MitM Attack Detection in BLE Networks using Reconstruction and Classification Machine Learning Techniques

Abdelkader Lahmadi [1], Alexis Duque[2], Nathan Heraief[3], Julien Francq[3]

[1]Université de Lorraine, CNRS, Inria, Loria, France

[2]Rtone - 120 rue de Saint-Cyr - 69009 Lyon - France

[3]Airbus CyberSecurity SAS, 78996 Elancourt Cedex, France

2nd Workshop on MACHINE LEARNING FOR CYBERSECURITY, September 14, 2020 - Ghent, Belgium

Overview

- Motivation
- Bluetooth Low Energy overview
- Methodology
- Proposed detection Approach
- Conclusions and future work

BLE-enabled connected devices

BLE (Bluetooth Low Energy) is a widely used radio technology by connected devices

- High end smartphones
- Sports / fitness devices
- Door locks
- Medical devices

Vulnerable to trivial attacks and can be easily compromised due their limited security features and lacking of secure development practices

- SweynTooth vulnerabilities [1]
- Man-in-the-Middle (MitM) attack: BTLEJuice, GAttack, Mirage tools

---

[1] E. Garbelini et al, SweynTooth: Unleashing Mayhem over Bluetooth Low Energy, USENIX ATC'20

IoT devices: attack detection and mitigation

Machine learning based methods

- Identify anomalies in network traffic through offline or online analysis [1]
- Detecting compromised IoT devices
- Devices specific communication patterns to detect anomalous behaviours deviation caused by attacks [2]
- Spoofing attacks in BLE enabled occupancy system

Focused on volumetric attacks, such as Mirai and few and rare work interested in attacks with sporadic network activity such as MitM

- Detecting MitM in BLE based eHealth care systems by using anomaly detection metrics [3]

---

[1] Hafeez et al: IoT-KEEPER: Detecting Malicious IoT Network Activity Using Online Traffic Analysis at the Edge. IEEE Transactions on Network and Service Management, (March 2020)

[2] Nguyen et al : DIoT: A Federated Self-learning Anomaly Detection System for IoT. In IEEE 39th International Conference on Distributed Computing Systems (ICDCS), 2019

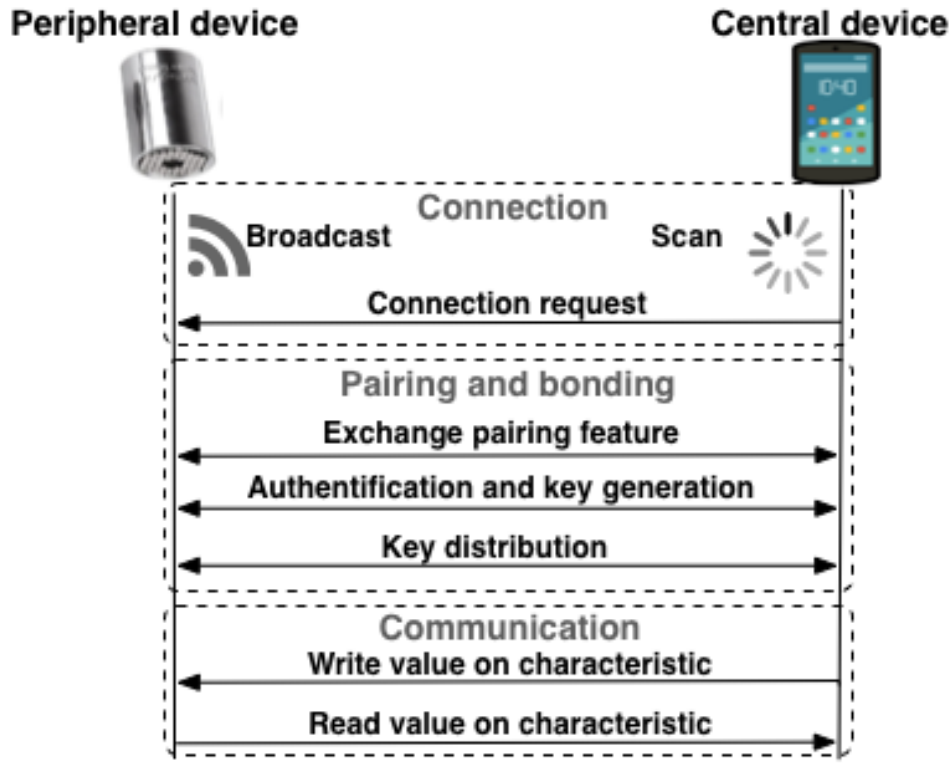
[3] Yassen et al: MARC: A novel framework for detecting MITM attacks in ehealthcare BLE systems. J.Medical Systems (2019)

BLE advertising and connection

Bluetooth Low Energy (BLE): battery-powered IoT

- 40 channels in 2.4Ghz ISM band
- Advertising: 3 channels and Data: 37 channels
- Two roles: Peripheral (e.g. sensors) and Central (e.g. smartphone)

- Enable Central devices to find devices and connect
- Advertising allows sending unidirectional but broadcast data
- Central device will listen for advertisements from Peripheral
- Connections allow the Central and Peripheral to exchange data bidirectionally

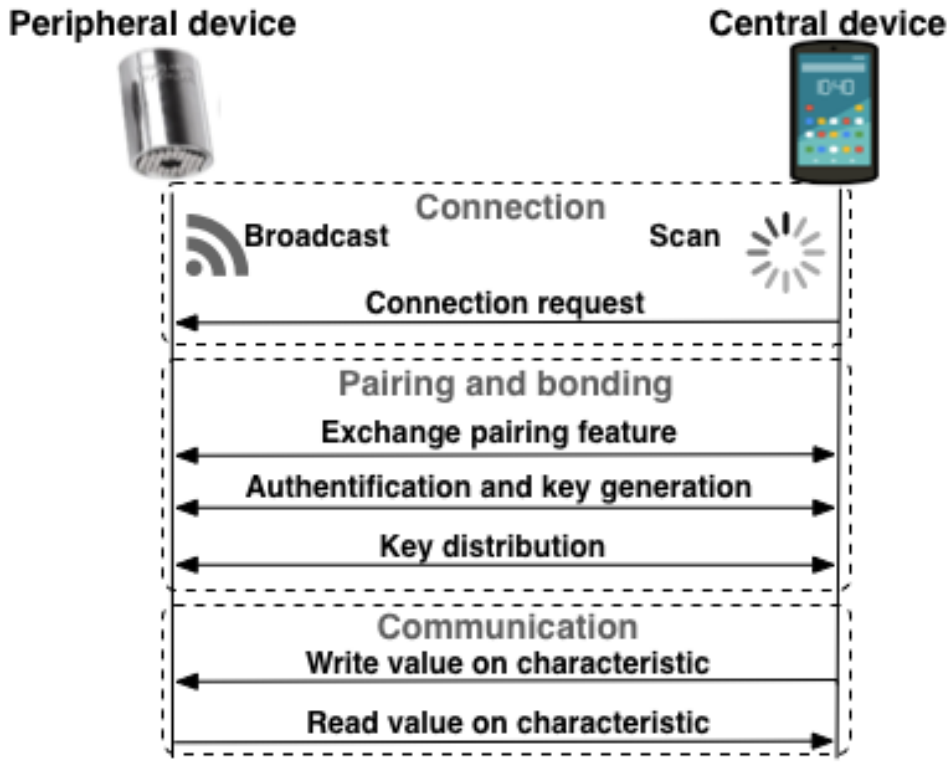


Data exchange

Bluetooth Low Energy (BLE): battery-powered IoT

- 40 channels in 2.4GHz ISM band
- Advertising: 3 channels and Data: 37 channels
- Two roles: Peripheral (e.g. sensors) and Central (e.g. smartphone)

- Data is transmitted on 37 data channels which are not used for advertising
- Data exposed by a Peripheral are presented in a GATT profile
- Attributes can be either services or characteristics
- Identified by a universally unique identifier (UUID)



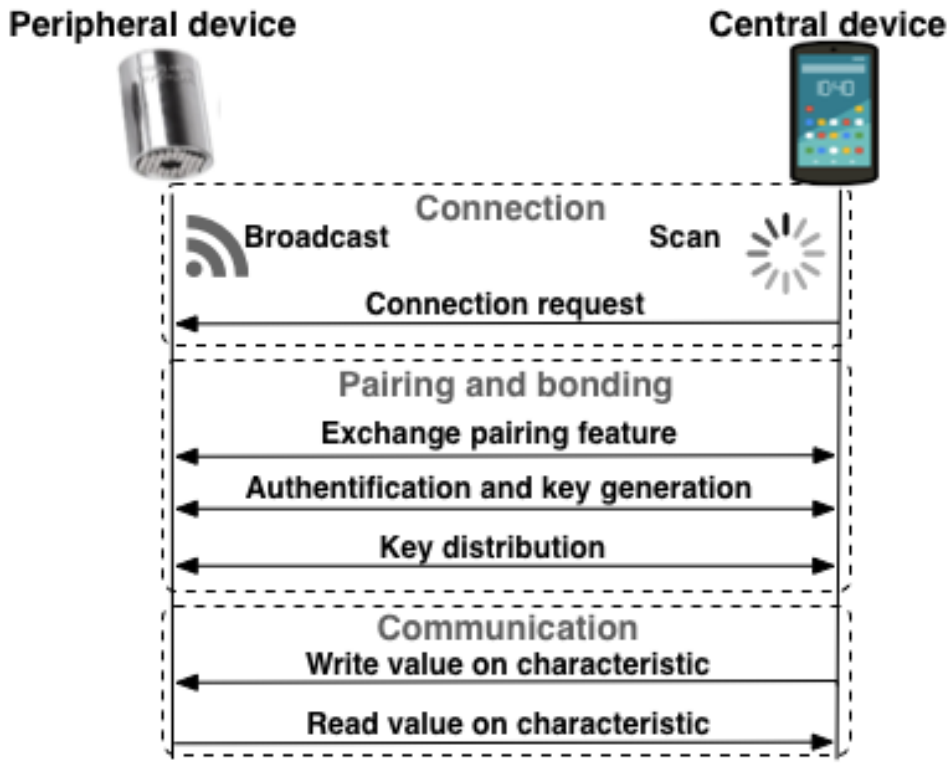
BLE security

Bluetooth Low Energy (BLE): battery-powered IoT

- 40 channels in 2.4GHz ISM band
- Advertising: 3 channels and Data: 37 channels
- Two roles: Peripheral (e.g. sensors) and Central (e.g. smartphone)

Pairing

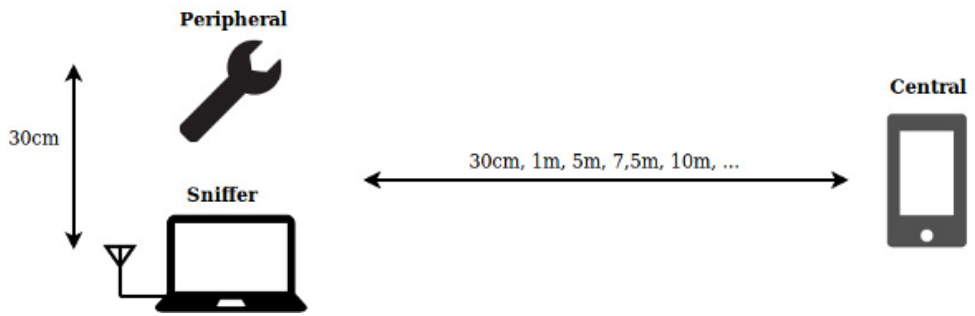
- Performed to establish keys which can then be used to encrypt a link
- Authenticating the identity of two devices
- Performed according to devices I/O capabilities
- Many BLE devices rely on the Just Works pairing method



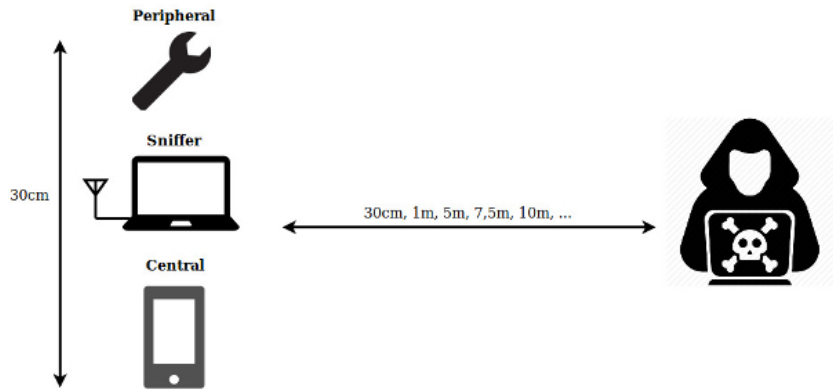
Experimental set-up

BLE-enabled torque wrench controlled by a smartphone

- Adjust and calibrate remotely with high precision the torque settings
- MitM attack to connect, pair, read and write to the device
- Attacker uses a clone to read, modify and write the settings of the torque wrench



(a) Normal scenario



(b) MitM scenario

Datasets building

Normal scenario

- Simulate a behaviour of the App running on the smartphone and generate different BLE packets including reading, writing and notifications

MitM scenario

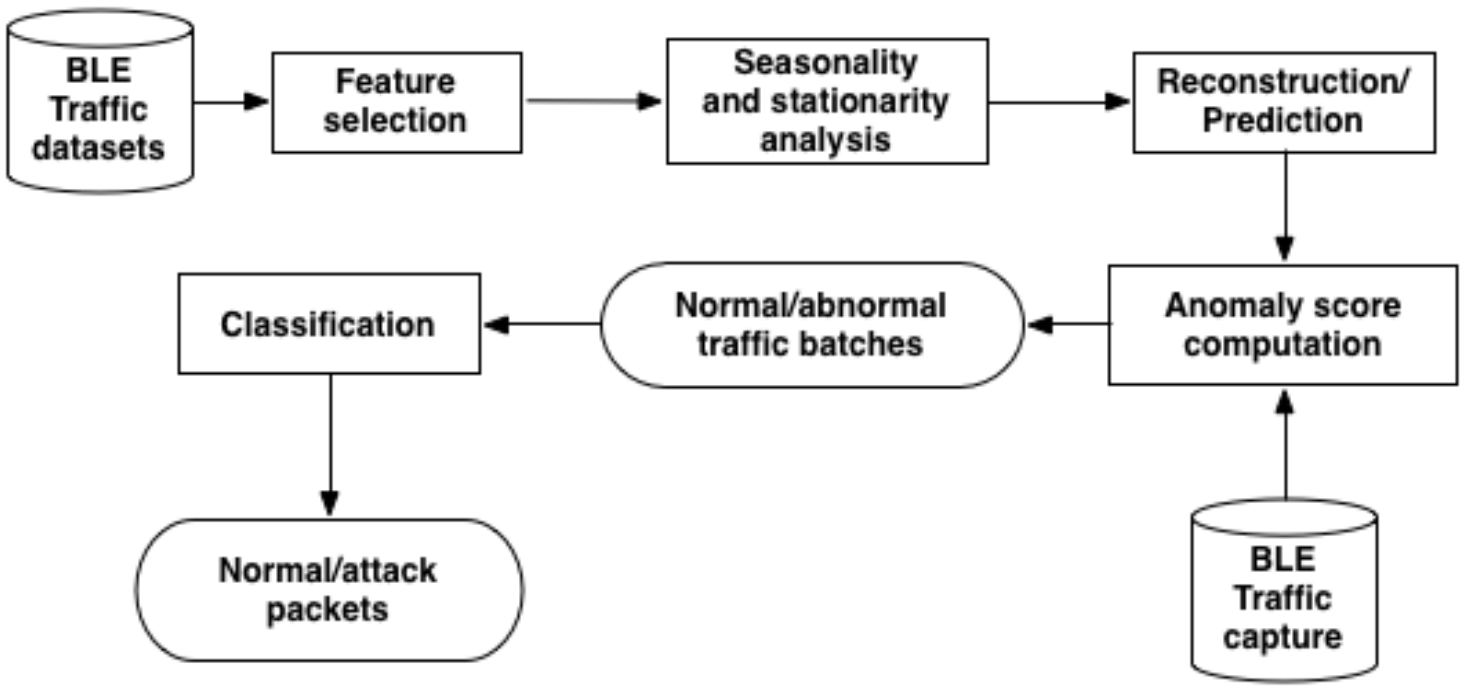
- The attacker will modify values written by the smartphone App on the BLE device

Datasets

- Varying the distance between the smartphone and the BLE device for the normal scenario
- Varying the distance between the attacker and the smartphone for the attack scenario
- Distance $\in \{30\text{cm}, 1\text{m}, 5\text{m}, 7.5\text{m}, 10\text{m}\}$
- 10 datasets that we merge in a single dataset with the distance value as a feature: attack or normal label

Proposed approach details

- Two Machine learning techniques applied jointly: reconstruction and classification
- Reconstruction: building a baseline model of normal patterns and then we measure deviations and errors from that model
- Classification: classify packets marked with attack features



Features extraction and analysis

Apply multiple feature selection methods : Variance, Chi2, Recursive Feature Elimination (RFE), Extra trees

- Let $\mathbb{F} = (f_1, f_2, ..., f_n)$ the set of features extracted from a BLE packet, with $n = 250$.
- Each method i provides a subset of features $F_{k,i}$ composed of k features
- $\mathbb{F}_{final} = \bigcap_{i=1}^{4} F_{k,i}$

4 features in a BLE packets dataset are the most important:

- Channel numbers: the channels used during the exchange of the BLE packets.
- Delta_time: the difference of time between two successive packets.
- Received Signal Strength Indication (RSSI): the signal-to-noise ratio value available in BLE packets.
- Distance: it denotes the distance between the mobile and the BLE device.

Model reconstruction

Learning the normal behaviour of the BLE packets exchange: minimise the error between the learned data and the original dataset

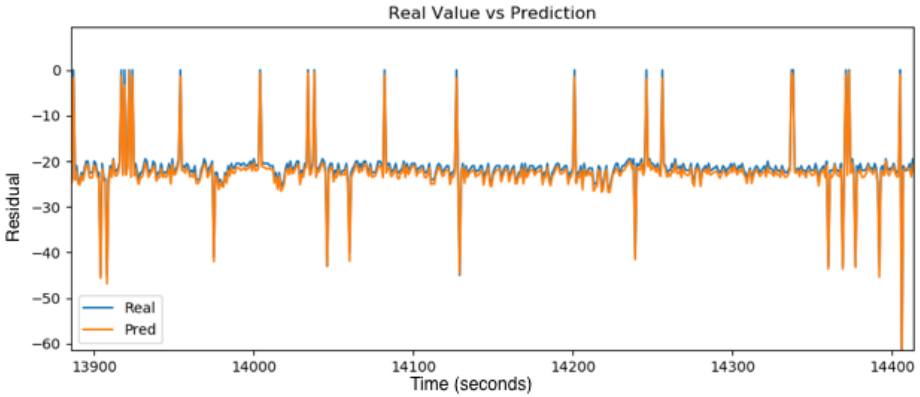
- Train the neural network on the dataset X_{train}
- Evaluate the obtained model on the $X_{validation}$ part while computing the reconstruction error
- Set a detection threshold to determine the presence of anomalies
- Residual defined as: $R(X_{train}, \widehat{X_{train}})$ with $\widehat{X} = f(X)$ and f represents the transformation of our auto-encoder.

Testing phase, compute the anomaly score α :

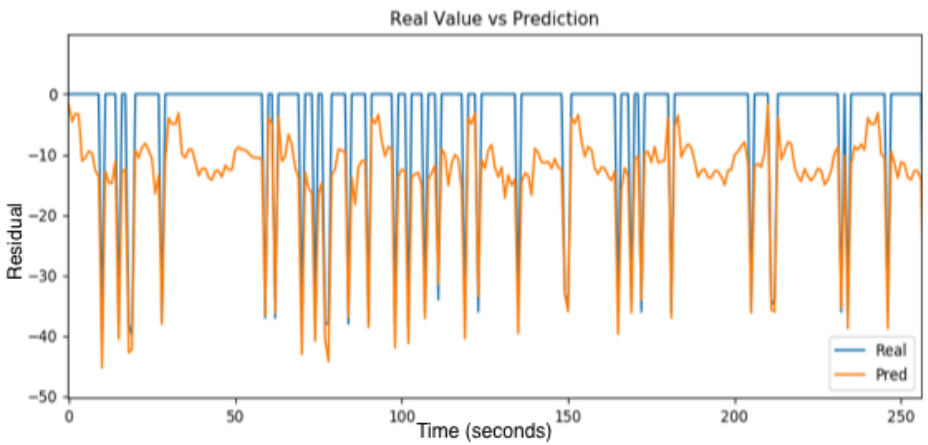
$$\alpha = \begin{cases} 0 \text{ if } |R(X_{test}, \widehat{X_{test}}) - \mu(R(X_{train}, \widehat{X_{train}}))| \leq 3 * \sigma(R(X_{train}, \widehat{X_{train}})) \\ 1 \text{ otherwise.} \end{cases}$$

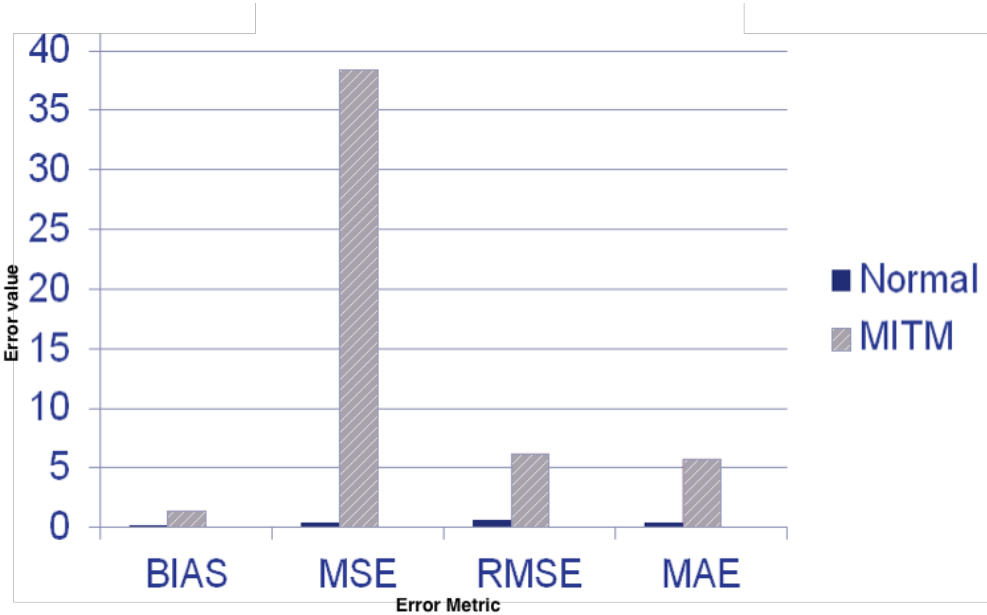
LSTM based model reconstruction: results

Testing of the normal model



Testing of the MitM attack model



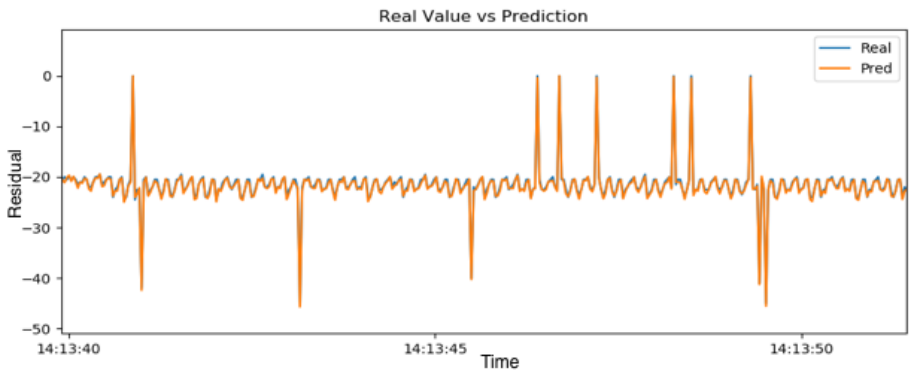


Reconstruction error between normal and attack patterns using LSTM

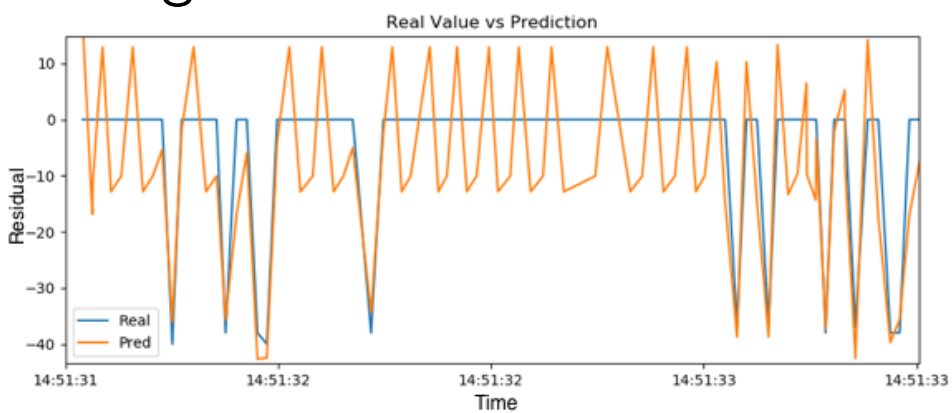
TCN based model reconstruction: results

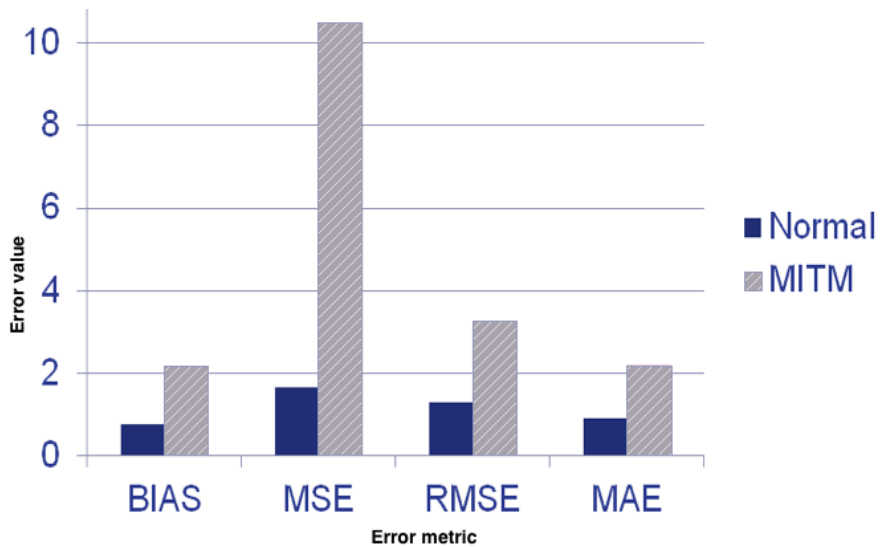
- Low value of the time-step in LSTM : low memory effect in the training neural network
- Using a Temporal Convolutional Network (TCN) instead of a LSTM: 30 time-step value

Testing of the normal model



Testing of the MitM attack model





Reconstruction error between normal and attack patterns using TCN

Model reconstruction: key takeaways

- Both LSTM and TCN models are able to detect suspicious batches using the the same anomaly score α
- TCN model has more accurate and lower reconstruction error with high memory effect compared to LSTM architecture
- But, without detecting packets involved in the attack

Classification of BLE packets

In suspicious batches of traffic, classify packets according to their class: "normal" or "attack"

- Jointly using Text-Convolutional Neural Network (Text-CNN) for feature extraction and a Random Forest algorithm for classification [1]
- Convert BLE packets payload into word embedding (Word2Vec) to extract salient features with Text-CNN
- Extract from BLE packets their traffic statistics
- Text-CNN based features are concatenated with statistical features and provided as input to Random Forest algorithm

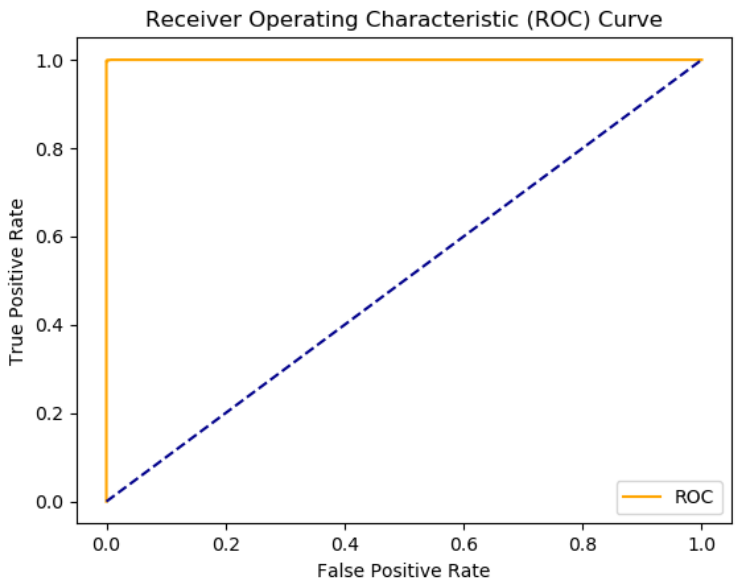
Statistical features
Number of packets per second
Number of bytes per second
Max, min and average packets length
Max, min and average time interval between 2 packets
Number of packets for each BLE packets type (ADV, DATA, etc.)

[1] Min, E., Long, J., Liu, Q., Cui, J., Chen, W.: TR-IDS: Anomaly-Based Intrusion Detection through Text-Convolutional Neural Network and Random Forest. Security and Communication Networks, 2018)

Classification of BLE packets: results

		Predicted labels: Normal	Predicted labels: Attack
Actual label	Normal	100% (9541/9543)	0% (2)
	Attack	0.3% (12)	99.7% (4207/4219)

- High classification performance
- Area Under the Curve (AUC) close to 1: good measurement of separability between "normal" and "attack" packets
- But our results are limited to the collected datasets and the experimental setup environment



ROC curve of the BLE packets classifier

Conclusions and future work

- Study on the use of machine learning techniques to detect MitM attack targeting BLE enabled IoT devices
- Feasibility of the attack in a real-world deployment while varying the distance between the BLE mobile and devices
- Apply jointly reconstruction and classification models based on neural networks to detect suspicious network packets
- High detection accuracy (≈ 0.99) and low false positive rate (≈ 0.03)

- Detecting more classes of BLE attacks including DoS and connection hijacking within various BLE environments
- Protection mechanisms for BLE networks

Thank you for your attention.

This work is funded by the French Government under grant FUI 23 PACLIDO (Protocoles et Algorithmes Cryptographiques Légers pour l'Internet Des Objets): `https://paclido.fr/`