The document reflects only the author’s view and the Commission is not responsible for any use that may be made of the information it contains.

InSecTT:
Intelligent Secure Trustable Things

Publishable Summary of WP2 Results

Document Type	Deliverable
Document Number	D2.4
Primary Author(s)	Johannes Peltola | VTT Pauli Räsänen | VTT
Document Version / Status	1.0 | Final
Distribution Level	PU (public)

Project Acronym	InSecTT
Project Title	Intelligent Secure Trustable Things
Project Website	https://www.insectt.eu/
Project Coordinator	Michael Karner | VIF | michael.karner@v2c2.at
JU Grant Agreement Number	876038
Date of latest version of Annex I against which the assessment will be made	2023-05-24

InSecTT has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 876038. The JU receives support from the European Union’s Horizon 2020 research and innovation programme and Austria, Sweden, Spain, Italy, France, Portugal, Ireland, Finland, Slovenia, Poland, Netherlands, Turkey

CONTRIBUTORS

Name	Organization		Name	Organization
Aurelien Ibanez	AKEO PLUS		John Barry	LIEBHERR
Baturay Dalgıç	ARCELIK		Yavuz Selim Bostanci, Mujdat Soyturk	MARUN
Markus Wolf	AVL		Tijana Markovic Miguel Leonortiz	MDH
Paulo Duarte	CAPGEMINI		Bernd-Ludwig Wenning	MTU
Christine Hennebert, Pierre-Alain Moellic	CEA		Ad Arts Ton Scheepers	NXP-NL
Luca Davoli, Laura Belli, Gianluigi Ferrari	CINI-UNIPR		Emir Tolga Demirel	PAVOTEK
Paolo Azzoni	ETH		Efi Papatheocharous	RISE
Marko Komssi	FSC		Magnus Isaksson	RTE
Mateusz Rzymowski	GUT			SETA
KHAN Rahat SERE Ahmadou	IDEMIA		Marcello Coppola	STM
Francisco Parrila, Alejandro Díaz	INDRA		Ashutosh Simha	TU-DEFT
Ramiro Robles	ISEP		Alessandro Chiumento, Kamran Zia	U-TWENTE
Drago Torkar	JSI		Liam O'Toole	UCC
Esa Piri	KAI		Jorge Portilla	UPM
Fernando Núñez	KLAS		Christoph Pilz	VIF
Andrii Berezovskyi, Nils Jörgensen, José Manuel Gaspar Sánchez, Kaige Tan	KTH		Johannes Peltola Pauli Räsänen Arttu Lämsä Ilkka Moilanen	VTT
Leander Hörmann	LCM		Veli-Pekka Salo	Wapice
Filippo Cerocchi	LDO			

FORMAL REVIEWERS

Name	Organization	Date
Magnus Isaksson	RTE	2023-06-17
Drago Torkar	JSI	2023-06-21

DOCUMENT HISTORY

Revision	Date	Author / Organization	Description
0.1	2023-04-14	F.Ademaj-Berisha/ SAL Pauli Räsänen / VTT	Draft version for input collection
0.2	2023-04-18	Christoph Pilz / VIF	VIF Content
0.3	2023-05-03	Luca Davoli, Laura Belli, Gianluigi Ferrari / UNIPR	UNIPR contribution
0.4	2023-05-03	Paulo Duarte	Capgemini Engineering Contribution
0.5	2023-05-04	Efi Papatheocharous / RISE	RISE contribution
0.6	2023-05-05	Pauli Räsänen, Johannes Peltola / VTT	VTT contribution + editorial updates
0.9	2023-06-26	Johannes Peltola / VTT	Draft for submission
1.0	2023-06-28	Johannes Peltola / VTT	Final release

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 EXECUTIVE SUMMARY

This deliverable is a public summary of the main results in terms of technical components achieved within InSecTT WP2: Reliable AI / Machine Learning for IoT. During the duration of three year the focus within WP2 has been on creating trusted, distributed, safe, explainable, and powerful AI methods and platforms for IoT systems, being one key enabler for implementations of the next generation Internet of Things. The WP2 activities have been closely aligned with WP3 and interaction between both work packages has ensured that solutions developed within WP2 make best use of secure, safe and reliable wireless systems developed in InSecTT WP3.

This deliverable summarizes the significant achievements and noteworthy elements attained throughout the three-year duration of the AI development in WP2. Given the extensive scope of the project and the involvement of numerous partners, it focuses on highlighting the key innovations and exploitable items. The exploitable items mentioned in this document are the outcomes of collaborative efforts across multiple Building Blocks and various components (see more about the project structure in Section 2).

This deliverable includes:

- A summary about the objectives and methodology developed in InSecTT WP2 and WP3.
- A description of the highlights in terms of novelty per partner summarized in a table.
- A more detailed description of exploitable item(s) per partner.

Throughout the project timeline, the partners in WP2 dedicated their efforts to developing both hardware and software solutions. These solutions have been integrated into diverse demonstrators, which serve as strong evidence of the project's value across multiple industrial use-cases. The technical accomplishments within WP2 specifically center around the following areas:

- Explainable and reliable application-level AI functionality that provide added intelligence and security for IoT devices and systems as a whole.
- Machine learning methods for supporting physical and MAC layers in wireless communications in order to achieve more robust and efficient wireless communication
- Secure and distributed models between cloud, edge, mobile device processing as well as efficient models for implementing machine learning in distributed IoT systems and devices.
- AI verification and validation framework for IoT applications as well as AI methods for managing complex V&V simulation environments
- Trust framework for evaluating and developing trusted and end user accepted AI solutions

In addition to this deliverable, there is also a book accessible that offers a thorough and comprehensive overview of the project's technical accomplishments and outcomes. This book serves as a valuable resource for interested readers who seek a more detailed description of specific topics, allowing them to delve deeper into the subject matter.

Keywords: Artificial Intelligence, machine learning, Internet of Things, explainable AI, cyber security, anomaly detection, wireless communication, distributed processing, dynamic load balancing, validation and verification, trustworthy AI.

2 OBJECTIVES AND METHODOLOGY

The overall objective of WP2 has been to increase the competitiveness of InSecTT partners in developing and applying AI and machine learning using safe, trusted and explainable models working in distributed IoT ecosystem. The specific objectives are intricately connected to the task structure defined within WP2:

- Enhance IoT devices and systems by developing AI functionality that is explainable, reliable, and capable of providing added intelligence and security at the application level (Task 2.1).
- Improve wireless communications by developing machine learning methods that support the physical and MAC layers, resulting in more robust and efficient wireless communication (Task 2.2).
- Establish secure and distributed models between cloud, edge, and mobile device processing, while also creating efficient models for implementing machine learning in distributed IoT systems and devices (Task 2.3).
- Develop an AI verification and validation framework specifically designed for IoT applications, along with AI methods for managing complex verification and validation simulation environments (Task 2.4).
- Establish a trust framework that enables the evaluation and development of trusted AI solutions, ensuring acceptance by end users (Task 2.5).
- Provide re-usable Building Blocks in line with the InSecTT architecture into Use Cases is an objective of all five tasks, as well as a public summary of technical achievements of the work package to foster deployment beyond project level (all Tasks).

Finally, and most important, WP2 activities have been defined in a way to fully support the achievement of overall objectives of InSecTT**,** which are to develop solutions for (1) Intelligent, (2) Secure, (3) Trustable (4) Things applied in (5) industrial solutions for European industry throughout the whole Supply Chain (6).

From a managerial perspective, activities have been organized in five Tasks (Building Blocks / BB's) as follows:

WP2	Building Block Name and Leader
Task BB 2.1	AI on application level, explainable and traceable AI – VTT (22 partners)
Task BB 2.2	AI for wireless transmission - ISEP (15 partners)
Task BB 2.3	AI on computational level (on device and edge) – UCC (26 partners)
Task BB 2.4	AI verification & validation – CEA (12 partners)
Task BB 2.5	Trustworthy AI – VIF (8 partners)

Table 1: Overview of WP3 BBs (Tasks)

From a **technical perspective** a three layer model has been developed in InSecTT and has proven to be very helpful for both, efficient execution and clear accountability, as well as facilitation of technical cooperation and utilization of synergies.

- **Layer 1: Building Blocks (BBs, equals Task layer)**

High-level structure as used in project submission and for managerial and reporting purposes. The BB level has been experienced as being too comprehensive and too diverse from a technical perspective to ensure clear ownership and to facilitate efficient cooperation. As a result, the Sub-BB concept has been developed and introduced.

- **Layer 2: Sub-Building Blocks (Sub-BBs)**

Main layer to organize and manage technical cooperation and alignment, offering the right granularity level for joint development targets and to exploit potential synergies. An overview for all WP3 Sub-Building Blocks is provided in Figure 1.

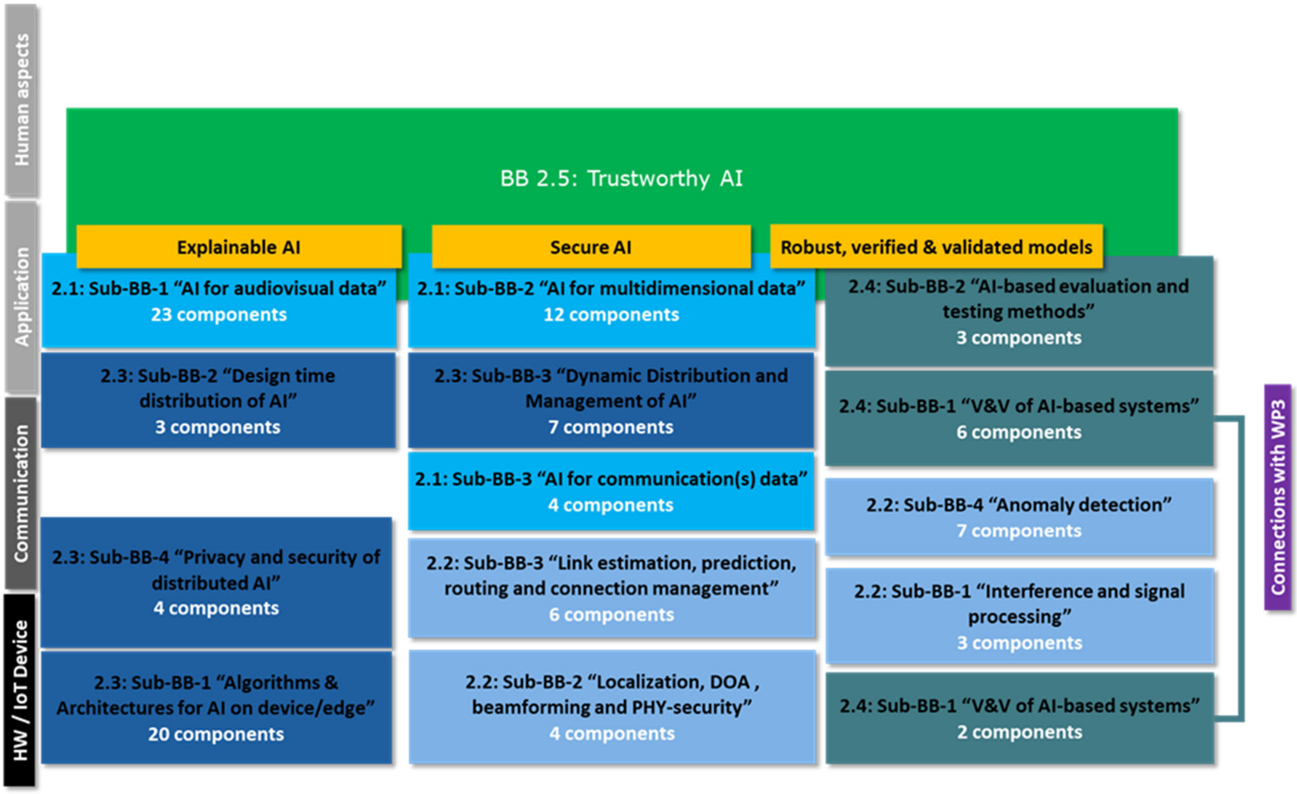

Human aspects
Application
Communication
HW / IoT Device
BB 2.5: Trustworthy AI
Explainable AI
Secure AI
Robust, verified & validated models
2.1: Sub-BB-1 "AI for audiovisual data" 23 components
2.1: Sub-BB-2 "AI for multidimensional data" 12 components
2.4: Sub-BB-2 "AI-based evaluation and testing methods" 3 components
2.3: Sub-BB-2 "Design time distribution of AI" 3 components
2.3: Sub-BB-3 "Dynamic Distribution and Management of AI" 7 components
2.4: Sub-BB-1 "V&V of AI-based systems" 6 components
2.1: Sub-BB-3 "AI for communication(s) data" 4 components
2.2: Sub-BB-4 "Anomaly detection" 7 components
2.3: Sub-BB-4 "Privacy and security of distributed AI" 4 components
2.2: Sub-BB-3 "Link estimation, prediction, routing and connection management" 6 components
2.2: Sub-BB-1 "Interference and signal processing" 3 components
2.3: Sub-BB-1 "Algorithms & Architectures for AI on device/edge" 20 components
2.2: Sub-BB-2 "Localization, DOA , beamforming and PHY-security" 4 components
2.4: Sub-BB-1 "V&V of AI-based systems" 2 components
Connections with WP3

Figure 1 WP2 sub-building blocks BB2.1 – BB2.5

- **Layer 3: Components**

This layer has been introduced to understand how the different partner level contributions do contribute to certain Sub-BBs and how they feed into the Use Case integration work. This also ensures clear accountability on partner level. "Component" in this context can mean any contribution, being it hardware (HW), software (SW), methodology, or any combination of them.

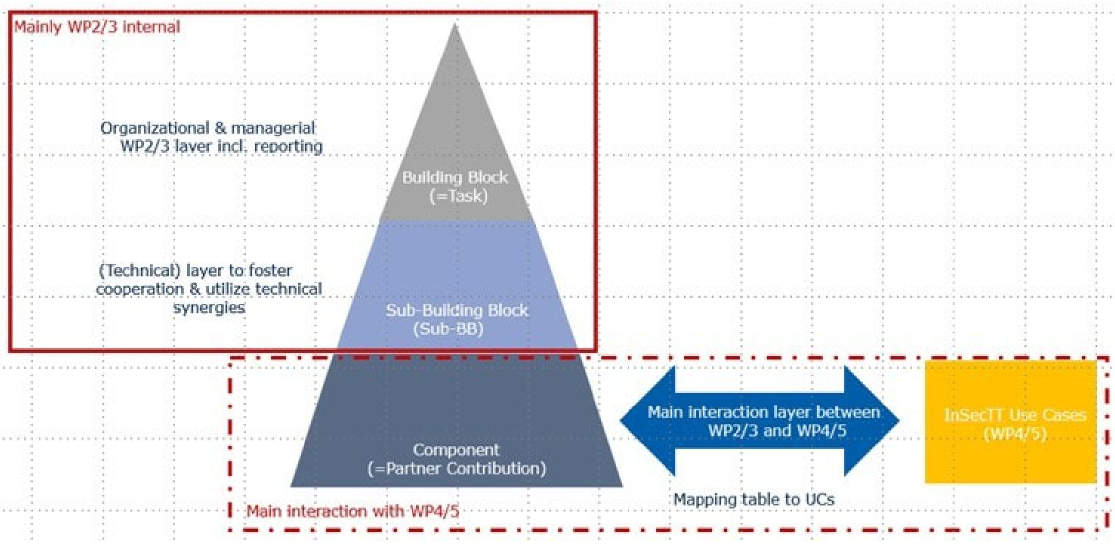

Mainly WP2/3 internal
Organizational & managerial WP2/3 layer incl. reporting
(Technical) layer to foster cooperation & utilize technical synergies
Building Block (=Task)
Sub-Building Block (Sub-BB)
Component (=Partner Contribution)
Main interaction layer between WP2/3 and WP4/5
InSecTT Use Cases (WP4/5)
Mapping table to UCs
Main interaction with WP4/5

Figure 2 InSecTT Three Layer Model

In conclusion the described three-layer model turned out to be an appropriate methodology to effectively address the significant size and the high complexity and diversity of a project like InSecTT. It motivates co-operation and synergies, and provides clear accountability. However, it also brings some challenges in terms of managerial effort (by Task- and WP leads), as well as discipline by all partners to think and execute in line with this model.

3 DESCRIPTION OF RESULTS

3.1 Introduction

As stated in Section 2, the activities carried out in WP2 are structured using a three-layer model comprising Building Blocks, sub-Building Blocks, and Components. The technical advancements and accomplishments in this work package are the result of efforts across various Building Blocks, sub-Building Blocks, and Components. To facilitate readers in comprehending the noteworthy aspects and to enhance document navigation, Section 3.2 presents a summary of partner contributions, outlining the novelty, target sectors/markets, and the achieved Technology Readiness Level (TRL) for each exploitable item and partner involved. For a more comprehensive understanding, Section 3.3 offers a detailed description of each exploitable item listed in Table 2.

3.2 Overview table

Partner	Exploitable Item	Novelty	Target Sectors/Markets	Achieved TRL
AKEO PLUS	Hardware-enabled secure generation of data proofs and traceability for embedded industrial systems	Generation of proofs and traceability of the data produced by an industrial application, compliant with the Ethereum blockchain, with the use of STM32MP1 devices.	Industrial manufacturing	5
ARCELIK	Cybersecurity assessment methodology	A questionaire-based assesment method for evaluating systems from cybersecurity aspect	Industrial communicationfor automation systems	4
AVL	AI-based automata learning algorithms (Wireless-Testbed)	AI-based automata learning algorithms for different protocols (BLE, BT, NFC)	Automotive Industry	3

Partner	Exploitable Item	Novelty	Target Sectors/Markets	Achieved TRL
CEA	Traceability and Trust of Embedded Neural Network Models	Complete device-level authentication process of an embedded AI inference using Ethereum blockchain.	Collaborative manufacturing	5
CINI-UNIPR	Multi-Interface GW (MIG) with short/mid/long-range communication technologies	Integration of multiple communication protocols in a multi-interface gateway for data collection in heterogeneous environments (and able to perform first-level AI-based processing, if required by the context)	Indoor (e.g., buses, public airports) and outdoor (e.g., for environmental monitoring)	5
CINI-UNIPR	Air quality sensing and prediction	Implementation of AI-based (i.e., ML and DL) algorithms to be executed in a COTS-based processing node collecting data for air quality monitoring and prediction	Indoor (e.g., buses, public airports) and outdoor (e.g., for environmental monitoring)	4

Partner	Exploitable Item	Novelty	Target Sectors/Markets	Achieved TRL
CINI-UNIMORE	Synthetic dataset on people surveillance Visual anomaly detection algorithms Explainability AI pipeline	Dataset has been completed and publicly release to the community and tested by multiple researchers. First and largest dataset on surveillance fully synthetic with testified transfer in real scenarios. Can be found here https://aimagelab.ing.unimore.it/imagelab/page.asp?IdPage=42 The visual anomaly detection algorithm has been tested and integrated on the NVIDIA Xavier target board. Composed by a novel training pipeline that can provide temporal anomaly location using video level annotation and an original architecture for real time performances on embedded boards. Exploiting a novel Multiple Instance learning loss for propagating gradient at frame level allowing Visual Explanation maps of anomaly with video level annotation	Safety, Smart City Safety, Public Transport Safety, Public Transport	6 7 6
CAP.ENG.	5G Platooning Control and Management System Simulation	Simulation Platform for 5G wireless platooning with monitoring and management following O-RAN concepts. Potential to accelerate the development of intelligent applications.	Telecom Industry, Research organizations	6

Partner	Exploitable Item	Novelty	Target Sectors/Markets	Achieved TRL
ETH	1 - people and flow mo 2 - identification of dangerous materials and of safety critical situations	1 – New line of flexible sensors with a very compact and tight profile which simplifies the integration of the sensor in the environment, capable to disappear in it. The cost of the sensor has been reduced of more than 100 times, allowing to install in the same space more sensors and obtain more precise results. The new sensor enables also multi applications. 2 – New low cost modular multisensor station for environmental monitoring: sensors for air properties monitoring can be chosen depending on the specific context and application. Small dimensions and low cost.	1 – Mobility domain, public transportation, public and private areas monitoring 2 – Physical security	1) 6 2) 7
FSC	Embedded AI-assisted security solution for consumer Wi-Fi routers	Applying the security solution within a customer environment in addition to running the feature directly on the router itself.	Global consumer markets together with router manufacturers and telecom operators	4 (5 in back end)

Partner	Exploitable Item	Novelty	Target Sectors/Markets	Achieved TRL
GUT	AI-based direction-of-arrival and localization algorithms Test generation and map exploration using ML-based approach	Using AI algorithms to calculate the direction of incoming jamming signal in real-world environments. Toolset for assessing effectiveness of the solution and comprehensive data collecting framework. Reducing time of validation testing in real-world and emulated environment by using . “smart” sampling approach.	Localization systems: healthcare, manufacturing, transport, security Autonomous surveillance, robotics, data collecting	6 4
IDEMIA	Adversarial attack and defense against Facial recognition systems (FRS)	Defense against adversarial attacks by dynamically adding autoencoders from a pretrained set to a base model.	Border control, access control	5
INDRA	Train positioning and train integrity based on ML	The novelty reached in these developments is the use of wireless systems to obtain accurate and reliable positioning and integrity of the train and its wagons making a sensor fusion of multiple GNSS signals, inertial sensors, RSSI and UWB through the use of ML algorithms.	Safety, Public Transport, Smart City, Railway, Localization systems, IoT	6

Partner	Exploitable Item	Novelty	Target Sectors/Markets	Achieved TRL
ISEP	1. Intelligent wireless MIMO layer 2. Spatial authentication platooning systems 3. Trustworthiness metrics calculation	1. New analysis of all processes related to MIMo based on AI. Optimization of AI according to the MIMO problem. 2. New spatial authentication for platoons based onmultiple DoAs of the vehicles of the platoon 3. New approach for system design based on trustworthiness metrics	1. Autonomous vehicles, platoons and wireless avionics 2. Vehicular 3. All	1. 6 2. 5 3. 5
JSI	Anomaly detection in ECG signals	Explainable AI/ML supported detection of anomalies in ECG signals based on trained autoencoder model.	Health, Consumers	6
KAI	Traffic data input for ML-based anomaly detection	Efficient real-time delivery of traffic data to ML-based anomaly detection algorithm.	Industry, critical communications, manufacturing	6
KLAS	TRX R6 compute gateway for transportation.	Combined hardware and software solution allowing a range of applications to be supported on a single platform. Modular and open platform allowing the operator to add features over time	Industrial, transportation and automotive	9

Partner	Exploitable Item	Novelty	Target Sectors/Markets	Achieved TRL
KTH	Edge-based CPS planning algorithm Federation of rule engines Algorithms for edge-based task offloading and resource management	An algorithm for safe planning considering occlusions was developed and evaluated against state-of-the-art approaches, both in simulation and through in-vehicle (prototypical) implementation and demonstrations. Automatic rule dependency detection allows multiple stakeholders to operate own rule engines yet achieve robust integration. Developing algorithms that leverage an edge-based scheme to address computational task offloading and resource management issues.	Safe autonomous driving Industry, Manufacturing, Intelligent Transportation System Intelligent transportation system, Soft robotics	6 5 5
LCM	Anomaly detection algorithms for embedded devices	Different approaches for anomaly detection of device and communication parameters of embedded devices.	Industrial IoT in general where wireless sensor nodes are developed/applied	4
LDO	1. Crowd Management System 2. Potholes detection 3. Privacy Preserving Audio monitoring	1. Integrated platform for crowd management (counting, social distancing, man-down detection) 2. New edge-deployable detector for road damages 3. Innovative solution for privacy preserving audio anomaly detection	1. Critical Infrastructures, Police Forces 2. Transportation 3. Critical Infrastructures, Police Forces	1. 7 2. 6 3. 7

Partner	Exploitable Item	Novelty	Target Sectors/Markets	Achieved TRL
LIE BHERR	1. Data centre & network architecture for container handling equipment 2. IoT platform for container handling equipment	1. Interfacing of container handling equipment machines and devices and a platform for AI-enabled security monitoring and optionally predictive maintenance methods 2. Real-time communications between cloud server and crane PLC and scada system for real-time analytics and predictive maintenance	1. Critical Infrastructures 2. Critical Infrastructures	1. 8 2. 5
MARUN	1. AI-based jamming detection on IoT end devices 2. Object detection and localization at the edge with V2X Day-2 (CPM) demonstrator	1. analyzing the effects of jamming attacks, self-aware IoT edge devices, edge computing, real-time detection 2. awareness of nearby vehicles in busy areas, 3D object detection model, V2X Day 2 CPM application	1. Manufacturing, Industrial IoT and Automation, Cybersecurity 2. Transportation and Automotive Industry, Smart Cities, Public Safety and Security	1. 6/7 2. 5/6

Partner	Exploitable Item	Novelty	Target Sectors/Markets	Achieved TRL
MDH	Federated Learning Framework for Network Attacks Detection and Classification Based on Random Forest Framework for Feature Encoding and Data Privacy based on Autoencoders and Optimization Algorithms	New horizontal federated learning approach based on random forest, which supports different methods of aggregating independent random forests in the server. Enhancement in the evaluation function used in the optimization algorithm by adding the performance of a machine learning algorithm, as well as support for single/multi-objective optimization.	Cybersecurity, industry and manufacturing systems (for both)	6 5
MTU	Connection monitoring and management plugin	Platform and front end for AI enabled connection management	Scenarios in any industry where a mobile router with multiple uplinks is required	6
NXP-NL- CCC&S	Offline explain-ability	NXP has taken the results from the first InSecTT year on offline explainability for imaging neural nets and turned these over to NXP business, NXP business has released eIQ toolkit for NXP MCUs containing explainability features and augmentation (robustness and strengthening against Adversarial eXamples).	Software toolkit for NXP automotive and industrial MCUs	9

Partner	Exploitable Item	Novelty	Target Sectors/Markets	Achieved TRL
PAVOTEK	AI-based anomaly detection software	Detect and report real-time anomalies using advanced data analytics algorithms	Manufacturing, Industrial IoT systems and Automation, Cybersecurity	6
RISE	Driver Monitoring System	Potential to increase safe driving together with partners.	Automotive industry, Cybersecurity, IoT, Industrial Control Systems	5
RTE	Preventive maintenance for industrial IoT	Classification and anomaly detection using various sensor types and ML algorithms for one-dimensional time-series applications in industrial IoT-system.	Industrial IoT systems	5
STM	Novel STM32 MCU	Increased security in novel STM32 MCU using the activity performed within the project and the collaboration with CEA	Industrial IoT systems	6
TIETO SE	Federated Learning for Network Intrusion and Attack Detection Driver Monitoring system with computer vision	New horizontal federated learning approach based on random forest, which supports different methods of aggregating independent random forests in the server. Extension of the work by including privacy preserving to the distributed model thanks to differential privacy on random forests. Joint work with MDH Increase safe driving using AI and ML technique	Cybersecurity, industry and manufacturing systems (for both) Automotive industry, Cybersecurity, IoT, Industrial Control Systems	6 5

Partner	Exploitable Item	Novelty	Target Sectors/Markets	Achieved TRL
TU-DELFT	AI based ECG anomaly detection Single anchor localization for cooperative multiagent systems	GMM vector field dynamical system which sythesizes and compresses ECG signals in real time has been developed. This is augmented with real-time, training free and storage free AI algorithms to classify anomalies, via computationally simple, edge-device compatible methods. The synthesizing framework is also extended to patient privacy preserving ECG dataset synthesis. Smart, real-time localization algorithm usin a single, possibly mobile anchor node has been developed, tested using UWB receiver. Applications to foration flight of fixed wing UAVs in leader-follower configuration was demonstrated	Healthcare IoT Systems and clinical pathways. Automotive IoT and Warehouse automation, formation of multiagent sytems	4 5
U-TWENTE	Autonomous Network Slicing and AI based QoS management in IoT Networks	AI network manager for network slicing in wi-Fi networks with automated traffic classification and uniform quality of service between wired and wireless networks.	Industrial IoT Networks and Wi-Fi based Solutions for enterprise	5
UCC	Remaining useful life prediction	Improved planning of maintenance operations	Industrial crane systems	6
	Video camera anomaly detection	Verification of camera position and orientation, and detection of tampering.	Scenarios in which video camera verification is critical, e.g., remote operation of machinery	6

Partner	Exploitable Item	Novelty	Target Sectors/Markets	Achieved TRL
UPM		Custom HW platform for Deep Learning (DL) at the Edge for LIDAR and camera with dedicated processing accelerator for DL tasks	Low power extreme edge IoT hardware platform for deep learning acceleration and processing, using LiDAR point-clouds as input data	5
VIF	vehicleCAPTAIN toolbox (routing core and ITS library) Trustworthiness Whitepaper on INSECTT Webpage	Development kit for fast entry and early-stage development of V2X technology	Research Organizations, Startups and Industry Partners	7
VTT	AI-Enriched Multimodal, Distributed Remote Vital Sign Analytics	Distributed AI-augmented sensor data processing in device-edge-cloud continuum; A specific example: Unobtrusive vital sign acquisition from noisy camera and milli-meter radar data sources with AI-based estimated value and confidence assessment.	Healthcare, IoT systems	7
WAPICE	Machine vision-based object detection	Potential to increase safety and overall situational awareness in smart areas	Smart cities, Industry, Logistics, Healthcare	7

Table 1 Overview of the expoloitable components in WP2.

3.3 Partner results

3.3.1 AKEO PLUS

Hardware-enabled secure generation of data proofs and traceability for embedded industrial systems

The objective of the developed solution is to securely attest data produced by an embedded system (e.g. embedded neural network) in an industrial scenario, in order to ensure the integrity, sequencing and authenticity of the generated data, while allowing to posterior verification.

At the device level, the attack surface is reduced through the use of hardware secure elements (STSAFE-A110, CmStick dongle). At the application level, the use of a blockchain (Ethereum) secures the proof history with a distributed topology, and thus allows to detect attacks such as data manipulation.

This architecture furthermore allows to integrate foreign actors in the application, guaranteeing each actor's sovereignty on its own produced data, while preserving complete transparency to cryptographic proofs on this data.

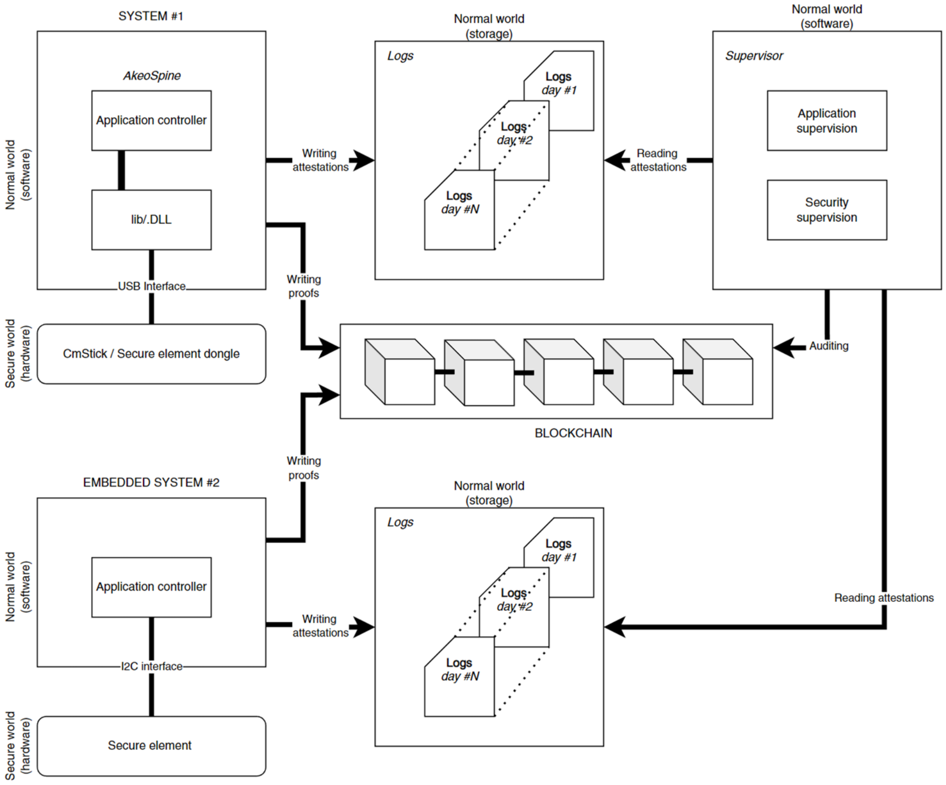


Figure 3 Illustration of the hardware-enabled secure traceability architecture developed

Contact person: Aurélien Ibanez (a.ibanez@akeoplus.com)

3.3.2 ARCELIK

Cybersecurity Assessment Methodology

ARCELIK has developed a questionnaire-based assessment methodology for evaluating cybersecurity level of a communication network used in a manufacturing system. In addition, the methodology can determine reliability of communication techniques used in a wireless or wired communication network and if the techniques are complaint with the industry standards and regulations. In this questionnaire, there are multiple choice and yes/no questions that analyse the system from various aspects of cybersecurity that a communication system used in manufacturing should have. By answering this questionnaire, cybersecurity level of the system can be determined in a scale ranging from 1 to 10. This questionnaire was developed in collaboration with the IT team.

Contact person: Baturay Dalgıç (baturay.dalgic@arcelik.com)

3.3.3 AVL

AI-based automata learning algorithms (Wireless-Testbed)

AVL has utilized existing automata learning algorithms to learn models of wireless communication protocols and states of target devices. By integrating automata learning methodology into the framework for the Wireless-Testbed of Use Case 5.03, we aim to subsequently derive test cases and therefore to increase the test coverage significantly.

By combining algorithms that utilize the LearnLib and AutomataLib libraries with a software interface as well as physical component (e.g., Nearfield Communication, Bluetooth Classic or Bluetooth Low Energy adapters), the corresponding connectivity interface on the target can be learned as shown in Figure 3. This allows us to infer a state model, which gives insights into the behavior and possible security flaws of the target.

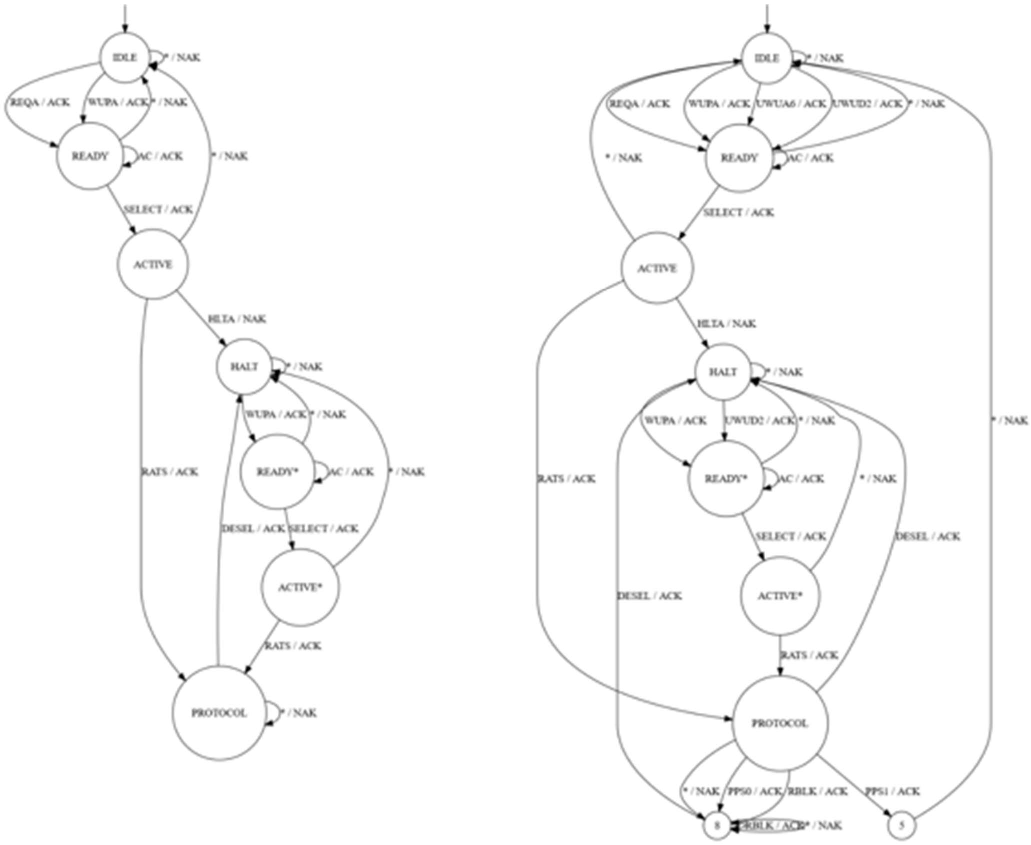


Figure 4 Learned NFC State model in comparison to protocol standard

The integration of these components into the software framework for the wireless testbed allows to easily parametrize and execute fuzz tests against targets and furthermore develop attacks based on their findings. For fuzz-testing, test cases may be generated by using different strategies e.g., using symbols that induce a state change and fuzz the respecting input fields. Furthermore, test cases can be generated by using model checking techniques on the inferred model and using specification violations as input for the test cases. Figure 4 displays the learner setup for NFC.

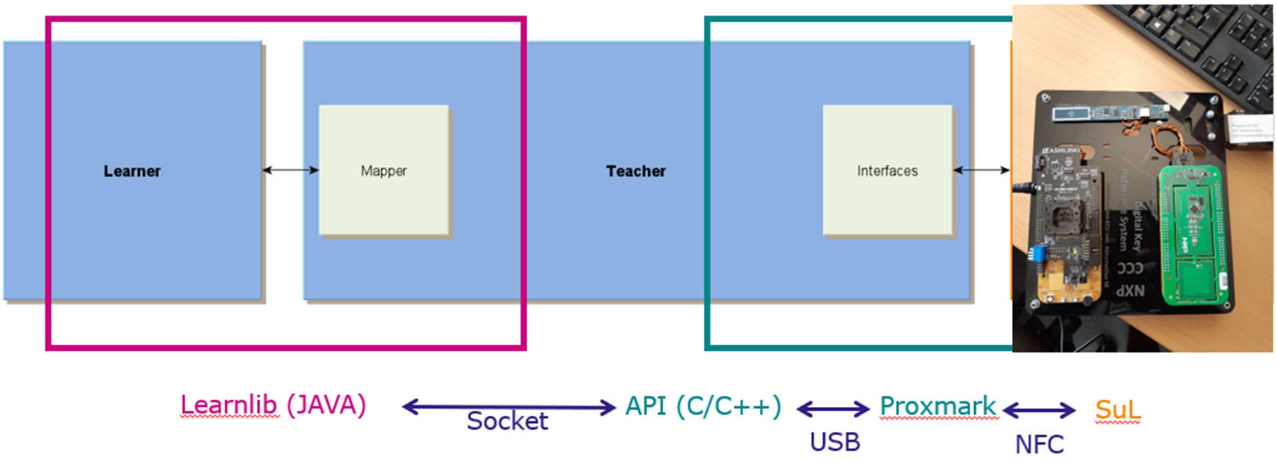


Figure 5 NFC Learner Setup

Contact person: Stefan Marksteiner (stefan.marksteiner@avl.com)

3.3.4 CAPGEMINI

5G Platooning Control and Management System Simulation

The development of autonomous driving technology requires reliable and low-latency communication networks that can support data transfer and coordination in real-time. 5G technology promises to meet these requirements and enable the deployment of autonomous vehicles. A software system simulation depicted in Figure 6 was developed for wireless networks based on 5G that allows the creation of a virtual environment that replicates real-world conditions. The virtual environment includes the User Equipment (in this case vehicles or platoons), communication networks and the traffic patterns. So, the system implements the RAN cells allowing the creation and validation of new propagation models, allowing to measure the conditions of the network considering the UE mobility. The system allows configuration and the validation of different propagation models, network slicing functionalities, generating as output performance measurements which can serve as synthetic data serving as input to different analytics applications and to train ML modules.

The system implements some features of O-RAN standardization, being integrated with non-RT RIC (or SMO in this case ONAP) and near-RT RIC, allowing the metrics collection via O1 interface (VES) and policy enforcement via A1 interface to optimize the network resources. Also, being integrated with non-RT RIC will allow the management and guidance of policies, to continuously optimize the network.

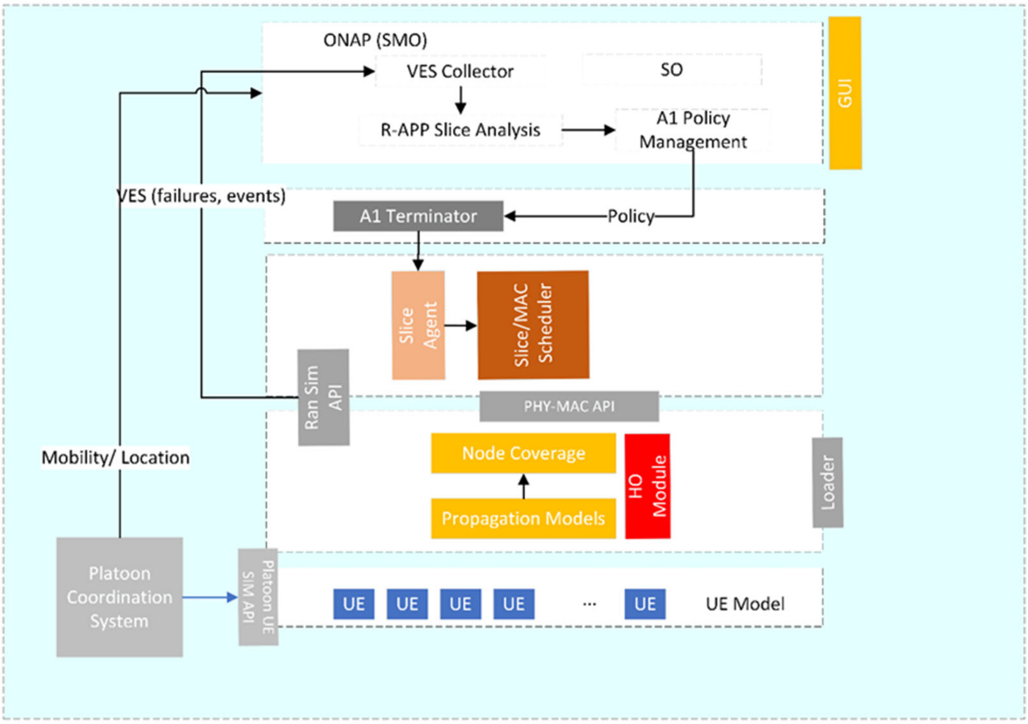


Figure 6 5G Platooning Control and Management System Simulation

Contact person: Paulo Duarte (pauloalexandre.duarte@capgemini.com)

3.3.5 CEA

HistoTrust: Traceability and Trust of Embedded Neural Network Models

With the large-scale deployment of machine learning models for many applications domains using a wide variety of hardware platforms, the security of AI-based systems is becoming an urgent need, highlighted with upcoming regulation, standardization and certifications actions that will shape the European approach for Artificial Intelligence (AI). Security of ML models encompasses a very complex attack surface because more and more models will be embedded in mobile devices, physically accessible by adversaries. In this context, an important challenge is to protect the integrity of model inferences to increase traceability and trust.

Thanks to InSecTT, CEA-LETI has developed the HistoTrust platform (see Figure 7 and Figure 8) that aims at authenticating the issuing device and protecting – at the device level – the integrity of an embedded AI (neural network model) by combining software and hardware security components using a high-end IoT platform proposed by STMicroelectronics (STM32MP1) that used a dual architecture (Cortex-M4 and Cortex A). HistoTrust takes benefit from blockchain technology to bring trust in the traceability of AI behaviour and help its explainability. HistoTrust attests in an Ethereum ledger all the relevant data produced by a physical device, especially the heuristics inferred by AI as well as some of its parameters. Thus, the audition of the ledger allows security verifications and AI behavior analysis. Importantly, the robustness of the embedded neural network model is strengthened against advanced integrity-based attacks that may leverage algorithmic and physical flaws as extensively studied in InSecTT and published in [22], [23], [24] and [25]. HistoTrust has been detailed in the journal *Annals of Telecommunications (2023)* [26].

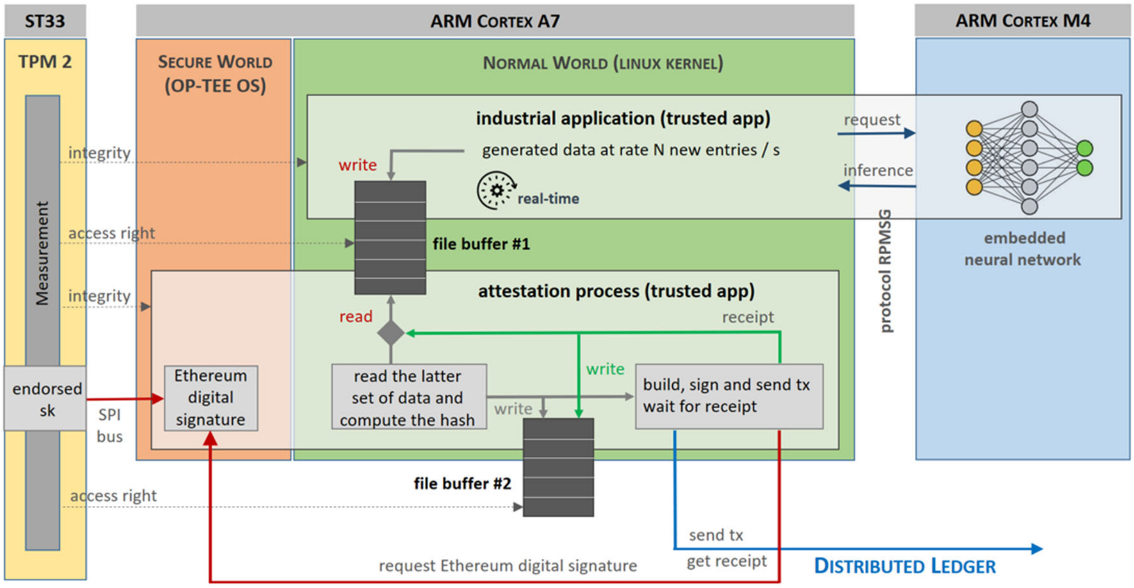

ST33
TPM 2
Measurement
integrity
access right
integrity
endorsed sk
SPI bus
access right
ARM Cortex A7
Secure World (OP-TEE OS)
Ethereum digital signature
Normal World (linux kernel)
industrial application (trusted app)
generated data at rate N new entries / s
write
real-time
file buffer #1
attestation process (trusted app)
read
receipt
write
read the latter set of data and compute the hash
write
build, sign and send tx wait for receipt
file buffer #2
request Ethereum digital signature
send tx
get receipt
request
inference
protocol RPMSG
ARM Cortex M4
embedded neural network
Distributed Ledger

Figure 7 HistoTrust platform on ARM-based processor HW.

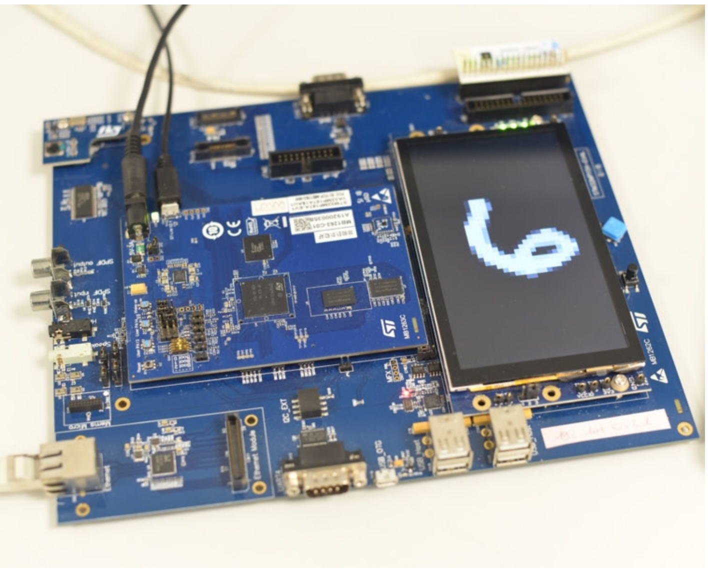


Figure 8 A prototype board with STM32MP1 running HistoTrust.

Contact persons: Christine Hennebert (christine.henneber@cea.fr), Pierre-Alain Moellic (pierre-alain.moellic@cea.fr)

3.3.6 CINI-UNIPR

Multi-Interface Gateway (MIG)

CINI-UNIPR has performed an experimental investigation of a prototypical Multi-Interface Gateway (MIG) developed to test its connectivity robustness. The prototypical hardware of the MIG is shown in Figure 9 and its internal representation is shown in Figure 10.

In particular, the MIG integrates heterogeneous networking technologies, supporting dependable wireless communications with focus on both short-range (i.e., BLE and WiFi) and long-range (i.e., LoRaWAN) connectivity ([6]), with an interface handler for each supported communication protocol, as well as a smart data broker in charge of applying enhanced techniques for smart routing policies and rules adoption ([7], [8]). Then, a state transition-like analytical model (shown in Figure 11) has been defined, where an additional aggregation mechanism is applied on the traffic packets to cope with possible transmission delays.

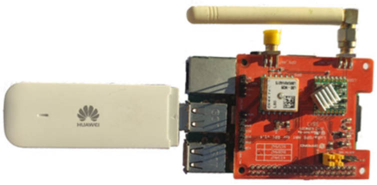


Figure 9 Prototypical MIG implementation equipped with different short- and long-range communication interfaces

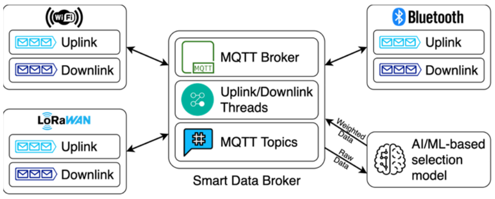


Figure 10 Block representation of the Multi-interface Gateway (MIG) proposed by CINI-UNIPR

As a reference example, in order to better estimate how the proposed MIG can perform, Figure 12 depicts a comparison among analytical and simulated interface server utilization ratio as a function of a parameter of the packets generator (for the BLE interface, among all those available in the MIG), while a comparison between analytical and simulated average service time, and number of information units per aggregated packet, as a function of the spreading factor allowed by the LoRaWAN protocol, is shown in Figure 13.

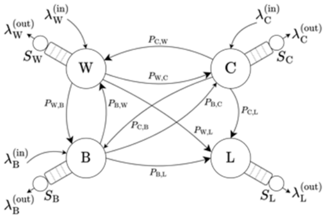


Figure 11 States transition representation of the Markov chain-based model regarding the MIG.

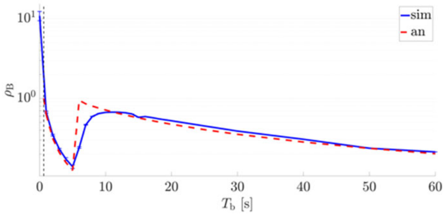


Figure 12 Comparison among analytical and simulated interface server utilization ratio as a function of a parameter of the packets generator, for the BLE interface.

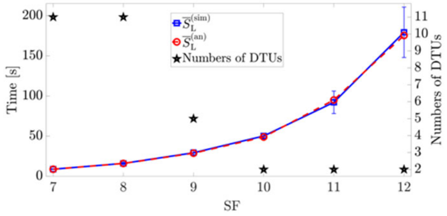


Figure 13 Analytical and simulated average service time and number of information units per aggregated packet as a function of the spreading factor allowed by LoRaWAN.

The MIG has been then applied in two reference scenarios in which CINI-UNIPR has been involved in the aim of the EU-funded InSecTT project, namely (i) inside a public transport bus moving during daily trips and (ii) in the pre-checks extra-Schengen area of an airport, even in order to analyze the latency experienced in daily activities, as well as the capacity to internally host enhanced data processing mechanisms (e.g., air quality sensing and prediction).

Contact persons: Gianluigi Ferrari (gianluigi.ferrari@unipr.it), Luca Davoli (luca.davoli@unipr.it), Laura Belli (laura.belli@unipr.it)

Air quality sensing and prediction through AI

CINI-UNIPR has implemented distributed processing strategies between the device level (with preliminary COTS-based environmental sensing nodes deployed in heterogeneous environments) and the edge level (even located near the environment in which on-field data are collected) in the scenarios of interest for the InSecTT project.

In particular, as shown in Figure 14, a preliminary prototypical air quality sensing node has been developed for air quality data collection, in detail composed by a COTS processing board, and by three I^2C- and SPI-connected COTS sensors, namely: (i) a MiCS5524 sensor sensitive to CO, ethanol, H_2, NH_3 (ammonia), and CH_4 (methane/propane/iso-butane) gases; (ii) a Sensirion SCD-30 sensor measuring CO_2, humidity, and temperature; and (iii) a Sensirion SPS30 sensor able to measure particulate matters (i.e., $PM_{2.5}$ and PM_{10}).

Then, in order to verify its workload, different replicas of this sensing node have been deployed in the use cases which CINI-UNIPR has been involved in, in the aim of the InSecTT project, namely (i) inside a public transport bus moving during daily trips (Figure 15) and (ii) in the pre-checks extra-Schengen area of a public airport (Figure 16).

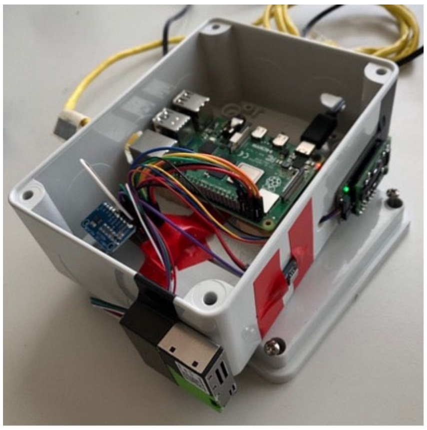


Figure 14 Prototypical air quality sensing node based on COTS processing board and featuring three COTS sensors.

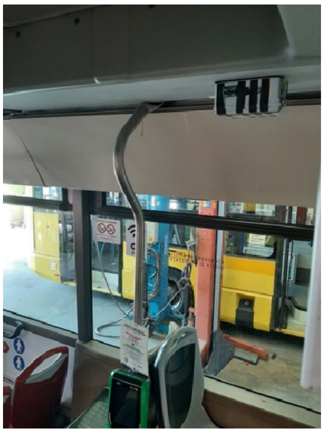


Figure 15 Experimental deployment inside a public transport bus.

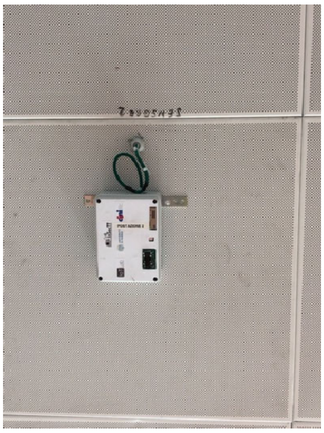


Figure 16 Experimental deployment inside the pre-checks extra-Schengen area in a public airport.

Then, given the data possibly being collected through these prototypical sensing nodes, CINI-UNIPR moved its focus on collecting air quality data to design efficient AI-based algorithms and proper deployment strategies ([9]). In detail, a comparative Artificial Intelligence of Things (AIoT)-oriented performance evaluation of different ML and DL algorithms for air quality prediction on embedded devices has been performed, focusing on the following algorithms:

- ML: k-Nearest Neighbors (k-NN), Decision Tree Regression (DTR), Support Vector Regression (SVR), Artificial Neural Networks (ANNs);
- DL: Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), Convolutional Neural Networks (CNN), Gated Recurrent Unit (GRU).

Therefore, CINI-UNIPR explored the possibility to embed the execution of these air quality prediction algorithms directly inside the MIG, targeting the deployment of edge computing-oriented AIoT-like systems. An example is shown in Figure 17, where a GRU-based processing has been applied on the $PM_{2.5}$ prediction.

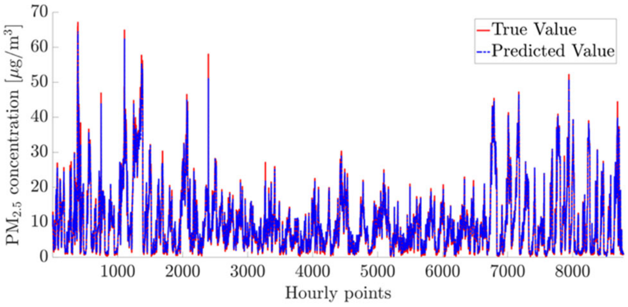


Figure 17 GRU-based air quality prediction.

Dataset:

CINI-UNIPR foresees the publication of an open dataset at the end of the InSecTT project, containing the air quality parameters values collected in the heterogeneous contexts of interest, namely (i) a public transport bus moving during daily trips and (ii) in the pre-checks extra-Schengen area of a public airport.

Contact persons: Gianluigi Ferrari (gianluigi.ferrari@unipr.it), Luca Davoli (luca.davoli@unipr.it), Laura Belli (laura.belli@unipr.it)

3.3.7 CINI-UNIROMA3

Biometrics on the Move

The BioMedia4n6 group of the UNIROMA3 subunit within CINI has developed a multi-biometric system relying on the joint use of face and hand vein pattern to deal with structured flows of passengers in an airport, within the activities carried out for the UC 5.16 Airport security - Structured and Unstructured People Flow in Airports.

In more detail, the developed system consists in a dedicated self-enrolment kiosk and a recognition gate, implemented through stand-alone boards, equipped with the hardware and software tools required to interact with the users and acquire their biometric traits, which are employed to perform biometric recognition.

Vein patterns are captured using cameras sensitive to near-infrared (NIR) light. Since this latter is absorbed to a greater amount by the oxygen in the blood, with respect to the surrounding tissues, imaging systems relying on this radiation are able to capture pictures where subcutaneous vein appear darker than the rest of the image. On the other hand, face images can be captured with standard webcams.

The use of two distinct and independent biometric characteristics, their acquisition through standard cameras, and the ability to control both the enrolment and verification process through a stand-alone board, allow to implement an automatic biometric recognition system robust against possible spoofing attacks, and cost-effective, since it requires extremely low-budget components. The developed prototype is shown in Figure 18, where the whole device required for the verification stage is shown, together with the NVIDIA Jetson Nano stand-alone board employed for the enrolment stage, connected with a NIR camera with a serial port, a monitor with an HDMI cable, a webcam and an input interface via USB, and to the paired board used for verification through an Ethernet connection.

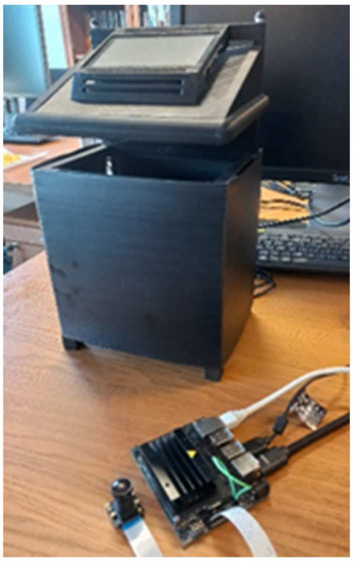


Figure 18 Prototype developed for the proposed multi-biometric system.

The lower part of the image shows the used stand-alone board, connected with a NIR camera through a serial port, and with connections for USB webcams, HDMI, and Ethernet.

The processing pipeline shown in Figure 19 has been developed to capture the hand vein patterns of users interacting with the devices with an on-the-fly modality, that is, users are simply asked to pass their hand over the deice to have their vein patterns acquired. Two cameras, set with different exposure times, have been employed in the proposed device, in order to capture more information and use it for verification purposes.

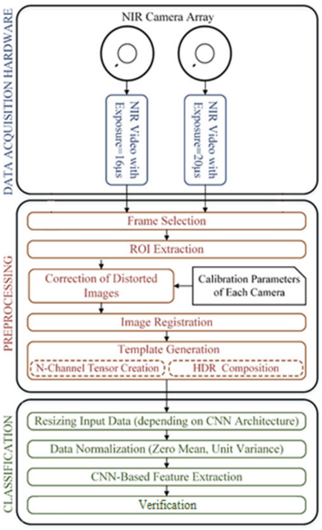

DATA ACQUISITION HARDWARE
NIR Camera Array
NIR Video with Exposure=16μs
NIR Video with Exposure=20μs
PREPROCESSING
Frame Selection
ROI Extraction
Correction of Distorted Images
Calibration Parameters of Each Camera
Image Registration
Template Generation
N-Channel Tensor Creation
HDR Composition
CLASSIFICATION
Resizing Input Data (depending on CNN Architecture)
Data Normalization (Zero Mean, Unit Variance)
CNN-Based Feature Extraction
Verification

Figure 19 Pipeline of the employed vein patter acquisition and processing.

The acquired face and vein traits have been processed with the aim of deriving discriminative information, allowing a subject to be recognized effectively. In order to train neural networks for such aim (see Figure 20), we have resorted to a training strategy relying on additive angular margin penalty (AAMP) instead of the traditional SoftMax. Such training strategy allows for deriving representations usable also for subjects not taken into account during the training stage. The trained networks are therefore employed only as feature extractors, with the obtained representations compared through a cosine distance to perform verification ([10]).

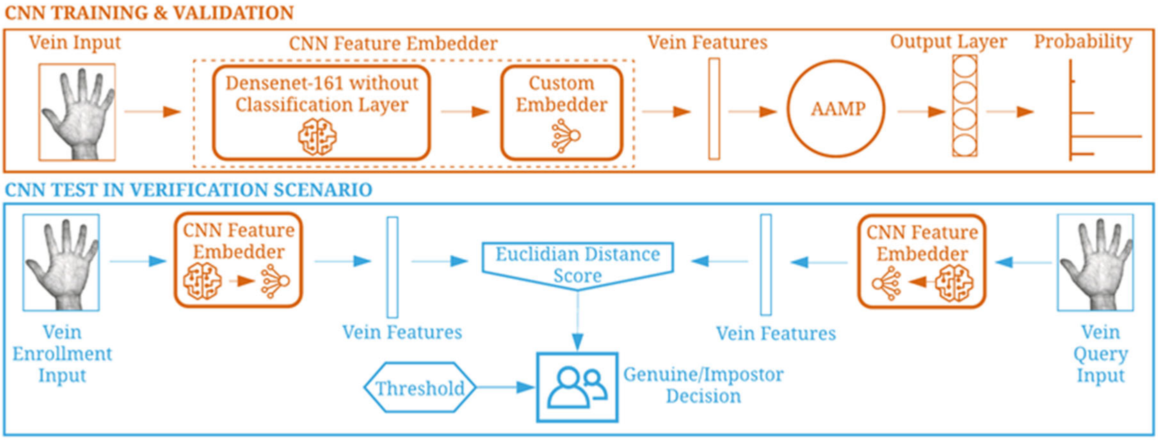


Figure 20 Training and operative phases of the employed neural networks for biometric verification.

The traits captured during the enrolment stage are processed on-board, and the generated representations are sent to the board dedicated to the verification stage through an ethernet connection. The original traits are not stored in the system for privacy reasons, only the embeddings are required at the verification stage.

On the explainability side, several aspects have been taken into account to shed light on specific characteristics of vein patterns, which are a trait still relatively underexplored, especially with respect to face.

In more detail, the similarity of vein structures in the finger, palm, and dorsal regions of the two hands of a person has been investigated ([11]). The results obtained indicate that significant similarities between corresponding vein patterns of different hands exist. Nonetheless, the performance achievable performing recognition using the trait belonging to a hand other than the one employed during enrolment is quite poor. It is yet possible to train a network in such a way that the vein patterns of different hands are considered as the same class. The obtained findings are similar to those given in literature for ear and in palmprint: similarities are found in left and right biometric traits, allowing to compute an equal error rate (EER) in the order of 10% for finger-vein data from SDUMLA database, and even lower over the dataset collected with the device developed within the INSECTT project.

The relevance of gender-specific characteristics in the exploitation of hand vein patterns for biometric recognition purposes has been also investigated ([12]). In more detail, tests have been performed to evaluate whether gender recognition can be performed through the analysis of hand vascular patterns. The obtained results testify that recognition rates similar to those achievable with face data can be actually accomplished. It has been also evaluated whether the anatomical specific characteristics of female and male populations could affect the discriminatory capabilities of the templates extracted from the considered traits and employed for user recognition purposes. Actually, it has been observed that the score distributions associated to vascular patterns from female subjects are characterized by larger intra-class and lower inter-class values, with respect to those related to male subjects. A novel gender-aware pipeline to be used for people verification has been therefore proposed.

Eventually, tests have been performed in order to evaluate whether radiation with wavelengths longer than those characterizing NIR radiation have been performed ([13]). In more detail, a biometric recognition system based on the electromagnetic interaction of a subject's hand with the emission of an antenna working in the X-band has been studied. The performed tests testify the existence of permanent discriminative characteristics within the measured signals.

Contact persons: Emanuele Maiorana (emanuele.maiorana@uniroma3.it), Patrizio Campisi (patrizio.campisi@uniroma3.it)

Synthethic Audio Dataset for Sound Event Detection

The COMLAB group of the UNIROMA3 subunit within CINI has developed a Coarse-to-Fine approach for both detecting and classifying anomalous events from an audio recording.

This framework can simultaneously detect if an audio recording is anomalous and, if this is the case, to identify which and where dangerous events are occurring. Specifically, this architecture is composed of two elements. The first is responsible for modelling the normal background of a target environment in an unsupervised fashion. If an anomalous audio is detected, a second element focuses on what and when the anomaly occurs.

This second block, i.e., the fine classifier, is a sound anomaly detection system based on a fully convolutional network which exploits image spatial filtering and an Atrous Spatial Pyramid Pooling module. To cope with the lack of datasets specifically designed for sound event detection, a dataset for the specific application of noisy bus environments has been designed. The dataset has been obtained by mixing background audio files, recorded in a real environment, with anomalous events extracted from monophonic collections of labelled audios. The performances of the proposed system have been evaluated through segment-based metrics such as error rate, recall, and F1-Score. An example of predicted and ground truth detections are depicted in Figure 21. Moreover, robustness and precision have been evaluated through four different tests. The analysis of the results shows that the proposed sound event detector outperforms both state-of-the-art methods and general-purpose deep learning-solutions [14]. The overall architecture is displayed in Figure 22.

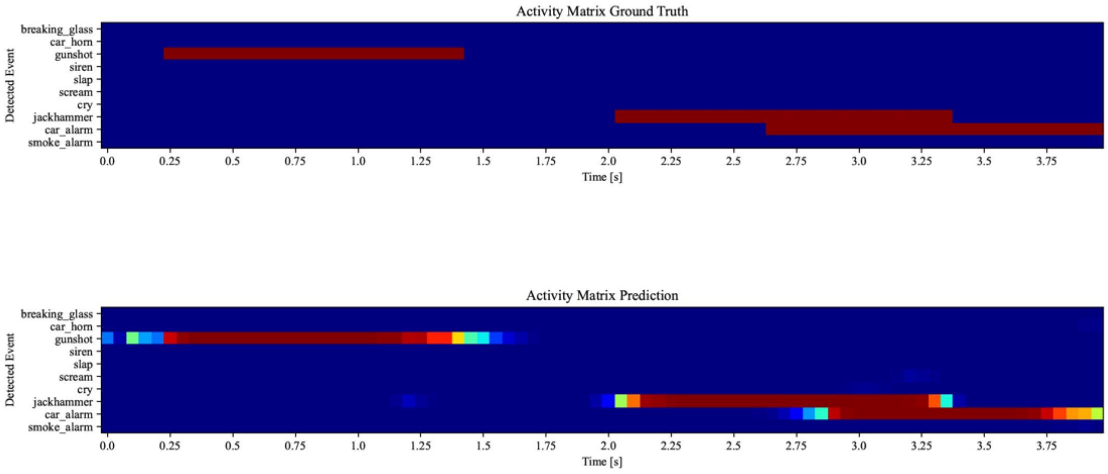


Figure 21 Example of predicted and ground truth detections in order to evaluate segment-based metrics.

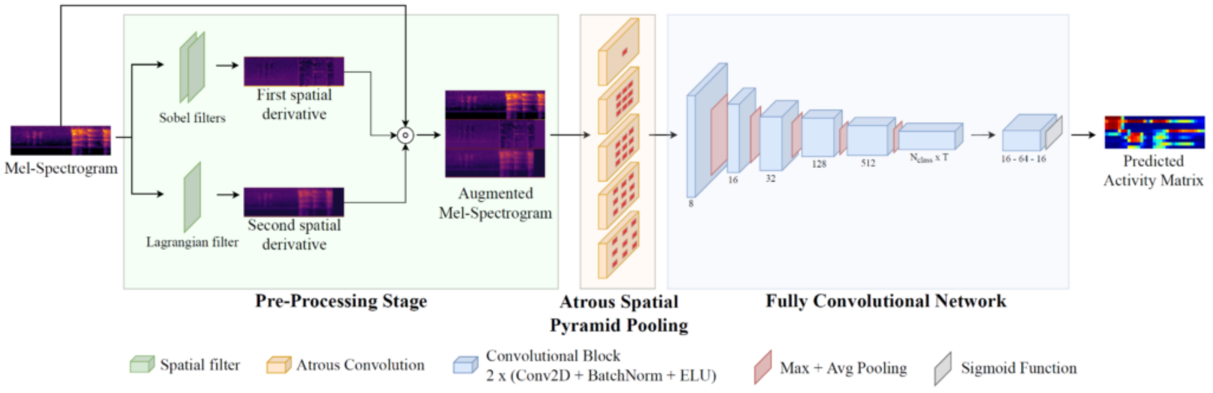


Figure 22 AuSPP model for fine Sound Event Detection.

Dataset:

An audio dataset, which is composed of real background recordings on the bus and anomalies' sounds from well-known state-of-the-art datasets, has been provided to the research community. The dataset with relevant source code is available open source at the following GitLab repository: https://gitlab.com/michael.neri/sound-event-detection-for-human-safety-and-security-in-noisy-environments.

Contact person: Michael Neri (michael.neri@uniroma3.it)

3.3.8 CINI-UNIMORE

Synthetic Dataset for People Surveillance

To overcome the lack of data of the currently available real datasets, the computer vision community proposed to extract multimodal synthetic data from videogames and computer-generated scenarios.

Our proposal is **MOTSynth** a large synthetic dataset extracted from the videogame Grand Theft Auto V (GTA-V), specifically designed for training models for pedestrian detection, tracking and segmentation (see Figure 23 and Figure 24). All the 764 videos are recorded in Full HD resolution at 25 FPS. In each video there are flows of virtual actors moving around in different paths while avoiding collisions among each other. The actors are random by varying generative attributes of 579 pedestrian models, changing the clothes, backpacks, bags, masks, hair and beard style, yielding over 9519 unique pedestrian identities, that is suitable for training a re-id model and multiple people tracking.

The dataset contains video clips, precise 3D annotations of visible and hidden body parts, consistent 2D bounding boxes, and segmentation masks for pedestrians. The dataset also provides depth maps that are valuable cues in multiple people tracking and may contribute to future advancements. MOTSynth is the most comprehensive and offers a high degree of variability in scenarios, number of entities, and types of annotations.

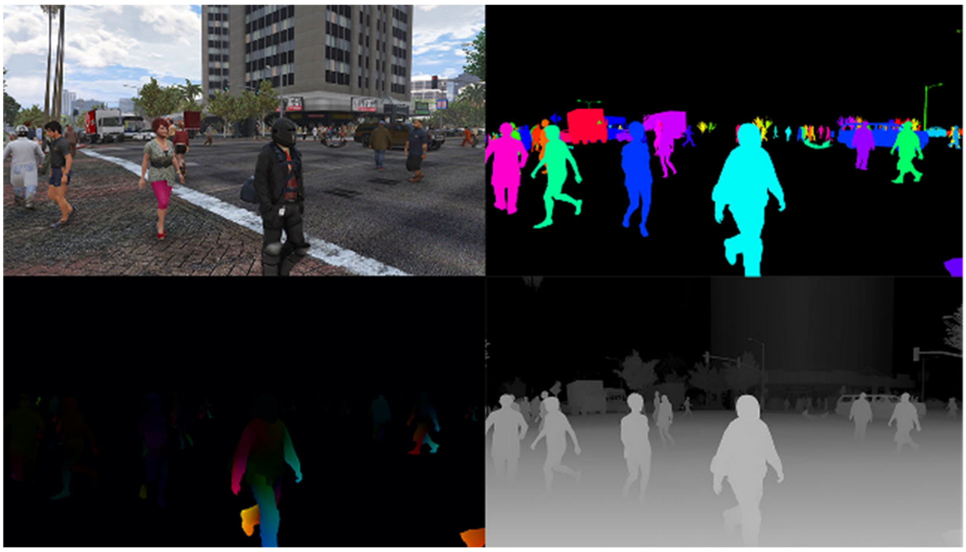


Figure 23 A frame of MOTSynth in a bustling urban scenario with mix of pedestrians and vehicles.

(top left)	The crowd of pedestrians of high variety, featuring individuals of different ages, genders, and ethnicities, making for a realistic representation of an urban environment.
(top right)	Segmentation view.
(bottom left)	Optical flow (near-distance) view.
(bottom right)	Depth map view.

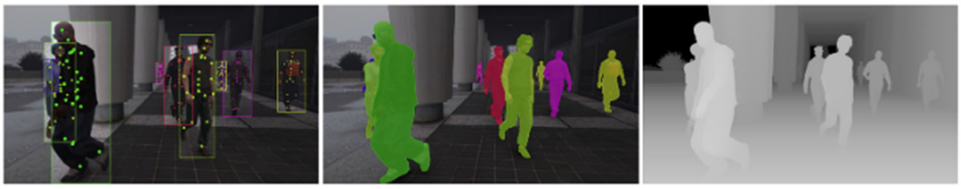


Figure 24 The visual representation of bounding box, segmentation, and depth map annotations in MOTSynth.

It contains over 40 million bounding boxes and over 1.3 million densely annotated frames, with an average of 29.5 people per frame and a maximum of 125 people. The actors in the dataset range in distance from 0 to 101 meters from the camera, resulting in bounding box heights between 0 and 1,080 pixels. The dataset was split into a training set of 576 clips and a validation set of 192 clips, with an effort to balance weather conditions, daytime, and density.

In terms of size, MOTSynth is superior to all other previously proposed datasets, with more instances and labels. In addition, MOTSynth volume of data, diversity of scenarios, and people variability allow it to bridge the synthetic-to-real gap in research. Overall, MOTSynth comprehensive annotations make it a valuable resource for researchers in the field. Further details have been published in [15].

Dataset:

As described in above paragraphs, **MOTSynth** ([16]) is a synthetic dataset for people surveillance in urban scenarios. The dataset has been presented and is available at consortium level (see [16]).

The dataset is first of a kind and has been proved to be as effective as normal data in training neural network architectures for common security task such as people detection, tracking and reidentification.

Contact persons: Simone Calderara (simone.calderara@unimore.it), Rita Cucchiara (rita.cucchiara@unimore.it)

Anomaly Detection

In terms of **video anomaly detection**, we developed a novel training technique to deal with video level annotation. In conventional anomaly detection training, when provided, the annotation is at video or clip level (e.g. normal/abnormal). The activity was to develop a novel multiple instance learning method specifically tailored for anomaly detection to exploit weak video level annotation using multiple instances learning on video anomaly detection deep backbones. The objective is to both classify the anomaly and extract the sub-clip in which the anomaly occurs in order to obtain a useful explanation of the model. By using weakly Supervised Anomaly detection ([17]), the predictor is allowed to learn not only from normal examples but also from a few labelled anomalies made available during training. In particular, we deal with the localization of anomalous activities within the video stream: this is a very challenging scenario, as training examples come only with video-level annotations (and not frame-level). Several recent works have proposed various regularization terms to address it i.e. by enforcing sparsity and smoothness constraints over the weakly learned frame-level anomaly scores. The idea is to ask the model to yield the same scores for different augmentations of the same video sequence as depicted in Figure 25.

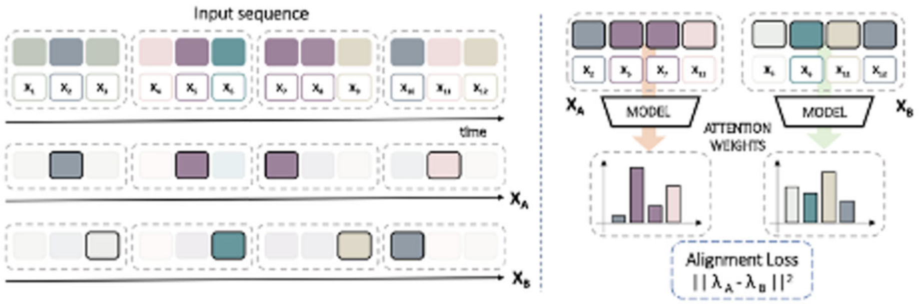

Input sequence
x_1 x_2 x_3 x_4 x_5 x_6 x_7 x_8 x_9 x_{10} x_{11} x_{12}
time
X_A
X_B
MODEL
MODEL
ATTENTION WEIGHTS
Alignment Loss
$||\lambda_A - \lambda_B||^2$

Figure 25 Overview of the proposed video anomaly detection framework.

(left) Augmentation function sampling two slightly different sequences out of a single one.

(right) The original sequence gets split into windows and for each, we randomly sample a single feature vector.

We test our contribution on:

- The XD Violence Dataset surpassing the state of the art ([18]).
- The Use Case 5.15 private dataset about public transport with the original backbone depicted in Fig. 3 specifically designed to be implemented and reach real time performances on NVIDIA Xavier boards. For the purpose we design and implement a deep learning model that carries

out the task at hand as a binary classification problem “normal vs. anomalous activity” (i.e., purely supervised learning). The model is based on a ResNet 2(D+1) structure instead of the traditional networks based on 3D convolutions. Such an architecture has proven to be easily optimizable and easier to deploy on the target hardware architecture (i.e., NVIDIA Xavier Board). The technical idea behind ResNet 2(D+1) is to approximate the 3D convolutions by a twofold sequential operation, the latter being composed of a 2D convolution in the spatial dimension followed by a 1D convolution in the temporal domain. Moreover, we use the Focal Loss as the objective of optimization, thanks to the beneficial effects it provides in presence of an imbalance between the number of examples per class. We conduct several experiments on proprietary private dataset of events in public transport and obtain a value of the ROC-AUC (which is a popular metric in the field) approximately equal to 0.81 over a maximum of 1.

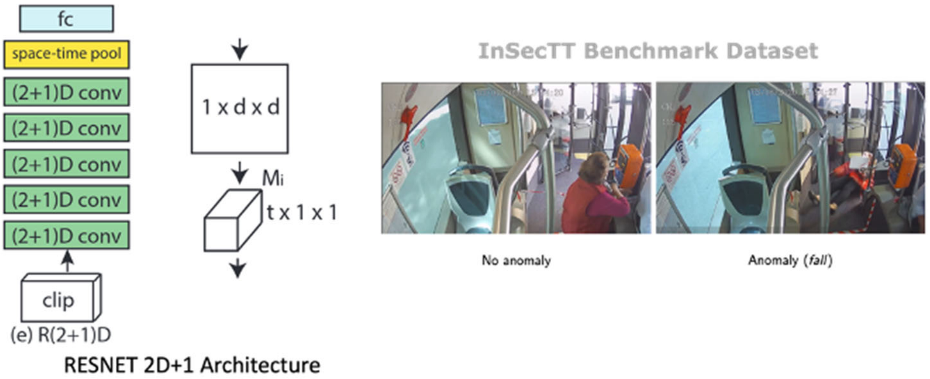


Figure 26 Overview of the Anomaly detection architecture designed to be implemented on the embedded device

Dataset:

For the task we collected a dataset on public transport anomalies in the context of both WP2 and U 5.15 activities. The dataset is composed by long, untrimmed real-world surveillance videos with 11 realistic anomalies recorded with a series of CCTV cameras placed inside SETA buses. The dataset has been acquired inside the INSECTT project and comprises of up to 5 cameras with multiple angles; 31 hours of video; divided in 182 sequences. Dataset is available to the public and to the consortium. Details and availability on the dataset can be found at AImageLab website (see [19]).

Contact persons: Simone Calderara (simone.calderara@unimore.it), Rita Cucchiara (rita.cucchiara@unimore.it)

Explainability AI pipeline

We integrated our weakly supervised Anomaly detection pipeline with **Explainability Class Activation Maps** developed in the previous stage of the project ([20]).

In particular during inference, we refine the anomaly scores by applying a post-processing step. Usually, two segments considered paramount by the model are interleaved by “holes”, mostly due to noisy acquisitions or poor representations. The purpose of this phase is therefore to merge temporally close detections in a single retrieved candidate. In particular, we initially take out from the candidate set all those time-steps whose corresponding attention scores are lower than a certain threshold (in our experiments, < 0.35). The remaining non-zero scores will be used to generate the temporal proposal.

To generate the proposals, we do not use the rough coefficients, but instead a more refined version. In particular, we compute a 1-d activation map in the temporal domain, called Temporal Class Activation Map (T-CAM), which indicates the relevance of the segment t in the prediction of one of the two classes involved (normal vs anomalous). Furthermore, we extract the Weighted T-CAM, which combines the MIL attention weight and the T-CAM activation values. Figure 27 shows a visual comparison by our MIL augmented T-CAM, original ones and ground truth temporal location of the anomaly.

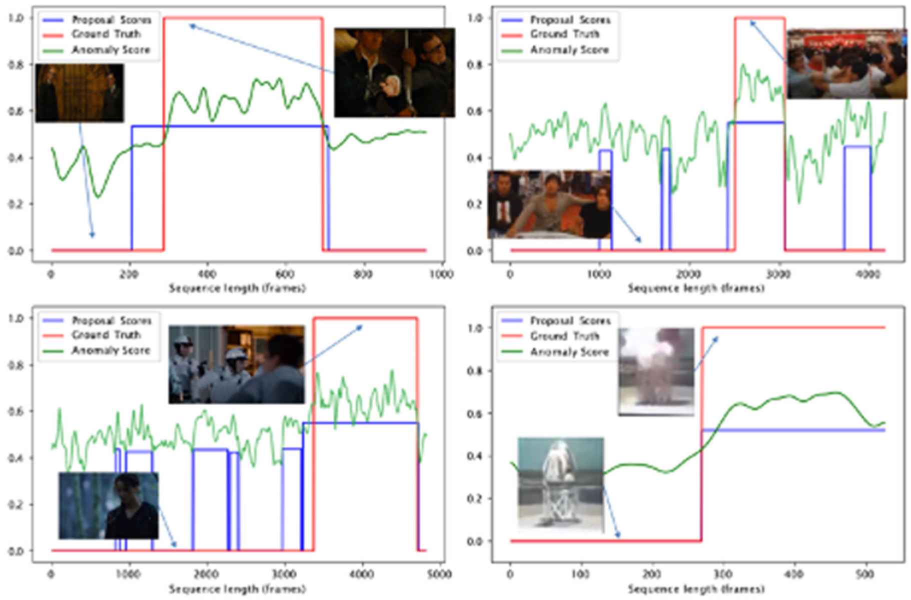


Figure 27 Qualitative examples on capabilities of the MIL augmented TCAM model to perform anomaly localization.

The temporal proposal scores are indicated with a blue line, while the weighted T-CAM scores and the ground truth are shown in green and red, respectively.

Figure 28 shows some visual results on the TCAM maps on the SETA dataset from UC 5.15 by using LayerCAM and enhancing it with our temporal attention scores. In the examples, we can see a man arguing and a person falling. In both cases we see that the network has higher activations where the anomalous action is taking place (red areas).

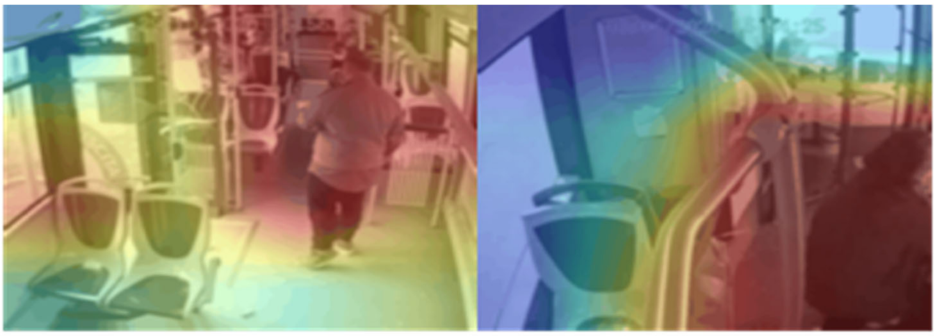


Figure 28 A visual example on T-CAM explanations on private public transport dataset.
(left) A person arguing
(right) A person falling.

To summarize the above contribution, the novelty of our technical solution can be summarized as follow:

- Design an efficient anomaly detection architecture that can perform in real time
- Design a novel training loss based on MIL that allows to exploit weak video level annotation for detecting frame level anomalies
- Integrate the MIL coefficient to enhance frame level explainability CAM maps obtaining a better frame level anomaly score and CAM map.

The aforementioned proposals can be applied to video anomaly detection and not limited to the public transport use case.

Contact persons: Simone Calderara (simone.calderara@unimore.it), Rita Cucchiara (rita.cucchiara@unimore.it)

3.3.9 ETH

The exploitation of the two project results indicated in Table 1 will affect two areas of the Eurotech product portfolio: sensing solutions and IoT integration solutions.

The two sensing solutions (people flow monitoring and environmental monitoring) developed in the project will contribute to increase the product offer of environmental monitoring stations, with smaller and less “intrusive” multi-sensor autonomous devices, and the product line of people counting and flow monitoring devices, that will result cheaper, smaller, easy to hide in the surrounding environment and, at the same time, more flexible, precise and efficient.

The sensor strips for people flow monitoring extends the Eurotech Passenger Counter product family: for this solution a patent has been submitted in the most important region of the world (Europe, USA, Japan, South Corea, etc.).

The environmental stations follow the same approach intended to reduce the dimensions and price of the device to increase the coverage and capillarity of the area covered by the monitoring services, without affecting the quality of the collected data and processing capabilities on the edge. The environmental station extends Eurotech Reliasense product family.

The IoT integration platform has been extended to support the new sensing solutions and to improve the capabilities and the autonomy of current software, in terms of data collection, data processing, device fleet monitoring and management. The results of the project in this area will influence the evolution of Eclipse Kura and Kapua, the open-source solutions developed and promoted by Eurotech and Red Hat for the IoT gateway management and system level integration. The results will directly influence also the Eurotech Everywhere Software Framework (ESF) and Eurotech Everywhere Cloud (EC), the commercial version of Eclipse Kura and Kapua respectively. The IoT platform deriving from the combination of the two software has already a large community of developers and adopters.

Contact person: Paolo Azzoni (paolo.azzoni@eurotech.com)

3.3.10 FSC

Embedded AI-assisted security solution for consumer Wi-Fi routers

F-Secure has studied the challenge to provide better security and threat detection for consumer IoT devices, increasing the cyber resilience of smart homes across society. We developed an embedded AI-assisted security solution (work carried out in WP2.2) along with the supporting backend infrastructure and data pipeline (work carried out in WP3.3). The key novel aspects of our work relate to applying the feature within a customer home environment in addition to running directly on the consumer home router itself.

Our work has focused on developing a technology that can learn normal versus abnormal traffic behaviour for IoT devices. The additional constraint specific to the consumer environment is that we only have visibility to network flow traffic data via the consumer home router. We developed several PoCs for a range of anomaly detection algorithms and have worked towards methods that would support running the feature directly on the router.

Figure 29 shows the continual scoring of network flows in time for some example devices in our test environment. Approximately halfway through the experiment, a series of flows corresponding to the Mirai IoT malware were injected into each device’s flow report to simulate a malware attack which are clearly visible in the anomaly scores.

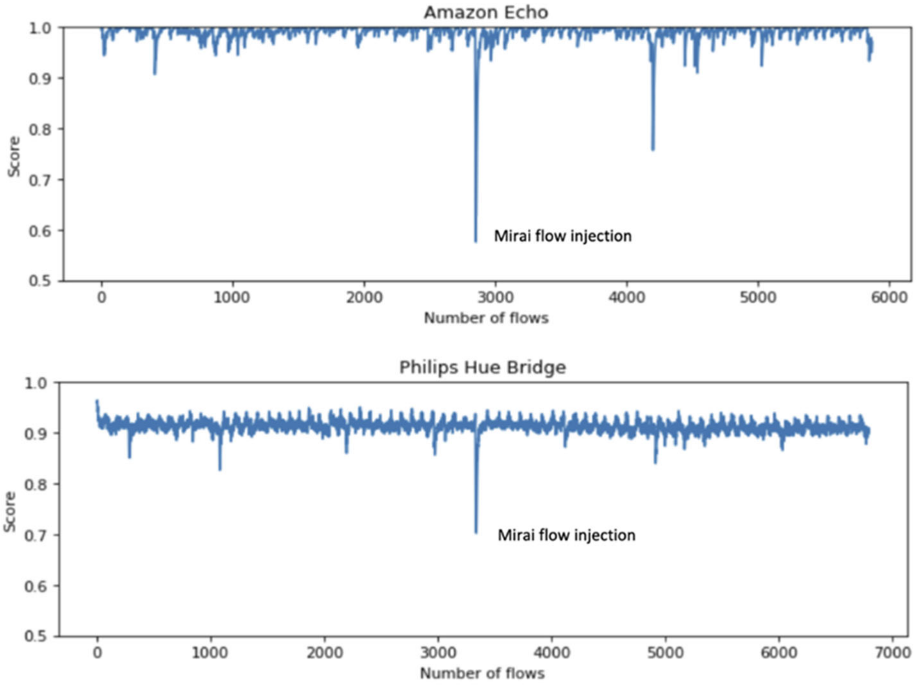


Figure 29 An example on anomaly scores for network flows from two devices in FSC test environment.

Our future aims for our exploitable item is to further develop the underlying algorithms with more customer pilot studies and estimate to launch the technology in Q2 2024. Hence, we will have achieved our aim of providing better security and threat detection for consumer IoT devices and expect this to contribute to an increased market share in the steadily growing consumer IoT market.

Contact person: Marko Komssi (marko.komssi@f-secure.com)

3.3.11 GUT

AI-based direction-of-arrival and localization algorithms

GUT has been developing AI-based algorithms for localizing WSN networks using specialized ESPAR antennas (see Figure 30). ESPAR antennas are electronically steerable and can perform beamforming in one plane with one active radiator and a few switchable passive elements. Their simple construction makes them appealing for commonly used WSN networks and in-demand applications such as indoor localization using DoA. One of the notable experiments involved the use of 8 ESPARs in an Airbus A321 aircraft in Hamburg to determine the positions of 30 Bluetooth Low Energy (BLE) transmitters in a highly reflective propagation environment. GUT investigated various AI-based approaches to address the challenging conditions, including Support Vector Machine, Decision Trees and Multilayer Perceptron algorithms. As a result, the average localization error across all nodes decreased from 8 to 2.5 meters, achieving 100% precision in terms of the number of localized nodes. These approaches and measurements are being utilized in UC5.2 in collaboration with project partners. Connecting the acquisition system (MPS) to the algorithm evaluation and analysis tool was a significant step in the development of the localization setup. While state-of-the-art DoA algorithms like PPCC have demonstrated their efficacy in controlled environments like anechoic chambers or free space simulations, real-world environments present challenges that necessitate more adaptable, AI-based solutions. In addition, GUT has developed visualization and analysis toolset, which enables to test the algorithms on various datasets and conveniently present the result.

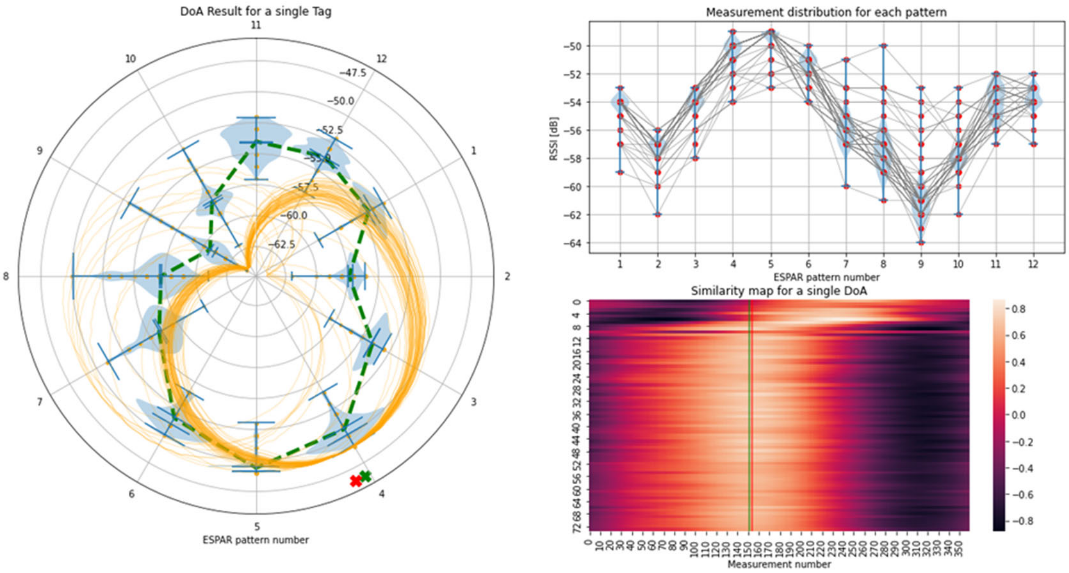


Figure 30 Direction-of-arrival algorithm analysis and visualization tool with an exemplary result.

Datasets:

- Anechoic chamber measurements comprising of 8 reference antennas with 50 interference source (jammer) positions.
- Indoor (room) measurements with 8 reference antennas and 20 transmitter positions.
- Outdoor (parking area) with 8 reference antennas and over 60 transmitter positions with various signal types.
- Airbus A320 measurements – real airplane high reflective, partially occluded environment with 8 reference antennas and 30 transmitter positions.

Contact person: Mateusz Rzymowski (mateusz.rzymowski@pg.edu.pl)

Test generation and map exploration using ML-based approach

In this work item, GUT explored the concept of using machine learning techniques for scenario generation, with a focus on map exploration to measure connectivity in the presence of radio interference sources, specifically jamming. The problem of exploration is a well-known task in robotics, where the aim is to minimize the number of steps taken to gain knowledge about a particular feature of an area or map. This feature is obtained by moving a robot and sampling the feature in each point. Traditionally, a popular approach is to sample the map in equally spaced points forming a grid, but this approach is slow and yields significant measurement error.

To address this issue, GUT worked on a more effective algorithm that can iteratively learn during the exploration procedure. The primary motivation was to speed up onsite interference measurements using robots and virtual channel simulators with the PhyWise framework. The specific application considered the measurement of connectivity between a transmitter and a moving receiver in the presence of a jamming source. The goal was to create a "connectivity map" with minimal robot measurement points (see Figure 31). Gaussian Process Classifier (GPC) algorithm was used, where each sampling point contributes to the final joint distribution over space, expressing the probability of connectivity with a binary observation variable (1 for connectivity, 0 otherwise). With every measurement step, the probability map was updated with knowledge about connectivity, and the next measurement point was calculated based on the most uncertain area on the map. The largest continuous area with probability close to 0.5 was selected using the Voronoi algorithm, and the robot was moved to its center to discover the next connectivity point. This process was repeated until the desired approximation was achieved.

Using this approach for scenario generation, GUT was able to reduce overall measurement time by five times compared to the brute-force grid search exploration scheme, while achieving the same relative approximation error of 5%.

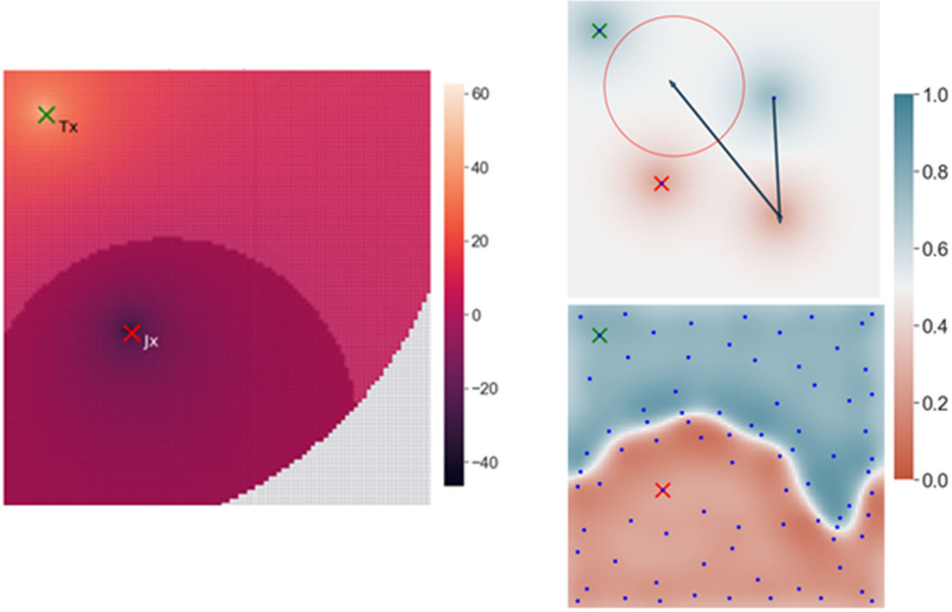


Figure 31 Connectivity map with minimal measurement points.

(left) Ground-truth connectivity map. Light red color represents area in which it is possible to receive signal from the transmitter “Tx” in a presence of interference source “Jx”.

(right) Connectivity map computed using the GUT proposed method. The dots represent consecutive positions of moving robot exploring the map. Heatmap is scaled with a calculated probability of connection.

Contact person: Mateusz Rzymowski (mateusz.rzymowski@pg.edu.pl)

3.3.12 IDEMIA

Fault injection simulator

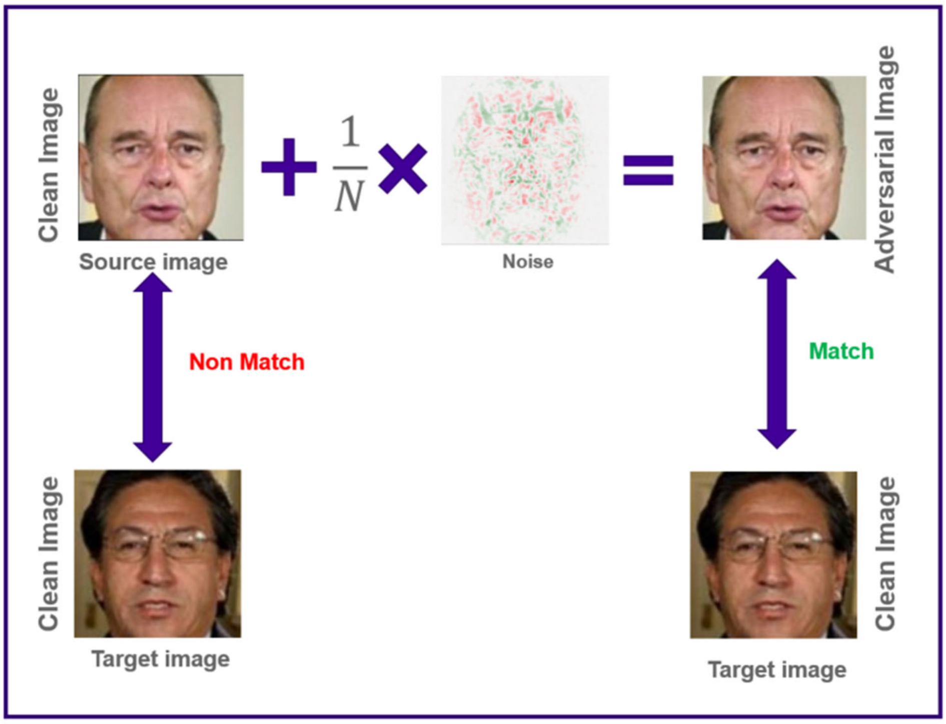
IDEMIA developed a fault injection simulator called CNNinja (**CNN** **Inj**ection **a**ttacker). The tool implements a “bit-flip” fault injection attack. CNNinja takes as input a program that uses an embedded CNN network (IDEMIA Internal) and runs on target (Raspberry Pi 3B with a Broadcom BCM2837 ARM Cortex-A53) or on emulator (Qemu) and instruments the execution to target specific memory displacement instruction. Two primary versions of the tool were created during the project. The first version utilized a very basic approach to fault injection attacks and served more as a proof-of-concept of what could be accomplished. This initial version of the program took a binary and attempted to alter any bit that made up the input program. However, this approach frequently resulted in crashes, as it was possible to modify the executable structure. Additionally, this version was relatively time-consuming, taking a week to attack even a simple program that performed basic addition, making it impractical.

The second version, which is still being finalized, appears to be a significantly more effective approach, at least on paper. This version enables modification of the program while it is executing memory operations and targeting memory instructions in assembly, particularly registers containing their results. The underlying logic has been implemented, but we are still working on developing heuristics to identify weight manipulation, which represents the ultimate goal. This poses a real challenge, given the internal nature of the CNN model.

CNNinja have been tested on an IDEMIA internal face recognition system that is CNN based and the idea will be to test the effectiveness of the tool on an Internal dataset.

Adversarial attacks on facial recognition systems:

In IDEMIA, we have developed a gray-box attack on facial recognition systems (FRS). Our method only needs to access to the templates. At each iteration, our algorithm optimizes the noise and re-calculate the template from the image with the added noise. The cosine distance between the source and target image template is used as the loss function. The loss function also contains an added term to limit the L2 distance between clean source image and adversarial source image.



[1]

Figure 32 Impersonation attack on facial recognition systems. Facial images are taken from publicly available LFW face dataset.

Adversarial attack protection scheme:

IDEMIA has also worked on protection schemes for adversarial attacks. There exist many approaches to defend against adversarial attacks, namely, detection, pre-processing of input images, robust

[1]

training etc. In IDEMIA, we have developed a method [1] to effectively defend against an adversary using a set of autoencoders. Below is a summary of the work.

Dynamic Autoencoders Against Adversarial Attacks:

Our approach involves adding autoencoders from a pretrained set dynamically to a base model, which serves as a countermeasure against attacks. This strategy involves modifying the underlying label regions of the model to protect them, thereby preventing adversaries from creating relevant adversarial perturbations. Our experiments demonstrate the effectiveness of this protection, particularly when the pretrained set contains enough elements. To accomplish this, we train a set of parasitic autoencoders independently using Gaussian noise and introduce them dynamically to thwart gradient-based adversarial attacks. By training the autoencoders independently, we gain greater flexibility when it comes to entropy in this random-based method. Our proposal involves adding different autoencoders simultaneously to multiple locations of the target model, which enhances its overall robustness. Future work could involve training the autoencoders on only one normalized channel, independently of the base model, so they can be generated once and placed anywhere in any model.

[1] Chabanne et al. Dynamic Autoencoders Against Adversarial Attacks, The 4th International Workshop on Data-Driven Security (DDSW 2023) March 2023, Leuven, Belgium

Contact person: Rahat Khan (rahat.khan@idemia.com)
SERE Ahmadou (ahmadou.sere@idemia.com)

3.3.13 INDRA

Train positioning and train integrity based on ML

INDRA has developed and deployed ML algorithms and methodologies for dynamic decision making about accurate train position, train integrity and length monitoring, and cargo identification and prioritisation for handling of goods.

Train positioning system has been improved including a GNSS selector that makes use of AI algorithms to dynamically choose the best GNSS source from multiple satellite networks at each time (GPS, differential GPS, Galileo or Glonass) considering signal conditions.

A Kalman Filter that receives values from IMU sensors has been also included to correct GNSS in areas with low GNSS signal.

Train integrity and length system has been developed using two main modules that make use of input measurement to process the values through ML algorithms: The first one receives measurements from accelerometers, RSSI sensors, train characteristics and the UWB (Ultra-Wide band) module, the second main module makes use of GNSS data from the positioning module to determine the train integrity. Both modules combine these pieces of information with the defined evaluation KPI’s to evaluate the train integrity in the final decision-making module.

The novelty reached in these developments is the use of wireless systems to obtain accurate and reliable positioning and integrity of the train and its wagons making a sensor fusion of multiple GNSS signals, inertial sensors, RSSI and UWB through the use of ML algorithms.

A brief summary of the results achieved during the project has been submitted on the 4rd Workshop on Management in Industry 5.0 (IFIP/IEEE NOMS_MFI50 2023, https://mfi50.icb.at/) “Internet of Things Technology for Train Positioning and Integrity in the Railway Industry Domain”.

Dataset:

> A novel dataset containing positioning and integrity train values has been generated with the data collected in the multiple pilots that have been performed to test the functionality of the developed systems, reaching therefore TRL6 in those systems. The dataset contains data from different sensors that are distributed along each wagon of real trains. The dataset is proprietary to the project consortium.

Contact person: Francisco Parrilla (fparrilla@indra.es)

3.3.14 ISEP

Intelligent wireless MIMO layer

Problem description

Wireless networks are rapidly evolving. They are the main enablers of a new technological revolution in which millions of embedded devices with sensing, networking, and actuation capabilities will be connected to edge/cloud infrastructure running artificial intelligence algorithms. The requirements on the new wireless layer are becoming more stringent and stricter every year. Critical applications such as autonomous vehicles or real-time object identification demand a wireless connectivity layer that is not only real-time and with ultra-low latency, but also secure, private, and trustworthy. The convergence between artificial intelligence and the internet of things paves the way for a previously untapped synergy between the wireless infrastructure and a new wave of intelligent algorithms. It is expected that in the coming years, the use of massive MIMO (Multiple-input multiple-output) systems will enable high speed and ultra reliable low latency services. This massive MIMO layer will be supported in large extent by a combination of multiple AI algorithms that improve not only the core applications but also the underlying wireless network infrastructure.

The use of massive systems brings further issues to be resolve such as the rise in the amount of training sequences to enable accurate channel equalization and signal decoding. The supporting AI algorithms for massive systems will have to do so with solutions to reduce the amount of training and signalling in the network while preserving reliability, low latency and security.

Contributions

Our contributions have been not only the testing of different types of AI for different wireless transmission operations for single and multiple antennas, but the optimality region of those algorithms. We have obtained bounds on the learning algorithms according to the characteristics of the wireless channel. We have also conducted our analysis under different types of impairments to the data sets that match the type of applications, as well as attacks or weaknesses of the learning infrastructure. The algorithms were tested in two types of use cases one for vehicle platoons and another one for wireless avionics intra-communications. We have also exploited the use of unsupervised learning algorithms in an attempt to reduce the required training bandwidth and thus make future massive MIMO networks more efficient.

Summary

We have created a library of intelligent wireless services for physical layer and medium access control design that will push the limit boundaries for future wireless networks, the algorithms are especially designed to be exploited with MIMO systems, due to their ability to increase wireless performance, at the expense of power and zero additional bandwidth except for that of signalling.

Contact person: Ramiro Robles (rasro@isep.ipp.pt)

Spatial authentication platooning systems

Problem description

In future vehicular applications, authentication of users/vehicles/operators and in general security features are critical due to the consequences for system operation. Any potential attack or issue that interrupts the efficient communication of commands and coordination between vehicles has the potential effect of causing a collision or another road incidents that poses threats to human lives and loss of infrastructure. The signal reliability has to be maintained at all costs. Therefore, MIMO (multiple-input multiple-output) systems are expected to be crucial in security of autonomous vehicles by rejecting sources of potential jamming and avoiding leaking information into eavesdropper directions.

Contribution

We have proposed a MIMO-based security system that is based on the unique spatial signature of each platoon. The signature can be used to multiplex information uniquely dedicated to the vehicles of the platoon using a massive MIMO transceiver at the BS. Each vehicle has a unique signature that can be translated into a type of direction of arrival information. Since a platoon is a collection of vehicles, the platoon has now a unique signature that is composed by the set of direction of arrivals of the vehicles and is translated into a MIMO subspace. The set of vehicles and their receivers create a MIMO system with the base station that has a unique spatial subspace structure. Using subspace signal processing and artificial intelligence tools to detect the components of the Platoon subspace, we can multiplex information into the incumbent subspace directions of the system, while reducing the information leakage to potential eavesdropping directions. In the opposite link direction, we can also detect signals that don't come from the incumbent spatial direction, thus allowing us not only to protect the integrity of the incumbent signals, but also detecting and rejecting interfering or jamming signals. The system can also be modified to detect anomalies of operation in the vehicle platoon.

Trustworthiness metrics calculation

Problem description

IoT system design has been rapidly evolving. After being considered a simple extension of sensor or Internet networks, IoT is now regarded as the new digital revolution. The impact of IoT is expected to be seen on almost everyday processes of human lives. Therefore, the metrics and performance indicators of this type of system will keep expanding even to non-technical areas.

Trustworthiness metrics are considered the new standard for IoT system design, as they gather that end-user impact that will characterize IoT in the years to come. However, their calculation is difficult to achieve due to the complexity of the system. There is no consensus so far how to create standard metrics for trustworthiness for IoT. However, we have to convergence quickly to an agreement otherwise devices and systems can be at risk.

Contribution

We have built system level simulators with some trustworthiness metrics to evaluate overall use case performance and AI performance. We have use the definition of trustworthiness metrics directly form standards for security and IoT design. We have made a connection between well-known formulae for security metrics such as CVSS and CWSS with an extended view of metrics for trustworthiness. We have also developed a framework for trustworthiness calculation based on scoring theory and the connections between scoring metrics across different layers and entities of the reference architecture. We have proposed a joint probabilistic method based on Markov chains that can help to calculate complex security and trustworthiness metrics for systems with multiple dependencies or with hybrid attacks.

Summary

We have proposed an architecture for decomposition of IoT systems and for trustworthiness metrics calculation based on an extension of some scoring metrics for IoT system design. The utter end of trustworthiness metrics is to improve IoT system design and when the metrics are trabnslated into labels for commercial regulatory, and standardization, acceptance and validation of products for difeenrte markets, the objective is to increase trust of the end users in the different products and systems developed.

Contact person: Ramiro Robles (rasro@isep.ipp.pt)

3.3.15 JSI

Anomaly detection in ECG signals

Traditional anomaly detection methods for 12-lead ECG signals often lack the ability to explain why an anomaly was detected. To address this limitation, we developed a new approach using a denoising autoencoder for explainable anomaly detection and denoising. Our approach involves reconstructing the original ECG signal from a noisy version of it and then using the reconstruction error to detect anomalies. We also provide a visual representation of the pinpointed anomalies on the ECG. We evaluated the performance of our approach on publicly available datasets and demonstrated its ability to detect anomalies with high accuracy and explain why they were detected. Additionally, we implemented our framework as a cloud-based service that allows for user-friendly ECG anomaly detection with minimal software and hardware requirements (Figure XX). Overall, our approach provides an effective and explainable solution for anomaly detection and denoising in 12-lead ECG signals, with potential applications in clinical settings as well as in consumer sector, for example in smart medical devices.

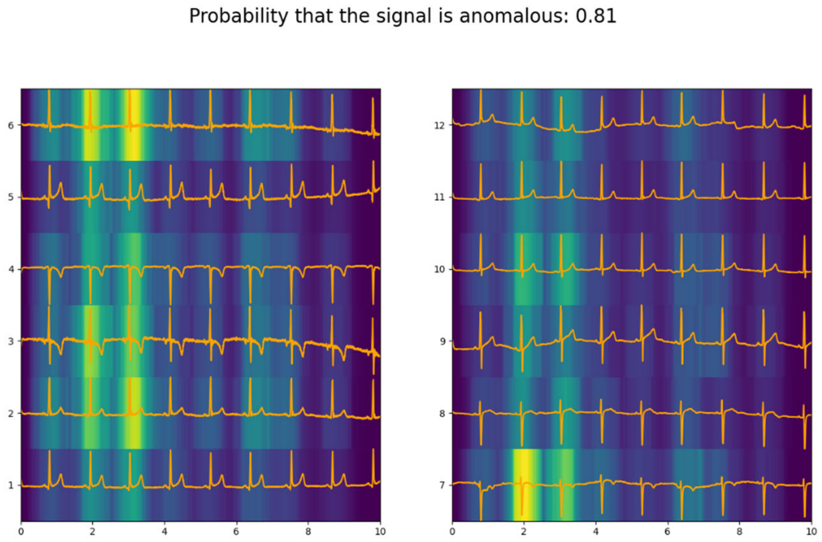


Figure 33 Result of an anomaly detection in 12-lead ECG signal.

Datasets:

> The data model was trained on a large dataset consisting of over 88,000 12-lead ECG samples obtained from https://moody-challenge.physionet.org/2021.

Contact person: Drago Torkar (drago.torkar@ijs.si)

3.3.16 KAI

Interfacing passive network QoS measurement with AI/ML solutions

Kaitotek focuses on a network quality and performance measurement solution. What makes Kaitotek's solution unique is the real-time passive QoS/QoE measurement, which is all software-based. The solution can measure the quality of the network without limitations on the applications being measured and network technologies. The passive solution does not burden networks with synthetic test traffic. This real-time measurement data is also useful for AI/ML solutions aiming at optimizing networks, connectivity management, and connected applications and services. It also enables innovations on new solutions on top of the network traffic analysis.

The interfaces against measurement software components were defined and implemented for easy integration with AI/ML solutions. Moreover, work on studying a new solution for network anomaly detection was initiated with RTE. A proof-of-concept study aims at finding new ways to detect anomalies in the network with continuous monitoring and learning of traffic patterns and characteristics. Pre-analyzing traffic, classifying them into flows, and feeding abstracted information about packets to the AI/ML solution can enable new ways to identify anomalies the currently available solution cannot see.

The future work will be to further study the potential of the solution together with RTE with different kind of application traffic and integrating the ML algorithm to the measurement solution.

Contact person: Esa Piri (esa.piri@kaitotek.fi)

3.3.17 KLAS

TRX R6 compute gateway

TRX R6 provides compute gateways for connected trains, light rail and buses. It is a combined hardware and software solution running KlasOS Keel allowing a range of applications to be supported on a single platform.

Supported applications include Passenger Wi-Fi, Infotainment and Passenger Information Systems, CCTV and TCMS access.

Open platform allowing the operator to add features over time.

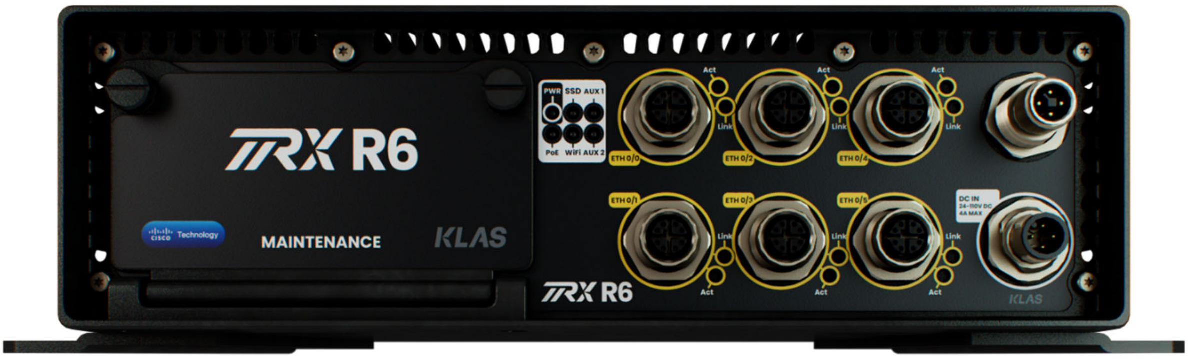


Figure 34 TRX R6 - Front

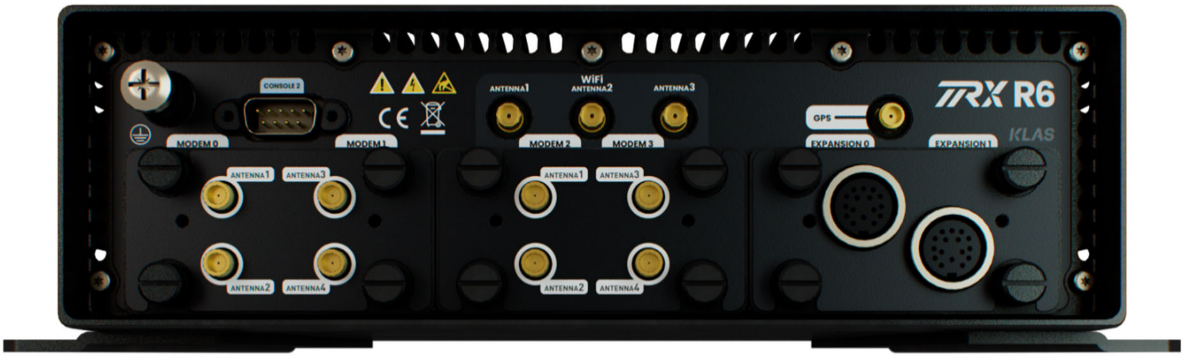


Figure 35 TRX R6 - Rear

Key Features:

- Combined router/server with multiple CPU options
- 6th Generation Intel® Core™ processors with dual-core i3, quad-core i5 and quad-core i7
- Support for six cellular modems (3G / 4G / 5G), Wi-Fi and Ethernet for connectivity
- Supports MVB, RS-422, RS-485 and GPS for telematics
- Integrated KlasOS Keel to supports third-party virtual machines
- Easily managed through integrated Keel CLI or SNMP via single IP address
- Firewall, captive portal and load balancing
- SDWAN tunnelling provides reliable connectivity over multiple WAN links with variable bandwidth
- Integrates with existing PKI infrastructure, secure data transfers using Keel private device keys

The TRX R6 is compliant with all the major railway certifications such as:

- EN50155:2017 Railway Applications – Electronic equipment used on rolling stock (EN50155: Heat, Cold, Insulation, Vibration and Shock), EN60068-2-1, EN60068-2-2, EN60068-2-30
- EN61373:2017 Railway Applications - Rolling stock equipment Shock and Vibration

- EN50121-3-2:2017 Railway Applications - Electromagnetic Compatibility - Part 3-2 Rolling Stock – Apparatus
- EN45545-2:2015 Railway Applications – Fire Protection of Railway Vehicles - Part 2
- EN55032 Electromagnetic Compatibility of Multimedia Equipment
- RED Compliant
- IP 54 Ingress protection
- RoHS 2 Compliant
- FCC/CE

Several TRX R6 have been provided to different InSecTT partners to develop, test, validate and qualify their solutions.

Contact person: Fernando Núñez (fernando.nunez@klasgroup.com)

3.3.18 KTH

Edge-based CPS planning algorithm

Autonomous vehicles need to anticipate potential hidden traffic participants, such as a hidden cyclist behind a vehicle or an unseen object behind a building. Existing methods need to assume a set of possible shapes and orientations and rely only on the current sensing to compute where hidden obstacles might be.

In the proposed method, KTH introduces a new modelling that allows to represent all possible hidden obstacles regardless of their shape and orientation. Also, making use of reachability analysis to combine previous sensing, the proposed method is able to produce less-conservative estimates of hidden obstacles while still guaranteeing safety.

The proposed algorithm has been extensively tested against state-of-the-art approaches in simulation and the results have been published in [28]. The implementation together with all the simulations are available publicly at the GitHub repository [29]. Finally, the proposed algorithm was also tested through in-vehicle (prototypical) implementation and demonstrations. The details about the experimental setup are described in [30].

Contact persons: José Manuel Gaspar Sánchez (jmgs@kth.se) and Martin Törngren (martint@kth.se)

Joint network- and task planning for trustworthy cloud robotics

With the emergence of Industry 4.0, comes an increasing need for multi-robot coordination and communication to efficiently complete joint tasks. A critical technology is the fifth generation (5G) mobile network, which enables cloud-controlled robots to execute tasks with differentiated quality-of-service (QoS) features. While there has been significant research on multi-robot planning, the integration with the capabilities of realistic network systems has been limited [31].

We present RoboPlan5G, a framework for 5G-aware robot planning. We propose a joint state-search model that includes task planning coordination in conjunction with 5G physical resource block (PRB) allocation. This process ensures trustworthiness by guaranteeing minimal throughput, based on technical specifications, and simultaneously providing efficient usage of the limited indoor 5G spectrum and the completion of tasks in the shortest possible time frame. Scenarios are generated in an Industry 4.0 environment, where the planner is shown to decrease the required 5G spectrum allocation by 50%, and on average improve the plan quality by 45% while maintaining a small computation time [32].

Repository containing files for planning, evaluation, simulation, and producing figures:
https://github.com/nilsjor/RoboPlan5G

Contact person: Nils Jörgensen (nilsjor@kth.se)

Algorithms for edge-based task offloading and resource management

We investigate and develop the algorithms by leveraging the edge-based scheme to solve computational task offloading and resource management problem. Two studies are performed in this topic in the field of intelligent transportation system and soft robotics:

(1) Vehicular Edge Computing (VEC) systems exploit resources on both vehicles and Roadside Units (RSUs) to provide services for real-time vehicular applications that cannot be completed in the vehicles alone. State-of-the-art optimization-based methods determine the optimal solutions in a centralized way at RSUs, but the complexity of solving the optimization problem is extraordinary, because the problem is not convex and has discrete variables. In this study, we decompose the joint optimization problem into two decoupled sub-problems: task offloading and resource allocation. Both sub-problems are reformulated for efficient solutions. Our new method efficiently achieves a near-optimal solution through decentralized optimizations, and the error bound between the solution and the true optimum is analysed. Simulation results demonstrate the advantage of the proposed approach. The details of the study are described in a journal publication [33].

(2) Soft robotics have the advantages of compliance and adaptability when working with vulnerable objects, but the deformation shape of the soft actuators is difficult to measure or estimate because of its highly nonlinear, hysteresis, and time-variant properties. A nonlinear and adaptive state observer is essential for the shape estimation from soft sensors. Current state estimation methods rely on complex nonlinear data-fitting models, and the robustness of the estimation methods is questionable. Our study investigates the soft actuator dynamics and the soft sensor model as a stochastic process characterized by the Gaussian Process (GP) model with the unscented Kalman filter (UKF) for more reliable variance adjustment during the sequential state estimation process. In addition, a major limitation of the GP model is its computational complexity during online inference. To improve the real-time performance while guaranteeing accuracy, we introduce an edge server to decrease the on-board computational and memory overhead. The experimental verification showcases a significant improvement in estimation accuracy and real-time performance compared to baseline methods. The details of the study are described in a journal publication [34].

Contact persons: Kaige Tan (kaiget@kth.se) and Lei Feng (lfeng@kth.se)

Integration architecture based on federation of rule engines

To increase the trust between stakeholders when integrating systems, KTH studied how to integrate such systems into a well-functioning System-of-Systems without centralizing essential components needed for the integration. The chosen approach involves the use of rule-based systems that are federated. The use of rules allows expressing policies declaratively, while the federation keeps each stakeholder in control of their system and the policies. In addition, we designed an algorithm to detect dependencies between rules across multiple federated nodes. The algorithm makes the proposed approach realistic without requiring all domain experts to have expertise in distributed systems, reducing risks of common fallacies in distributed computing. A prototype implementation was developed and evaluated; a journal publication [27] is under review.

Dataset(s) (if applicable):

NA (associated benchmark data in the journal paper will be published on Zenodo as part of the peer review)

Contact person: Andrii Berezovskyi (andriib@kth.se)

3.3.19 LCM

Anomaly detection algorithms for embedded devices

Wireless sensor networks (WSNs) are prone to different types of attacks since all participants in a certain range are sharing the same communication channel and organizational measures are not feasible to be applied in order to prevent such attacks. Especially jamming attacks, in which the full or parts of the channel are made unusable. Thus, the sensor network has to be able to detect and react on communication anomalies. Since the wireless sensor nodes are supplied either by batteries or by energy harvesting, the implemented algorithms must be energy efficient. For a target application we considered the energy efficient communication protocol EPhESOS [21].

We have looked both, the inside and outside view on the event of wireless jamming attacks on the sensor nodes. For the outside view, the received signal strength indicator (RSSI) is measured on all relevant channels of the 2.4GHz ISM band. This gives an overview of the communication situation on the whole ISM band. For the inside view we are analysing the RSSI values as well as the number of values received packet. Using our communication protocol EPhESOS, this value corresponds to missed received packets at the base station. Figure 36 shows a repeated jamming scenario where a certain time slot corresponding to the communication period of a certain sensor node is jammed. Figure 37 shows an exemplary diagnostic sequence based on the quantiles of repeated occupancies of frames.

The evaluated algorithms are suited to be implemented in embedded devices due to their simplicity. Besides the usage for detecting communication-related anomalies, the algorithms can also be used to detect device-related parameters. One example would be the targeted measurement value itself in order to detect a disconnection of the sensor element itself. In general, the anomaly detection algorithms can be used for application and to monitor different kind of parameters by adapting the threshold values accordingly.

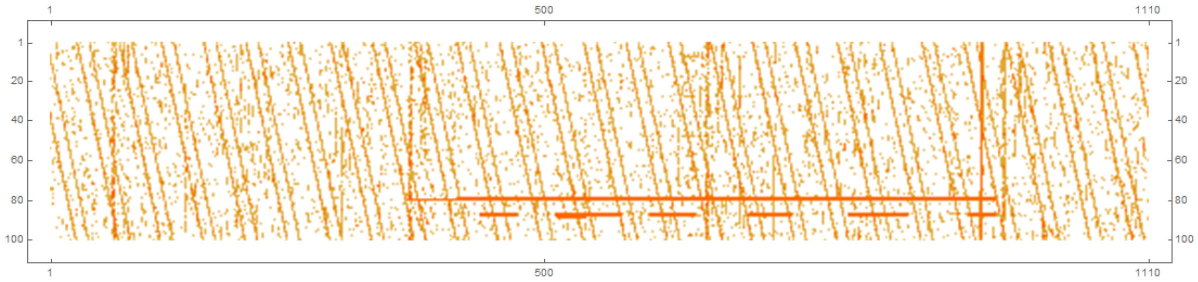


Figure 36 Depiction of an experiment with repeated jamming.

Each vertical line represents a single super-frame and the colour indicates the RSSI values of the 100 time slots. The first horizontal line corresponds to the beacon signal, while the dashed line visualizes the repeated jamming attempts.

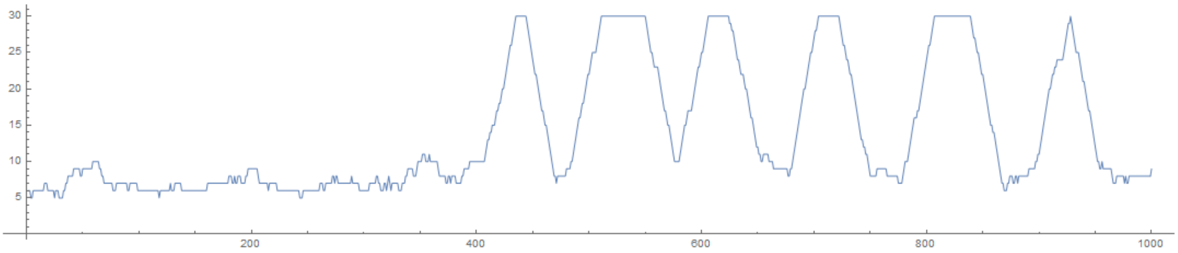


Figure 37 Diagnostic sequence for detecting jamming attacks based on quantiles of repeated occupancies of frames.

Contact person: Leander Hörmann (leander.hoermann@lcm.at)

3.3.20 LDO

Crowd Management System

LDO has developed a suite of three different algorithms to tackle the following problems: counting people in a given area, measuring social distancing and detecting situation of overcrowding and a detector for man-down situation.

The three algorithms developed during InSecTT are based on YOLOv5 detector, which was the latest version of the YOLO family available at the beginning of the InSecTT project in 2020, and the first of the YOLO family to be a complete rewrite of the original DarkNet model in the PyTorch framework.

People counting and social distancing also include a tracker, which is needed for the creation and association of tracks out of a sequence of frames where a given object is detected (in the case of interest for InSecTT UC5.16 "persons"). The tracker involved in those algorithms are based on a simple algorithm, SORT (https://arxiv.org/pdf/1602.00763.pdf), which leverages on Kalman's filter and the Hungarian algorithms; these allow for real-time performanes of the algorithms. Man-Down detection instead it is a YOLOv5 based binary classifier which does not make use of any tracking algorithm in its pipeline, and directly infer man-down situation based on the ratio between the orthogonal edges of the bounding box identifying the RoI.

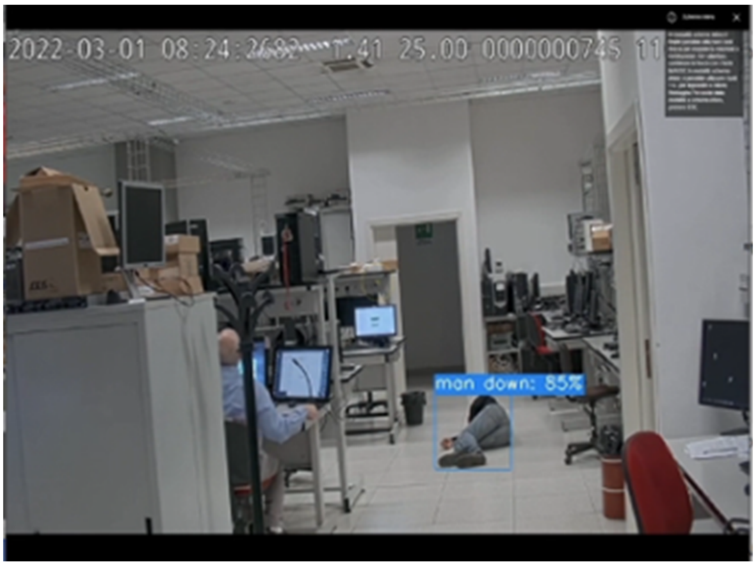
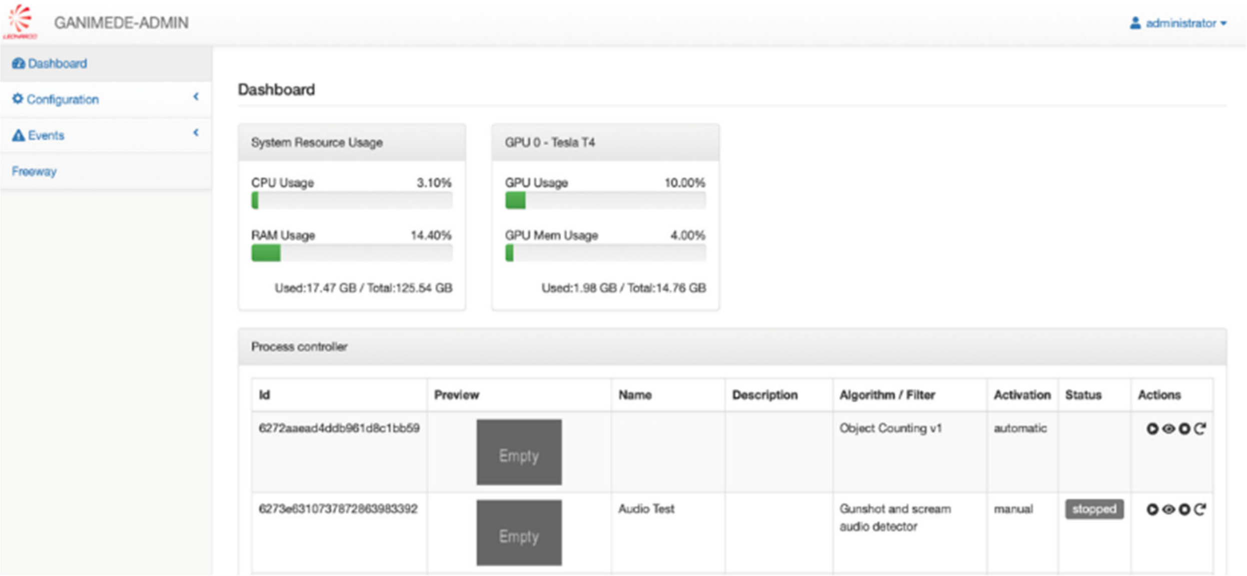


Figure 38 Example of person down detection

The people counting algorithm also allow the user to set one (or more) “virtual fence(s)” to count the number of people on the two sides of the fence, this allow for example to control incoming and outgoing flow of people from a given area.

The social distancing has been originarily conceived to provide control operators with a tool allowing them to check whether social distancing measures are respected or not by the persons in a given area. This could be useful during pandemics but it also applies in context of public security, as well as industrial plants monitoring. The setup requires a calibration procedure, which is performed taking reference points on a grid (the denser the better) and mutual distances between them on the area surveiled by the camera. The grid will allow a projection on a cartesian reference system where bounding boxes are first approximated by their barycenter.

The three algorithms have been integrated in Leonardo’s GANIMEDE platform, a video content analysis platform available via LDO’s SC2-Security Management platform.



Dashboard

System Resource Usage

CPU Usage 3.10%

RAM Usage 14.40%

Used:17.47 GB / Total:125.54 GB

GPU 0 - Tesla T4

GPU Usage 10.00%

GPU Mem Usage 4.00%

Used:1.98 GB / Total:14.76 GB

Process controller

Id	Preview	Name	Description	Algorithm / Filter	Activation	Status	Actions
6272aaead4ddb961d8c1bb59	Empty			Object Counting v1	automatic		
6273e6310737872863983392	Empty	Audio Test		Gunshot and scream audio detector	manual	stopped	

Figure 39 GANIMEDE platform UI

In particular through GANIMEDE it is possible to add video (and audio) streams (both from IP cameras or from other sources) and to execute different processes i.e algorithms which may be run onto the given video stream.

Dataset(s) (if applicable):

- YOLOv5 been trained on COCOdataset (https://cocodataset.org/#home) and dataset from the MOT Challenge (https://motchallenge.net/).

Contact person: filippo.cerocchi@leonardo.com

Potholes detection module

Detection of potholes and more generally detection of road damages is of interest for both public transportation companies and the municipalities. The purpose of road damage early warning and road surface monitoring is twofolded: on one hand, especially for public transport dedicated lanes, it enables prompt intervention of maintenance, avoiding unnecessary mechanical stress for the vehicles of the fleet, thus enhancing the quality of service and the safety of people onboard. On the other hand, when bus rides pass through not on dedicated lanes, this may help municipality in the monitoring and maintenance of the routes where citizens also drive.

The potholes detection module is an AI based module for the detection of potholes, designed to run on the edge (i.e. directly on-board of the bus). The module deploys YOLOv5 detector. YOLOv5 comes with different sizes of the underlying neural network. The lighter models exhibit faster inference with a reasonable loss of accuracy, thus being well suited for the use case, where the vehicle is usually moving at 30-50 Km/h and inference must happen in real-time. For what concerns the tracker, it is based on SORT (see the Crowd Management System) though slightly easier to construct and guaranteeing better results due to the simplified motion available (potholes are always moving from the top of the frame to the bottom).

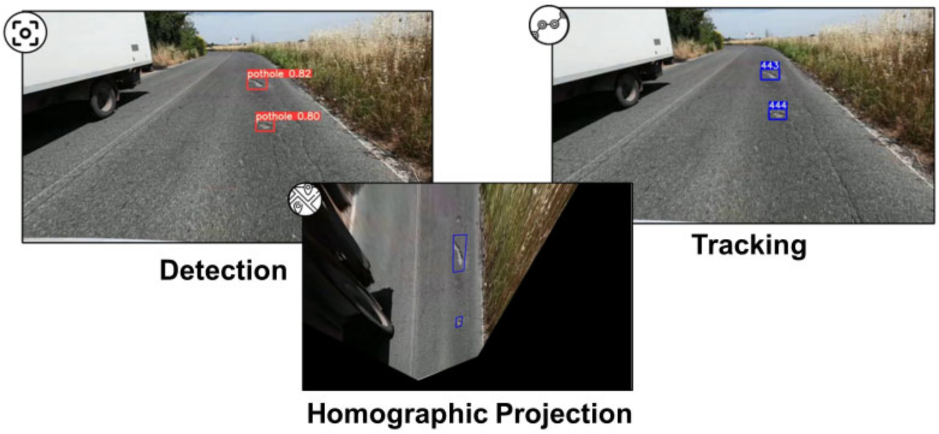


Figure 40 Pothole detection image processing

From the hardware point of view the Potholes detection module has been implemented and demonstrated using an IP camera (blue box in the figure below) connected via ethernet to and a NVIDIA Jetson Xavier (red box in the figure below), dedicated to the processing and where a MQTT client has been instantiated to send alerts in case a previously set threshold for the detected pothole size is exceeded. In the green box below LDO's nOBU and an auxiliary screen (yelow box).

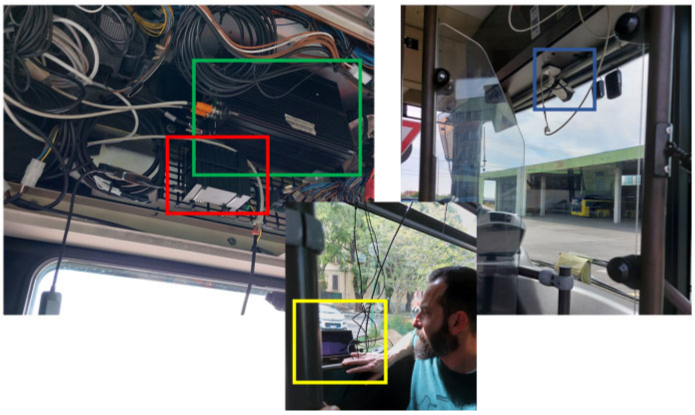


Figure 41 Pothole detection HW setup

Dataset(s) (if applicable):

- RDD2020: https://github.com/sekilab/RoadDamageDetector
- Kaggle: https://www.kaggle.com/chitholian/annotated-potholes-dataset

Contact person: filippo.cerocchi@leonardo.com

Privacy Preserving Audio Monitoring

Audio monitoring is complementary to video monitoring and can enhance the situational awareness of control operators (such as security control operators in airport terminals), allowing them to get information when the field of view is occluded or complementing information from video streams. Unfortunately, audio monitoring raises serious privacy issues as security operators could possibly ear personal information which are not relevant to their monitoring task.

LDO's Privacy Preserving Audio Monitoring algorithm tackle the problem as follows: the algorithm estimates the spectrogram corresponding to the voice data; voice data is then subtracted from the original spectrogram and added again after aggressive subsampling.

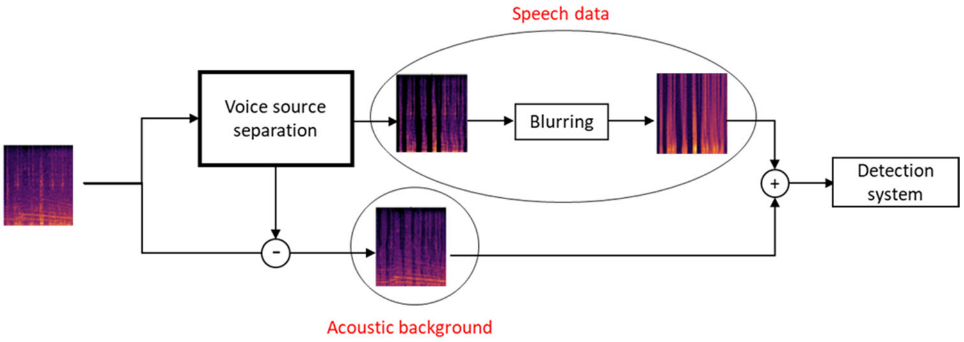


Figure 42 Audio monitoring processing flow

Commercially, audio-based monitoring solution are rare, and, up to our knowledge *no commercially available audio-monitoring system possess the capability of preserving privacy*. Indeed, the standard solution for voice anonymization simply perform a *pitch shift* of the audio data, which unfortunately *is an invertible operation, so the original voices can be recovered thus resulting in a vulnerability of privacy with potential damage for the data subjects and the company holding the data*. Besides this privacy aspect, it is worth to mention that a pitch shift affects the audio signal as a whole, possibly degrading the detection performances.

The algorithm is based on [1]; multiple variation of the original UNet has been tested before validating the current version of the algorithm. In contrast with [1] the model has been trained, tested and validated using heterogeneous voice durations mixed with the FSD50k dataset.

The algorithm has been added as a feature of LDO's GANIMEDE platform.

References:

1. Cohen-Hadria, M. Cartwright, B. McFee and J. P. Bello, "Voice Anonymization in Urban Sound Recordings," *2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)*, 2019, pp. 1-6, doi: 10.1109/MLSP.2019.8918913.
2. Fonseca, Eduardo, et al. "Fsd50k: an open dataset of human-labeled sound events." *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 30 (2021): 829-852.

Dataset(s) (if applicable):

- FSD50K dataset (reference [2]) more than 200 classes of sounds, comprising gunshot and screams.
-

Contact person: filippo.cerocchi@leonardo.com

3.3.21 LIEBHERR

Configurable Industrial Data Centre for Container Handling Equipment

Liebherr has expanded a container terminal and crane network and industrial data centre architecture for remote operating desk to crane real-time communications which also provides a platform for AI-enabled security monitoring and optionally predictive maintenance methods. The IDC hosts core switches to manage network connections between multiple cranes and crane types and remote desks and includes active back-up server and manual fail over procedure.

Liebherr has provided proprietary data from multiple cranes (equivalent to 36 years of operation) and has undertaken specific on-crane testing to support the development of videofeed verification by InSecTT partners UCC. Liebherr has also provided calculation methods for steel wire rope remaining useful life for an algorithm development by InSecTT partners UCC.

Contact person: John Barry (LCC) (John.Barry@liebherr.com)

IoT Platform for Container Handling Equipment

Liebherr has developed an internet of things platform to support the provision of digital services (including for example, asset management and operational insights), providing near-real time communication between PLC and crane SCADA system, and cloud database and analytics processing. The device (PLC) streaming API and the edge (crane IPC) event variable service communicate with the cloud through the container terminal DMZ, providing flexibility in the future division of processing between edge and cloud (and device) to support analytics and predictive maintenance methods, such as those developed by InSecTT partners UCC.

Contact person: John Barry (LCC) (John.Barry@liebherr.com)

3.3.22 MARUN

AI-based jamming detection on IoT end devices

The integration of IoT systems into the manufacturing process in Industry 4.0 creates a bridge between digital and physical environments which leads to increased productivity and reduces costs in general. Integration of IoT has also introduced new security challenges such as jamming attacks to damage industrial systems. These attacks are significant threats to the security of manufacturing systems, therefore taking countermeasures is crucial. The manufacturing domain requires specific attention to address the security threats of cyberattacks and protect against their possible outcomes.

It is important to take action on the presence of jamming attacks especially as intelligent devices become more popular. Self-awareness is important because when the communication link is interrupted, it is not possible to remotely manage or stop industrial processes. Marmara University proposes a solution that introduces a novel approach by using AI-based models to analyse the effects of jamming attacks on the upper layers of the protocol stack on an IoT end device. A testbed was developed in the VeNIT Lab to demonstrate different types of jamming attacks. By using the network parameters, the effects of attacks are observed in transport and application layer traffic. Different anomaly/jamming detection models including a state-of-the-art stacked LSTM were implemented as a solution. Several methods are being investigated to detect jamming on the IoT devices. These methods allow certain measures to be taken to address the issue, even when communication is disrupted due to jamming attacks. The changes in network indicators, such as anomalous changes, lack of data, or increased delay, are being evaluated to detect the presence of jamming. This solution aims to improve the security of the IoT network against jamming attacks and ensure self-awareness of IoT devices. Furthermore, the study emphasizes the significance of the effects on the upper layers of the protocol stack and offers an effective solution for detecting and preventing jamming attacks in the manufacturing domain at the edge device.

Contact persons: Mujdat Soyturk (mujdat.soyturk@marmara.edu.tr), Yavuz Selim Bostanci (yavuz.bostanci@venit.org)

Object detection and localization at the edge with V2X Day-2 (CPM) demonstrator

One of the major challenges of providing information about surrounding vehicles and objects in busy places such as airports is to create awareness for nearby vehicles even though they can't receive or determine this information individually. To address this problem, MarUn developed an object detection application that uses a camera placed at intersections. MarUn measured the performance of 2D and 3D object detection algorithms using recently developed frameworks with their edge performance.

MarUn also developed a 3D object detection model for localizing objects in the smart intersection area and investigated a new method for generating a 3D object dataset. To achieve this, a realistic intersection simulation of the Marmara University Campus is developed. Surveillance cameras are mounted in the simulation to collect realistic vehicle and pedestrian traces. The 3D object detection model was then trained using data from the simulation environment. By comparing the performance of models trained on different datasets, we have observed poor performance of the model trained on available open-source datasets for intersection images, but improved performance of the model trained on simulation images for object localization in the simulation environment. Ultimately, MarUn is able to test intersection area algorithms in a simulation environment while using V2X standards, enhancing the safety and security of intersections.

Along with that, MarUn has implemented a V2X Day 2 application that is compliant with established standards for Collective Perception Message (CPM). This beyond-state-of-the-art application serves as a critical enabler for comprehensive demonstration of the object detection system, ensuring the identification of nearby vehicles and objects with high accuracy and real-time processing capabilities.

Contact persons: Mujdat Soyturk mujdat.soyturk@marmara.edu.tr, Yavuz Selim Bostanci yavuz.bostanci@venit.org

3.3.23 MDH

Federated Learning Framework for Network Attacks Detection and Classification Based on Random Forest

MDH are working on distributed machine learning with the aim of building a shared (global) learning model via decentralized learning using data generated from local devices or processes. We expect that this learning paradigm will be highly relevant for realization of the edge learning algorithms in the relevant scenarios in this case study.

We proposed a federated learning framework for network attacks detection and classification based on Random Forest (RF) ([35]). The main idea of the framework is to train independent RFs on clients using the local data, merge independent models into a global one on the server and send it back to the clients for further use. The architecture of the proposed framework is presented in Figure 43. Previously the framework was evaluated only for attack detection using four well-known intrusion detection datasets, including KDD ([36]), NSL-KDD ([37]), UNSW-NB15 ([38]), and CIC-IDS-2017 ([39]).

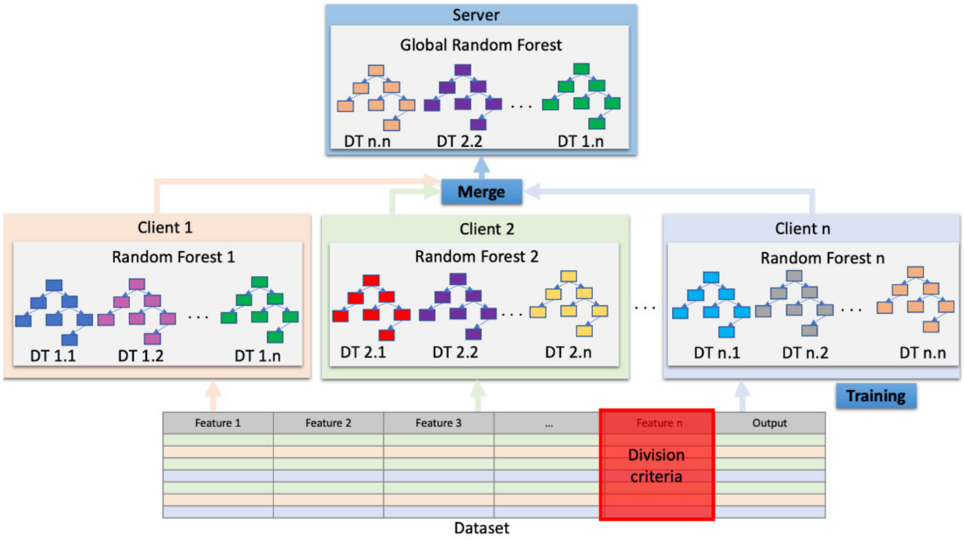
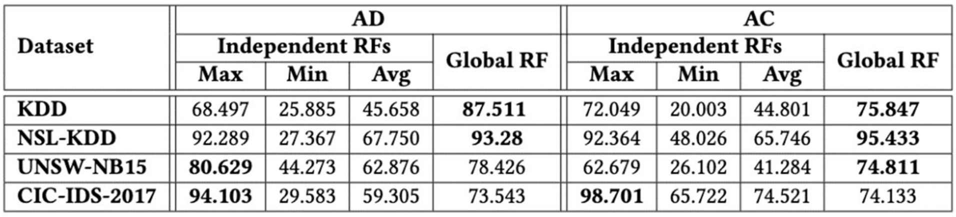


Figure 43 The architecture of federated learning framework for network attacks detection and classification based on random forest

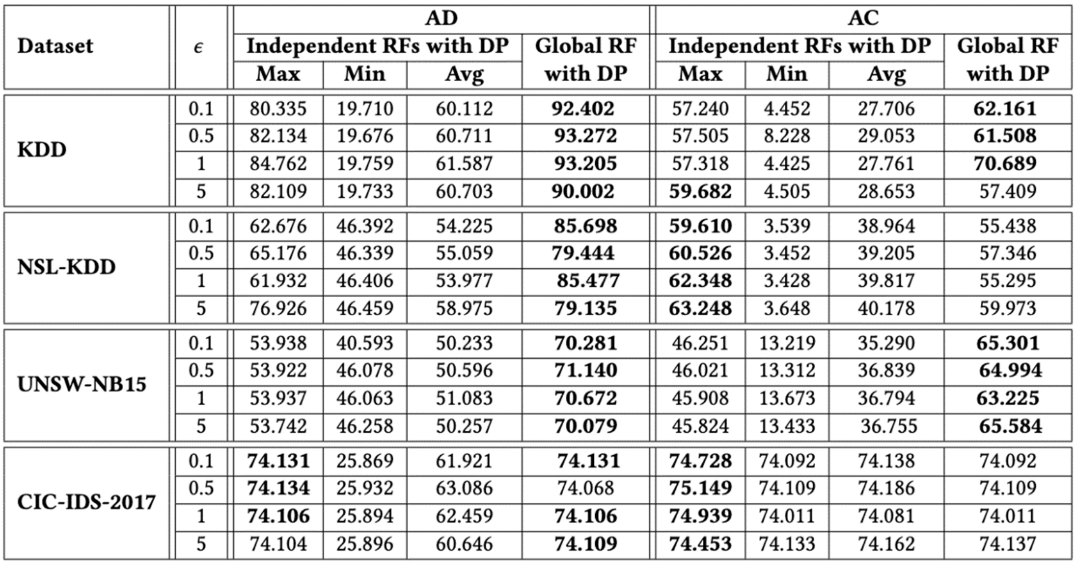
The evaluation of the proposed framework is extended for attack classification using the same datasets as in the previous study. As it can be seen in Figure 44, the results showed that the framework outperforms the average performance of independent RFs on clients for both Attack Detection (AD) and Attack Classification (AC). For AD the global RF improved the maximum accuracy of individual RFs for KDD and NSL-KDD, and it was very close for UNSW-NB15. Also, for AC the global RF improved the maximum accuracy of individual RFs for three out of four datasets.



Dataset	AD				AC			
	Independent RFs			Global RF	Independent RFs			Global RF
	Max	Min	Avg		Max	Min	Avg	
KDD	68.497	25.885	45.658	**87.511**	72.049	20.003	44.801	**75.847**
NSL-KDD	92.289	27.367	67.750	**93.28**	92.364	48.026	65.746	**95.433**
UNSW-NB15	**80.629**	44.273	62.876	78.426	62.679	26.102	41.284	**74.811**
CIC-IDS-2017	**94.103**	29.583	59.305	73.543	**98.701**	65.722	74.521	74.133

Figure 44 Comparison of AD and AC accuracy of global RF with maximum, minimum and average accuracy of independent RFs on the entire testing set of KDD, NSL-KDD, UNSW-NB15 and CIC-IDS-2017 dataset

Additionally, it was evaluated how adding Differential Privacy (DP) into RF, as an additional protective mechanism, affects the framework performances. Independent RF with differential privacy was trained for each client. Four different values of **ε** parameter were tested: 0.1, 0.5, 1 and 5. The global RF was tested on the entire testing set and the results and performances of global RF were compared with the performances of independent RFs with DP. From the results presented in Figure 45 we can conclude that adding DP penalizes the performance of RF, as expected, but the use of the proposed framework still brings benefits in comparison to the use of independent local models.



Dataset	ϵ	AD				AC			
		Independent RFs with DP			Global RF with DP	Independent RFs with DP			Global RF with DP
		Max	Min	Avg		Max	Min	Avg	
KDD	0.1	80.335	19.710	60.112	**92.402**	57.240	4.452	27.706	**62.161**
	0.5	82.134	19.676	60.711	**93.272**	57.505	8.228	29.053	**61.508**
	1	84.762	19.759	61.587	**93.205**	57.318	4.425	27.761	**70.689**
	5	82.109	19.733	60.703	**90.002**	**59.682**	4.505	28.653	57.409
NSL-KDD	0.1	62.676	46.392	54.225	**85.698**	**59.610**	3.539	38.964	55.438
	0.5	65.176	46.339	55.059	**79.444**	**60.526**	3.452	39.205	57.346
	1	61.932	46.406	53.977	**85.477**	**62.348**	3.428	39.817	55.295
	5	76.926	46.459	58.975	**79.135**	**63.248**	3.648	40.178	59.973
UNSW-NB15	0.1	53.938	40.593	50.233	**70.281**	46.251	13.219	35.290	**65.301**
	0.5	53.922	46.078	50.596	**71.140**	46.021	13.312	36.839	**64.994**
	1	53.937	46.063	51.083	**70.672**	45.908	13.673	36.794	**63.225**
	5	53.742	46.258	50.257	**70.079**	45.824	13.433	36.755	**65.584**
CIC-IDS-2017	0.1	**74.131**	25.869	61.921	**74.131**	**74.728**	74.092	74.138	74.092
	0.5	**74.134**	25.932	63.086	74.068	**75.149**	74.109	74.186	74.109
	1	**74.106**	25.894	62.459	**74.106**	**74.939**	74.011	74.081	74.011
	5	74.104	25.896	60.646	**74.109**	**74.453**	74.133	74.162	74.137

Figure 45 Comparison of AD and AC accuracy of global RF with DP with maximum, minimum and average accuracy of independent RFs with DP on the entire testing set of KDD, NSL-KDD, UNSW-NB15 and CIC-IDS-2017 dataset

The proposed framework is recommended in the applications where the data cannot be centralized and the goal is to apply machine learning, while protecting the data as much as possible. This work is currently under review ([40]). The framework was developed in collaboration with Tietoevry.

Datasets:

Publicly available dataset has been utilized:

- KDD ([36])
- NSL-KDD ([37])
- UNSW-NB15 ([38])
- CIC-IDS-2017 ([39])

Contact persons: Tijana Markovic (tijana.markovic@mdu.se) and Miguel Leonortiz (miguel.leonortiz@mdu.se)

Framework for Feature Encoding and Data Privacy based on Autoencoders and Optimization Algorithms

MDH was working on feature encoding techniques that can help to preserve data privacy, reduce the network overload and speed up the machine learning algorithms without affecting the accuracy to a high degree. We evaluated an approach based on an autoencoder trained with differential evolution for feature encoding of network data with the goal of improving security and reducing data transfers. We created a general model where any optimization algorithm could be used, including multi-objective optimization algorithms (Figure 46). One of the novel elements used in differential evolution for intrusion detection is the enhancements in the fitness function by adding the performance of a machine learning algorithm ([41]).

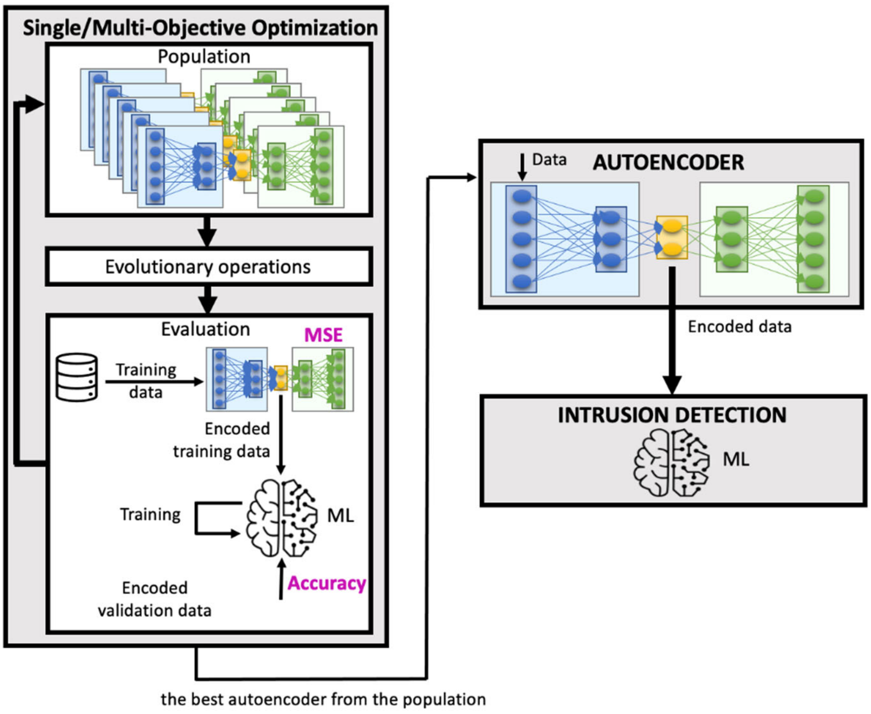
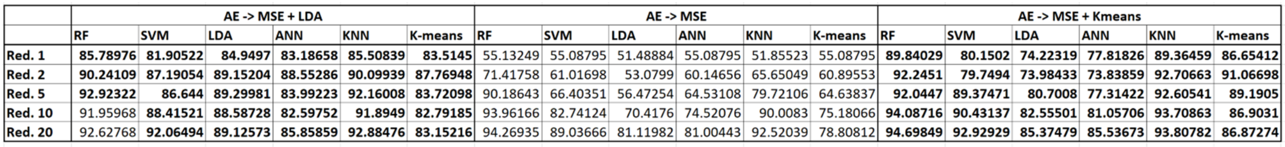


Figure 46 Architecture of the proposed framework for feature reduction and data privacy.

Two ML methods were considered in the evaluation function of Differential Evolution: Linear Discriminant Analysis (LDA) and K-means, together with the Mean Squared Error (MSE). After, creating the new features with the framework, we tested the performance of 6 different ML algorithms: Random Forest (RF), Support Vector Machine (SVM), LDA, Artificial Neural Network (ANN), K-Nearest Neighbour (KNN) and K-means. The results can be found in Figure 47. It can be observed how including the ML algorithm helps to obtain a better performance. However, there are no statistical differences between using LDA or K-means in the evaluation function. The paper that presents this work will be submitted in the next months.

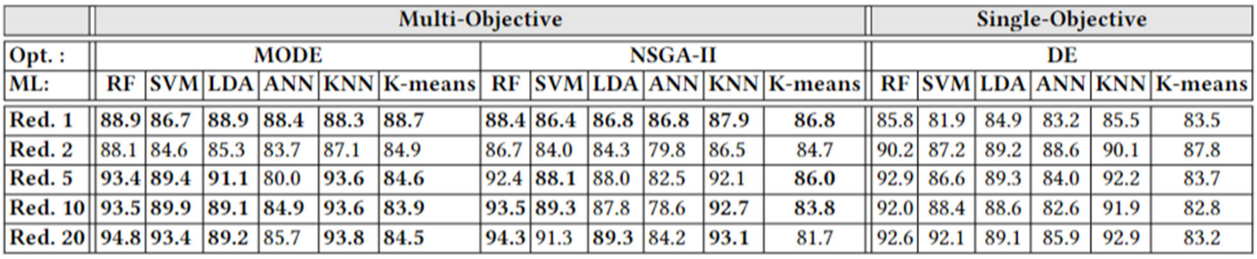


	AE -> MSE + LDA						AE -> MSE						AE -> MSE + Kmeans					
	RF	SVM	LDA	ANN	KNN	K-means	RF	SVM	LDA	ANN	KNN	K-means	RF	SVM	LDA	ANN	KNN	K-means
Red. 1	**85.78976**	**81.90522**	**84.9497**	**83.18658**	**85.50839**	**83.5145**	55.13249	55.08795	51.48884	55.08795	51.85523	55.08795	**89.84029**	**80.1502**	**74.22319**	**77.81826**	**89.36459**	**86.65412**
Red. 2	**90.24109**	**87.19054**	**89.15204**	**88.55286**	**90.09939**	**87.76948**	71.41758	61.01698	53.0799	60.14656	65.65049	60.89553	**92.2451**	**79.7494**	**73.98433**	**73.83859**	**92.70663**	**91.06698**
Red. 5	**92.92322**	**86.644**	**89.29981**	**83.99223**	**92.16008**	**83.72098**	90.18643	66.40351	56.47254	64.53108	79.72106	64.63837	**92.0447**	**89.37471**	**80.7008**	**77.31422**	**92.60541**	**89.1905**
Red. 10	91.95968	**88.41521**	**88.58728**	**82.59752**	**91.8949**	**82.79185**	93.96166	82.74124	70.4176	74.52076	90.0083	75.18066	**94.08716**	**90.43137**	**82.55501**	**81.05706**	**93.70863**	**86.9031**
Red. 20	92.62768	**92.06494**	**89.12573**	**85.85859**	**92.88476**	**83.15216**	94.26935	89.03666	81.11982	81.00443	92.52039	78.80812	**94.69849**	**92.92929**	**85.37479**	**85.53673**	**93.80782**	**86.87274**

Figure 47 A comparison of ML algorithms after using Autoencoder (AE) with MSE+LDA and MSE + Kmeans against AE using only MSE.

Boldface values mean that the results of the algorithms using ML methods are better than not using them.

Furthermore, we tested two multi-objective algorithms: Multi-Objective Differential Evolution (MODE) and Non-sorting Genetic Algorithm II (NSGA-II). The results (Figure 48) show that MODE is better than NSGA-II and single-Objective DE. More information can be found in [42].



	Multi-Objective												Single-Objective					
Opt. :	MODE						NSGA-II						DE					
ML:	RF	SVM	LDA	ANN	KNN	K-means	RF	SVM	LDA	ANN	KNN	K-means	RF	SVM	LDA	ANN	KNN	K-means
Red. 1	**88.9**	**86.7**	**88.9**	**88.4**	**88.3**	**88.7**	**88.4**	**86.4**	**86.8**	**86.8**	**87.9**	**86.8**	85.8	81.9	84.9	83.2	85.5	83.5
Red. 2	88.1	84.6	85.3	83.7	87.1	84.9	86.7	84.0	84.3	79.8	86.5	84.7	90.2	87.2	89.2	88.6	90.1	87.8
Red. 5	**93.4**	**89.4**	**91.1**	80.0	**93.6**	**84.6**	92.4	**88.1**	88.0	82.5	92.1	**86.0**	92.9	86.6	89.3	84.0	92.2	83.7
Red. 10	**93.5**	**89.9**	**89.1**	**84.9**	**93.6**	**83.9**	**93.5**	**89.3**	87.8	78.6	**92.7**	**83.8**	92.0	88.4	88.6	82.6	91.9	82.8
Red. 20	**94.8**	**93.4**	**89.2**	85.7	**93.8**	**84.5**	**94.3**	91.3	**89.3**	84.2	**93.1**	81.7	92.6	92.1	89.1	85.9	92.9	83.2

Figure 48 Accuracy of ML algorithms for AD using the encoded features from the autoencoder train by MODE, NSGA-II and single-objective DE.

Boldface values mean that the results of multi-objective algorithms are better than single-objective DE.

Datasets:

A publicly available dataset has been utilized:

- UNSW-NB15 ([38])

Contact persons: Miguel Leonortiz (miguel.leonortiz@mdu.se) and Tijana Markovic (tijana.markovic@mdu.se)

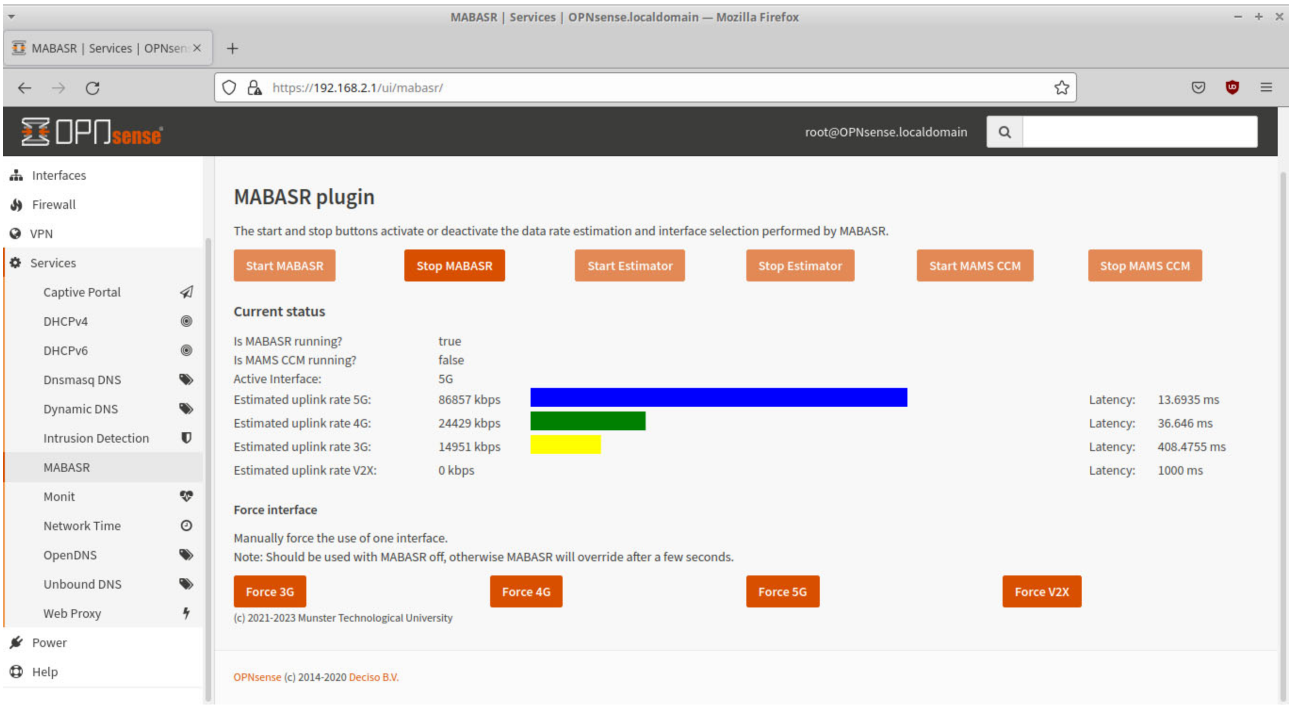
3.3.24 MTU

Connection monitoring and management plugin

The aim of the development that led to this monitoring and management plugin for a routing and firewall platform was that in order to be useful in practice and beyond the scope of the project, connection monitoring and management solutions should be integrated into a platform that is relevant in target scenarios. The target scenario in the scope of this project is an onboard system on a train that has to maintain uplink connections while the train is moving.

Based on advice from project partner KLAS, the routing and firewall platform OPNsense was chosen as target platform for implementation as KLAS also use OPNsense on their hardware in some of their commercial offerings for the railway sector.

MTU's plugin for OPNsense has been named "MABASR plugin" after one of its underlying AI algorithms, MABASR [4], that was developed in the course of this project. Its front end, as shown in Figure 49, features status readings, estimated data rates and latencies for the available interfaces, as well as a number of control buttons to start and stop underlying algorithms or force the device to use a specific interface. The AI algorithms behind it estimate the uplink data rates that are displayed and select the active interface accordingly. The user has the option to use local estimation and decision (MABASR buttons), or to only perform local estimation (Estimator buttons) and send estimation results to a cloud entity where a decision is made (MAMS CCM buttons).



MABASR plugin

The start and stop buttons activate or deactivate the data rate estimation and interface selection performed by MABASR.

Current status

Is MABASR running?	true		
Is MAMS CCM running?	false		
Active Interface:	5G		
Estimated uplink rate 5G:	86857 kbps	Latency:	13.6935 ms
Estimated uplink rate 4G:	24429 kbps	Latency:	36.646 ms
Estimated uplink rate 3G:	14951 kbps	Latency:	408.4755 ms
Estimated uplink rate V2X:	0 kbps	Latency:	1000 ms

Force interface

Manually force the use of one interface.
Note: Should be used with MABASR off, otherwise MABASR will override after a few seconds.

(c) 2021-2023 Munster Technological University

OPNsense (c) 2014-2020 Deciso B.V.

Figure 49 Screenshot of MTU's plugin front end in OPNsense

Through integration in a platform that is used in relevant markets, MTU's AI algorithms for connectivity resource estimation and management enhance that platform and are themselves enhanced with a front end for improved usability.

Dataset:

In the course of MTU's development efforts, multiple measurements of real-life 5G channel parameters and uplink throughput were conducted in Cork City and suburbs. These datasets were used as training and testing data for the data rate estimation and interface decision AI algorithms.

The collected datasets include timestamps, channel parameters as well as measured uplink throughput. The datasets are publicly available on GitHub [5].

Contact person: Bernd-Ludwig Wenning (berndludwig.wenning@mtu.ie)

3.3.25 NXP-NL

eIQ toolkit – Explainability and augmentation plug-ins

As a result of the first year InSecTT investigations, NXP-NL-CCC&S has turned over the InSecTT explainability research (InSecTT UC5.2, UC5.5, UC5.6) and adversarial example research (InSecTT BB2.4) to the NXP Business Units, which resulted in the creation of augmentation and explainabilty plug-ins in the eIQ Toolkit. The eIQ Toolkit ([43]) can be used in conjunction with NXP MCUs.

eIQ toolkit – Augmentation

The “Augmentation tool” has been inserted in the eIQ product, containing 10+ generic counter measures against adversarial attacks. These counter measures were studied in the first year of the InSecTT project.

NXP released the eIQ toolkit and ensured the studied generic countermeasures, such as random Blur augmentation, gaussian noise, flips, etc. were integrated in the eIQ toolkit. See Figure 50 below.

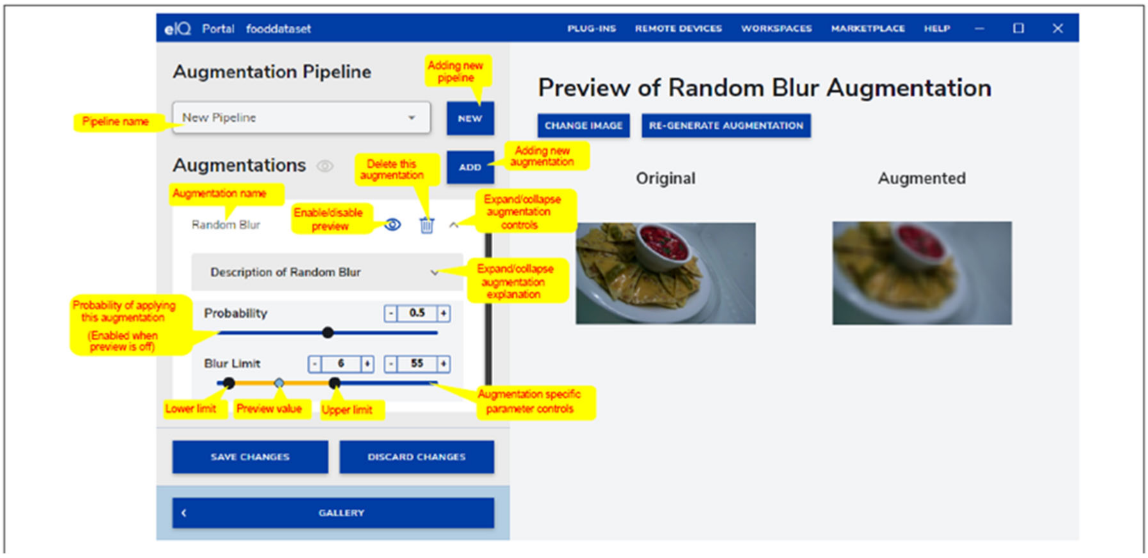
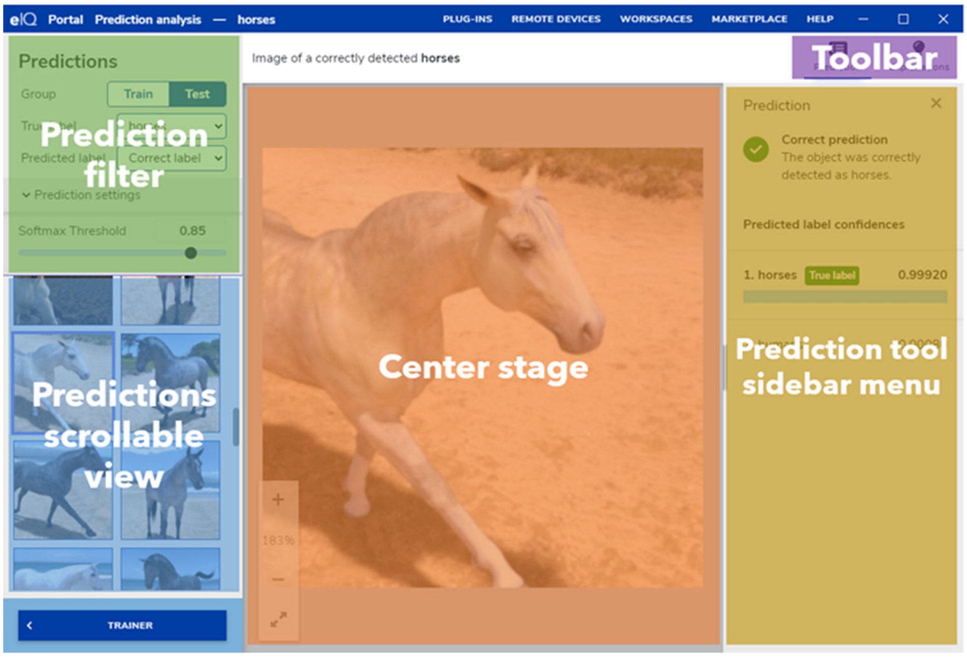


Figure 50 Augmentation Pipeline workspace in eIQ Portal

eIQ toolkit – Explainability

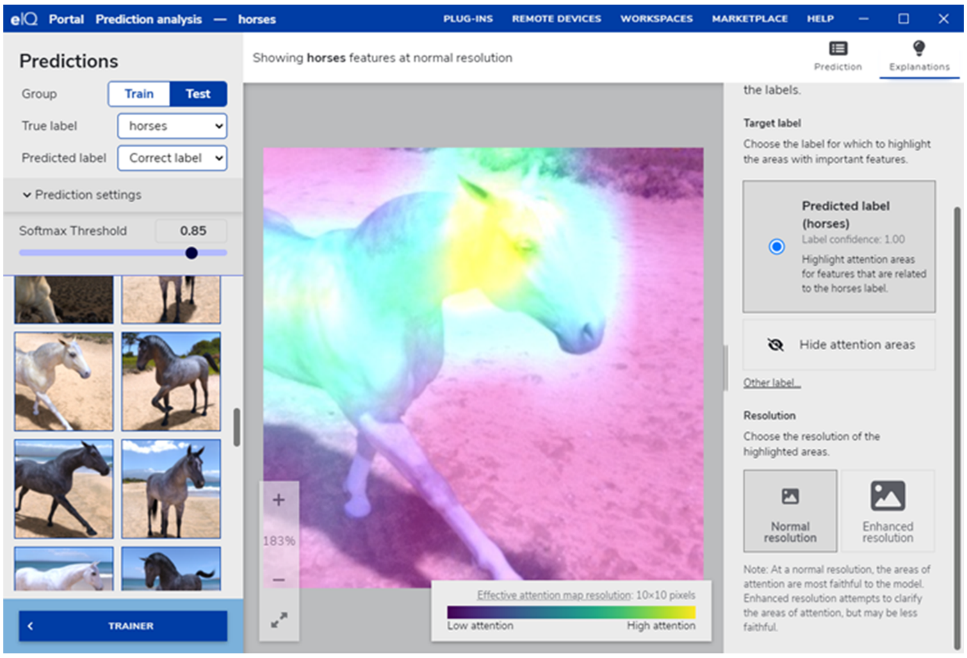
The toolkit contains default datasets. In the example in Figure 51 below, one of the default datasets (horses and people) was used. After training an image-based model (by a customer), the toolkit allows to do a prediction analysis. The prediction tool gives access to the explainability methods.



The object was correctly detected as horses.

Figure 51 An example on dataset in eIQ Portal

Within the prediction tool (see Figure 52), the heatmap can be recalled, which signifies the relevant features the model used to determine the output class ("horse") thus enabling detailed analysis on the model (mis-) behaviour.



Showing **horses** features at normal resolution

the labels.

Choose the label for which to highlight the areas with important features.

Highlight attention areas for features that are related to the horses label.

Choose the resolution of the highlighted areas.

Note: At a normal resolution, the areas of attention are most faithful to the model. Enhanced resolution attempts to clarify the areas of attention, but may be less faithful.

Figure 52 A heatmap annotation on relevant features in predictions in eIQ Portal.

Also based on the InSecTT results, NXP has combined certain explainability methods and created a so-called "enhanced resolution" (see Figure 53), giving a finer detail on how the neural network looks at data. The methods used to create "enhanced resolution", in a freely available toolkit, are beyond state of the art.

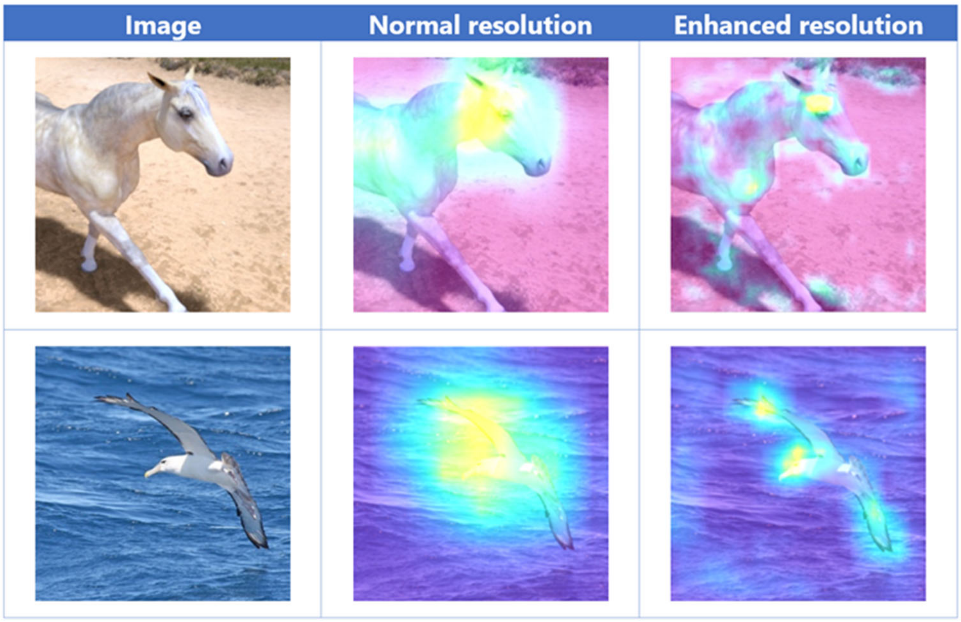


Figure 53 Explainability annotations for predictions in eIQ Portal.

eIQ Explainability is derived from the first-year investigations InSecTT investigations (as was shown in the "explainability webserver". The throughput from research to product takes 3-5 years. The year 2/3 InSecTT investigations on sensor data (based on the partner's input data and used AI models) will progress into the eIQ toolkit over the next years.

Dataset(s):

Dataset embedded in the eIQ Toolkit (as of 21-Apr-2023) ([43])

Contact person: Ton Scheepers (ton.scheepers@nxp.com)

3.3.26 PAVOTEK

AI Based Anomaly Detection Software

AI-based anomaly detection software is an innovative solution that uses artificial intelligence and machine learning techniques to detect abnormal behavior in data. This software learns patterns of normal behavior by analyzing data from many different data sources and then identifies abnormal patterns. In this way, users can take preventive measures by ensuring early detection of potential threats. The software offers flexible configuration options for real-time or retrospective analysis and provides easy use with a user-friendly interface.

Dataset(s):

The datasets used include real-time data from various network traffic sources. The datasets relies on diverse datasets obtained from various sources of network traffic. These datasets include real-time data collected from different network environments, such as corporate networks, research networks, and public network traces. The datasets are used for training and testing the attack detection algorithms, as well as for analyzing the impact of known attacks and identifying new attack patterns.

Contact person: Raziye Armağan Keskin (armagan.keskin@pavotek.com.tr)

3.3.27 RISE

Driver Monitoring System

The developed system is a smartphone-based smartwatch companion app that utilizes the sensors of both devices to gather data on the usage of smart devices during driving. The primary objective is to enhance road safety by minimizing driver inattention and distraction. This is achieved by alerting the driver about tasks that involve diverting attention away from driving or essential driving activities towards competing activities, which are mainly related to the usage of smart devices like smartphones and smartwatches. Smartphones are a significant contributor to road accidents due to activities such as reading, texting, calling, and checking activity. Therefore, the system's focus is on using smartphones and wearables as sensor devices to detect distraction and communicate the detected distraction to the driver.

Figure 54 below presents the conceptual diagram of the components developed to collect sensor data and perform Artificial Intelligence (AI) experimentation, using Machine Learning (ML) to predict driving distractions using smart devices. For more details, the interested reader may refer to the publication [44].

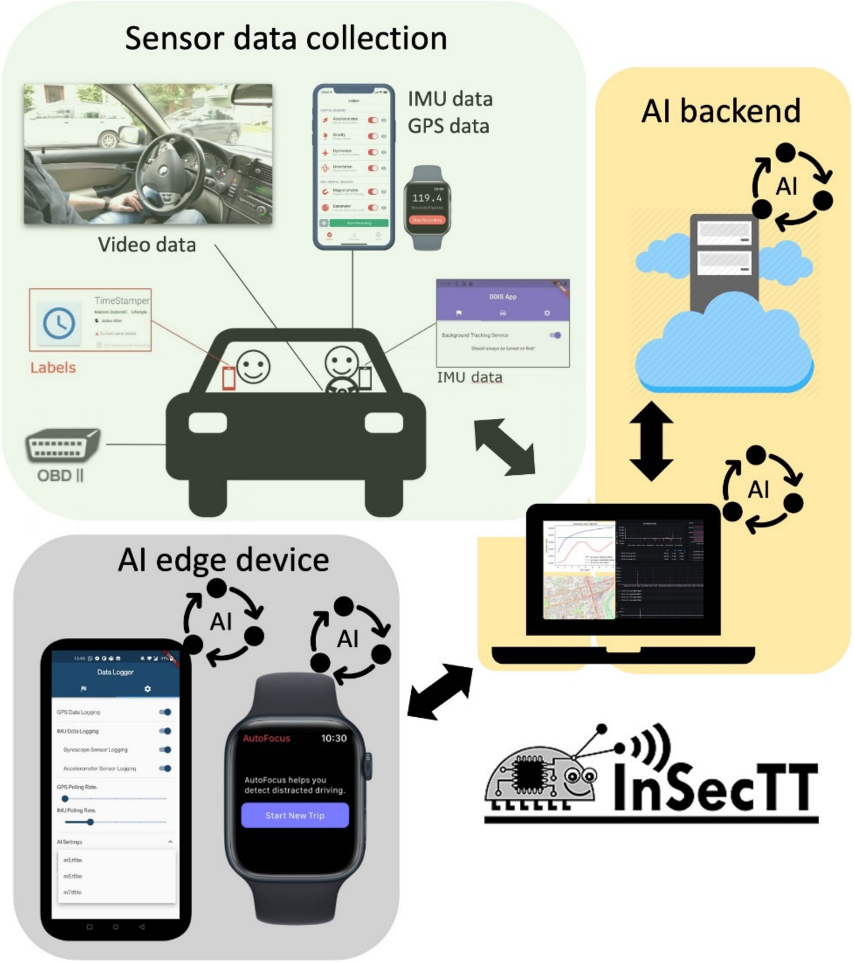


Figure 54 Driver Monitoring System Conceptual Diagram

Datasets:

A labelled dataset has been produced containing the timestamp, seconds elapsed, IMU (gyroscope, gravity and accelerometer), GPS latitude and longitude of an activity containing driving distractions or normal driving (non-distracted). The dataset is proprietary to the project group.

Contact persons: Efi Papatheocharous (efi.papatheocharous@ri.se) and Anders Wallberg (anders.wallberg@ri.se)

3.3.28 RTE

Preventive maintenance for industrial IoT

We have developed an ML-based platform for anomaly detection in industrial IoT systems. The target for anomaly detection is predictive maintenance. More specifically, the motorboat m/v Tucana in Gdansk harbor has been equipped with eleven sensors. RTE has during the project developed various machine learning-based algorithms for anomaly detection and classification that can be used for predictive maintenance. Anomaly detection in this context, mainly regards identifying certain predefined states for the onboard systems on the boat. For practical purposes, a simplification was done that will show that the principles work as intended, a proof of concept. This is done to shortcut the otherwise time-consuming task of labelling the data that is received from the onboard sensors. Finally, end-to-end communication has been implemented from boat to RTEs ML-based platform and from there to the user interface as handled by Vemco.

Performing near real-time predictions in a live system affects choice of algorithms, communication protocols and overall software design. Embedded designs need to be reliable and include a sufficient amount of error handling, which extends to the parts that implement ML. Hence, the choice of tools, frameworks and standards chosen for this project have been made with embedded applications in mind, although the ML part itself is platform independent. In stage 1 of the project, the ML model is a sequential neural network, implemented by using the TensorFlow framework. The model is saved as a protbuf (.db) file, which makes it easy to deploy on different hardware devices.

Stage 2 includes training a neural network in a cloud environment and export it as a TensorFlow Lite (.tflite) model. TensorFlow Lite is optimized for on-device machine learning and has multiple platform support, covering Android and iOS devices, embedded Linux and a range of microcontrollers. The second development stage also includes an exploratory approach, in which an unsupervised ML model is trained (clustering).

Figure 55 below depicts an architectural overview on the deployed set-up in data acquisition for the activities completed in the project.

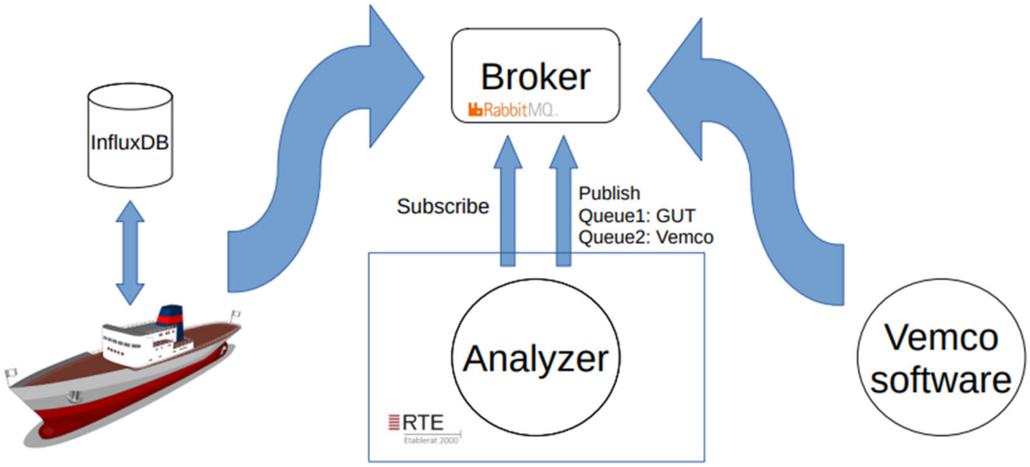


Figure 56 Data acquisition set-up for anomaly detection for predictive maintenance purposes

Dataset:

A labelled dataset has been created using real sensor measurements from a motorboat for professional use in the harbour of Gdansk.

Contact persons: Magnus Isaksson (magnus.isaksson@rte.se) and Christina Gratorp (christina.gratorp@rte.se)

3.3.29 STM

STM32 MCU/MPU security

STM has supported che CEA in the development of the HistoTrust platform that aims at authenticating the issuing device and protecting – at the device level – the integrity of an embedded AI (neural network model) by combining software and hardware security components available in the STM32 MP1. Thanks to this cooperation, has allowed STM to improve the design robustness of the novel STM32 MCU that ST is currently designed. This new STM32 has been enhanced by an increased robustness against side channel attack. It is based on the Arm® Cortex®-M7-based STM32H7 MCU series leverages on latest silicon technologies to achieve high performance. As mentioned, this new product is designed with a comprehensive set of security features one of these features is the integration of 2 AES accelerators. One is targeted to get high performance while the second is side-channel attack protected. Another improvement thanks to this project is the introduction of secure firmware install feature. This feature is an immutable secure service embedded by STMicroelectronics in this device. This feature allows secure and counted installation of OEM firmware in untrusted production environment (such as OEM contract manufacturer). The confidentiality of the installed images written either in the internal flash memory or encrypted in an external flash memory, is also protected, using the AE. Last but not Least, the HistoTrust technology developed by CEA that attests in an Ethereum ledger all the relevant data produced by a physical device will available, thanks to the security extensions, to the STM32 novel device to strengthened against advanced integrity-based attacks that may leverage algorithmic and physical flaws.

Contact person: Marcello Coppola (marcello.coppola@st.com)

3.3.30 TIETO SE

Federated Learning for Network Intrusion and Attack Detection

Tieto SE worked on a federated learning approach for detecting potential network intrusions and attacks on network devices. To achieve that we used Random Forest techniques that we extended to the Federated Learning setting. Random Forests are a good candidate for Network Intrusion and Attack detection as they have good prediction performance and relatively small in terms of size and inference time. These reasons make them adequate to use them in a Federated environment where they would be deployed on low-resource devices. Federated Learning has the aim of building a global model from the aggregation of multiple local models. A local model is trained on local data present on a device and then sent to a server. The server combines the local models to form a global model, which can be sent and deployed to the devices.

By design, it enables additional data privacy property to the system. However, we were interested in going further by adding the differential privacy guarantee to our models. We carried out experiments on openly available datasets where we showed the effectiveness of our approach.

This work has been made in collaboration with MDH and part of the work has been published or submitted.

For more details, please look at the item from MDH called **Federated Learning Framework for Network Attacks Detection and Classification Based on Random Forest** on page 62.

Datasets:

In this context, publicly available dataset have been utilized:

- KDD ([36])
- NSL-KDD ([37])
- UNSW-NB15 ([38])
- CIC-IDS-2017 ([39])

Contact person: David Buffoni (david.buffoni@tietoevry.com)

Driver Monitoring system with Computer Vision

Tieto SE developed a Computer Vision based system for detecting distracted drivers. A recurrent cause of road fatalities is due to distracted drivers. Our aim is to reduce these distractions by proposing an agnostic approach of automatically detecting them and educate the drivers of the potential consequences on their driving. Smart devices are more and more used and can be also a distraction source. We focused on detecting smart devices usage such as smartwatch or smartphone with computer vision techniques.

To achieve that, we combined an open image dataset with some data collected on our own, to build a computer vision system capable of detecting three categories: driving safely, using phone or using watch. We also investigated techniques for overcoming to generalisation performance challenges such as data augmentation, oversampling... Then, as data privacy may be a challenge, we consider of developing models which can be deployed on low-resource devices. To that end, we focused on quantization techniques for optimizing size of the model. We were able to divide by a factor of 10, the size of the model without losing accuracy performance.

Datasets:

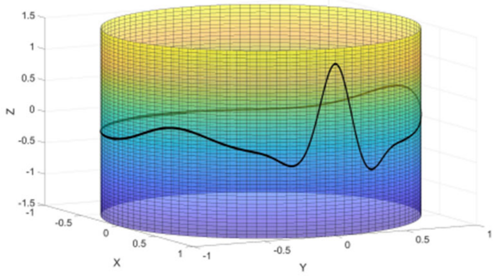
- State Farm Distracted Driver Detection ([45])
- Additional proprietary data manually collected in the InSecTT project.

Contact person: David Buffoni (david.buffoni@tietoevry.com)

3.3.31 TU-DELFT

AI based ECG anomaly detection

TUDelft worked on a dynamical systems approach to synthesizing and analyzing ECG signal data in real time, and algorithms which are deployable on edge devices. As many patients are in remote places and in many countries the patients to doctors ration is very poor which calls for a miniature hardware that remotely captures ECG and transmits data to the doctors. However, the exact reproduction of ECG requires high bit rate and thus requires transmitting a compressed set of parameters. Further, sending large volumes of annotated raw data to train diagnostic models also compromises the patients privacy. We design and present a system that generates synthetic ECG signals from clinical data in real-time using a highly minimized set of parameters. The system comprises a nonlinear dynamical model whose parameters are trained in real-time to synthesize a signal which matches clinical data with high accuracy. The parameters of the trained system are then transmitted in each cycle of the ECG wave to reconstruct the original signal using the same model at the medical practitioners' location. The parameter learning problem is highly complicated as one needs to solve a nonlinear, non-convex dynamic optimization problem, which usually only converges to local optima. To address this issue, we propose a novel two-stage algorithm that automatically chooses an initial set of parameters in the vicinity of the global optimum and then performs stochastic gradient descent iterations. We perform experiments to demonstrate the accuracy and real-time performance of the system. We show that on average our system processes clinical data of one second in 0.68 s on a microcontroller, with an RMSE error of 0.0038 the average, and 17 parameters per ECG cycle. Our system is also easy to implement, requires minimal storage i.e. only one ECG cycle at any given time, and does not depend on offline training, unlike existing methods.



$$
\begin{aligned}
\dot{x} &= (1-\sqrt{x^2+y^2})x - \omega y \\
\dot{y} &= (1-\sqrt{x^2+y^2})y + \omega y \\
\dot{z} &= -\sum_{i\in\{P,Q,R,S,T\}} a_i \Delta\theta_i \left(-\frac{\theta_i^2}{2b_i^2}\right) - (z - z_0)
\end{aligned}
$$

Figure 57 Gaussian mixture ODE on a unit cylinder describes the PQRST ECG wave

We demonstrate our contributions via extensive experiments that: 1) Our system does not rely on any training data, and uses only one ECG cycle at any given time, thereby minimizing storage and enabling real-time operation. 2) The learning algorithm takes sufficiently less time than the ECG signal rate, which allows enough buffer time for implementing wireless communication protocols for real-time transmission. 3) We show a clear improvement of the learning error (RMSE) over existing methods, while simultaneously minimizing the algorithmic complexity. 4) We also propose a hardware system called heart watch which comprises data-acquisition.

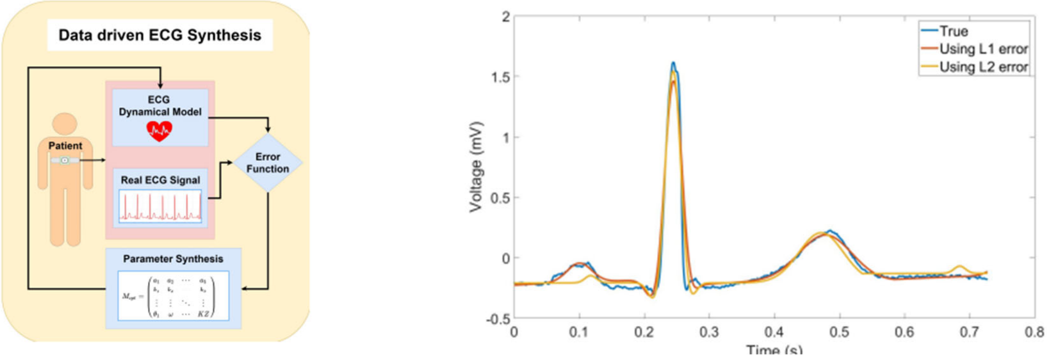
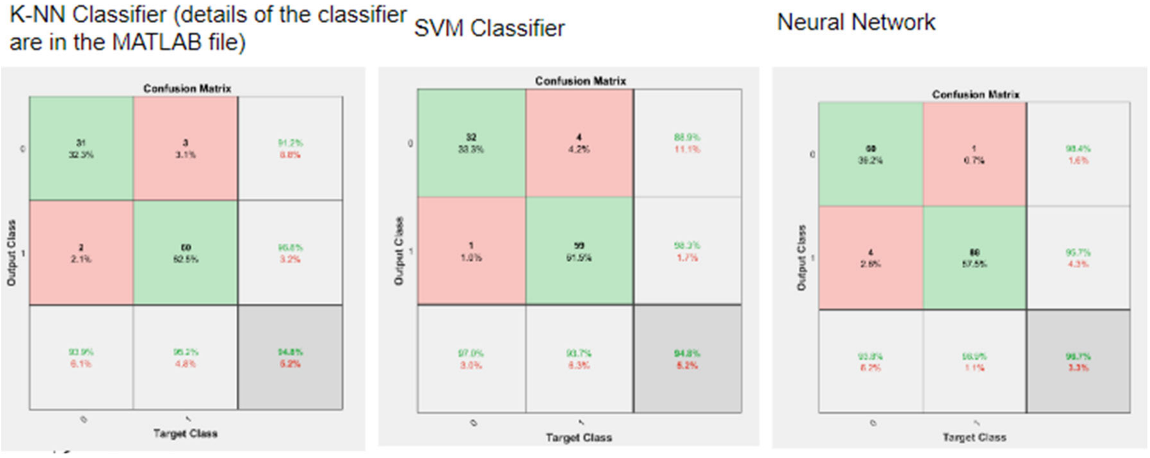


Figure 58 Real-time parameter learning based data driven ECG synthesis

Anomaly classification: The parameter vector extracted by the dynamical system were used as features for classifying morphological anomalies (I.e. obtained within a single cycle). Typical AI methods such as CNN based ones are computationally and storage-wise heavy, as the input layer takes the entire ECG signal. We found that inserting the parameter layer here instead, highly reduces computation. Further, we found that a novel way of classifying anomalies was to reconstruct the ECG using reduced parameter dimensions (via PCA) and classifying the RMSE vector as belonging to an anomalous or healthy signal.


K-NN Classifier (details of the classifier are in the MATLAB file)

SVM Classifier

Neural Network

Figure 59 Anomaly classification results

Link to datasets used (Physionet): https://www.physionet.org/content/mitdb/1.0.0/

Reference: Simha, Ashutosh, et al. "Heart Watch: Dynamical Systems Based Real Time Data Driven ECG Synthesis." *2021 IEEE 7th World Forum on Internet of Things (WF-IoT)*. IEEE, 2021.

Smart localization and formation of multiagent systems:

The context of this research is that there is a single anchor, which is possible mobile and one or many mobile agents which are relatively localized. Majority of existing single-anchor localization algorithms make use of antenna arrays or special antenna systems. However, the need for specialized antenna systems incurs higher costs, complexity and power consumption. This paper presents a novel single-anchor localization algorithm, which does not require antenna arrays or special antennas. The algorithm is implemented in a two-robot system, where one of the robots acts as the anchor and the other acts as the target. The localization algorithm uses velocity measurements of each robot and distance measurements between robots. Using the change in distance between the robots and the velocity of the target robot relative to the anchor robot, the target robot can be localized relative to the anchor robot. The localization algorithm uses a Kalman filter as a state estimator of the movement of the target robot and a Savitzky-Golay filter to filter the distance measurements. Several simulations and laboratory experiments, as well as experiments in real warehouse environments (indoor) have been performed.

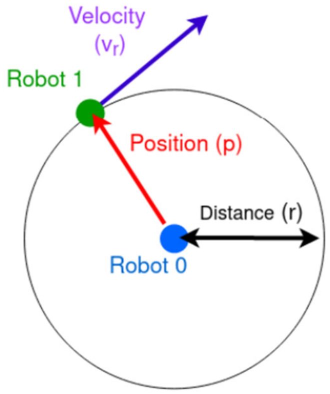


Fig. 1: A model of the problem setup. Given are two robots (R_0, R_1). Robot R_0 acts as a mobile anchor and R_1 is the target. Two given measurements are the velocity v_r of R_1 relative to R_0 and the distance r between the two robots. The desired result is the position p of R_1 relative to R_0 .

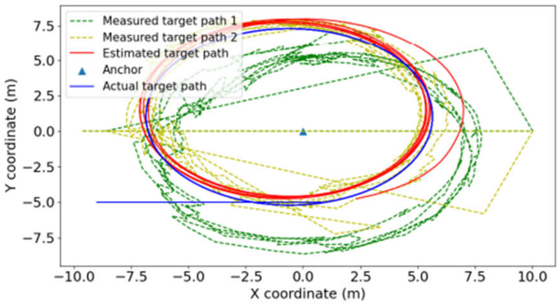
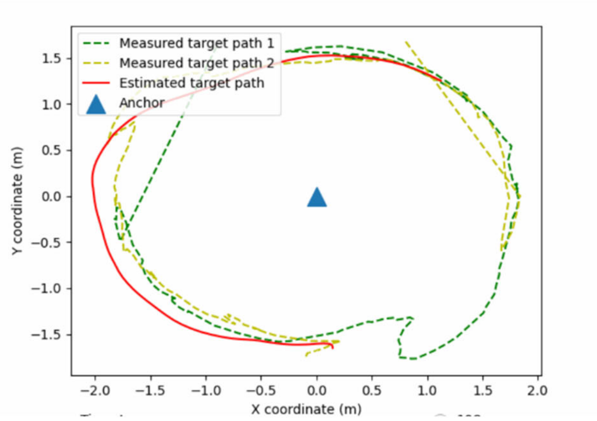


Fig. 10: Simulation with Gaussian noise added to the measurements, combining Kalman filter and Savitzky-Golay filter to filter the distance measurements ($\sigma_r = 1.0$, $\sigma_v = 0.1$).



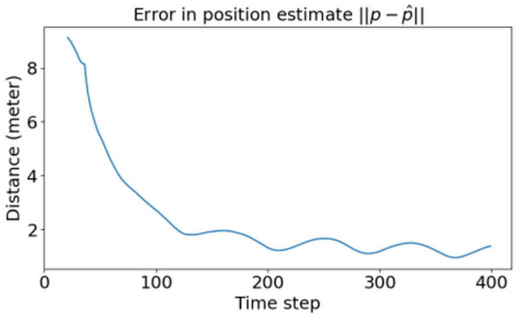


Fig. 12: The error in estimated position over time. In this case, the algorithm initially localized to the incorrect measured position (same scenario as Figure 11). The error in estimated position gradually decreases over time and converges.

Figure 60 Model and results for smart localization and formation of multiagent systems

Application**:** Formation flight in leader-follower configuration of fixed wing UAV systems:

The above mentioned algorithms were applied to the formation flight problem of fixed wing UAVs with a leader-follower configuration. The leader was the mobile anchor and the follower had to localize itself globally, in a relative path with respect to the leader, as well as control the flight dynamics to maintain formation. The developed algorithms were verified in hardware-in-loop as well as real outdoor flight experiments.

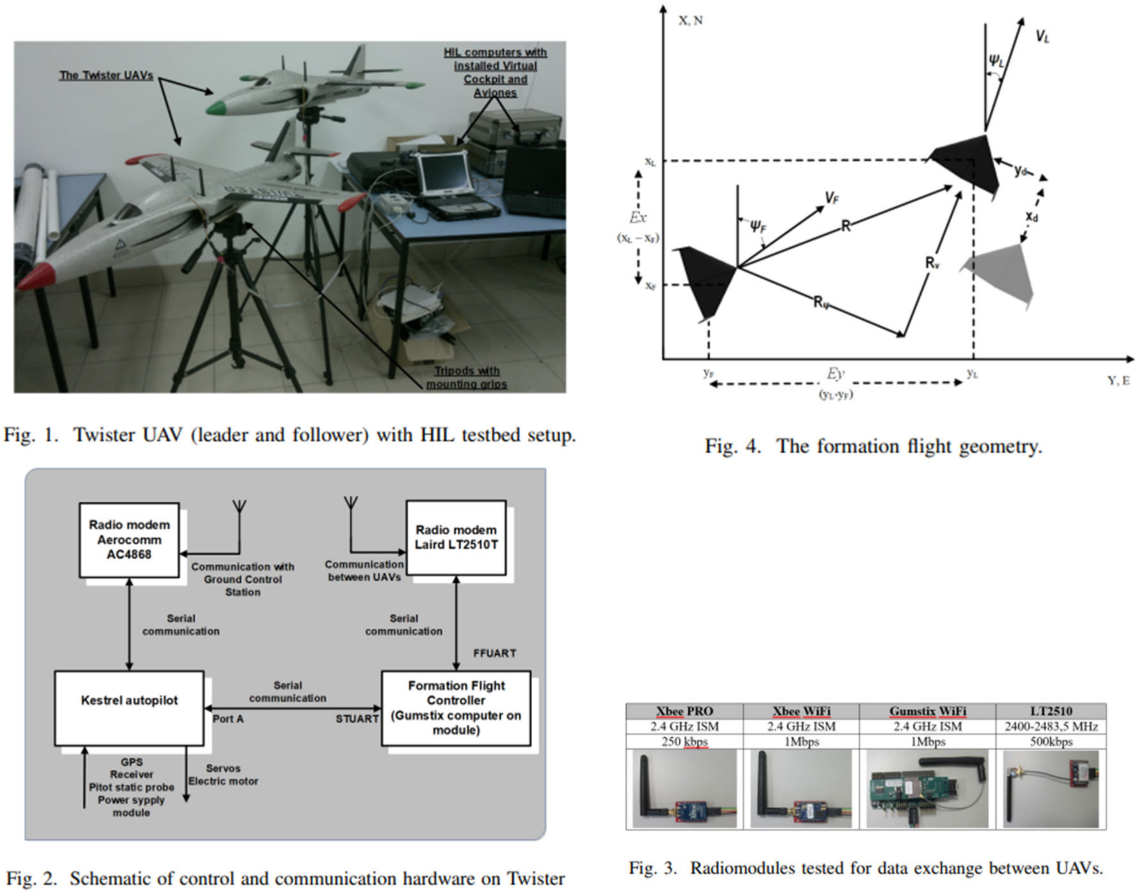

The Twister UAVs
HIL computers with installed Virtual Cockpit and Aviones
Tripods with mounting grips
X, N
V_L
ψ_L
x_L
Ex
$(x_L - x_F)$
x_F
ψ_F
V_F
R
R_x
R_y
y_d
x_d
y_F
Ey
$(y_L$-$y_F)$
y_L
Y, E
Radio modem Aerocomm AC4868
Communication with Ground Control Station
Communication between UAVs
Radio modem Laird LT2510T
Serial communication
Serial communication
FFUART
Kestrel autopilot
Serial communication
Port A
STUART
Formation Flight Controller (Gumstix computer on module)
GPS Receiver Pitot static probe Power sypply module
Servos Electric motor

Fig. 1. Twister UAV (leader and follower) with HIL testbed setup.

Fig. 4. The formation flight geometry.

Fig. 2. Schematic of control and communication hardware on Twister

Xbee PRO	Xbee WiFi	Gumstix WiFi	LT2510
2.4 GHz ISM	2.4 GHz ISM	2.4 GHz ISM	2400-2483,5 MHz
250 kbps	1Mbps	1Mbps	500kbps

Fig. 3. Radiomodules tested for data exchange between UAVs.

Figure 61 Formation flight control setup

The smart localization and formation algorithm was tested in HIL setup and hardware in the following configuration:

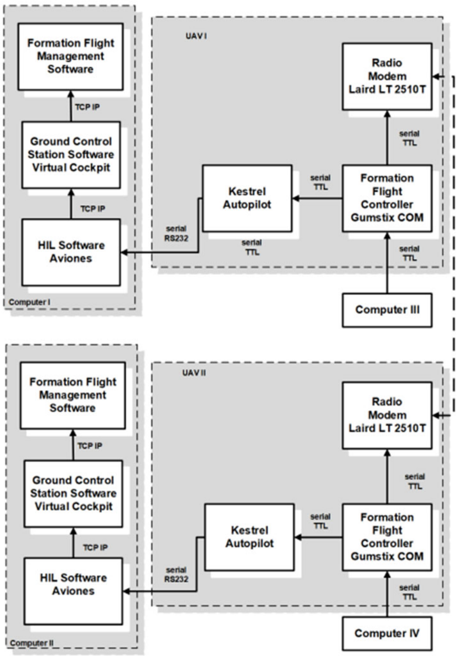

Formation Flight Management Software
TCP IP
Ground Control Station Software Virtual Cockpit
TCP IP
HIL Software Aviones
Computer I
UAV I
Radio Modem Laird LT 2510T
serial TTL
Kestrel Autopilot
serial TTL
Formation Flight Controller Gumstix COM
serial RS232
serial TTL
serial TTL
Computer III
Formation Flight Management Software
TCP IP
Ground Control Station Software Virtual Cockpit
TCP IP
HIL Software Aviones
Computer II
UAV II
Radio Modem Laird LT 2510T
serial TTL
Kestrel Autopilot
serial TTL
Formation Flight Controller Gumstix COM
serial RS232
serial TTL
Computer IV

Fig. 7. The HIL testbed scheme.

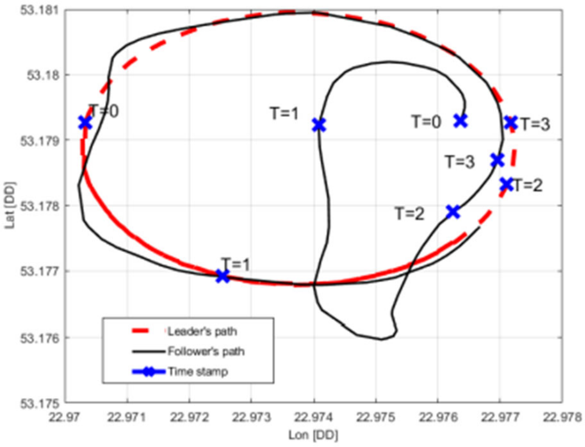


Fig. 11. The UAV trajectory during in-flight tests of switching formation flight controller.

Figure 62 Formation flight control results

Reference: Ambroziak, L., C. Kownacki, and A. Simha. "Switched Control Strategy for Robust Formation Flight with HIL and In-Flight Validation." *ICC 2022-IEEE International Conference on Communications*. IEEE, 2022.

Contact person: Ashutosh Simha(<a.simha@tudelft.nl)

3.3.32 U-TWENTE

Autonomous Network Slicing and AI based QoS management in IoT Networks

IoT devices in any IoT network rely on reliable and efficient connectivity with each other and gateways to exchange data. Connectivity entails more than just the presence of a link, rather IoT applications have specific throughput, latency and reliability demands, also known as Quality of Service (QoS) requirements. Meeting these requirements is crucial to the success of IoT use cases for example haptic feedbacks and video not being delivered on time to a remote surgeon will lead to failure of surgery. To meet QoS requirements efficiently and dynamically in challenging wireless environments, Utwente has developed a Software Defined Networking (SDN) based autonomous network slicing solution in a real-world Linux-powered Access Point. The SDN controller is connected through persistent TCP connections to the Access points (AP) in the IoT network and collects network statistics and performance KPIs to control and manage the network. The network statistics and KPIs are stored in the database and read from there continually by the SDN applications to manage QoS in the network. Our proposed SDN framework in shown in Figure 63 below ([46]).

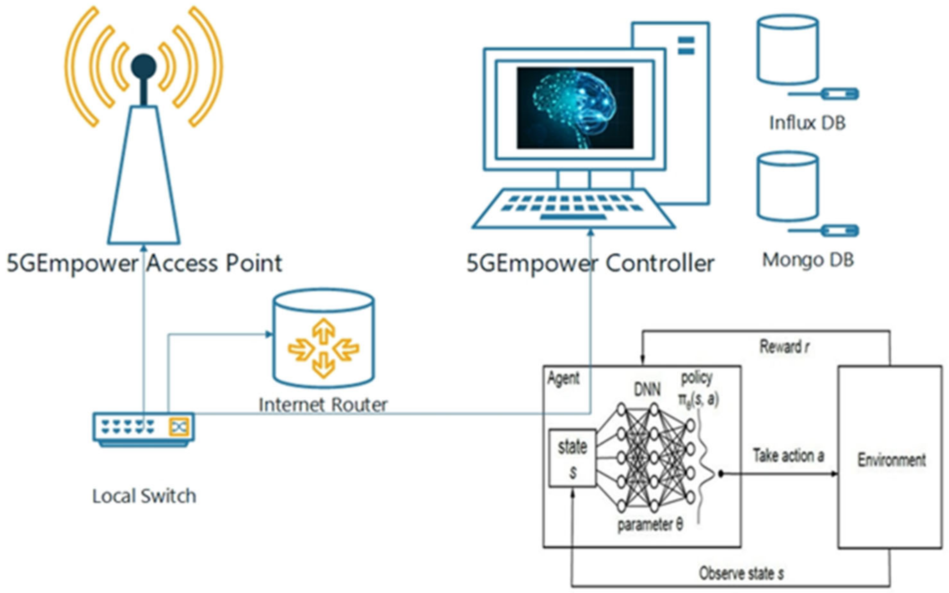


Figure 63 SDN framework architecture for AI based QoS management.

In order to meet QoS requirements, the APs need to understand the QoS requirements of the traffic flows present in the network. The applications and services in an IoT network are classified into QoS classes based on Differentiated Services Code Point (DSCP), representing traffic characteristics, and subsequently network slices are autonomously created. The IoT traffic flows, representing a QoS class, are matched to the correct network slices using traffic classifier and then resources are assigned to the slices to meet the QoS requirements. The workflow of the developed solution is shown in Figure 64.

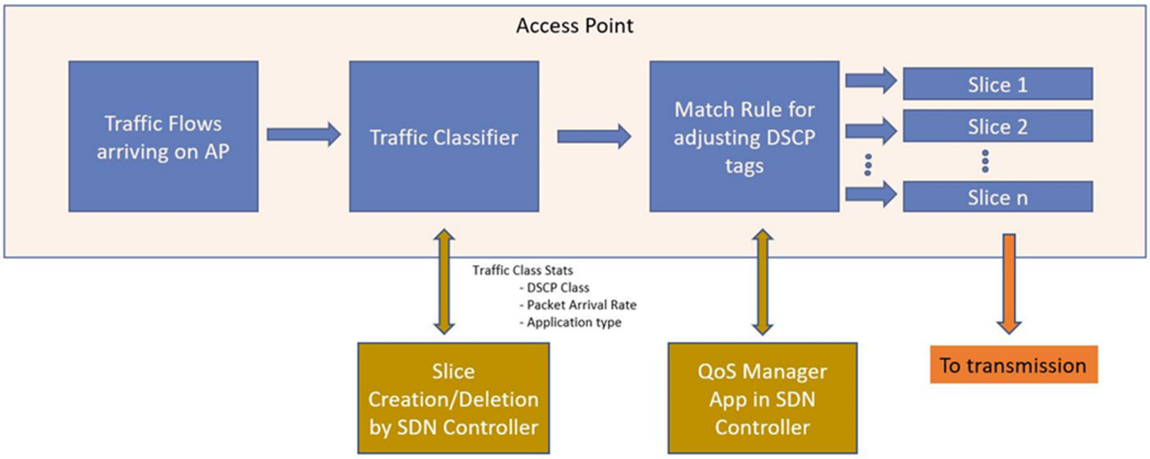

Access Point
Traffic Flows arriving on AP
Traffic Classifier
Match Rule for adjusting DSCP tags
Slice 1
Slice 2
Slice n
Traffic Class Stats
- DSCP Class
- Packet Arrival Rate
- Application type
Slice Creation/Deletion by SDN Controller
QoS Manager App in SDN Controller
To transmission

Figure 64 An overview on processing flow in AI based QoS management.

Since wireless channel conditions is a continually changing phenomenon, the resource management of network slice to meet QoS requirements needs to be dynamic and adaptive to the channel conditions. For this purpose, UTwente has developed an AI based slice resource management solution and employed a Deep Reinforcement Learning (DRL) model to allocate slice resources (Airtime). The DRL-based SDN application assign resources (Queue Quantums) to slices based on slice requirements, which are estimated using packet arrival rates and packet sizes in each of the slice, to meet their throughput requirements. The DRL agent continually adapts the slice resource management depending upon wireless channel conditions and QoS requirements so that IoT sensors and applications get the required QoS all the time. The developed system works in completely autonomously and can prioritize slices over one another by adjusting quantum assignments. The proposed solution is also able to prioritize IoT sensors over one another by moving their traffic flow from one slice to the other using traffic rules created in the SDN controller.

The DRL based autonomous network slicing solution is able to meet the QoS requirements of the IoT devices and applications in a complete autonomous manner and adapts to the changing requirements and channel conditions ([47], [48]).

Dataset:

> The system was trained in real world environment using experiments run in real time in a IEEE 802.11n based IoT network using 5GEmpower framework. The developed SDN applications and 5GEmpower agent is made available to public on Github.

Contact persons: Alessandro Chiumento (a.chiumento@utwente.nl) and Kamran Zia (k.zia@utwente.nl)

3.3.33 UCC

Remaining Useful Life Prediction

UCC has identified bending cycles in the operation of Liebherr crane ropes and, through the application of the Freyer equation for the calculation of tensile stress, has arrived at a prediction of rope damage as a function of position along the rope. See Figure 65 below.

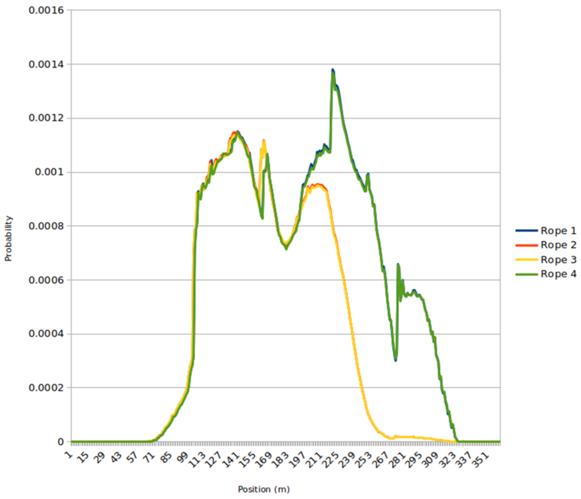


Figure 65: An example of probability of rope damage according to position along several ropes

Our technology is being incorporated into a new digital platform under development by Liebherr, allowing better planning of maintenance operations.

Contact person: Ken Brown (k.brown@cs.ucc.ie)

Video Camera Anomaly Detection

UCC developed a deep learning model for verifying the physical camera source of video feeds. The approach is based on identifying a unique noise fingerprint generated by the camera pixel sensors. Our method improves on the state-of-the-art on the public Vision dataset [49]. We achieve 98% accuracy on a specific application dataset provided by Liebherr, in which frames have had minimal encoding or compression.

We also developed a twin neural network model for detecting anomalies in live video feeds from a moving camera (see Figure 66 below). Two frames are passed in as input, and the observed differences are compared to the predicted differences based on the known camera displacement. Our experiments have demonstrated that we can predict camera movement from the images with RMSE 0.8(mm) and an R^2 value of 98%

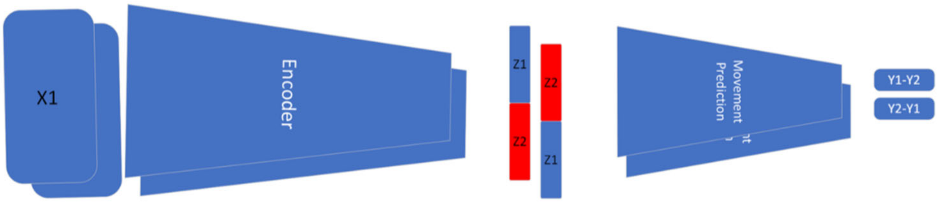


Figure 66 UCC's twin-neural network structure for detecting anomalies in live video feeds

These two approaches were combined into a general video feed classifier for the detection of anomalous video feeds from video cameras. Our software is currently being evaluated using video feeds obtained from a Liebherr crane in an industrial setting.

The above work was performed on proprietary data sets obtained from Liebherr, namely crane rope data and digital video camera recordings.

Contact person: Ken Brown (k.brown@cs.ucc.ie)

3.3.34 UPM

Custom modular IoT HW platform for LiDAR point-cloud processing at the edge

In the last years, new demands for processing capabilities along with low power consumption and high efficiency have appear due to new application scenarios. On e of these scenarios involves the use of sensors that produce high amount of that as LIDAR, and that require to process 3D information.

Deep learns has demonstrated to be efficient and well performing for processing 3D point clouds for object detection and classifications. However, the requirements for processing these DLs have made the impossible to be executed in low end processors as the ones traditionally used at the edge in IoT applications.

In InSecTT, a custom HW platform for processing LiDAR point-clouds at the edge has been designed, manufactured and tested satisfactory. The platform includes a processor with ARM Cortex A5 architecture and a Coral coprocessor for intensive DL processing at the edge (see Figure 67).

Additionally, a simulation environment has been developed to be able to synthetically generate dataset form different LiDAR models with even different operating principles. The datasets can be used to train different neural networks and to select the most appropriate LiDAR for a specific application. This approach, combined with the custom platform represents a powerful element to successfully develop and deploy IoT edge application based on 3D point-cloud LiDAR information

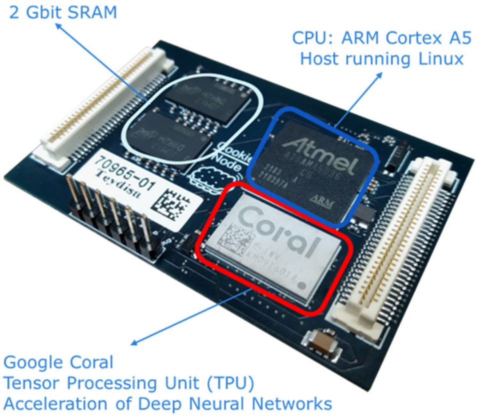

2 Gbit SRAM
CPU: ARM Cortex A5
Host running Linux
Google Coral
Tensor Processing Unit (TPU)
Acceleration of Deep Neural Networks

Figure 67 HW platform for DL processing at the edge

More details regarding the platform and the synthetic data generation can be found in [51] and [52]. Also, part of the work has been included in Cristian Wisultschew's PhD Thesis [53].

Contact person: Jorge Portilla (jorge.portilla@upm.es)

3.3.35 VIF

VehicleCAPTAIN Toolbox: Routing Core and ITS library

The vehicle communication platform to anything (vehicleCAPTAIN) was created to deal with the multitude of soon available V2X communication technologies, such as 802.11p, 802.11bd, C-V2X (LTE), C-V2X (5G) and V2N in general. Each of the technologies is provided by multiple hardware vendors, each with a different API.

The vehicleCAPTAIN toolbox, shown in Figure 68,solves this problem by providing a single interface for communication, while using a routing software to distribute and receive V2X messages via each of the connected interfaces. The vehicleCAPTAIN toolbox is free and open source software (FOSS), available on Github [2]. The software (developed in WP2) consists of:

- An ITS message library that provides most basic V2X functionality for early testing and development of new message formats, without the need to buy inflexible closed source packages.
- A Routing Core that provides a single interface which connects to multiple connected hardware interfaces. (Authors Note: the software is not yet FOSS, because a way has to be found to mask licensed libraries, before uploading, i.e. the libraries from each communication interface)
- A ROS2 interface for V2X in general and for the routing core in specific.

- A library of ROS2 type ITS messages. This part has also been proposed to the Autoware community [3], as future addons to the automated driving software.

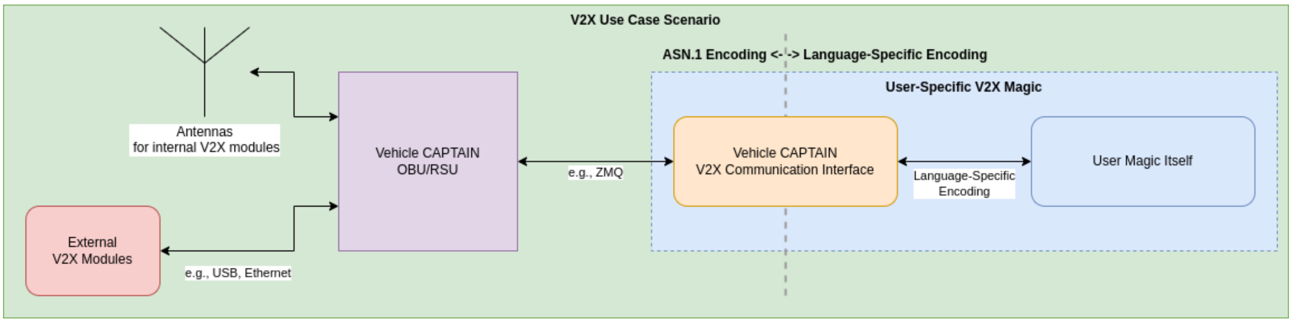


Figure 68 High Level Architecture of the vehicleCAPTAIN toolbox.

Antennas and Modules on the left are controlled by the vehicleCAPTAIN Routing Core (OBU/RSU). The single communication interface connects to the routing core via ZMQ and provides language specific encoding with ITS libraries.

The vehicleCAPTAIN toolbox is part of the automated driving demonstrators at VIF and is continuously improved and used across projects. The vehicleCAPTAIN toolbox was also already provided to partners in other projects, to define a common interface.

Finally, we are proud to say that the vehicleCAPTAIN has its own chapter in the InSecTT book. And as will be discussed in the accompanying WP3 document [1], two papers validating the concept are on its way of publication.

Contact person: Christoph Pilz (christoph.pilz@v2c2.at)

Trustworthiness Whitepaper

The development of artificial intelligence (AI) technologies experiences worldwide an ongoing challenge to become trustworthy and ethical for users and the general public. This challenge currently stands between the promise of AI to create immense societal and individual impact and its realization. Because of this, possible large marketplaces still remain hesitant or closed. We have investigated this problem and identified potential solutions in the InSecTT project, a large international EU research and development project that investigates ethical smart technologies1. Thereby, working with industrial and research partners, we assert that developing trustworthy AI technologies is not foremost a technical challenge but increasingly an organizational and structural challenge that results from applying traditional ways of conceiving, designing, and selling technologies to new types of problems. Because AI technologies can shift the role that humans and society see as acceptable, traditional development processes that rely on strict separation of specialties are overburdened. In our view a research and development approach for trustworthy AI systems should put human concerns and needs at the center of the development process to effectively integrate humans and systems. Also, such approach should be based on EU guidelines for developing ethical AI. Our proposed approach to develop trustworthy AI systems combines these two directions and centers around the assessment of trustworthiness risks through intensive user involvement prior to the elicitation of system requirement that are then managed throughout the system’s life-cycle. Also, the approach includes concrete recommendations to establish the organizational prerequisites that would enable organizations to implement the design process recommendations. All these aspects have been pulled together in the Trustworthiness Whitepaper [54].

In the white paper, the Human-Systems Integration (HSI) approach is described and motivated. The white paper intends to inform managers of technical organizations and product managers as well as principal investigators to set up the prerequisites for trustworthy AI. The white paper also wants to solicit inputs for further refinement and discussion.

Contact Person: Peter Moertl (peter.moertl@v2c2.at)

3.3.36 VTT

AI-Enriched Multimodal, Distributed Remote Vital Sign Analytics

VTT has developed signal processing and AI algorithms to measure bio signals about stationary and semi-stationary patients using RF-based milli-meter radar (60GHz) and regular COTS web camera, respectively. The main novelty is on machine learning models for robust extraction of signals from noisy sensor data with limited amount of training data, use of explainable AI for providing insight of the decision making for the end user and use of distributed ML deployment architecture. Figure 69 presents the developed multi-sensory data regression architecture for vital sign analysis.

In web camera, vital signs can be observed as slight changes in the colour of the skin. Face tracking is applied to a video feed to extract specific areas of the face where, for example, e.g., the pulsation of the heart can be detected as minor changes in the tone of the skin. Through signal processing steps this phenomenon can be presented as a time series rPPG signal. By applying further signal processing, the rPPG signal is transformed into an image presenting the frequency domain presentation of the signal. This is followed by running the image formatted data through analysis with neural network models which extract the heart rate value. As the neural network models are black box by nature (data is fed in, results are retrieved back), novel methods have been developed to provide more insight how the models actually work and what do they pay attention to. Two approaches for giving more information of the reasoning of the neural network were implemented for the purpose. GRADCAMs highlight the areas in the image formatted vital sign signal where the neural network pays attention to. Confidence estimation provides more straight-forward information from the end user point-of-view: it gives an estimate of the reliability of the given vital sign estimate.

Multimodal sensing is applied for Improving the robustness of the vital sign analysis. Milli-meter radar allows vital sign monitoring from stationary subjects in a field of view of about 160° in a radius of about 20 meters. The method is based on the monitoring of mechanical vibrations of the chest wall using reflected signals. The vibration causes small variations in phase between the signals reflected from consecutive transmissions. At the radar, reflections are analysed using a signal analysis method, called Cepstrum, that finds the periodicity of signal components, such as the respiration and heart rates. Both camera and radar-based methods are non-invasive, meaning that it does not require sensors or other measurement devices attached to the skin. Patients can be in different positions and unobstructive measurement is still possible. The method can be extended for other health application contexts, including, for example, automatic sleep quality and efficiency monitoring and respiratory sleep disorder measurement. As depicted in Figure 69, the method also supports simultaneous monitoring of several persons.

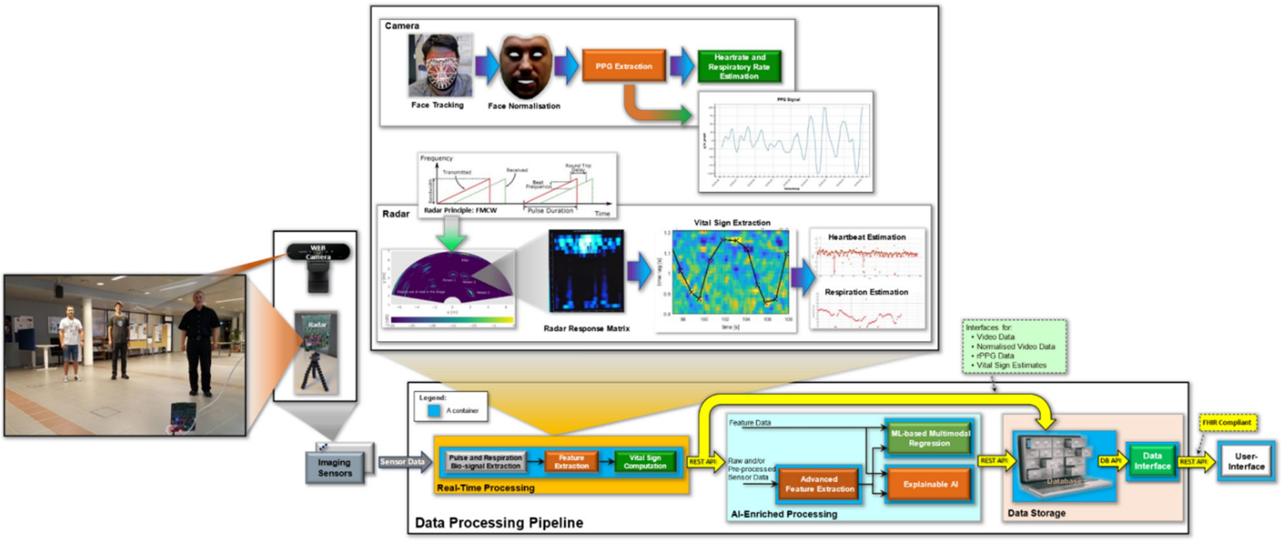


Figure 69 Multi-sensory data regression architecture for unobtrusive vital signs acquisition

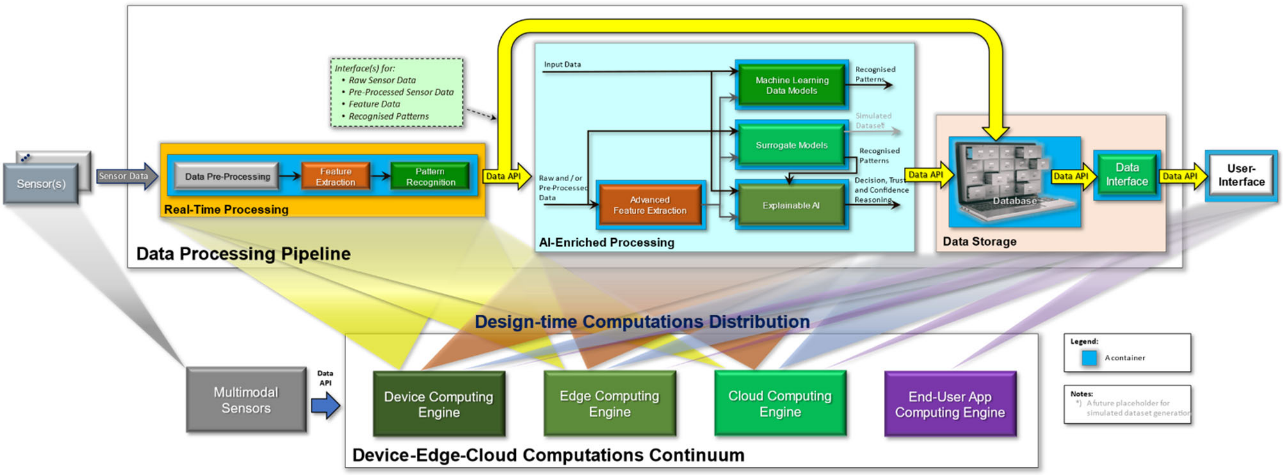


Figure 70 Data processing pipeline for distributed device-edge-cloud continuum framework

The processing of the multi-sensor vital sign analysis is deployed using the developed distributed processing framework for design time adaptive load balancing of AI processing between device, edge and cloud. Figure 70 presents the overall processing environment based on containerised nodes interacting via well-defined data APIs. The framework guides towards an optimal run-time operating environment set-up by considering various constraints, including but not limited to communications frequency and latency, available network bandwidth and deployable processing resources at hand, and data security. Resources can be re-allocated manageably between processing layers and nodes for reaching full benefit of distributed computing capabilities for IoT systems.

In the context of InSecTT use case about smart hospital, VTT has implemented a reference platform about the architecture, providing efficient processing of camera and radar data for extracting reliably vital signs information, and making data available at various abstraction levels ranging from low-level sensor data to AI-enriched vital sign information over well-defined RESTful data APIs.

Datasets:

VTT has utilized three public datasets for PPG data as follows:

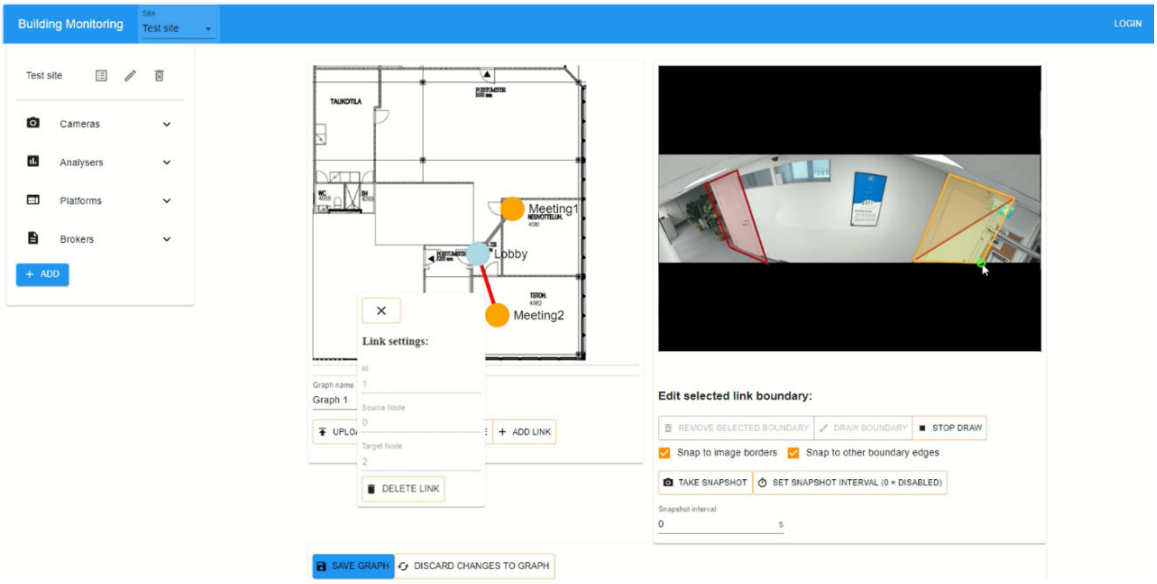
- PPG DaLiA ([55])
- PURE ([56])
- LGI-PPGI-Face-Video-Database ([57])

Contact person: Johannes Peltola (johannes.peltola@vtt.fi)

3.3.37 WAPICE

Machine vision-based object detection

Wapice has been working on a situational awareness concept suitable for tracking object movements in smart areas. The basic idea is to utilize the increasing number of cameras located in industrial zones or urban environment. These often have surveillance cameras installed and there already is monitoring based on these cameras, but automatic processing of camera streams is not used to full potential. Our computer-vision based monitoring solution can be used for tracking critical information concerning assets, vehicles or persons. The solution relies on video cameras that are part of existing infrastructure or mounted for purpose. The video stream can be used from single camera or combined from several cameras to the same computation unit, which then tracks and analyses certain areas of interest in video. The computation unit then sends the processed information to the back-end server, where a cloud application can be developed using Wapice's IoT and AI back-end tools. This solution can be used both in indoor and outdoor locations. Using a configuration tool, it is possible to draw areas of interest to the image captured from the video stream. Configuration data based on critical areas as drawn is then sent to the edge node that takes care of the video stream analysis. All information containing sensitive GDPR data is obfuscated in the edge node and only non-sensitive information is forwarded to the cloud. Figure 71 and Figure 72 below show the tool in action used in an indoor scenario at Wapice office. In the example in Figure 71, the critical areas are drawn to doors so that people entering or exiting the room can be tracked.



Link settings:

Id: 1

Source Node: 0

Target Node: 2

Edit selected link boundary:

☑ Snap to image borders ☑ Snap to other boundary edges

Snapshot interval: 0 s

Figure 71 WAPICE dashboard with marked areas of interest for building monitoring.

In Figure 72, the data send from edge node to cloud is demonstrated. Video image is only for demonstration and debugging purposes (normally all sensitive data is obfuscated in edge node).

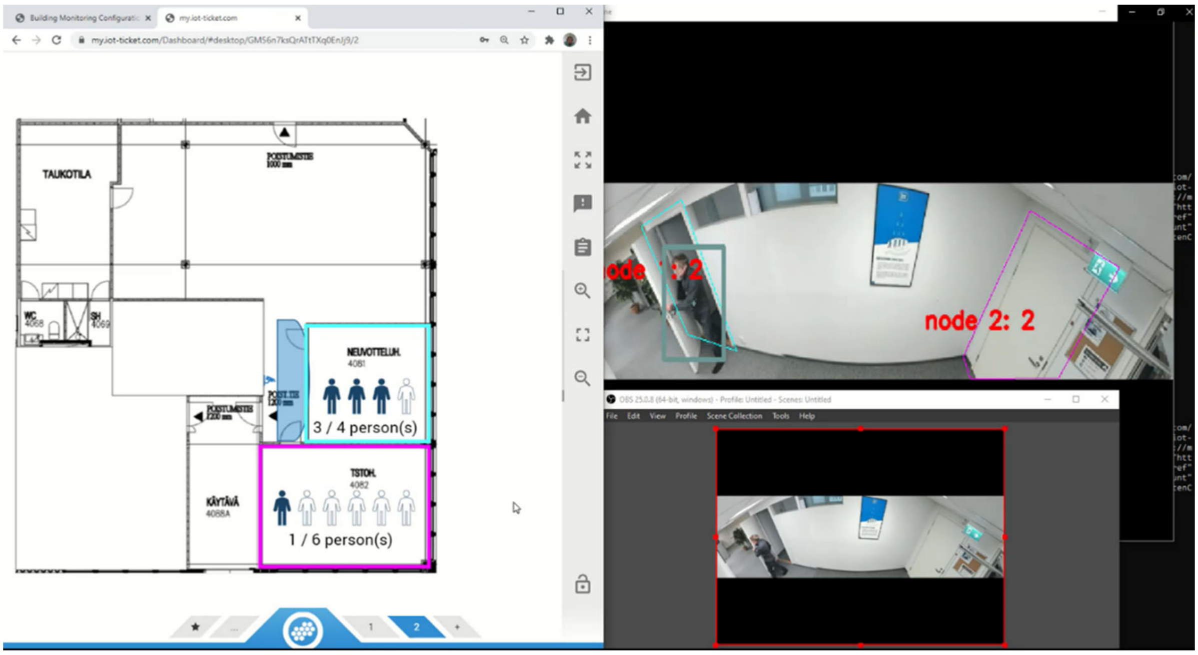

Building Monitoring Configurati
my.iot-ticket.com
my.iot-ticket.com/Dashboard/#desktop/GMS6n7ksQrATtTXq0EnJ/9/2
TAUKOTILA
POISTUMISTIE 1000 mm
WC 4068
SH 4069
NEUVOTTELUH. 4081
3 / 4 person(s)
TSTOH. 4082
1 / 6 person(s)
KÄYTÄVÄ 4068A
ode : 2
node 2: 2
OBS 25.0.8 (64-bit, windows) - Profile: Untitled - Scenes: Untitled
File Edit View Profile Scene Collection Tools Help

Figure 72 Machine vision-based object detection in WAPICE Dashboard

(left) Data send from edge node to cloud.

(right) Input video image is for demonstration and debugging purposes only.

Contact person: Veli-Pekka Salo (veli-pekka.salo@wapice.com)

4 CONCLUSIONS AND OUTLOOK

This deliverable presents a summary of the overall results accomplished within WP2: Reliable AI / Machine Learning for IoT. Over the course of the three-year project, WP2 has focused on developing trustworthy, distributed, secure, explainable, and powerful AI methods and platforms specifically tailored for IoT systems. The activities in WP2 have closely aligned with the efforts undertaken in WP3 to enhance wireless communication by integrating novel communication methods. As a result, the solutions devised in WP2 benefit from these advancements.

To streamline the content and account for the significant number of partners and their respective contributions, this deliverable primarily highlights the key exploitable items per partner. It offers a concise description of the main innovative components developed within WP2. Each exploitable item includes information on the target sectors or markets as well as the achieved Technology Readiness Level (TRL), demonstrating the technical readiness of the achievements and potential application sectors in the future.

Throughout the three-year duration of the InSecTT project, the partners involved in WP2 dedicated their efforts to developing both hardware and software solutions. These solutions have been successfully integrated into diverse demonstrators, which serve as significant highlights of the project. These demonstrators specifically cater to multiple industrial use-cases, showcasing the practical applicability and value of the developed solutions within various industries.

The technical achievements within WP2 are focused on:

- Anomaly detection methods can analyse different data streams and automatically learn the properties of normal behaviour and detect when data contains patterns that deviate from learned typical conditions.
- Parametric modelling and prediction methods estimate the statistics of the data and create predictions about data or data source behaviour in the future.
- Explainable and Interpretable AI tools can provide additional information about the neural network processing stages and importance of different properties of the input data. Explainable AI methods contribute to the trust of the AI systems as they can provide detailed information about the reasoning in AI decision making process.
- AI Model Security development provides tools how learning systems can protect their integrity from learning maliciously targeted incorrect data and how to protect the information stored in the neural network model structures.
- Link estimation, prediction, routing, interference detection and connection management
- Localization, Direction Of Arrival, beamforming and PHY-security
- Algorithms and architectures for AI on device / edge
- Design and run -time dynamic distribution of AI
- Tools and methodologies dedicated to the evaluation of complex AIoT systems
- Automatically (or semi-automatically) generate testing scenarios to improve the coverage of V&V
- Automatically select critical parameters – and appropriate metrics – and predict their impact at a module-level or at the global system level for specific modules or at the global level of the system.

- AI trustworthiness measures for end users and methodologies for improving trustworthiness of AIoT systems.

5 REFERENCES

[1] Ademaj-Berisha, F. (2023). D3.4 – Publishable Summary of WP3 Results. InSecTT.

[2] Pilz, C., Sammer, P., Neubauer P. (2023). vehicleCAPTAIN toolbox, https://github.com/virtual-vehicle/vehicle_captain

[3] Autoware Foudation (2023). Welcome to the Autoware Foundation, https://www.autoware.org/

[4] Nikolov, G., Kuhn, M., McGibney, A., Wenning, B.-L. (2022). MABASR – A Robust Wireless Interface Selection Policy for Heterogeneous Wireless Networks. IEEE Access, 10, 26068-26077.
https://doi.org/10.1109/ACCESS.2022.3156597

[5] Munster Technological University. (2022). 5G Measurement Data [Data set]. https://github.com/MTU-Insectt/Measurements5G

[6] Cagliaris, E., Davoli, L., Cilfone, A., & Ferrari, G. (2020). *A Modular Multi-interface Gateway for Heterogeneous IoT Networking*. 2020 International Symposium on Advanced Electrical and Communication Technologies (ISAECT), Marrakech, Morocco, 2020, pp. 1-6. doi:10.1109/ISAECT50560.2020.9523689.

[7] Davoli, L., Moreni, M., & Ferrari, G. (2022). *A Sink-oriented Routing Protocol for Blue Light Link-based Mesh Network*. Ch. 2 in *Wireless Mesh Networks for IoT and Smart Cities: Technologies and Applications*, pp. 21-31. doi:10.1049/PBTE101E_ch2.

[8] Cilfone, A., Davoli, L., Belli, L., & Ferrari, G. (2022). *Seamless IoT Mobile Sensing through Wi-Fi Mesh Networking*. Ch. 4 in *Wireless Mesh Networks for IoT and Smart Cities: Technologies and Applications*, pp. 67-80. doi:10.1049/PBTE101E_ch4.

[9] Codeluppi, G., Davoli, L., & Ferrari, G. (2021). *Forecasting Air Temperature on Edge Devices with Embedded AI*. Sensors, 21(12). doi:10.3390/s21123973.

[10] R. Salih Kuzu, E. Maiorana and P. Campisi, "Loss Functions for CNN-based Biometric Vein Recognition," 2020 28th European Signal Processing Conference (EUSIPCO), Amsterdam, Netherlands, 2021

[11] R.S. Kuzu, E. Maiorana, P. Campisi, "On the intra-subject similarity of hand vein patterns in biometric recognition," Elsevier Expert Systems with Applications, Vol. 192, 2022.

[12] R. S. Kuzu, E. Maiorana and P. Campisi, "Gender-Specific Characteristics for Hand-Vein Biometric Recognition: Analysis and Exploitation," in *IEEE Access*, vol. 11, pp. 11700-11710, 2023

[13] E. Maiorana, D. Ramaccia, L. Stefanini, A. Toscano, F. Bilotti and P. Campisi, "Biometric Recognition using Microwave Reflection Spectroscopy," *2022 24th International Microwave and Radar Conference (MIKON)*, Gdansk, Poland, 2022

[14] M. Neri, F. Battisti, A. Neri and M. Carli, "Sound Event Detection for Human Safety and Security in Noisy Environments," in IEEE Access, vol. 10, pp. 134230-134240, 2022, doi: 10.1109/ACCESS.2022.3231681.

[15] Fabbri, M., et al., "MOTSynth: How Can Synthetic Data Help Pedestrian Detection and Tracking?," 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 2021, pp. 10829-10839, doi: 10.1109/ICCV48922.2021.01067.
https://doi.org/10.1109/ICCV48922.2021.01067

[16] AImageLab, University of Modena and Reggio Emilia. (2022). MOTSynth [Dataset]. https://aimagelab.ing.unimore.it/imagelab/page.asp?IdPage=42

[17] Boschini, M., Bonicelli, L., Buzzega, P., Porrello, A. and Calderara, S. (2022). Class-Incremental Continual Learning Into the eXtended DER-Verse. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 5, pp. 5497-5512, 1 May 2023, doi: 10.1109/TPAMI.2022.3206549.

[18] Panariello, A., Porrello, A., Calderara, S., Cucchiara, R. (2023). Consistency-Based Self-supervised Learning for Temporal Anomaly Localization. In: Karlinsky, L., Michaeli, T., Nishino, K. (eds) Computer Vision – ECCV 2022 Workshops. ECCV 2022. Lecture Notes in Computer Science, vol 13805. Springer, Cham. https://doi.org/10.1007/978-3-031-25072-9_22

[19] AImageLab, University of Modena and Reggio Emilia. (2022). INSECTT Anomaly detection SETA Dataset. https://aimagelab.ing.unimore.it/imagelab/researchActivity.asp?idActivity=79

[20] Cornia, M., Baraldi, L., Cucchiara, R. (2022). Explaining transformer-based image captioning models: An empirical analysis. AI Communications, Volume 35, Issue 201 January 2022, pp 111–129. https://doi.org/10.3233/AIC-210172

[21] H.-P. Bernhard, A. Springer, A. Berger, and P. Priller, “Life cycle of wireless sensor nodes in industrial environments,” in 13th IEEE Int. Workshop Factory Commun. Sys., Trondheim, Norway, May 2017.

[22] Joud, R., Moëllic, P. A., Bernhard, R., & Rigaud, J. B. (2021, June). A Review of Confidentiality Threats Against Embedded Neural Network Models. In 2021 IEEE 7th World Forum on Internet of Things (WF-IoT) (pp. 610-615). IEEE. doi: 10.1109/WF-IoT51360.2021.9595434.

[23] Dumont, M., Moëllic, P. A., Viera, R., Dutertre, J. M., & Bernhard, R. (2021, June). An overview of laser injection against embedded neural network models. In 2021 IEEE 7th World Forum on Internet of Things (WF-IoT) (pp. 616-621). IEEE. doi: 10.1109/WF-IoT51360.2021.9595075.

[24] Joud, R., Moëllic, P. A., Pontié, S., & Rigaud, J. B. (2023, January). A Practical Introduction to Side-Channel Extraction of Deep Neural Network Parameters. In Smart Card Research and Advanced Applications: 21st International Conference, CARDIS 2022, Birmingham, UK, November 7–9, 2022, Revised Selected Papers (pp. 45-65). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-25319-5_3

[25] Dumont, M., Hector, K., Moellic, P. A., Dutertre, J. M., & Pontié, S. (2023). Evaluation of Parameter-based Attacks against Embedded Neural Networks with Laser Injection: 42st International Conference on Computer Safety, Reliability and Security, SAFECOMP 2023, Toulouse, France, September 19–22, 2023.

[26] Paulin, D., Joud, R., Hennebert, C., Moëllic, P. A., Franco-Rondisson, T., & Jayles, R. (2023). HistoTrust: tracing AI behavior with secure hardware and blockchain technology. Annals of Telecommunications, 1-15., 2023, https://doi.org/10.1007/s12243-022-00943-6

[27] A. Berezovskyi, R. Inam, J. El-khoury, L. Mokrushin, E. Fersman, “Integrating Systems of Systems with a federation of rule engines”. Submitted to the Journal of Industrial Information Integration.

[28] T. Nyberg, J. M. Gaspar Sánchez, C. Pek, J. Tumova and M. Törngren, "Evaluating Sequential Reasoning about Hidden Objects in Traffic," 2022 ACM/IEEE 13th International Conference on Cyber-Physical Systems (ICCPS), Milano, Italy, 2022, pp. 306-307, doi: 10.1109/ICCPS54341.2022.00044.

[29] KTH RPL Planiacs. (2022). Foresee the unseen [Simulation results]. KTH Royal Institute of Technology. Repository for the simulation setup and results. https://github.com/KTH-RPL-Planiacs/foresee-the-unseen

[30] J. M. G. Sánchez, T. Nyberg, C. Pek, J. Tumova and M. Törngren, "Foresee the Unseen: Sequential Reasoning about Hidden Obstacles for Safe Driving," 2022 IEEE Intelligent Vehicles Symposium (IV), Aachen, Germany, 2022, pp. 255-264, doi: 10.1109/IV51971.2022.9827171.

[31] N. Jörgensen, A. Kattepur, S. Mohalik, A. Vulgarakis and E. Fersman, "Towards 5G-Aware Robot Planning for Industrial Applications," *2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA)*, Stuttgart, Germany, 2022, pp. 1-8, doi: 10.1109/ETFA52439.2022.9921449.

[32] N. Jörgensen, A. Kattepur, S. Mohalik, A. Vulgarakis and E. Fersman, "RoboPlan5G: Coordinating Cloud-Controlled Mobile Robots with 5G Network Configuration," submitted to: *2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA)*, Sinaia, Romania, 2023.

[33] Tan, K., Feng, L., Dán, G., & Törngren, M. (2022). Decentralized Convex Optimization for Joint Task Offloading and Resource Allocation of Vehicular Edge Computing Systems. IEEE Transactions on Vehicular Technology, 71(12), 13226-13241.

[34] Tan, K., Ji, Q., Feng, L., & Törngren, M. (2023). Edge-enabled Adaptive Shape Estimation of 3D Printed Soft Actuators with Gaussian Processes and Unscented Kalman Filters. IEEE Transactions on Industrial Electronics.

[35] T Markovic, M Leon, D Buffoni, S Punnekkat. (2022). Random Forest Based on Federated Learning for Intrusion Detection. In Artificial Intelligence Applications and Innovations: 18th IFIP WG 12.5 International Conference, AIAI 2022, Hersonissos, Crete, Greece, June 17–20, 2022, Proceedings, Part I (pp. 132-144). Cham: Springer International Publishing.

[36] University of California Irvine. (1999). KDD Cup 1999 Data Set [Dataset]. https://kdd.org/kdd-cup/view/kdd-cup-1999/Data, https://archive.ics.uci.edu/ml/datasets/kdd+cup+1999+data

[37] Canadian Institute for Cybersecurity, University of New Brunswick. (2009). NSL-KDD Dataset [Dataset]. https://www.unb.ca/cic/datasets/nsl.html

[38] Cyber Range Lab of UNSW Canberra. (2015). The UNSW-NB15 Dataset [Dataset]. https://research.unsw.edu.au/projects/unsw-nb15-dataset

[39] Canadian Institute for Cybersecurity, University of New Brunswick. (2017). Intrusion Detection Evaluation Dataset (CIC-IDS2017). https://www.unb.ca/cic/datasets/ids-2017.html

[40] T. Markovic, M. Leon, D. Bufoni and S Punnekkat (2023). Random Forest with differential privacy in a federated learning framework. Submitted to IEEE Transactions on Sensors Network.

[41] M. Leon, T. Markovic, S. Punnekkat, “Feature Encoding with Autoencoder and Differential Evolution for Network Intrusion Detection using Machine Learning”, In Proceedings of the Genetic and Evolutionary Computation Conference Companion (pp. 2152-2159).

[42] M. Leon, T. Markovic and S Punnekkat (2023) .Multi-Objective Optimization on Autoencoder for Feature Encoding and Attack Detection on Network Data. Accepted in GECCO

[43] NXP Semiconductors. (2023). eIQ® Toolkit for End-to-End Model Development and Deployment [ML Deployment Platform & Dataset]. https://www.nxp.com/design/software/development-software/eiq-ml-development-environment/eiq-toolkit-for-end-to-end-model-development-and-deployment:EIQ-TOOLKIT

[44] Papatheocharous, E., Buffoni, D., Maurer, M., Wallberg, A., and Ezquerro, G., Driver Distraction Detection Using Artificial Intelligence and Smart Devices, Springer Book Chapter, 2023.

[45] Kaggle, Inc. (2016). State Farm Distracted Driver Detection [Dataset]. https://www.kaggle.com/c/state-farm-distracted-driver-detection

[46] Amur, S.H., Zia, K., Chiumento, A., Havinga, P. (2023). Autonomous Network Slicing and Resource Management for Diverse QoS in IoT Networks. IEEE PerSASN 2023 (IEEE Percom Workshop).

[47] Zia, K., Chiumento, A., Havinga, P., Riggio, R., Huang, Y.(2023). QoS Aware Slice Resource Management using Deep Reinforcement Learning in IoT Networks. IEEE DCOSS 2023.

[48] Zia, K., Chiumento, A., Havinga, P. J. M. (2022). AI-Enabled Reliable QoS in Multi-RAT Wireless IoT Networks: Prospects, Challenges, and Future Directions. IEEE Open Journal of the Communications Society, vol. 3, pp. 1906-1929, 2022, doi: 10.1109/OJCOMS.2022.3215731

[49] Shullani, D., Fontani, M., Iuliani, M. et al. (2017). VISION: a video and image dataset for source identification. EURASIP J. on Info. Security 2017, 15 (2017). https://doi.org/10.1186/s13635-017-0067-2, https://jis-eurasipjournals.springeropen.com/articles/10.1186/s13635-017-0067-2, https://lesc.dinfo.unifi.it/VISION/

[50] Shullani, D., Fontani, M., Iuliani, M., Alshaya, O., Piva, A. (2017). *VISION: a video and image dataset for source identification.* EURASIP Journal on Information Security, 15. 10.1186/s13635-017-0067-2.

[51] Wisultschew, C., Hernández, R., Pastor, C. and Portilla J. (2022). Synthetic LiDAR Labeled Dataset Generation to Train Deep Neural Networks for Object Classification in IoT at the Edge. IEEE Internet of Things Journal, 2022, doi: 10.1109/JIOT.2022.3194716

[52] Wisultschew, C., Pérez, A., Otero, A., Mujica G., Portilla, J. (2022). Characterizing Deep Neural Networks on Edge Computing Systems for Object Classification in 3D Point Clouds. IEEE Sensors Journal, 2022, doi: 10.1109/JSEN.2022.3193060

[53] Wisultschew, C. (2022). Point Cloud-Based Object Detection and Classification at the Edge of the Internet of Things [PhD Thesis]. Universidad Politécnica de Madrid. https://oa.upm.es/71520/1/CRISTIAN_WISULTSCHEW_PUIGDELLIVOL.pdf

[54] Moertl, P., Ebinge, N. (2022). The Development of Ethical and Trustworthy AI Systems Requires Appropriate Human-Systems Integration: A White Paper. InSecTT Project. https://www.insectt.eu/wp-content/uploads/2022/11/Trustworthiness-Whitepaper-InSecTT-Format-v02-1-1.pdf

[55] Reiss, A., Indlekofer, I., Schmidt, P., & Van Laerhoven, K. (2019). Deep PPG: Large-scale Heart Rate Estimation with Convolutional Neural Networks. MDPI Sensors, 19(14). https://www.mdpi.com/1424-8220/19/14/3079, https://archive.ics.uci.edu/ml/datasets/PPG-DaLiA, https://doi.org/10.5281/zenodo.3902728

[56] Stricker, R., Müller S., & Gross, H. -M. (2014) Non-contact video-based pulse rate measurement on a mobile service robot. The 23rd IEEE International Symposium on Robot and Human Interactive Communication, Edinburgh, UK, pp. 1056-1062. https://doi.org/10.1109/ROMAN.2014.6926392

[57] Pilz, C.S., Zaunseder, S., Krajewski J., Blazek, V. (2018). Local Group Invariance for Heart Rate Estimation from Face Videos in the Wild. The IEEE Conference on Computer Vision and

Pattern Recognition (CVPR) Workshops, pp.1254-1262, Salt Lake City, 2018. https://github.com/partofthestars/LGI-PPGI-DB

A. ABBREVIATIONS AND DEFINITIONS

Term	Definition
5G	The 5th generation mobile network
AAMP	Additive Angular Margin Penalty
AC	Attack Classification
AD	Attack Detection
AE	Autoencoder
AI	Artificial Intelligence
AIoT	Artificial Intelligence of Things
AMQP	Advanced Message Queuing Protocol
ANN	Artificial Neural Network
API	Application Programming Interface
ATAM	Architecture Trade-off Analysis Method
BFA	Bit-Flip Attack
BiLSTM	Bidirectional LSTM
BLE	Bluetooth Low Energy
BPM	Beats per Minute (heart rate in vital signs)
CCM	Client Connection Manager
CNN	Convolutional Neural Network
CO	Carbon monoxide
CO_2	Carbon Dioxide
CoAP	Constrained Application Protocol
COTS	Commercial Off-the-Shelf
CPM	Collective Perception Message
CPS	Cyber-Physical Systems
DAC	Digital-to-Analog Converter
DDoS	Distributed Denial of Service
DDS	Data Distribution Service
DL	Deep Learning
DE	Differential Evolution (a type optimization algorithm in machine learning)
DNN	Deep Neural Networks
DoA	Direction-of-Arrival (in radio signal processing)
DP	Differential Privacy
DRL	Deep Reinforcement Learning
DTR	Decision Tree Regression
ECG	Electrocardiogram

Term	Definition
EM	Electro-magnetic
EN50155	A European standard covering electronic equipment used in rolling stock for railway applications (https://standards.iteh.ai/catalog/standards/clc/b34c4e1c-23a6-4c2c-91ee-85f2f8ee9e3a/en-50155-2021)
EER	Equal Error Rate
ESPAR	Electronically Steerable Passive Array Radiator (in radio signal processing)
FFT	Fast Fourier Transform
FHIR	Fast Healthcare Interoperability Resources (a standard for health care data exchange, published by HL7, https://hl7.org/fhir/)
FIR	Finite Impulse Response
FOSS	Free and open-source software
FPGA	Field-Programmable Gate Array
FPS	Frames Per Second
FRS	Facial Recognition System
FTP	File Transfer Protocol (IEFT standard RFC959)
GeoJSON	An open standard geospatial data interchange format that represents simple geographic features and their nonspatial attributes
GMM	Gaussian Mixture Model (an unsupervised ML algorithm)
GNSS	Global Navigation Satellite System
GP	Gaussian Process
GPS	Global Positioning System
GPU	Graphical Processing Unit
GP-UKF	Unscented Kalman filters with Gaussian process (http://dx.doi.org/10.1109/IROS.2007.4399284)
GRU	Gated Recurrent Unit
GUI	Graphical User Interface
HD	High-Definition
HDMI	High-Definition Multimedia Interface
HR	Heart Rate
HIS	Human-Systems Integration
HTTPS	Hypertext Transfer Protocol Secure
HTTPS REST	Representational state transfer (REST) over HTTPS
I^2C	Inter-Integrated Circuit
ICS-Flow	An anomaly detection dataset for intrusion detection in industrial control systems
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IMU	Inertial Motion Unit
IoT	Internet of Things

Term	Definition
IoU	Intersection over Union
IPM	Inverse Perspective Mapping
IQ	In-phase and Quadrature components in radio modulation techniques
ISM	Industrial, Scientific, and Medical radio band (a group of radio bands or parts of the radio spectrum which is internationally license free)
ITS	Intelligent Transport Systems
JSON	JavaScript Object Notation
k-NN, KNN	k-Nearest Neighbors
KPI	Key Performance Indicator
LDA	Linear Discriminant Analysis
LiDAR	Light Detection and Ranging
LTE	Long-Term Evolution (the fourth generation (4G) wireless standard for mobile cellular devices)
LTSM	Long Short-Term Memory (a subtype of RNN)
MABASR	Multi-Armed Bandit Adaptive Similarity-based Regressor
MAE	Mean Average Error
MAMS	Multi-Access Management Service
MAPE	Mean Absolute Percentage Error
MCU	Microcontroller Unit
MIG	Multi-Interface Gateway
MIL	Multiple Instance Learning
MINLP	Mixed-Integer Nonlinear Programming
ML	Machine Learning
MODE	Multi-Objective Differential Evolution (a type of machine learning algorithm)
MQTT	Message Queuing Telemetry Transport
MSE	Mean Squared Error
NIR	Near Infrared Radiation
Near-RT RIC	Near-Real-Time RAN Intelligent Controller
NOMA	Non-Orthogonal Multiple Access
Non-RT RIC	Non-Real-Time Radio Intelligent Controller
NSGA-II	Non-sorting Genetic Algorithm II (a type of machine learning algorithm)
OFDM	Orthogonal Frequency-Division Multiplexing (a type of digital transmission used in digital modulation for encoding digital data on multiple carrier frequencies)
ONAP	Open Network Automation Platform (https://docs.onap.org/en/latest/platform/overview/index.html)
OPC UA	Open Platform Communications Unified Architecture
QR	Quick Response (a two-dimensional barcode)
O-RAN	Open Radio Access Network (https://www.o-ran.org/)

Term	Definition
OS	Operating System (software)
OSS	Operational Support System
PC	Personal Computer
PCA	Principal Component Analysis
PM	Particulate Matter
PoC	Proof-of-Concept
QoE	Quality of Experience
QoS	Quality of Service
RadCom	Joint Radar and Communication
RAN	Radio Access Network
ReLU	Rectified Linear Unit (an activation function in ML algorithms)
REST	REpresentational State Transfer and an architectural style for distributed hypermedia systems
RFC	Request for Comments (a formal document from the IETF)
RF	Random Forest (a type of machine learning algorithm)
RF	Radio Frequency
RGB	Red-Green-Blue, a colour model based on additive colour primaries.
RIC	RAN Intelligent Controller (https://wiki.onap.org/display/DW/A1+Adapter+in+ONAP)
RNN	Recurrent Neural Network
ROS2	Robot Operating System generation 2
rPPG	Remote Photoplethysmography, a non-contact video-based method for heart rate monitoring
RSSI	Received Signal Strength Indicator
RSU	Restricted Stock Unit
RT	Real-Time
RTE	Run-Time Environment
RTAD	Real-Time Anomaly Detection
SBC	Single Board Computer
SDK	Software Development Kit
SDN	Software Defined Networking
SDR	Software Defined Radio
SIC-RD	Successive Interference Cancellation, Retransmission Diversity
SMO	Service Management and Orchestration
SPI	Serial Peripheral Interface
SQuaRE	Systems and software Quality Requirements and Evaluation
STD	Standard Deviation
SVM	Support Vector Machine (a type of deep learning algorithm)

Term	Definition
SVR	Support Vector Regression
t-SNE	t-distributed Stochastic Neighbour Embedding (a method for dimensionality reduction)
TCP	Transport Control protocol (IETF RFC 9293, https://datatracker.ietf.org/doc/rfc9293/)
TPU	Tensor Processing Unit (an application-specific integrated circuit (ASIC) used to accelerate machine learning workloads)
TRL	Technology Readiness Level
UAV	Unmanned Aerial Vehicle
UC	Use Case
UE	User Equipment
UKF	Unscented Kalman Filter
UWB	Ultra-wideband (a short-range, wireless communication protocol that operates through radio waves)
V2N	Vehicle-to-Network (in automated traffic and self-driving cars, connection via internet)
V2X	Vehicle-to-Everything (in automated traffic and self-driving cars)
vehicleCAPTAIN	Vehicle Communication Platform to Anything
V&V	Verification and Validation
VEC	Vehicular Edge Computing
VES	Virtual Event Streaming (https://docs.onap.org/projects/onap-dcaegen2/en/latest/sections/services/ves-http/index.html)
VPU	Vision Processing Unit (a specific type of AI/ML accelerator hardware designed to accelerate machine vision tasks)
WD	Weight Decay (a staple regularization technique for DNNs)
WSN	Wireless Sensor Network
WWW	World Wide Web
XAI	eXplainable AI
XMPP	Extensible Messaging and Presence Protocol
YOLO	You Only Look Once (a ML algorithm)