# Passban IDS: An Intelligent Anomaly-Based Intrusion Detection System for IoT Edge Devices

Mojtaba Eskandari, Zaffar Haider Janjua, Massimo Vecchio, and Fabio Antonelli
OpenIoT Research Area, FBK CREATE-NET, 38123 Trento, Italy

IEEE Internet Of Things, 2020

# Outline

- Introduction
- Passban IDS
  - Packet flow discovery
  - Feature extraction
  - Training and prediction
  - Action & Web manager
- Evaluation
- Conclusion

# Introduction

- Internet of Things (IoT) is a constantly evolving umbrella of technologies aiming at connecting diverse devices and everyday objects

- Embracing such a paradigm shift in our daily lives increases the risk of data privacy breaches and cyber-security attacks

➔ Various IDSs have been suggested about IoT

# Major approaches for IDS

**Signature based approach**

- Identify attack using pattern (signature)
- Can only detect already-known attacks
- Attacks should have characteristics
- Increases of attack types → Increases of signatures ➔ low performance
- Human experts are needed to study, analyze, and craft signatures

**Anomaly based approach**

- Attacks are identified by ML trained by benign traffic
- Can address limitations of signature based approach
- One data source can make a mixture of underlying varying behavior → Hard to model

# IDS for IoT

- Signature based IDS is very hard to efficiently deployed
  - Unknown attacks cannot be detected
  - Various types of new attacks are introduced for IoT environment
  - IoT gateway is usually low-cost → Update for new signatures is difficult

➔ Anomaly based IDS is more suitable for IoT

# Goals of Passban IDS

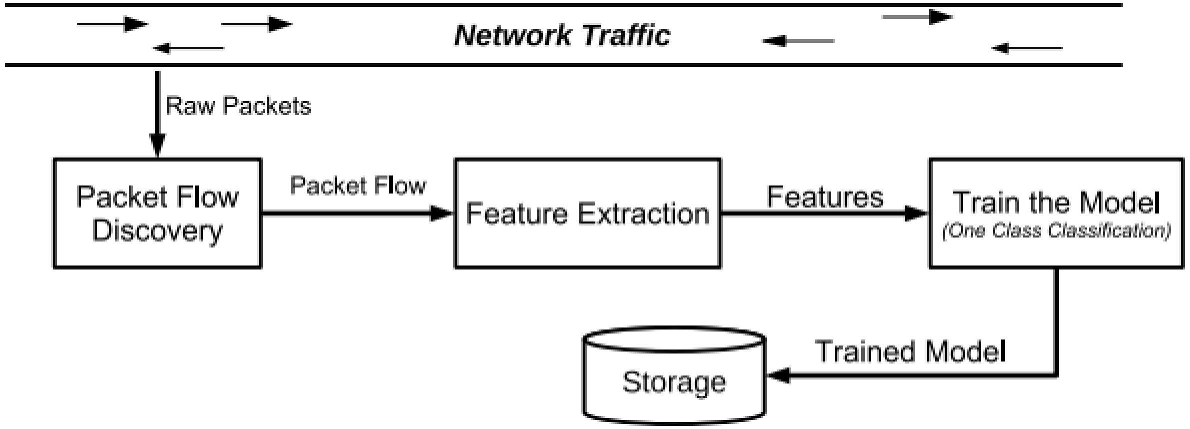
- Ensure data protection near the IoT data sources
- Scalability (in terms of new threats)
- Reduce FP for satisfying detection accuracy requirement

# Contributions

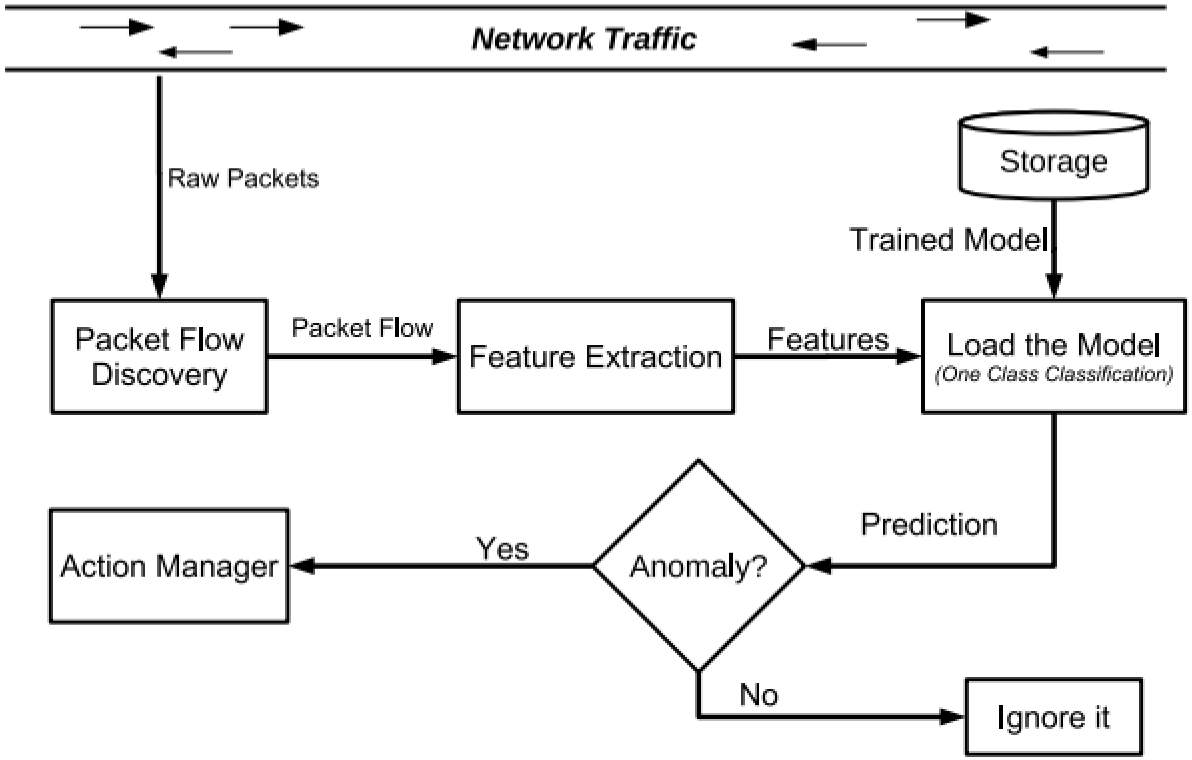
- Suggest a platform-independent anomaly based IDS (Passban) working on edge devices
- Implement Passban in AGILE framework
- Deploy real IoT testbed, collect dataset, and evaluate Passban
- Pack Passban into a Docker container for public

# Passban IDS

# Overview




(a)

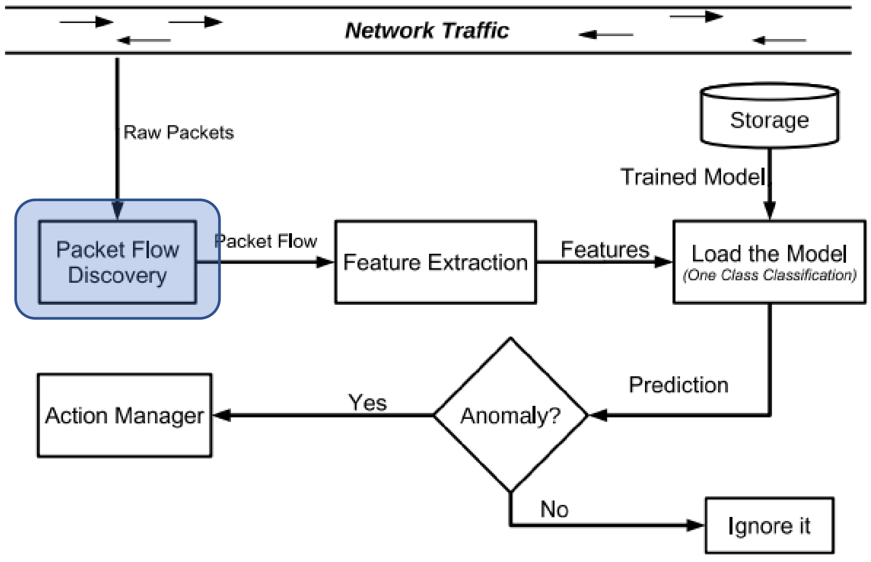



(b)

- Training phase (a)
  - Packet capture → Feature extraction → Train model → Save
- Prediction phase (b)
  - (Load model) → Packet capture → Feature extraction → Prediction → Action

# Packet flow discovery

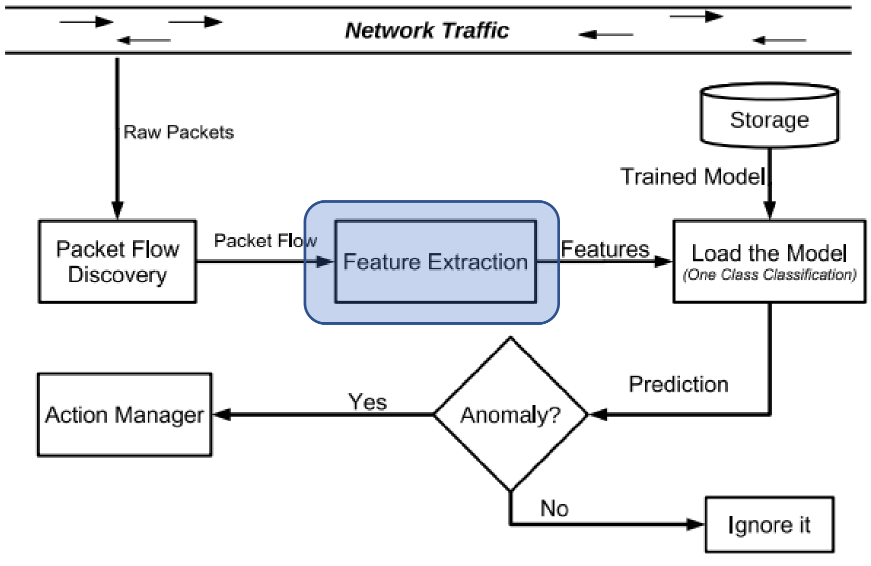
- Constantly observe network traffic
- Capture network raw packets
- Send them to feature extraction block



(b)

# Feature extraction

- Calculate network flow statistics
- Build features to feed train/predict block



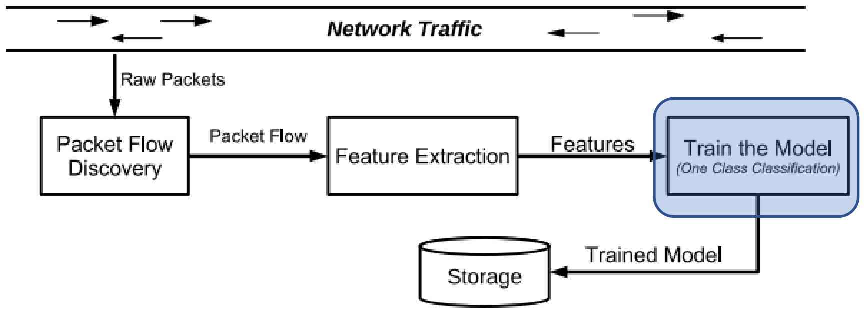
(b)

| Type | Features | Descriptions |
|---|---|---|
| Traffic volume | 12 features: max_fpktl, max_bpktl, mean, min, sflow, etc.. | Size of largest packet (in forward/backword), mean/min packet size, number of bytes, etc.. |
| Packet statistics | 4 features: sflow_fpackets, sflow_bpackets, total_fpackets, total_bpackets | Average number of packets, total packets |
| Time statistics | 8 features: mean_active, mean_fiat, max, min, duration | Mean active time, mean time interval between two packets in forward, etc.. |

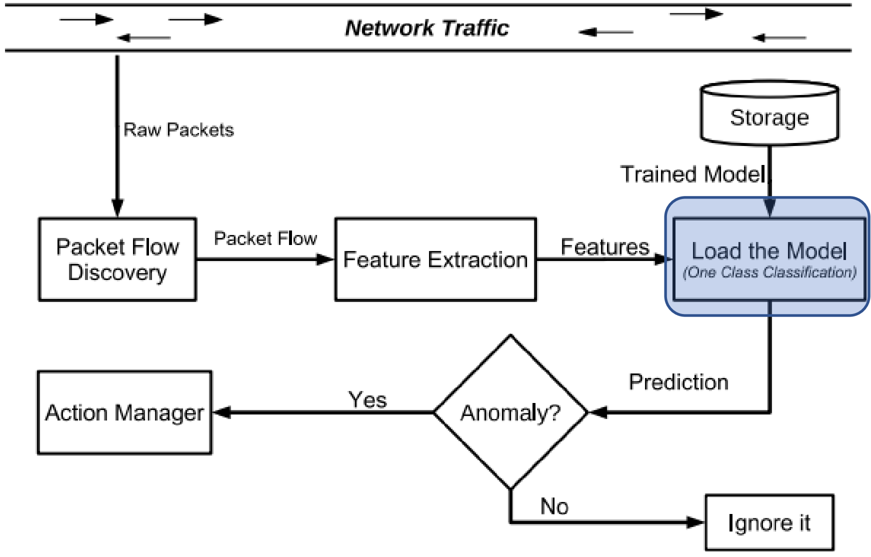
# Train/prediction

- During training phase
  - ML algorithm is trained to learn normality traffic
  - Trained model is stored in the local memory

- During prediction phase
  - Model is loaded from local storage
  - Predict captured flow as “anomaly” or “benign”
  - Anomalies are sent to Action manager




(a)




(b)

# ML in Passban

- Supervised learning is hard to be applied for Passban

**Supervised learning needs well-defined labeled data** — **Lack of well-defined labeled data in IoT**

**Supervised learning can identify trained classes** — **High frequency of emergence of new attacks in IoT**

**Unsupervised learning is more suitable for Passban**

# ML in Passban: Isolation Forest

- Two unsupervised learning algorithm is used
- Isolation Forest (iForest)*
  - Anomalies are few and characterized by attribute values which are quite different from normal
  - Generate forest of data induced random trees
  - Each tree is built by recursively partitioning the instances until all the instances are isolated

  → The instances having anomalies are represented by shorter paths in the tree



* F. T. Liu, et. Al., "Isolation-based anomaly detection," ACM Trans. Knowl. Discov. Data, vol. 6, no. 1, pp. 1–39, 2012.

# ML in Passban: Local Outlier Factor

- Local Outlier Factor (LOF)*
  - Density-based method for identifying outliers
  - Density estimation is based on a comparison between distances measured of a point with its k-nearest neighbors

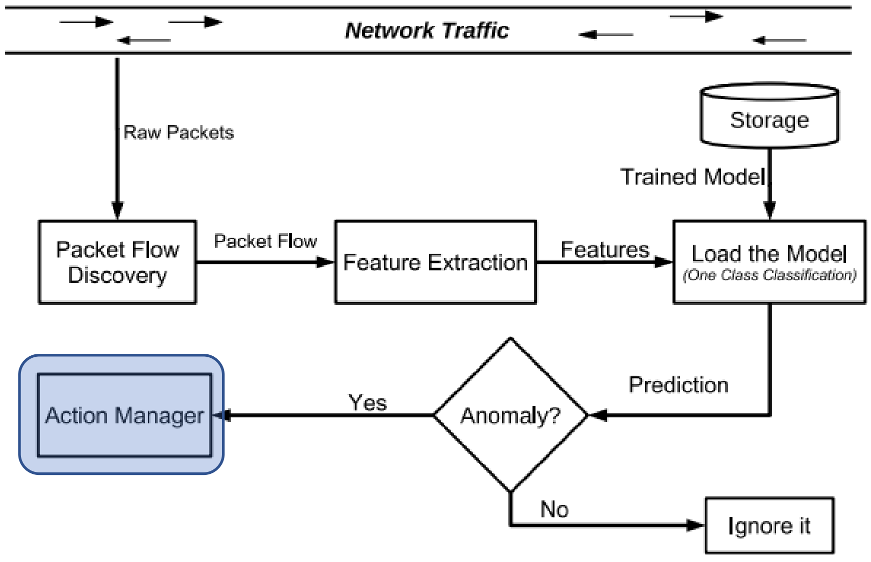
  → Data points belonging to denser regions having similar density are considered normal

  → Data points occurring in the lower density regions which are considered outliers



* M. M. Breunig, et. Al., "LOF: Identifying density-based local outliers," SIGMOD Rec., vol. 29, no. 2, pp. 93–104, May 2000

# Action manager

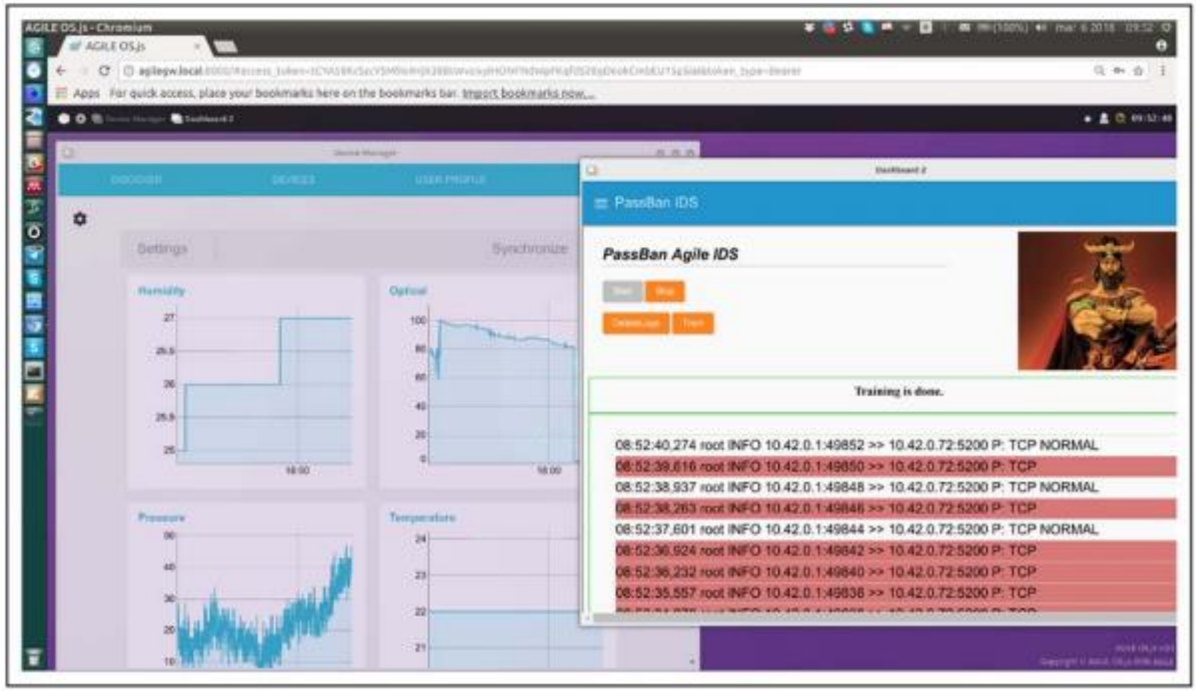


(b)

- Take proper actions to traffic predicted as "anomaly" by prediction

- Several actions are defined
  - Log details about packets
  - Block the flow
  - Send notification to network administrator
  - Switch off critical devices

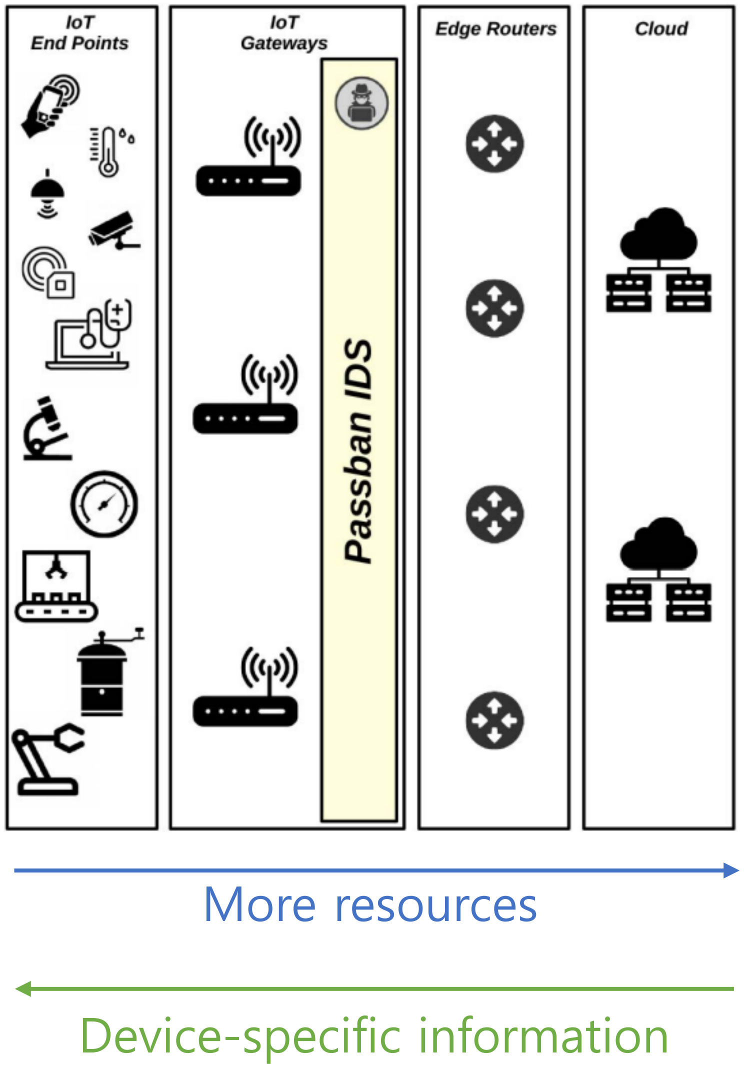
# Web manager




- User interface for network administrator
- Functions
  - Show status of the IDS
  - Start/stop IDS
  - Change phases (training/prediction)
  - Manage logs of anomalies

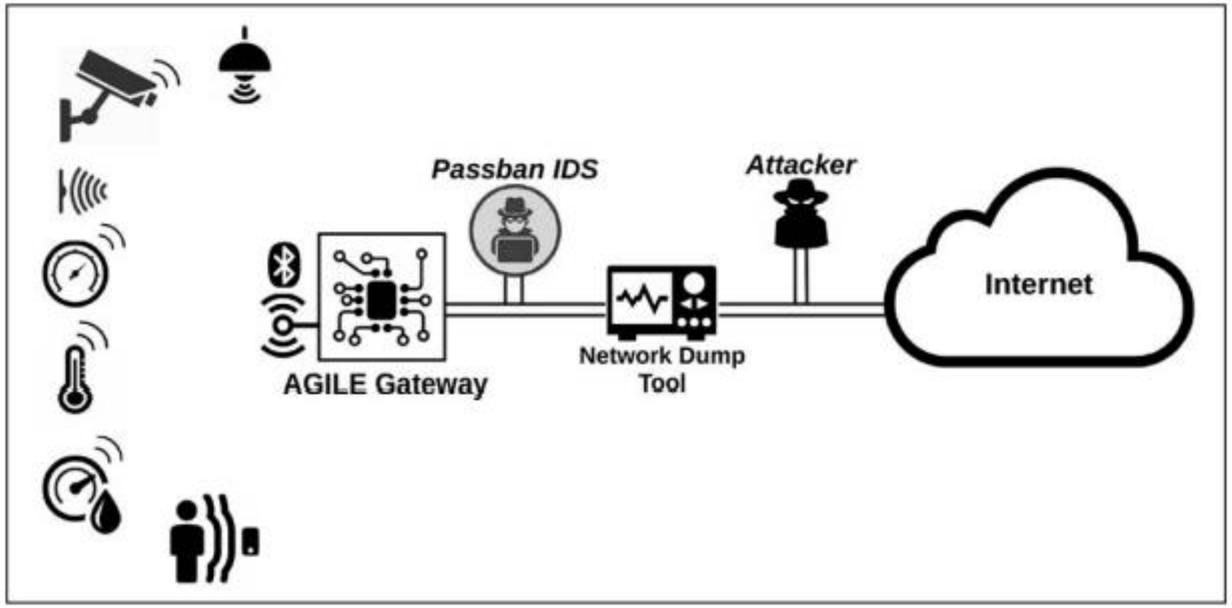
# Analysis about Passban

- Passban gets advantage from 'locality'
  - Aggregation of multiple streams occurs at each level; IoT gateway, edge router, cloud
  - High levels may exhibit more generic characteristics, rather than device-specific characteristics
    - May reduce the performance of an IDS when detecting threats
- Passban limitation
  - 'Benign phase' is necessary, false positive, network change leads to new training phase, resource exhaust due to 'SYN flood'



# Evaluation

# Testbed setup



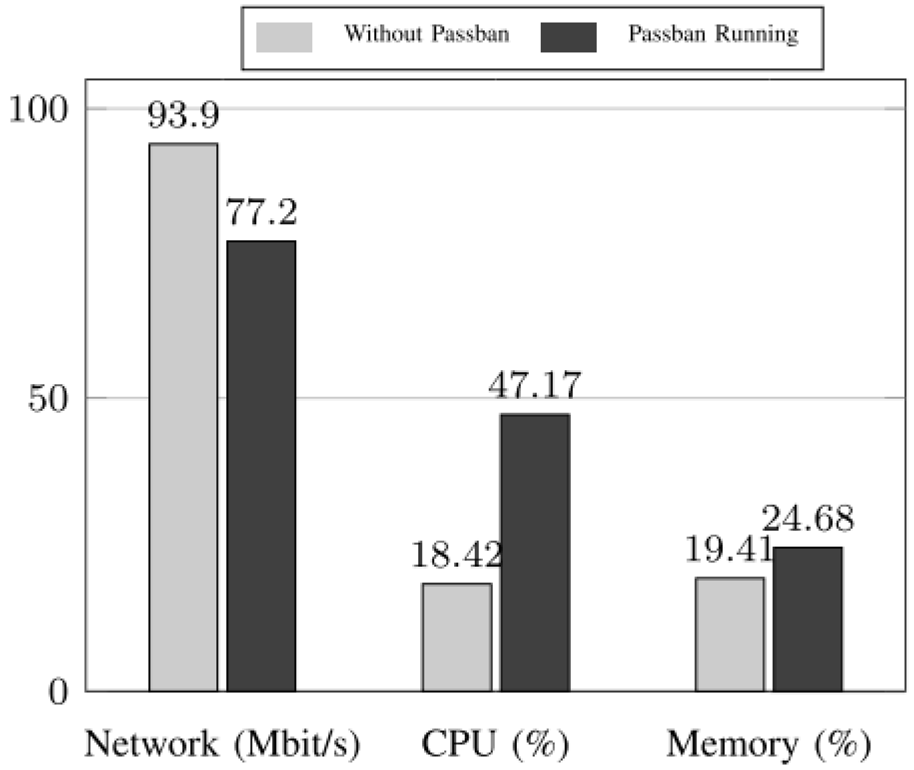

- IoT devices
  - Texas instrument BLE SensorTag endowed with
    - a) TMP007: Temperature sensor
    - b) BMP280: Altimeter/Air pressure sensor
    - c) OPT3001: Ambient light sensor
    - d) DHC1000: Humidity sensor
    - e) MPU-9250: 9-axis motion sensor
  - FosCam FI8910W as WiFi IP Camera

# Attacks towards IoT devices

- 4 types of attacks are introduced
  - Port Scanning
    - Enables reconnaissance on the target system to discover possible vulnerable points
  - HTTP Brute Force
    - Almost every IoT gateway provides a Web interface to interact with various IoT devices
    - Web interface is usually protected via a pair of username/password credentials
  - SSH Brute Force
    - SSH protocol is usually used by a system administrator to communicate with the gateway
  - SYN Flood
    - Try to consume enough server resources in order to make the system unresponsive to legitimate traffic
    - Especially harmful to IoT gateways

# Resource utilization

- Memory usage
  - 24.68% when Passban is executing
  - 19.41% when it is not executing
  - Passban requires 54 MB
- Average CPU load
  - 47.17% when Passban is executing
  - 18.42% when it is not executing
- Network throughput
  - Raspberry Pi can handle max 93.9 Mb/s
  - With Passban, this bandwidth is reduced to 77.2 Mb/s



# Performance evaluation

| *Attack* | *Technique* | *#Normal* | *#Attack* | *FP* | *TP* | *FN* | *TN* | *Precision* | *Recall* | *F1* |
|---|---|---|---|---|---|---|---|---|---|---|
| Port Scanning | iForest | 148 | 57 | 1 | 57 | 0 | 147 | **0.98** | **1** | **0.99** |
| | LOF | 148 | 57 | 10 | 52 | 5 | 138 | 0.84 | 0.91 | 0.87 |
| HTTP Brute Force | iForest | 106 | 36 | 2 | 35 | 1 | 104 | **0.95** | **0.97** | **0.96** |
| | LOF | 106 | 36 | 7 | 35 | 1 | 99 | 0.83 | **0.97** | 0.89 |
| SSH Brute Force | iForest | 870 | 389 | 9 | 370 | 19 | 861 | **0.98** | **0.95** | **0.96** |
| | LOF | 870 | 389 | 7 | 302 | 87 | 863 | **0.98** | 0.78 | 0.87 |
| SYN Flood | iForest | 117 | 31 | 2 | 27 | 4 | 115 | **0.93** | **0.87** | **0.9** |
| | LOF | 117 | 31 | 5 | 27 | 4 | 112 | 0.84 | 0.87 | 0.85 |

- LOF and iForest are able to detect all the tested attacks with satisfactory accuracies
  - iForest reaches always the best values in terms of both precision and recall, hence also in terms of F1: 0.99, 0.96, 0.96, and 0.90 for Port Scanning, HTTP Brute Force, SSH Brute Force, and SYN Flood, respectively

# Conclusion

- Authors presented Passban, an intelligent anomalybased IDS purposely designed to be directly hosted and executed by a typical edge device

- Authors built an IoT testbed able to resemble a typical smart home automation environment

- Passban is evaluated against four common attacks (namely, port scanning, HTTP brute force, SSH brute force, and SYN flood)