IoT Privacy, Security and Safety Supervision Engines

Magdalena Kacmajor
Senior Applied Researcher
IBM Ireland
IPSE: IoT Privacy, Security and Safety Supervision Engines

- A set of novel runtime components acting in concert to understand and monitor the cyber-physical ecosystem
  - **Privacy Engine**: privacy by design
    -> handling data encryption policies based on blockchain technologies
  - **Security Engine**: firmware authentication
    -> identification of security vulnerabilities, rule-based filtering and validation with blockchain
  - **Safety Supervision Engine**: safety policies enforcement:
    -> monitoring data streams with machine learning deployed on the edge
- Topology service and IoT Language
  - Enable functionality of the Privacy and Safety Supervision Engines
- Predictive Analytics for anomaly detection
Topography Service and IoTL

**IOTL - state definition language**
- Define the topology of an IoT system (IoT State)
- Define privacy and safety rules and policies
- Share and sync IoT state across Fog Network

**Topology service**
- Provide REST services for Chariot Engines
- Implement IoT State
- Handle persistence

**IoTL: Efficient tool for communicating IoT state**
- Concise but comprehensive representation of current state
- Easy to share across the Fog Network
- Easy to sync to ensure consistent state
- Easy to store and recover
- Easy to interact with via REST interface
CORE SPECIFICATIONS

- **Entities**
  - Zones
  - Gateways
  - Sensors

- **Relations**: Defined between two components in the system.
  - Dependency, correlation, equality, delayed condition...

- **Safety policy definition**
  - Enforcements
  - Plugs

- **Privacy policy definition**:
  - Access Control Lists,
  - Schemas
  - Anonymization
Safety Supervision Engine

Safety policy enforcement

Static (rule-based) policies
- Leverage human experience

Dynamic (ML-based) policies
- Near-real time anomaly detection

Monitor IoT State

Machine learning deployed on edge

Support for custom user-defined ML models

'Developer’ deep learning anomaly detector supplied

Detailed Deployment Guide provided

Stream Listener: Monitor, assess and enforce

- Web interface for registering and enforcing safety policies
- Detect & Predict safety policy violations with associated Alert Generation
- Integration of dynamic (ML-based) policies and user-defined rules

Plug & Play Machine Learning: easily upload custom models

- Safety supervision without manual effort – does not require time or expert knowledge
- Machine Learning deployed on edge
- Of-the-shelf Deep Learning anomaly detector provided
Safety Supervision Engine and Anomaly Detection

Fog Node

Zone

Gateway

Sensors

[temp < 30]

Plugs

Custom ML models

IoTL

State definition language

MQTT

Alert

Northbound dispatcher

CHARIOT Dashboard

Machine learning deployed on the edge

Deep learning: near-real time anomaly detection

Southbound dispatcher

CHARIOT – 3rd Workshop, 22 October 2020
Safety Supervision Engine and Anomaly Detection

- **Integration with CHARIOT Dashboard**
  - Complete REST API provided for safety policy management and anomaly detectors configuration
  - CHARIOT Dashboard provides user-friendly GUI
  - Alternatively, safety policies can be managed through IoT Manager UI

```
enforce [device_52806c75c3fa_Sensor1.hum < 15]
enforce [device_52806c75c3f2_Sensor5.temp]
```

```
plug device_52806c75c3f2_Sensor5.temp BASIC_LSTM (device_52806c75c3f2_Sensor5)
```
The project CHARIOT has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 780075

Contact Details

IBM Ireland
Magdalena Kacmajor
magdalena.kacmajor@ie.ibm.com
Privacy Engine and Data Encryption

Konstantinos Skianis PhD
Senior Researcher
CLMS
Privacy Engine and Data Encryption - Intro

Main goals
• Protect private and sensitive data
• Identify types of sensors and services with regards to privacy
• Components communicate without exposing sensitive information

Novel aspects
• Anonymization methods
• Cognitive: use machine learning models for disseminating messages
• Provides insight on privacy threats based on topology information
• Self-contained service deployed on a Fog node

Main benefits
• Create value from IoT sensor messages by training specialized dissemination classification models
• A complete framework for managing private data in industrial IoT environments
Privacy Engine and Data Encryption

- Two types of checks enables passive and active safeguarding
- Inspection checks helps administrator of IoT network to actively map all privacy related information during configuration setup
- Filtering safeguards information exchange with other parties by encrypting and anonymizing information

<table>
<thead>
<tr>
<th>IoTTL Statement</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>define SENSOR S1 --params {&quot;privacySensitive&quot;: 1.0}</td>
<td>Mark a sensor as privacy sensitive.</td>
</tr>
<tr>
<td>acl BMS S2 DENY</td>
<td>Safeguard access to sensor messages</td>
</tr>
<tr>
<td>acl BMS S2 ALLOW</td>
<td></td>
</tr>
<tr>
<td>schema EmployeeID --pattern &quot;\d{4}-\d{4}-\d{4}\d{4}&quot; --private</td>
<td>Manually define privacy sensitive formats.</td>
</tr>
<tr>
<td>expect S1 EmployeeID</td>
<td></td>
</tr>
<tr>
<td>anonymize S1 age SHA256</td>
<td>Enable privacy engine to anonymize age on message originated from S1 Sensor.</td>
</tr>
</tbody>
</table>
Privacy Engine and Data Encryption

**Anonymization**
- Administrator defines message fields to be anonymized
- Engine applies anonymization logic on messages originating from specific tables
- Anonymization replaces value with random sized string of ‘*’
- MD5 & SHA256 pseudo-anonymizes data by returning hashed value

**Encryption**
- Prevents sensitive information leakage to unauthorized users
- Public Key encryption adds end to end encryption between Fog Node and External services preventing MitM attacks
- Access control lists defined by the CHARIOT by using IoT-L guards user data
Privacy Engine and Data Encryption - Standalone

**Manual private data guard**
- Provides insight on privacy threats based on topology information
- Topology information can be pulled by API
- Information can also be pulled by local file created by Administrator, to achieve standalone functionality (without the platform API)
- This version can be installed in single board Linux PC and connected to external MQTT broker to receives messages per configuration

**Cognitive - Detect privacy violation by using dissemination level classifier**
- Collection of messages from every sensor is used to produce datasets for model training
- Message types stemming from private sensors are used to compose attributes of training instance
- Machine learning to produce Dissemination level classifier
- Fully automated process, variable reliability
K. Skianis
CLMS
k.skianis@clmsuk.com
Predictive Analytics for Out-of-Bounds Behaviour

Kostas Zavitsas PhD
VLTN
Predictive Analytics for Out-of-Bounds Behaviour

- Technical objectives:
  1. Identify sources of variation in a monitored system
  2. Datasets of varying dimensions capturing a stochastic real-world processes
  3. Calculate bounds of normal behavior

- Business objectives:
  - robust/ context – agnostic
  - efficient/ no human intervention

- All 3 Chariot case studies offer ample datapoints and opportunities to train accurate ML based predictive models

Locomotive / Fleet – DMMS

Smart Building/ Technology campus – BMS & Security IoT

Airport – BMS
Predictive Analytics for Out-of-Bounds Behaviour

- Anomaly Detection component pipeline:
  - Part 1: Training
    - Data preprocessing
      - Temporal resampling
    - Normalization and regularization to avoid overfitting to one feature
  - Cross validation algorithm used with k=10
  - Unsupervised machine learning clustering models
    - Elliptic Envelope (EE)
    - Isolation Forest (IF)
    - One Class Support Vector Machine (OSVM), and
    - Density-based spatial clustering of applications with noise (DBSCAN)
  - model evaluation assessed with the Fowlkes-Mallows index (FM)
    \[
    FM = \frac{\frac{TP}{TP + FP} \cdot \frac{TP}{TP + FN}}{\sqrt{\frac{TP}{TP + FP} \cdot \frac{TP}{TP + FN}}}
    \]
  - Update dashboard information
  - Upload model to Security Engine

- Part 2: Prediction
  - Collect live data
  - Check if out of bounds behaviour
Predictive Analytics for Out-of-Bounds Behaviour

- Unsupervised AD modelling

- Best performing model:
  - Elliptic Envelope with 97% prediction accuracy for incorrect Indoor temperature readings
Contact Details

VLTN
Kostas Zavitsas
k.zavitsas@vltn.be