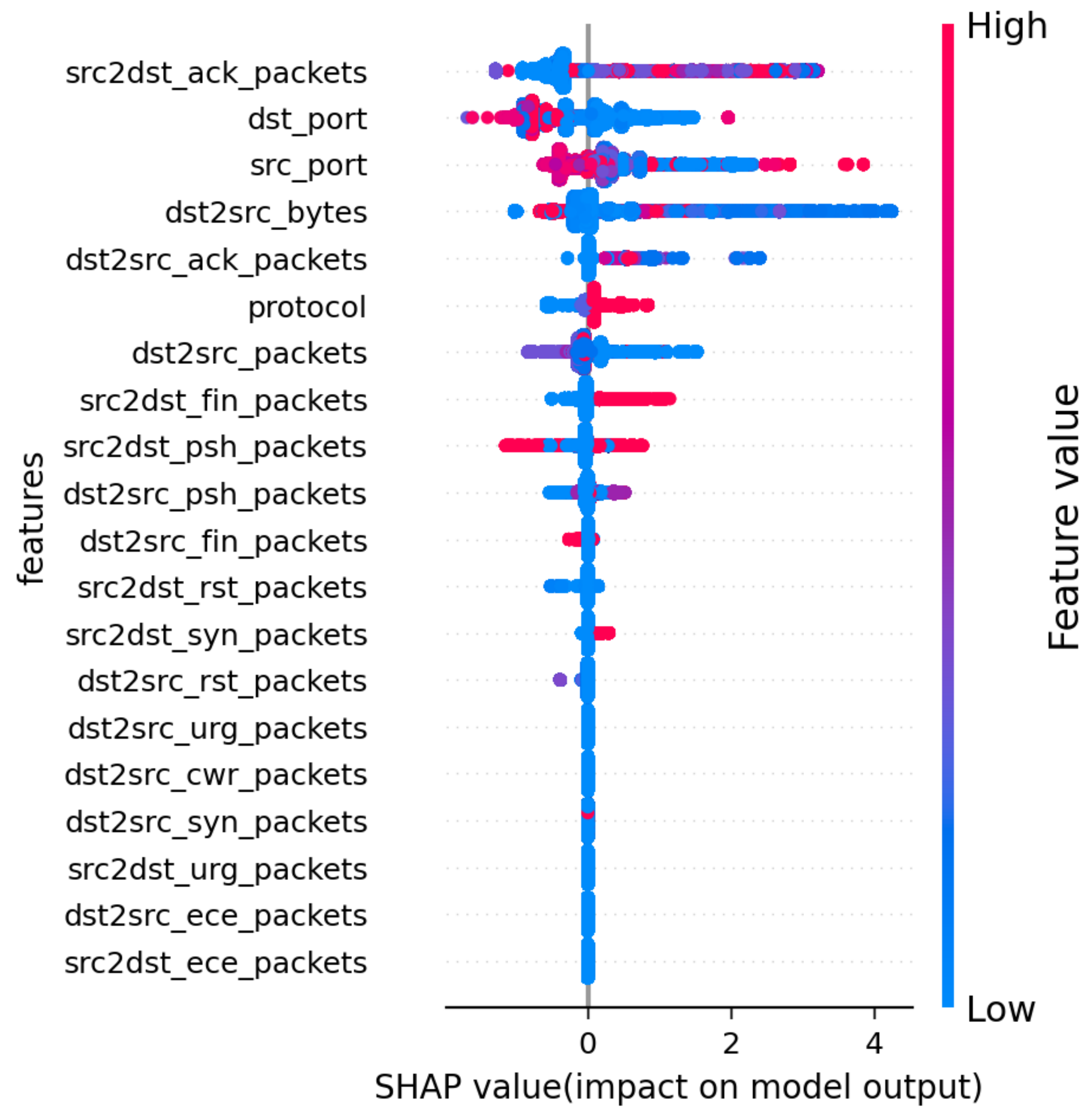
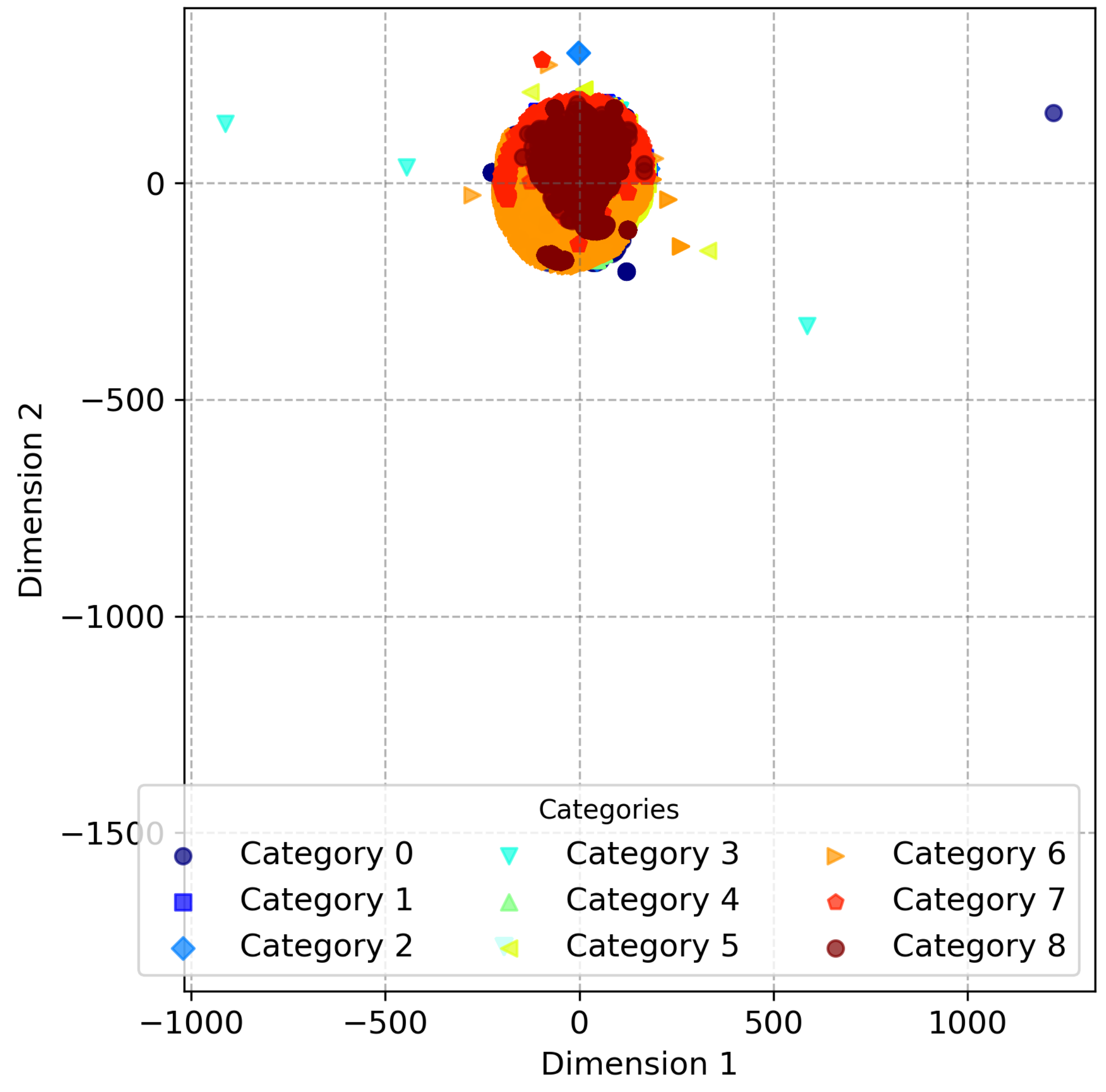


Figure.1: Meid20, LightGBM, incoming bytes, incoming packets, TCP flags, Port, IPv4 add , idle

How can the practicality of ML-based IoT device identification be improved?

t-SNE



SHAP

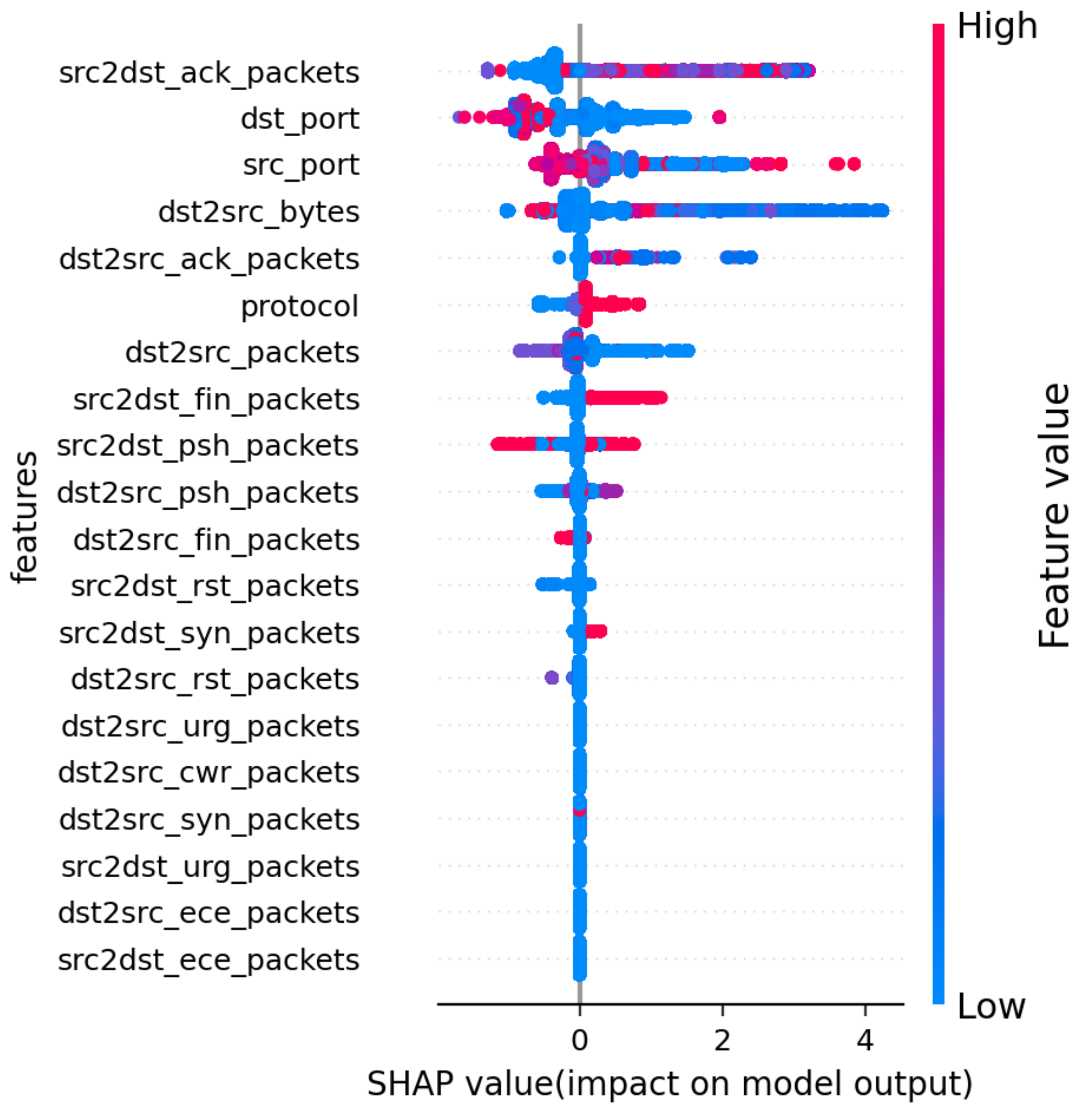
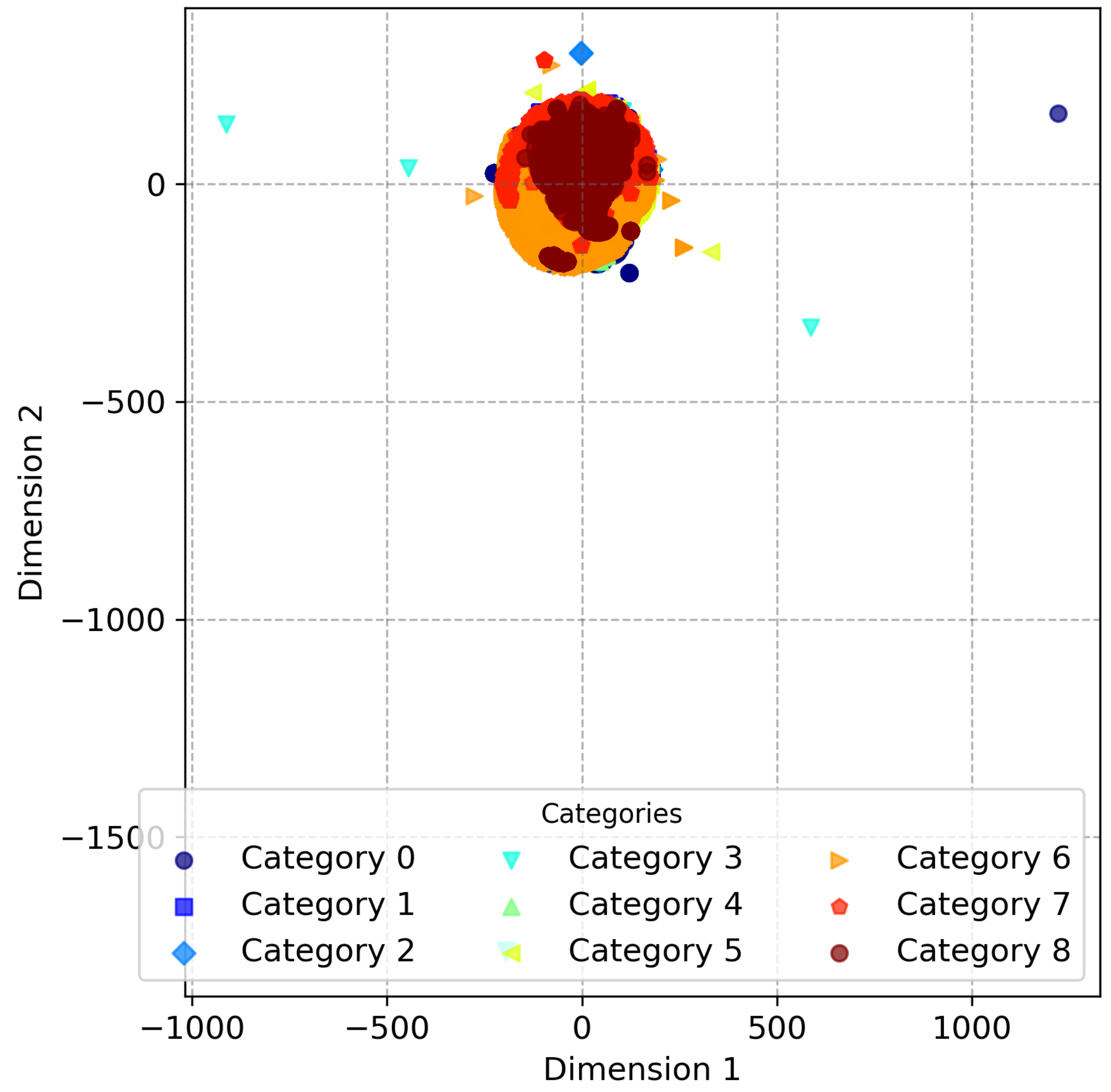


Figure.1: Meid20, LightGBM, incoming bytes, incoming packets, TCP flags, Port, IPv4 add , idle

How can the practicality of ML-based IoT device identification be improved?

t-SNE



SHAP

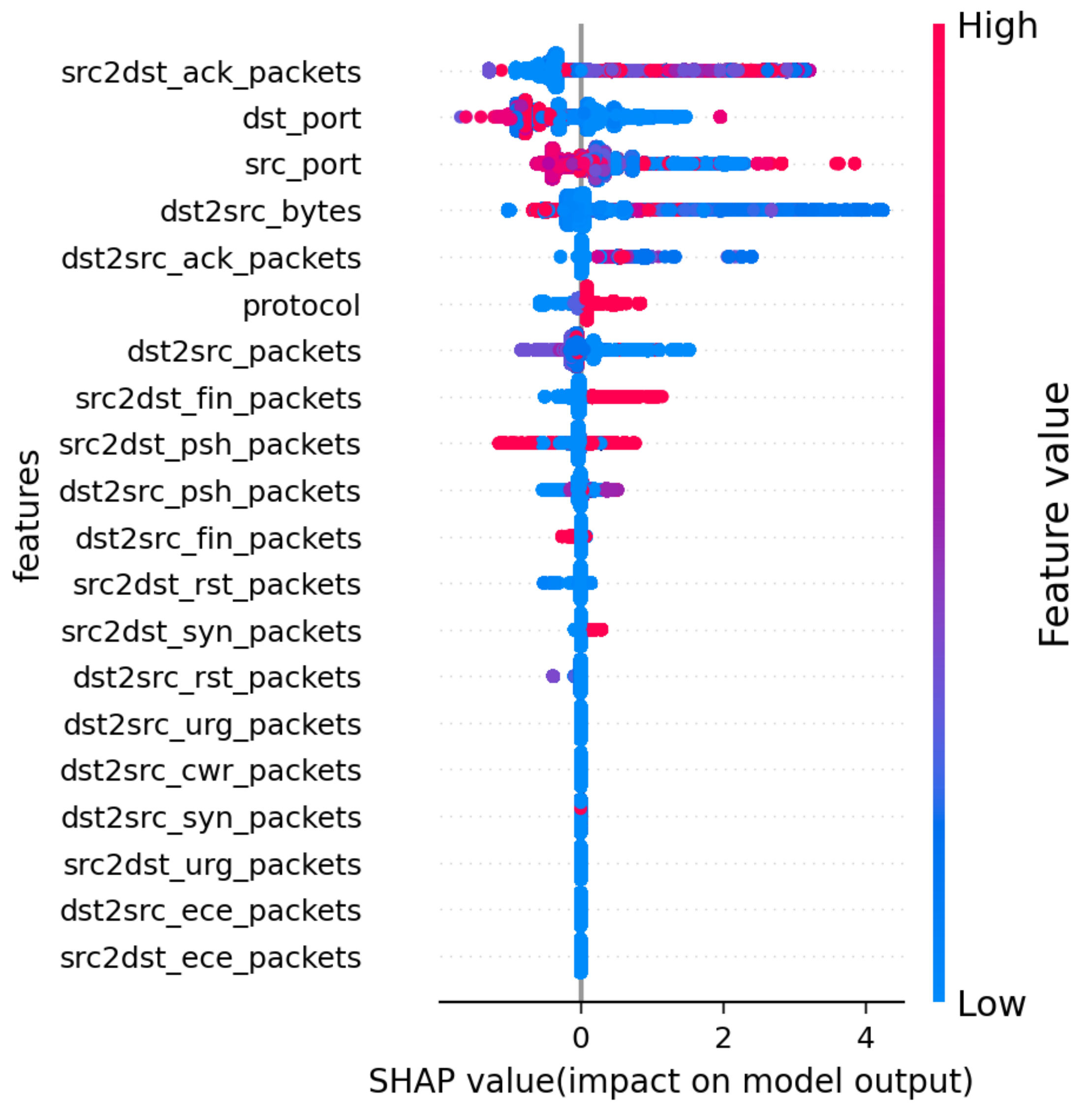
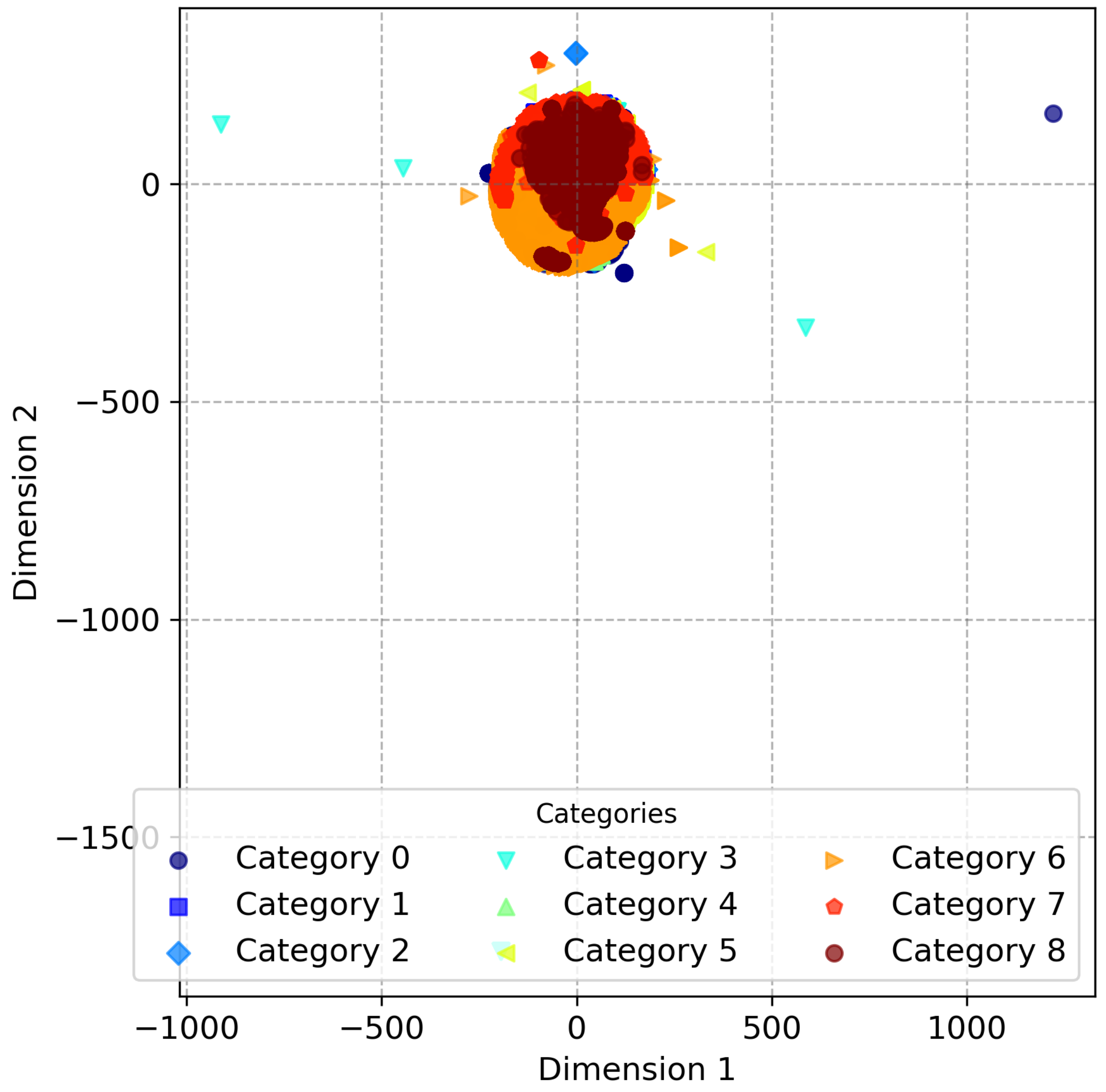


Figure.1: Meid20, LightGBM, incoming bytes, incoming packets, TCP flags, Port, IPv4 add , idle

Key Findings: Simple features used in the training datasets (e.g., mean, variance) fail to capture distributional characteristics effectively.

How can the practicality of ML-based IoT device identification be improved?

t-SNE



SHAP

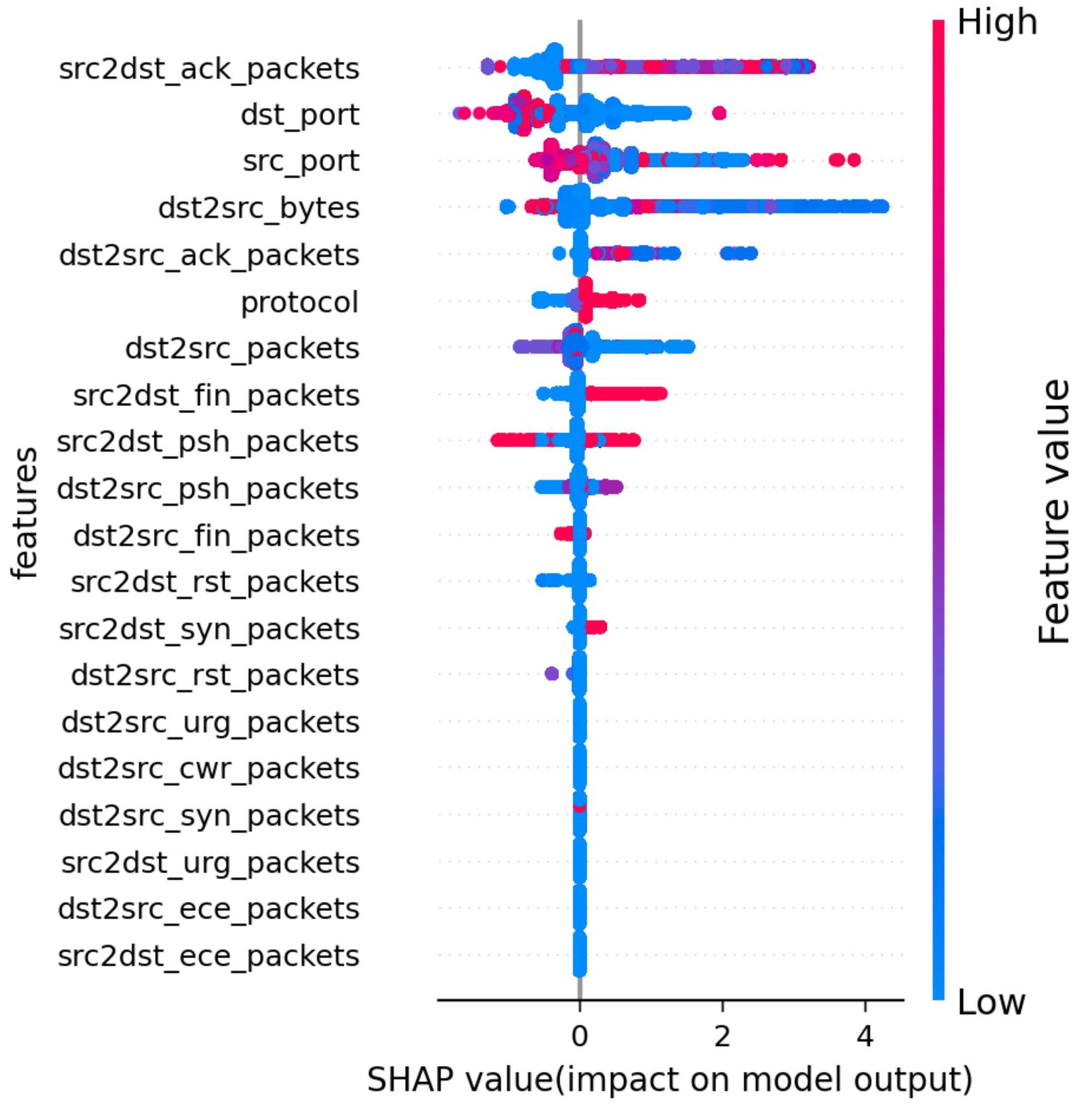


Figure.1: Meid20, LightGBM, incoming bytes, incoming packets, TCP flags, Port, IPv4 add , idle

Key Findings: Simple features used in the training datasets (e.g., mean, variance) fail to capture distributional characteristics effectively.

Recommendation: Avoid simple first-order statistical features (eg., mean, variance), and instead, features such as entropy are more suitable.

Thank You!

Evaluating Machine Learning-Based IoT Device Identification Models for Security Applications

Eman Maali, Omar Alrawi, Julie McCann
e.maali19@imperial.ac.uk

<code>