Anomaly Detection in the Internet of Medical Things: Design and Evaluation of a Cross Layer Dataset

University of Beira Interior
Instituto de Telecomunicações

Rui P. Pinto | Bruno M. C. Silva | Pedro R. M. Inácio

Outline

- **Introduction**
- **Related Work**
- **Dataset Design and Methodology**
  - IoMT Topology
  - Normal & Anomalous Traffic Profiling
  - Traffic Capture Process
  - Dataset Structure and Label Assignment
- **Experimental Setup and Evaluation**
  - Machine Learning Evaluation Setup
  - Classification Tasks (Binary, Layer-Wise, Multiclass)
  - Experimental Results
- **Conclusions & Future Work**
  - Conclusions
  - Future Work

Introduction

IoMT

- The **Internet of Medical Things (IoMT)** integrates smart medical devices, sensors, and cloud platforms to support remote monitoring and personalized treatment.
- It enables **real-time data transmission** and **decision-making**, enhancing care for elderly, chronically ill, and mobility-restricted patients.

Introduction

Security & Reliability Challenges

- While offering transformative potential, **IoMT** systems are **vulnerable to cyberattacks, device malfunctions, and data anomalies**.

- **Compromised data** can delay **diagnoses** and **treatment**, potentially **harming patients**.

Introduction

Anomaly Detection

- Effective **anomaly detection** is **critical** for ensuring **data integrity** and **system resilience**;

- It can identify **malicious intrusions**, **device failures**, and **unusual health data patterns** before they escalate;

- However, the availability of **datasets to train or test ML models remains limited** and **lacking diversity in anomaly types**.

Introduction

Contributions

- Design and present a realistic **multi-layer IoMT anomaly detection dataset**, showing construction and characteristics (**X-IoMT**);
- Conduct a **evaluation across multiple ML tasks** using representative models;
- Discuss the **challenges** and **opportunities** of the dataset.

Related Works

Limitations of Current Datasets:

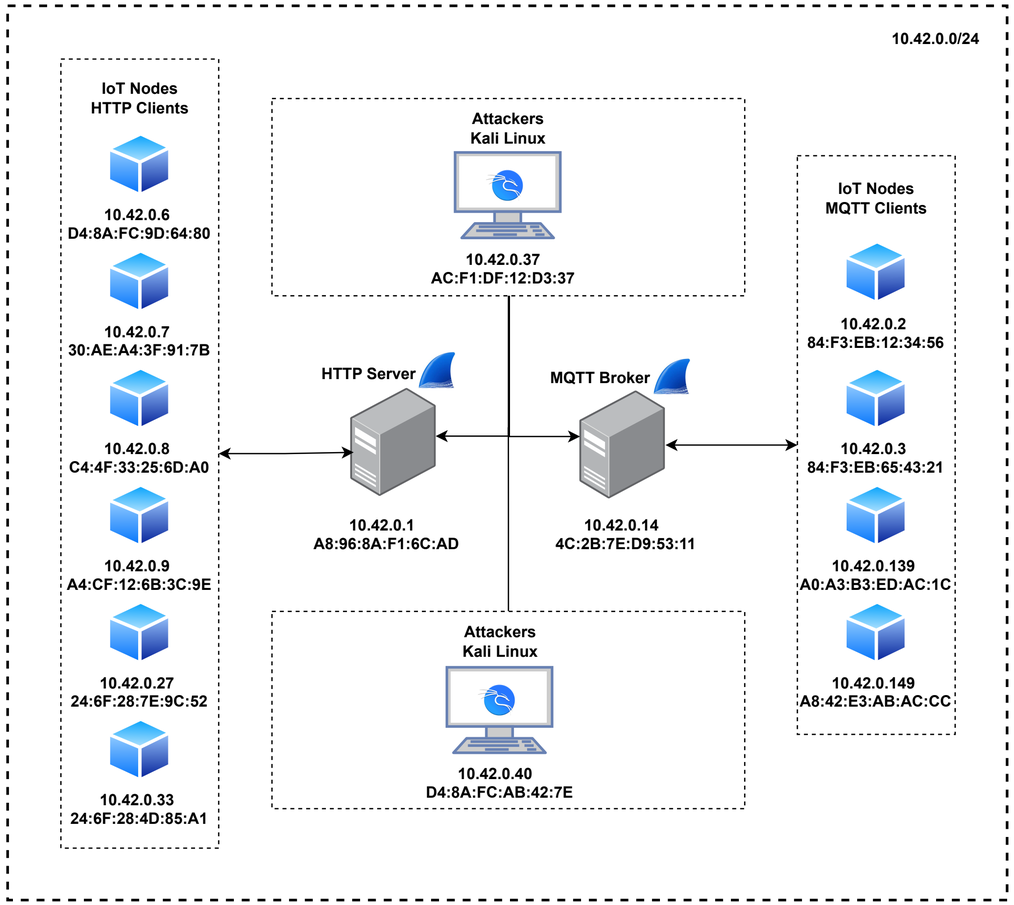
- Most **existing** IoMT datasets use **synthetic data** or **focus** narrowly on **specific anomaly** types;
- They often **lack multi-layer context**, **reducing** the **realism and generalization** capabilities.

Advantages of the X-IoMT Dataset:

- X-IoMT is built from real **IoT devices** and **spans perception, network, and application layers**;
- It incorporates **device failures**, **cyberattacks**, and **clinical anomalies**, providing a richer, multi-dimensional dataset.

Dataset Design and Methodology

IoMT Topology



Topology of the experimental IoMT environment used to create the X-IoMT dataset, showing the deployed devices, their network segregation, and communication flows established via Wi-Fi using HTTP and MQTT protocols.

Dataset Design and Methodology

Normal & Anomalous Traffic Profiling - Perception Layer

Feature	Value Range
CPU	160 to 220 megahertz(MHz)
RSSI	-95 to -51 decibel-milliwatts (dBm)
Internal Temperature	51.7 to 90.°C

Value ranges of selected features used for detecting perception-layer anomalies. Anomalous values were simulated, whereas normal ranges were obtained from the regular operation of the devices.

Dataset Design and Methodology

Normal & Anomalous Traffic Profiling - Network Layer

Target	Attack/Label	Description/Objective
HTTP Server	SYN_Flood	Floods to exhaust resources
HTTP Server	FIN_Flood	Floods to exhaust resources
HTTP Server	Slowloris	Slow HTTP request to keep connection
MQTT Broker	MQTTBruteForce	Attempt to crack credentials
MQTT Broker	MQTT_DoS	Malformed and excessive requests
HTTP & MQTT	ReconAttack	Scans to gather information

Types of Network-layer attacks used for anomaly generation in X-IoMT.

Dataset Design and Methodology

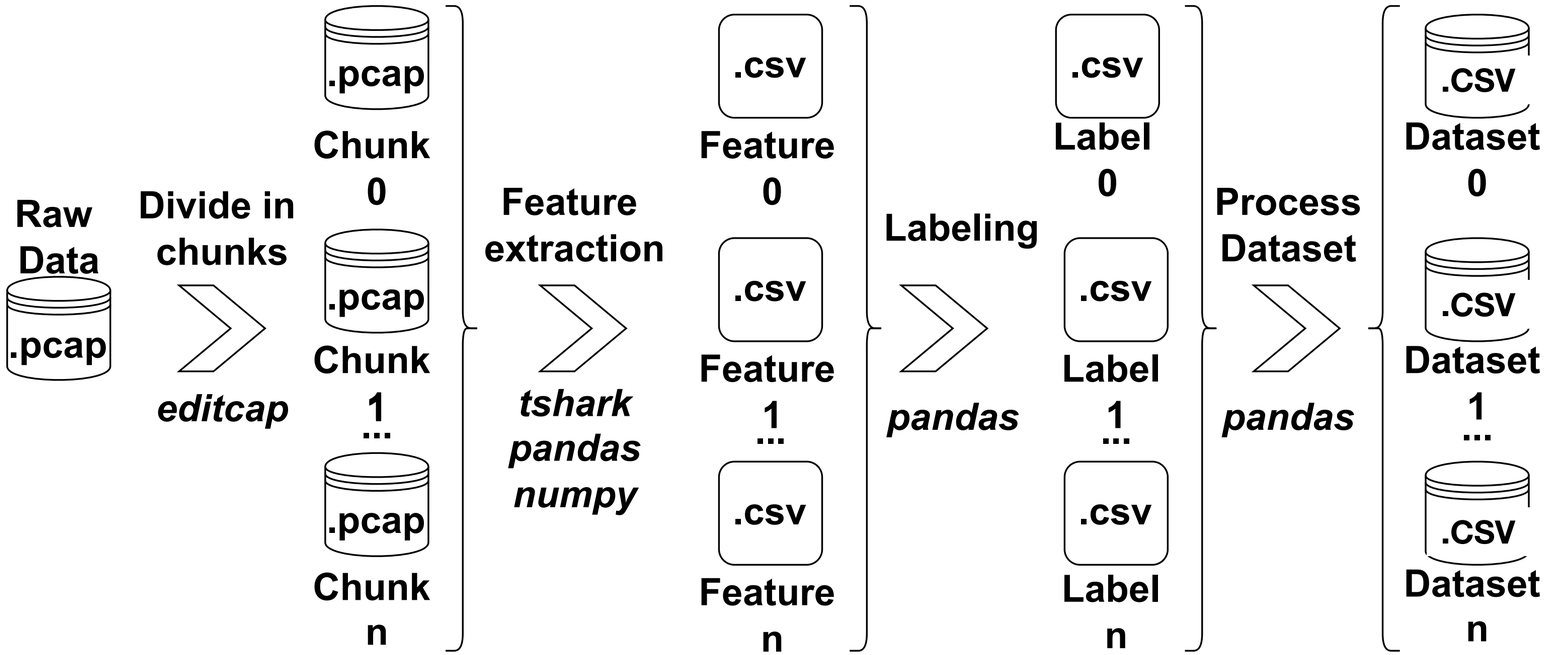
Normal & Anomalous Traffic Profiling - Application Layer

Sensor/ Feature	Normal Range	Anomalous Values	Objective
Body Temperature	35.7 to 37.2.ºC	< 35.4º, > 38ºC	Verify if clinical data anomalies can be detected

Application-layer anomaly simulated in the X-IoMT dataset using body temperature data

Dataset Design and Methodology

Traffic Capture Process

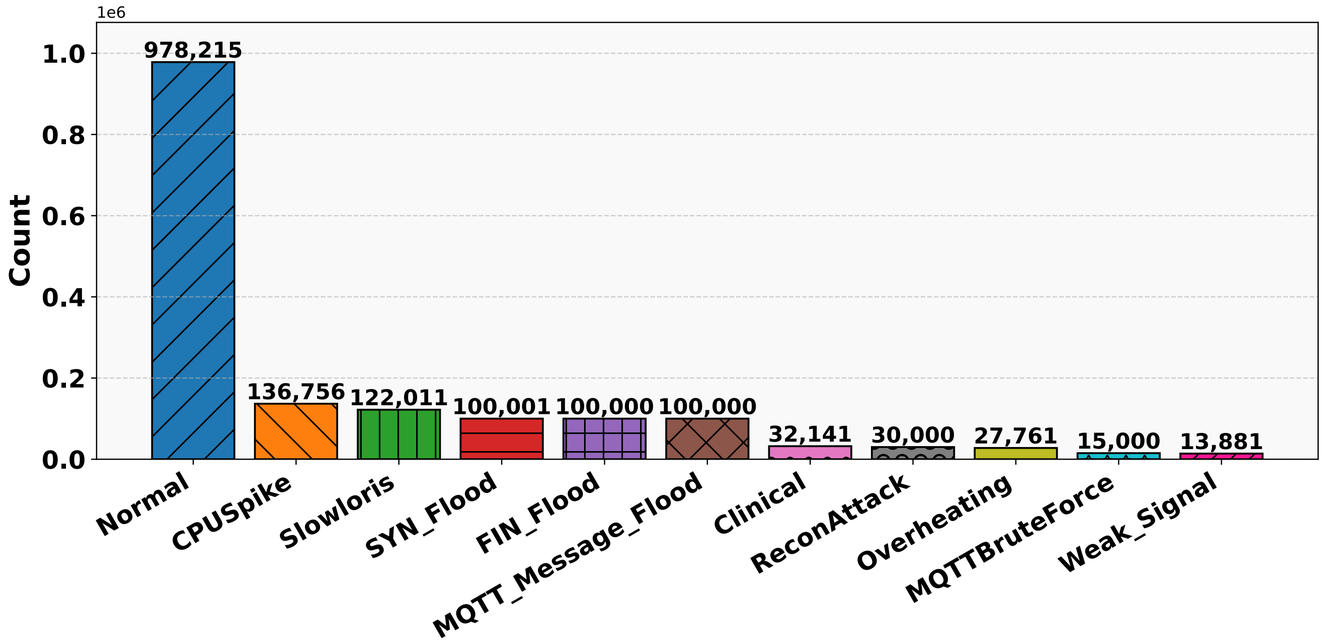


Stages of the X-IoMT dataset generation: raw traffic capture, feature extraction, labeling, and final dataset preparation.

Dataset Design and Methodology

Dataset Structure and Label Assignment

Labeling Task	Number of Classes	Labels
Binary	2	0 - Normal 1 - Anomaly
Layer-wise	4	0 - Normal 1 - Perception 2 - Network 3 - Application
Multiclass	11	Normal, Overheating, WeakSignal,CPUSpike, SYN_Flood, FIN_Flood, Slowloris,MQTTBruteForce, MQTTDoS, ReconAttack and Clinical



Experimental Setup & Evaluation

Machine Learning Evaluation Setup

Model	Key Hyperparameters
LR	max_iter=2000, class_weight='balanced'
RF	n_estimators=300, class_weight='balanced',n_jobs=-1, random_state=42
AB	n_estimators=300, learning_rate=0.5, random_state=42
DNN (Binary)	Layers: 128 → 64 → 1, activation=relu/ sigmoid, optimizer=Adam, epochs=8, batch_size=1024
DNN (Multiclass)	Layers: 256 → 128 → K, activation=relu/ softmax, loss=sparse_categorical_crossentropy, epochs=12, batch_size=1024

- **Accuracy:** measuring the proportion of correctly classified instances out of the total number of samples;
- **Recall:** capturing the ability of the model to detect anomalies by measuring the proportion of correctly identified anomalous instances;
- **Precision:** quantifies how many of the instances predicted as anomalous are actually anomalies;
- **F1-Score**: harmonic mean of precision and recall.

Experimental Setup & Evaluation

Classification Tasks Binary Results

Model	Accuracy	Macro Precision	Macro Recall	Macro F1
LR	0.9530	0.9534	0.9493	0.9512
RF	0.9898	0.9891	0.9899	0.9895
AB	0.9836	0.9832	0.9829	0.9831
DNN	0.9744	0.9754	0.9717	0.9734

Binary Classification:

- All models **performed well**, with **Random Forest** achieving the highest accuracy (98.9%)

Experimental Setup & Evaluation

Classification Tasks Layer-Wise Results

Model	Accuracy	Macro Precision	Macro Recall	Macro F1
LR	0.9232	0.8015	0.9587	0.8212
RF	0.9859	0.8952	0.9884	0.9293
AB	0.8553	0.7008	0.6390	0.6514
DNN	0.9599	0.8316	0.9818	0.8643

Layer-wise Classification:

- **Performance declined slightly** as models identified not just anomalies, but their **origin within the IoMT stack.**

Experimental Setup & Evaluation

Classification Tasks Multiclass Results

Model	Accuracy	Macro Precision	Macro Recall	Macro F1
LR	0.6587	0.8342	0.9325	0.8147
RF	0.9845	0.9565	0.9866	0.9669
AB	0.6163	0.6690	0.6600	0.5851
DNN	0.9004	0.8537	0.9433	0.8668

Multiclass Classification:

- Most challenging task. **Random Forest** and **DNN** significantly **outperformed** simpler models.

Conclusions & Future Work

Conclusions

- X-IoMT enables the real-world deployment of **multi-layer**, **device based dataset**;
- Provides a **benchmark** for **evaluating ML models** in complex healthcare environments;
- **Sets a foundation for future studies** in robust **anomaly detection in IoMT** systems.

Conclusions & Future Work

Future Work

- **Expand clinical data** by including more physiological signals like **heart rate**, **SpO2** and **respiratory rate** for richer anomaly contexts;
- Simulate sophisticated threats including adversarial anomalies (**adversarial augmentation**) for **stress testing models**;
- Promote **dataset availability** by **releasing X-IoMT to the public** to encourage reproducibility in **anomaly detection research**.

Thank You!
Questions?