Deliverable D4.3

Analysis and Evaluation of system performance and scalability

Editor:	Vangelis Angelakis, LiU
Deliverable nature:	Report (R)
Dissemination level: (Confidentiality)	Public (PU)
Contractual delivery date:	29 Feb 2016 (extended to 30 Apr 2016)
Actual delivery date:	8 May 2016
Suggested readers:	Researchers, IERC, application developers, system administrators
Version:	1.0
Total number of pages:	174
Keywords:	

Abstract

This document presents the results of the work within Task 4.4 “Performance and Scalability Analysis”. Analytical investigations were conducted to characterize the performance of various mechanisms which can be used to securely interconnect the devices in potential deployments of the Use Cases of RERUM. Furthermore, algorithms, schemes and mechanisms developed here aim to increase the reliability, availability and scalability of a RERUM Device deployment and to improve system performance at large. Most of the research is applicable on the gateways, while a good part of the work presented can be applied to the RDs. Performance trade-offs and scalability have been addressed from a broad array of viewpoints, namely: i) the properties of sensors used in the RDs and the trade-offs that led to their selection. ii) trade-offs of (ii.a) security vs energy consumption under DTLS and CS, (ii.b) network performance vs energy and overheads for the BMFA and CADC. (ii.c) network performance vs. fairness. iii) cooperation of heterogeneous networks supporting RD deployments and their performance under a wide array of relevant scenarios and metrics. iv) scalability analyses for the system and for specific security and reliability aspects of RERUM such as the use of CS, Trust-based routing, message authentication and connectivity both in large scale WLANS and in underlay CR-based networks supporting the communication infrastructure. Results in relevant areas from the UC-O1 have already been presented within lab environment studies conducted in WP5, while some of the mechanisms theoretically evaluated here will be further evaluated in a lab environment and RERUM’s field trials as part of WP5.

Disclaimer

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 609094.

Impressum

Full project title	Reliable, resilient and secure IoT for smart city applications
Short project title	RERUM
Number and title of work-package	WP4 - Reliability, availability, robustness and scalability
Number and title of task	T4.4 - Network interconnectivity
Document title	Analysis and Evaluation of system performance and scalability
Editor: Name, company	Vangelis Angelakis, LiU
Work-package leader: Name, company	Elias Tragos, FORTH
Estimation of person months (PMs) spent on the Deliverable	

Executive summary

This deliverable presents analyses performed and techniques developed within the RERUM project for investigating the performance and scalability of a RERUM use-case deployment by different networking and security mechanisms deployed in RERUM Devices, gateways. We have analysed performance trade-offs pertaining to the sensors candidates focusing on scalability and measurement reliability. This deliverable (D4.3) is the output of the activities of Task 4.4 “Performance and scalability analysis” within Work Package 4 (WP4) “Reliability, availability, robustness and scalability”.

Some of the techniques discussed have been developed as part of previous RERUM Tasks (T3.2, T4.1, T4.2 and T4.3), whereas some of them are novel. RERUM’s requirements in terms of scalability and performance have been laid out in D2.2, and the content of that deliverable informed the progress of Task 4.3 documented herein. In brief the deliverable addresses:

- **Measurements and Sensor selection criteria** to perform a characterizations and measurements of the sensors used for the trial of the RERUM use cases.
- **Trade-offs between Security and energy consumption using DTLS** evaluating the computation (handshake) latency vs energy consumption
- **Trade-offs between network performance, signalling overheads and energy consumption with the bi-directional multicast forwarding algorithm** (BMFA) showing comparative results between the BMFA and its rival Trickle Multicast / Multicast Protocol for Low power
- **Trade-offs between network performance and energy consumption of the congestion-aware duty-cycling algorithm** focused on the duty-cycling and 6LoWPAN performance
- **Trade-offs in trustworthiness in sensing** to efficiently and effective detect attacks under constrained computational resources with relation to the RD network density and scale.
- **Trade-offs in compressive sensing encryption** against common attacks versus the energy consumption and the execution time required to run the CS encryption in an RD
- **Achieving Trade-off of fairness versus networking performance** taking a 5G standpoint and we resume to present under the same view
    - **Overhead reduction** of D-MIMO based techniques
    - **Timeliness of information for dense networks of RDs** through the AoI metric for urgent notifications in RERUM deployments
- **A Load Coupling Characterization** addressing the notion of fairness
- **Scalability of the leak-resilient message authentication codes** to quantify how well the secure communication is supported
- **Trusted Routing Scalability** to analyse our trust-based routing scheme for different scales of attack intensities and RD network sizes.
- **Adaptive CS scalability** analysis investigating the performance of the CS, in terms of the reconstruction error, when the data sparsity changes.
- **Networking Performance and scaling properties for**
    - **Cognitive 802.15.4 WPANs**
    - **802.11 WLANS,**
    - **classic 802.11 networks**
    - **HetNets** and
    - **underlay networks** supporting RDs have been conducted through analytical models and numerical studies and system – level simulations

The highly technical and deeply scientific nature of this deliverable may create difficulties to non-expert readers; thus, an introductory part is included at the beginning of each section describing briefly (i) the motivation for developing each technique, (ii) the relation with the RERUM Use Cases and the practical problem the technique tries to solve, while in each section we deliver also a brief discussion on lessons drawn for RERUM and relevant Smart-City IoT based project at large.

List of authors

Company	Author	Contribution
FORTH	Elias Tragos Alexandros Fragkiadakis Pavlos Charalampidis George Stamatakis Apostolos Traganitis	Compressive sensing encryption Adaptive CS scalability analysis Trust-based routing scalability analysis Performance of large scale cognitive IEEE 802.15.4 WPANs Performance of large scale cognitive IEEE 802.11 networks
UNIVBRIS	Georgios Z. Papadopoulos, George Oikonomou, Konstantinos Maraslis	Trade-offs for DTLS Trade-offs for the BMFA multicast forwarding algorithm Trustworthiness Problems in sensing
LiU	Evangelos ANGELAKIS Niklas DANIELSSON Ioannis AVGOULEAS Sesanka KATURI Ludvig KRATS	Load Coupling Characterization for Offloading Dense Networking for RD timeliness and scalability with underlay networks
Cyta	Athanasios Lioumpas	Fairness vs Throughput Trade-offs Cooperative HetNets for RDs and their overhead Scalability performance of HetNets
SSRL	George Moldovan	Scaling of the leak-resilient message authentication codes
Zolertia	Marc Fàbregas Bachs Antonio Jesús Liñán Colina Àitor Mejías Sotorra	Sensor and measurement performance analysis Sensors centric trade-offs

Table of Contents

List of Figures

List of Tables

Abbreviations

6LoWPAN	IPv6 over Low power Wireless Personal Area Networks
ACS	Adaptive Compressive Sensing
AC	Attack Coefficient
ADC	Analog to Digital converter
AES	Advanced Encryption Standard
AOI	Age of Information
API	Application Programming Interface
ARL	Average Run Length
AWGN	additive white Gaussian noise
BMFA	Bi-Directional Multicast Forwarding Algorithm
BWHT	Block Walsh-Hadamard Transform
COA	Ciphertext-only Attack
CADC	Congestion Aware Duty Cycle
CBR	Constant Bit-Rate
CCR	Channel Check Rate
CDF	Cumulative density function
CoAP	Constrained Application Protocol
CPA	Chosen Plaintext Attack
CPM	Change poing method
CR	Cognitive Radio
CRSO	Cognitive Radio Smart Object
CS	Compressive Sensing
CSMA	Carrier Sense Multiple Access
DAG	Directed Acyclic Graph
dB	Decibel
DCT	Discrite Cosine Transform
DIN	(DIN Rail) Deutsches Institut für Normung, metal rail of a sandard type
DODAG	Destination-Oriented Directed Acyclic Graph
DS	Dempster-Shafer
DSA	Dynamic Spectrum Access
DSRC	Dedicates Short Range Communications
DTLS	Datagram Transport Layer Security
DTMC	Discrete Time Marcov Chain
EC	European Comission
ECC	Elliptic Curve Cryptography
EEA	European Environment Agency
EM	Electro-Magnetic
ETX	Expected Transmission Count
FFR	Fractional Frequency Reuse
FPGA	Field Programmable Gate Array
HB	Home Box
HBHO	Hop-By-Hop Option
I2C	Inter-Integrated Circuit protocol communication
ICMPv6	Internet Control Message Protocol version 6
IEEE	Institute of Electrical and Electronics Engineers

IERC	Internet of Things European Research Cluster
IETF	Internet Engineering Task Force
IMSI	international mobile subscriber identity
INR	Interference to Noise Ratio
IoT	Internet of Things
IP	Integer Program
IP65	International Protection Marking
ISM	Industrial Scientific Medical
ITS	Intelligent Transportation Systems
kbps	Kilobits per second
KPA	Known Plaintext Attack
KS	Kolmogorov-Smirnov
LLN	Low-Power, Lossy Network
LLSEC	Link-Layer Security
LPN	Low Power Nodes
LTE	Long Term Evolution
M2M	Machine to Machine
mA	Milli Amperes
MAC	Medium Access Control
MAPE	Mean Absolute Percentage Error
MBR	Misbehaviour Rate
MCU	Micro-Controller Unit
MDP	Markov Decision Process
MILP	Mixed Integer Linear Program
MLD	Multicast Listener Discovery
MLME	MAC Sublayer Management Entity
MOP	Mode of Operation
MPL	Multicast Protocol for Low power and Lossy Networks
MPR	Multi-packet reception
MQTT	Message Queuing Telemetry Transport
MSPS	Mega Samples per Second
MTU	Maximum Transfer Unit
mV	milli-Volts
NAA	Network Aware Applications
NAPS	naming/addressing/profile server
NCO	Numerically Controlled Oscillator
ND	Neighbour Discovery
NRM	Network Resource Manager
OFDMA	Orthogonal Frequency Division Multiple Access
ONS	object name service
OS	Operative System
PGFL	Probability Generating Functional
PLCP	Physical Layer Convergence Procedure
PM	Particle Matter
PMR	Packet Modification Rate
PR	Primary Receiver

PT	Primary Transmitter
PP	Partitioning Problem
PPB	Parts per Billion
PPM	Parts per Million
PPP	Poison Point Process
PSK	Pre-Shared Key
PU	Primary User
QoE	Quality of Experience
QoS	Quality of Service
RD	RERUM Device
RD	RERUM Devices
RDC	Radio Duty Cycle
REST	Representational Transfer State
RFC	Request For Comments
RH	Relative Humidity
RL	Reinforcement Learning
RPL	Routing Protocol for Low power and Lossy Networks
RSSI	Received Signal Strength Indicator
RSU	Road Side Units
RTSA	Real-Time Spectrum Analyzer
SDR	Softare Defined Radio
SDU	Service Data Unit
SIA	Service to Interface Assignment
SIMD	Single Instruction, Multiple Data
SINR	Signal to Interference-plus-Noise Ratio
SMDP	Semi Markov Decision Process
SMRF	Stateless Multicast RPL Forwarding
SOAP	Simple Object Access protocol
SOS	Sensor Observations Service
SPI	Serial Peripheral Interface
SPS	Sensor Planning Service
SR	Secondary Receiver
SRC	Sample Rate Conversion
SRM	Structurally Random Matrix
SSDF	Spectrum Sensing Data Falsification
SSN	Semantic Sensor Networks
SSP	Subset-sum problem
ST	Secondary Transmitter
SXCS	SensomaX Companion Simulator
TCP	Transmission Control Protocol
TI	Texas Instruments
TLS	Transport Layer Security
TM	Trickle Multicast
TM / MPL	Trickle Multicast / Multicast Protocol for Low power and Lossy Networks
TNDRP	topology and network resource discovery protocol
UART	Universal Asynchronous Receiver Transmitter

UC	Use Case
UCB-E	Upper Confidence Bound Exploration
UDGM	Unit Disk Graph Medium
UDP	User Datagram Protocol
UE	User Equipment
URI	Uniform Request Identifier
USRP	Universal Software Radio Peripheral
V	Voltage
VBR	Variable Bit-Rate
VCAN	virtualized content-aware network
VDC	Voltage Direct Current
VOC	Volatile Organic Compounds
WSN	Wireless Sensor Network
WWRF	Wireless World Research Forum

1 Introduction

1.1 Scope

This document presents the results of the EU-FP7-SMARTCITIES-2013 project RERUM [RERUM] with regards to analytical investigations to characterize the performance trade-offs of various mechanisms which can be used to securely interconnect the devices in potential deployments of the Use Cases of RERUM. Furthermore, algorithms, schemes and mechanisms developed here aim to increase the reliability, availability and scalability of a RERUM Device deployment and to improve system performance. Most of the research is applicable on the gateways, while a good part of the work presented can be applied to the RDs. Specifically, this document is the output of Task 4.4 "Performance and scalability analysis", which ran from M13 of the project.

As discussed in D4.1 [RD4.1], up until now, the IoT world focused mostly on enabling device interconnectivity through the virtualisation of physical devices and objects and the centralized management of their virtual counter-parts. However, developing only IoT platform-side mechanisms without any focus on the devices themselves does not solve availability and reliability issues, because it does not solve efficiently the problems arising due to the resource-constrained nature of IoT devices and networks formed among them. Within RERUM, the devices have a very important role in the system architecture and the goal is to embed intelligence on them so as to improve overall system reliability and to increase overall device availability. In doing so, device resources can be delivered on-time whenever they are requested by the RERUM Middleware. D4.2 Presented techniques with very low program and data memory requirements, in order to ensure the low-power operation of RERUM Devices and increase their lifetime towards improving overall system availability. Our use-case scenarios call for several RDs and gateways to be deployed within the city running on batteries. Their energy consumption characteristics were there analysed.

In this report we address four broad topics: (1) Sensor performance, (2) Performance Trade-offs, (3) RD Networking, and (4) Scalability. To this end we report on work provide our analyses, methods and outcomes on investigations conducted for:

- **Measurements and Sensor selection criteria** to perform a characterizations and measurements of the sensors used for the trial of the RERUM use cases in terms of what can affect the scalability of the system, and reliability and performance in terms of data gathering.
- **Trade-offs between Security and energy consumption using DTLS** through a simulation evaluation of the computation (handshake) latency and energy consumption of a specific implementation of DTLS, applicable for our RERUM implementations.
- **Trade-offs between network performance, signalling overheads and energy consumption with the bi-directional multicast forwarding algorithm** (BMFA) investigated via experiments targeting not only showing comparative results between the BMFA and its rival Trickle Multicast / Multicast Protocol for Low power and Lossy Networks (TM / MPL), but also to identify their overall behaviour when they are subjected under different configurations.
- **Trade-offs between network performance and energy consumption of the congestion-aware duty-cycling algorithm** focused on the duty-cycling and 6LoWPAN performance, since increased channel sampling significantly increases the performance of the network in terms of goodput, delay and packet loss at a cost of higher energy consumption when the network is idle
- **Trade-offs in trustworthiness in sensing** investigating, through game-theoretic approaches, how lightweight methods can be to efficiently and effective detect attacks under constrained computational resources also with relation to the RD network density and scale.

- **Trade-offs in compressive sensing encryption** presenting results regarding the security strength against common attacks versus the energy consumption and the execution time required to run the CS encryption in an RD
- **Achieving Trade-off of fairness versus networking performance** investigating with a 5G standpoint how multiple access technologies can provide an optimum trade-off between network throughput and individual RD throughout to avoid connectivity loss due to severe interference, following a similar standpoint we resume to present
- **Overhead reduction** of D-MIMO based techniques on the overhead signalling within the network.
- **Timeliness of information for dense networks of RDs** is presented through the AoI metric and using a MAC protocol employing a method for buffer management appropriate for urgent notifications in RERUM deployments
- **A Load Coupling Characterization** study is also given to derive the optimal offloading of demands to a supporting network of WiFi AP-acting gateways or small cells, addressing the notion of fairness
- **Scalability of the leak-resilient message authentication codes** to quantify how well the secure communication supported by the low-powered devices can scale with increasing numbers of RDs and how it adapts to various mobility ratios within the underlying network of RDs
- **Trusted Routing Scalability** to analyse our trust-based routing scheme for different scales of attack intensities and RD network sizes.
- **Adaptive CS scalability** analysis investigating the performance of the CS, in terms of the reconstruction error, when the data sparsity changes.
- **Networking Performance and scaling properties for**
    - **Cognitive 802.15.4 WPANs**
    - **802.11 WLANS,**
    - **classic 802.11 networks**
    - **HetNets** and
    - **underlay networks** supporting RDs have been conducted through analytical models and numerical studies and system – level simulations

1.2 Intended audience

This deliverable presents analyses performed and techniques developed within the RERUM project for investigating the performance and scalability of a RERUM use-case deployment by different networking and security mechanisms deployed in RERUM Devices, gateways. The deliverable has a very narrow target audience: It aims mainly for researchers that are working in the areas of IoT Networking, network routing, IPv6 networking for constrained devices and the link layer of the network stack for wireless embedded devices. Solutions in the document are explained in detail so that the respective readers can with relative ease apply them and test them on their systems. The document also aims at other IoT related projects and the Internet of Things European Research Cluster (IERC) members to provide them with the RERUM solutions on improving the lifetime of their network deployments. In this respect, a dialogue with other projects for integrating the RERUM device-oriented solutions with the middleware-oriented solutions of other projects can start, in order to develop jointly an optimised IoT framework with emphasis (but not exclusively focused) on smart city applications.

1.3 Position within the project

1.3.1 Relation with other tasks and WPs

This deliverable (D4.3) is the output of the activities of Task 4.4 within Work Package 4 (WP4).

D4.3 used as input the results of the following tasks of other WPs, as shown also in Figure 1:

- all of WP2 deliverables D2.2-D2.3, D2.5;
- all of WP3 deliverables D3.1, D3.2 ,D3.3;
- WP4, Tasks 4.1, 4.2, 4.3 and deliverables D4.1, D4.2.

Input from D2.2 relates to system requirements on scalability and performance. Those requirements had been used as basis for the design and development of the respective security, privacy, tryst and networking technologies within the Tasks in WP3 and WP4. The trade-offs on performance characteristics of some of those technologies are investigated in this task, gathering input from the respective deliverables of WP3 and WP4, where those technologies are described. From D2.3 and D2.5, the input to this deliverable is related to the specific modules of the system architecture that have to be investigated for their scalability and performance, as well as to the deployment model of the architecture, the scalability of which has to be investigated too.

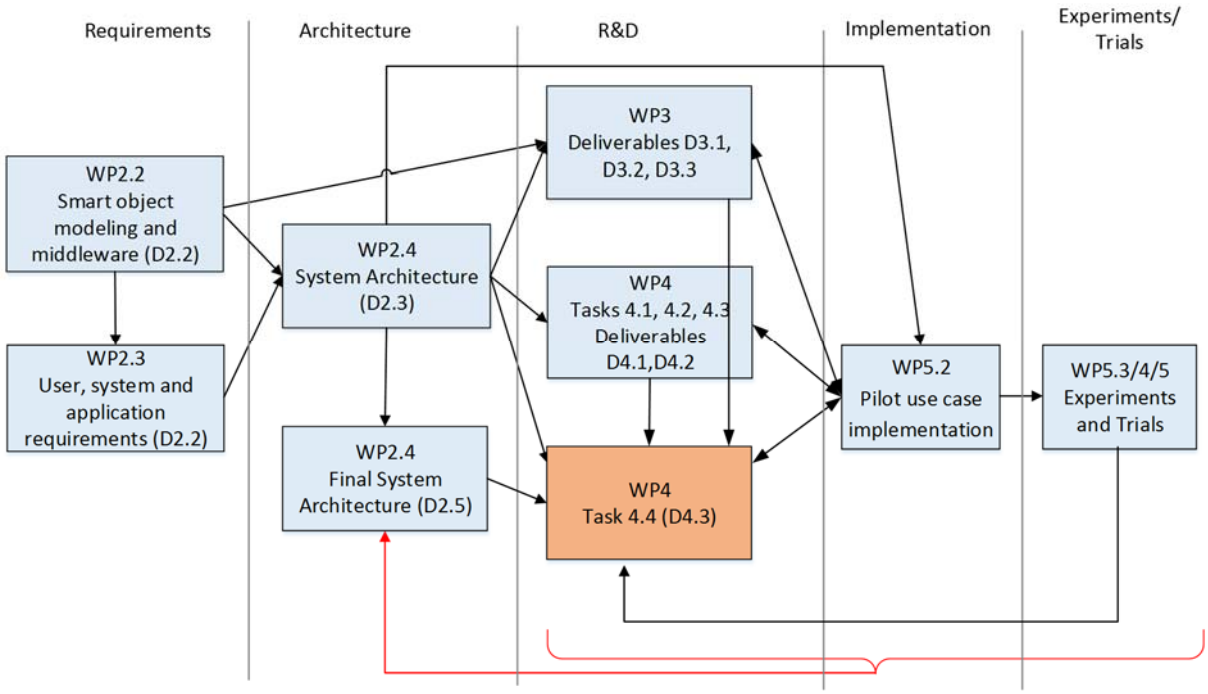


Figure 1: Position of D4.3 within the project

The outputs of D4.3 (as they were gathered during its execution period) provided output to Task 5.2 (deliverables D5.2 and D5.5) for the scalable implementation of the mechanisms for running the use cases within experiments and trials in Tasks 5.3, 5.4 and 5.5. The early results from the experiments and trials were used for refining and optimizing the developed mechanisms within WP3 and WP4 that would provide results for the optimization of the implemented mechanisms within WP5 (as it is depicted in the feedback loop shown with the red lines in the figure).

1.3.2 Relation with the use cases

Overall, results from this deliverable can be used for the implementation of the use cases. The efficient networking of the devices is of key importance for the performance of any sensing device in order to increase the scale of the resulting systems, lifetime and reliability of the overall system, in terms of correct measurements arriving timely at the middleware. All solutions presented in this deliverable can be applied to all cases. In each section of the document we have provided detailed relation to the Use Cases.

1.4 Structure of the document

This deliverable continues as follows:

- Section 2 presents a performance evaluation for the sensors utilized in RERUM, other than those of UC-O1 which have already been discussed in D5.3.
- Subsequently, in Section 3, we discuss trade-offs of security mechanisms and network performance
- Section 4 focuses on the pure networking aspects of the RD deployments focusing on performance characterizations of mechanisms that can be employed in RERUM
- Section 5 discusses scalability of networking and security mechanisms
- Section 6 concludes the deliverable with an overview of the key results.

2 Performance evaluation for sensors and RD networking

The scope of this section is to perform a characterizations of the sensors that are being used in the use cases in terms of what can affect the scalability of the system, mainly regarding the amount of data each sensor should gather to have a reliable measurement, susceptible of becoming useful information.

2.1 Measurement types

The sensor's measurements must be always understood from the use cases point of view. Therefore, this performance analysis will start with a review, UC per UC, of the measurements, putting emphasis on the purpose of each one and the quality we should expect for that purpose. Use case O1, the Smart Transportation, will be skipped because does not include any sensor, only position data.

2.1.1 Use Case-O2 – Environmental Monitoring

The purpose of this use case is to perform a set of outdoor measurements related to the quality of the environment. Different kind of sensors and different devices will be spread-out on the city to do so. The main parameters that are being measured are:

- Meteorological: to control the weather using temperature and relative humidity sensors, atmospheric pressure meter, and an external support for a rain gauge, a wind speed meter, or anemometer, and a wind direction indicator.
- Air quality: to monitor the quality of the breathed air through the concentration of gases such as SO_2, NO_X, O_3, the pollution in terms of particle matter in suspension smaller than 10µm, known as PM_{10}, and the presence of strong smells or other volatile organic compounds, using a VOC sensor.
- Noise: the measure and control that the noise level is under the comfort standards.

The purpose in all cases is to detect anomalous situations that might annoy the citizens; therefore the quality required for the measurement just need to detect those situations, no asses the environmental situation.

These sensors are installed in outdoor proof enclosures, granting IP65 protection to the core of the RD device, the RE-Mote, while the sensor parts need to be exposed to the environment, sometimes placed in a separated, non-IP box, witch is attached to the main enclosure.

Indeed, some of the weather sensors, namely the wind and rain , are placed externally with a small mast support, connected via communication cables and water-proof wall-through holes on the main box, reducing the cost.

Noise level meter, performed with a microphone, is placed outside the core box but in a protected area, to reduce the effect of the rain in the measurement, despite, according to the standards, is not required to measure while raining.

Outdoor IP65 gateways located nearby the sensors complete the use case solution.

2.1.2 Use Case-I1 – Energy management

This use case include some indoor devices with different configurations, all them measuring parameters related to the energy consumption in buildings. In general the use case intends to measure, and detect situations of air-conditioning and heating system misuse, that can convert, for instance, into an overheating situation, excess of the relative humidity in the air, etc., and/or excess (or lack,

meaning some fault) of energy consumption, either from the mains or the lighting system. Some environmental sensors has been added to the RD, mixing this use case with the I2 one.

In general, the detection does not require a very precise sensor's system but the measurement could require it, depending on the use that city will do with the data; in the context of the RERUM project, however, the cities would not require a precision system because the data will be used for indicative measures.

The Table 1 shows the different types of RD created for this use case.

Table 1: Setup types for use case I1.

UC-I1 setup configurations	
Type 1	Indoor RD with tri-phasic current and voltage sensors
Type 2	Light sensor, temperature/humidity sensor and PM_{10} meter
Type 3	Current sensor, light sensor, temperature/humidity sensor, and PM_{10}
Type 4	Light, temperature and humidity sensors

2.1.3 Use Case-I2 – Comfort quality

The objective of this use case is to provide a set of sensors to holistically determine the comfort quality index in indoor spaces.

Diverse measurement should used to determine this multi-parametric index and control the indoor environment according to improve it; for these reason, multiple combinations of sensors are used for this use case.

It also includes few outdoor sensors, with meteorological sensors and air quality, to correlate the data measured indoor with the actual situation outdoor, on the same building; in other words and as a concrete example, the temperature indoor could not be detached from the outdoor one, and a certain low value indoor could be normal and understood as related to a low outdoor one. Table 2 shows the different device types for this use case.

Table 2: Setup types for use case I2.

UC-I2 setup configurations	
Indoor	
Type 1	Temperature and humidity sensor
Type 2	Temperature and humidity sensor, relay actuator
Outdoor	
Type 3	Temperature and humidity sensor, noise meter
Type 4	Temperature and humidity sensor, noise meter, gas and air quality sensors (SO_2+NO_X+O_3+VOC+PM_{10})
Type 5	Weather station

2.2 Commercial sensors used in UCs for each type of measurement

In this sub-section the reader can find a detailed description of the exact commercial sensors that has been chosen for each measurement, describing also its main characteristics and the selection criteria we have used for each one. The goal is to set the basis for a performance analysis of them.

Table 3: Sensor's Specifications Summary

Measurement type	Part number	Manufacturer/ Provider	Input voltage	Maximum range device	Relative error accuracy
Current	i-Snail-VC-50	ELKOR	N/A	0-50A	N/A
Voltage	CE-VJ03-32MS2	IRELEC	12V	0-250V	2%
Light	TLS2563	TAOS	3V	0-40000lux	N/A
Temperature/humidity	SHT25	SENSIRION	3V	-40 to 123,8°C 0 to 100 RH	±0,2°C 1,8%
Atmospheric pressure	BMP085	BOSCH	3V	300-1100hPa	±0,5hPa
Noise	ZNK14XXX	ZOLERTIA	3V	34-95dB$_A$	±3dB$_A$
Rain gauge	80422 assembly pack	ARGENT	3V	0,28mm/tick	N/A
Anemometer			3V	2,4km/h/tick	N/A
Wind direction			3V	0 to 360°	N/A
Particle matter, PM_{10}	GP2Y1010AU	SHARP	5V	0-900 ppm	N/A
Gases - SO_2	4-SO2-20	Honeywell (1)	3V	0 to 150 ppm	±0,1 ppm (2)
Gases - NO	4-NO-250	Honeywell (1)	3V	0 to 1000 ppm	±0,5 ppm (2)
Gases - NO_2	4-NO2-20	Honeywell (1)	3V	0 to 250 ppm	±0,1 ppm (2)
Gases - O_3	OX-A421	Alphasense (1)	3V	0 to 50 ppm	±0,015ppm(2)
Volatile Organic Compounds (VOC)	IAQ-CORE	AMS	3V	CO_2: 450-60000 ppm TVOC: 125-600 ppb	N/A
(1) Provided and calibrated by Libelium™, (2) Ideal conditions					

2.2.1 Selection Criteria

When choosing a particular sensor for a given application, there are many factors to be considered. The main factors are economical and related to the sensor's performance. In most cases, we have to balance these factors and try to select the best trade-off for each application.

2.2.1.1 Current Sensor

The i-Snail-VC-50 current sensor was chosen for many reasons. One critical point was the factory calibrated feature, that guarantee the better performance on the acquisition.

The 50 Amperes range is another important parameter, selected following to the range specified by the municipalities according to the power consumptions to measure. The analog output range, in this case from 0 to 5 Volts, is also good to have to make easier the implementation, the interface with the RE-Mote device: only a resistor divider is needed to convert the signal to the 3 Volts needed.

2.2.1.2 Voltage Sensor

CE-VJ03-32MS2 is an industrial isolated voltage transducer (2500VDC electromagnetic isolation) only mounted in UC-I1, together with current sensor to calculate the power consumption. It has been selected following the request to implement a high current/voltage measurement from the power lines, where this sensor exactly accommodates and is designed for. The power supply input has a range from 4 to 24V, ideal to be powered with the RE-Mote's 5V and the DIN rail offers much better deployment, especially in industrial installation. Finally, it offers a high voltage precision from 0 at 250V monitoring, with 2% accuracy.

2.2.1.3 Pressure Sensor

BMP085 is a low-power calibrated I^2C digital pressure sensor that can extend the range of measurement from 500 meters under the sea to 9000 meters, from 300 to 1100hPa, with a few power consumption of 0,1µA in stand by. It is also a relatively cheap sensor, which is acceptable for the project's specifications.

2.2.1.4 Light Sensor

TLS2563 is an I^2C light to digital converter that has as a main feature a system based in two photodiodes to get the visible light approximately as the human eye do. Additionally, the I^2C interface also makes simple the connection of this sensor with the RE-Mote.

2.2.1.5 Temperature/humidity Sensor

The digital SHT25 sensor has been chosen because it has many features, among others, the fact that:

- it is a factory calibrated sensor with 1,8% of relative humidity and 0,2°C error,
- it has an I^2C compatible communication,
- it is low-power,
- and it gives 2 parameters, temperature and relative humidity, in a single small package.

2.2.1.6 Noise Sensor

The noise sensor used in the trials is a Zolertia's design, called Zonik, that implements a calibrated and cheap noise meter designed for environmental measurements. The full-scale range of 34dB min and 95dB max is acceptable to measure standard noises, also according to the European normative. And the cost of the sensor board is really cheap compared to other options with demonstrated performance on outdoor measurements. The communication is also I^2C, and the sensor automatically calculates the average sound level, known as equivalent level, with a standard 1/8 seconds average window.

2.2.1.7 Weather Station

The inexpensive and well-known weather sensor kit assembly 80422 offers a rain gauge, wind vane and anemometer in a single pack and ready to install. The most important feature to select this part is its simple operation without using active components. The principle of the sensor use sealed magnetic reeds to make measurements. By construction the wind vane and rain gauge provides a pulse train that is very simple to be acquired by the RE-Mote board. The anemometer is simply measured as an analogic input.

2.2.1.8 PM_{10}

GP2Y1010AU is a cheap and low-power optical dust sensor from Sharp, easy to be implemented with a simple ADC input. We can get the current particles in ppm in the air with a fine precision versus cost of the sensor. This relationship offers one of the best selection criteria to choose it.

2.2.1.9 Gas sensors

There are many sensors included in this section. SO_2, NO, NO_2, O_3 and VOC sensor is used to acquire the air compounds concentration, related to the air quality.

The selected set of gas sensors is a packed solution from a single provider. The fact of having a certified calibration as well as the good error rating were the main criteria to select this.

The iAQ-Core VOC sensor was selected to provide a TVOC and/or CO_2 data at a low cost and acceptable precision: this sensor has an internal auto-calibration method with temperature adaption.

2.3 Sensors specification study

In this section the datasheets of each sensor or component have been studied in depth in terms of its performance. Some parameters such as accuracy, resolution, range, stability, lifetime and price from are listed and shown here as a basis for the sensor's trade-off rationale that would be found in the section 3.

The study of the features, and the selection of sensors as well, has been driven by the use case requirements, not by the sensor's characteristics itself; in other words, the sensors has been selected to fit on the trials requirements in spite of not being the bests of each class.

2.3.1 Accuracy

Sensor accuracy is known as the capacity of a measuring sensor to give results close to the true value of the measured quantity.

Usually datasheets from all sensors show relative error accuracy and is unusual to find the absolute error accuracy. Testing accuracy is not possible unless it compares with another calibrated and homologated device. Therefore, measuring the accuracy value is impossible in many cases, as we only know the error accuracy by datasheets.

The following table compares the sensors accuracy values chosen for RERUM:

Table 4: Sensor's Accuracy Error

Measurement type	Sensor	Absolute error accuracy	Relative error accuracy
Current	i-Snail-VC-50	N/A	0,5% of full scale 0,1% ripple
Voltage	CE-VJ03-32MS2	N/A	2%
Light	TLS2563	N/A	N/A
Temperature/humidity	SHT25	±1,5°C ±5% RH	±0,2°C 1,8% RH
Pressure	BMP085	±4hPA	±0,5hPa
Noise	ZNK14XXX	5dB@100Hz	±3 dB_A
Weather	80422 assembly pack	N/A	N/A
PM_{10}	GP2Y1010AU	N/A	N/A
SO_2	4-SO2-20	N/A	±0,1 ppm
NO	4-NO-250	N/A	±0,5 ppm
NO_2	4-NO2-20	N/A	±0,1 ppm
O_3	OX-A421	N/A	±0,015ppm
VOC	iAQ-CORE	N/A	N/A

Absolute error is the amount of physical error in a measurement, normally is an absolute error.

$$Absolute\ error = \frac{Result}{True\ Value}$$

Relative error gives an indication of how good a measurement is relative to the size of the thing being measured.

$$Relative\ error = \frac{Absolute\ error}{True\ value}$$

Environmental effects can increment the error accuracy on sensors that could not be calibrated.

2.3.2 Resolution

Resolution implies the minimum unit that the sensor can measure. It depends on the construction and the acquisition method.

Table 5: Sensor's Resolution

Measurement type	Sensor	Resolution
Current	i-Snail-VC-50	12,2mA
Voltage	CE-VJ03-32MS2	61mV
Light	TLS2563	1 Lux
Temperature/humidity	SHT25	±0,1 ºC ±0,1% RH
Pressure	BMP085	0,03hPA
Noise	ZNK14XXX	0,12 dB_A (average)
Weather	Rain Gauge	0.2794mm/pulse
	Anemometer	2,4Km/h/pulse
	Wind Vane	22,5°
PM_{10}	GP2Y1010AU	0,5V/(0,1mg/m3)
SO_2	4-SO2-20	0,014ppm
NO	4-NO-250	0,013ppm
NO_2	4-NO2-20	0,013ppm
O_3	OX-A421	0,015ppm
VOC	iAQ-CORE	CO_2: 1 ppm TVOC: 1 ppb

Analog sensors resolution, such as current, wind vane or PM_{10} is restricted by the ADC resolution. This means that each resolution is defined by the minimum quantization value per bit conversion.

2.3.3 Range

The range in terms of a sensor is the maximum and minimum value of a physical units that a sensor can measure with a certain performance. When a sensor works outside this bounds it could mean 3 things:

1. It needs a calibration to keep performing a good measurement.
2. It can not measure or keep its response at its minimum or maximum value.
3. It's damaged permanently.

In general, values outside from a manufacturer's recommended range should be considered as wrong measurements.

The following table compares the range values of sensors, taken from datasheets:

Table 6: Sensor standard and recommended range

Measurement type	Measurement range	Recommended range	Comments
Current	0 to 50A	N/A	Depend on the device to measure
Voltage	0 to 250V	230V	Standard input voltage of the network
Light	0 to 40000 lux	0 to 500lux	Level standard in light home, or an office.
Temperature/humidity	-40 to 123,8°C 0 to 100% RH	-10 to 50°C 0 to 80% RH	See references
Pressure	300 to 1100hPa	> 600hPA	600hPA = 4000 meters
Noise	34 to 95dB_A	< 85dB_A	Referenced in scale with a bus at 3m
Rain Gauge	0 to 10mm	O to 8 mm	10mm@35ticks/s
Anemometer	0 to 96km/h	0 to 50km/h	96km/h@40ticks/s
Wind Vane	0 to 360°	N/A	We need to detect all directions.
PM_{10}	0 to 900ppm	0 to 400ppm	
SO_2	0 to 150ppm	0 to 20ppm	As Datasheet linear measurement
NO	0 to 1000ppm	0 to 250ppm	As Datasheet linear measurement
NO_2	0 to 250ppm	0 to 20 ppm	As Datasheet linear measurement
O_3	0 to 50ppm	0 to 20ppm	As Datasheet linear measurement
VOC: CO_2	<450 to 60000ppm	<450 to 2000ppm	Normal ventilation applications limits are “2000” as value. Inferior limit of 450 can be measured.
VOC: TVOC	125 to 600 ppb	N/A	

Range recommendations are based on the study case of each parameter given by meteorological data. It is important to note that these limits reflect the real measurements in environmental air, and is the real case study. Although all sensors can measure higher values, it’s not useful to calculate the limits of the sensor.

2.3.3.1 Indoor Light Levels illuminance

The lux is a common use level unit that measures the light level. To choose our limit we can used the next standard table [SSS01]:

Table 7: Illuminance table reference

Condition	Lux level
Full Daylight	10752
Overcast Day	1075
Very Dark Day	107
Twilight	10
Deep Twilight	1
Full Moon	0,108

As the light sensor is assembled inside an enclosure, we can determine that the maximum limit of light that we can give is approximately half full daylight. To use all the available range, the sensor should be placed with a direct sight of the light source, which is complicated in most of the cases.

2.3.3.2 Average temperatures

To study the last average temperatures, we get the average natural temperatures in few regions around of Europe. [SSS01]

2.3.3.3 Barometric measures

To set the recommendation limit in pressure, we can get the maximum mountain level in Europe, Mont Blanc in the Alps, at 4810 meters. We consider that over 4000m no weather station will be placed.

2.3.3.4 Anemometer limits

To define local wind speed limits, we have relied on climatological results [SSS01], getting the maximum wind speed rating. Figure 2 shows an example of wind speed map for northern Europe. Finally the decision was to establish a limit of 50km/h, which will fit in almost all possible scenarios.

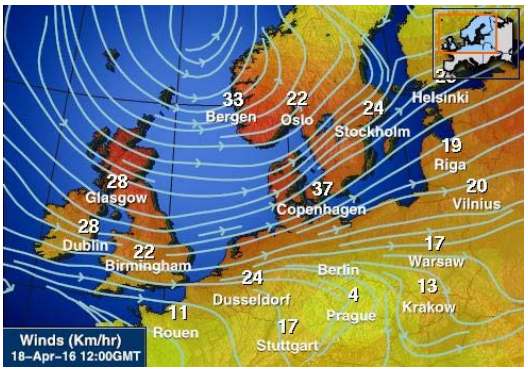


Figure 2: Wind Speed in North Western Europe [SSS01]

2.3.3.5 Noise limits

The average noise limits as standard in street is approximately between 40-90dB$_A$. A typical value during the night time is 45dB$_A$, and a urban noise of 85dB$_A$ is common during the day. Figure 3 shows a standard city noise pollution [SSS02]. However, the used noise sensor expanded 5dB$_A$ the upper and lower bounds.

Figure 3: Noise Map from Dublin City Council. [SSS02]

2.3.4 Stability

There are two different sources of instability in a sensor. The first one, and the most common, is due to the internal physical limitations that any sensor has by its nature. The other one due to the environment and external factors.

Any sensor has its own stability, which is know from the very first moment, when a sensor is installed; knowing this, the measurement can be calibrated, improving a lot of the precision within a wide range.

External variables are the most common source of deviations on the measurements. Many external factors can cause instability in sensors, most of them are environmental; the most common ones are:

- Temperature changes
- Humidity effects
- Ruggedness
- Corrosion
- Susceptibility to EM interferences

In order to give good stability in the measurement, is important to take into account the sensor's manufacturer's specifications. Factory calibrated sensors provides specific data to counteract most of the mentioned external factors without needing to add calibration processes into the firmware: it's just need to add calibration tables and take them into account on every measurement. Our calibrated sensors are listed in Table 8.

Table 8: Sensor certifications and calibration.

Measurement type	Part number	Certification/Calibration
Temperature/humidity	SHT25	SHT25 were tested according to AEC-Q100 Rev. G. Sensors specifications are tested to prevail under the AEC-Q100 temperature grade 2 test conditions
Pressure	BMP085	Calibration in factory by internal EEPROM
Noise	ZNK14XXX	Calibration results in section 0 of this document
SO_2	4-SO2-20	Asked to provider for this
NO	4-NO-250	**Figure 4:** NO Gas sensors calibration certification
NO_2	4-NO2-20	Asked to provider for this
O_3	OX-A421	**Figure 5:** O3 Gas sensors calibration certification
VOC	iAQ-CORE	Calibrated with an internal self-maintain heater calibration

Calibration Protocol

RAE SYSTEMS by Honeywell

Customer: Solidsense Gmbh

RAE Model No.: CLE-0522-400 **Sales Order No.** SSB-150703-01

Gas: NO **Date Shipped:** 7/20/2015

Quantity Shipped: 20EA **Tester:** OQC1

Span gas conc. [ppm]: 35 **Baseline range [μA]:** -0.2~1.5

Sensitivity range [μA /ppm]: 0.32~0.48 **T90 [sec]:** ≤45

Exposing procedure No.: Exposed to the air 1min – Through span gas 4min

Comment:

Serial No.	Baseline μA	Sensitivity μA /ppm	T90 sec	—	Status
CLEW00001-S7	0.33	0.40	14		PASS
CLEW00002-S7	0.34	0.40	12		PASS
CLEW00003-S7	0.29	0.37	14		PASS
CLEW00004-S7	0.31	0.37	17		PASS
CLEW00005-S7	0.31	0.41	12		PASS
CLEW00006-S7	0.37	0.44	13		PASS
CLEW00007-S7	0.33	0.37	16		PASS
CLEW00008-S7	0.29	0.39	14		PASS
CLEW00009-S7	0.31	0.42	13		PASS
CLEW00010-S7	0.32	0.40	13		PASS
CLEW00011-S7	0.37	0.36	24		PASS
CLEW00012-S7	0.37	0.42	17		PASS
CLEW00013-S7	0.35	0.40	14		PASS
CLEW00014-S7	0.38	0.41	15		PASS
CLEW00015-S7	0.27	0.42	10		PASS
CLEW00016-S7	0.28	0.35	17		PASS
CLEW00017-S7	0.27	0.35	17		PASS
CLEW00018-S7	0.37	0.36	20		PASS
CLEW00019-S7	0.30	0.41	12		PASS
CLEW00020-S7	0.27	0.41	12		PASS

OQC PASS

Figure 4: NO Gas sensors calibration certification

Performance Data
OX-A421
diciembre 9, 2015

Libelium

Serial No	Zero Current (nA)	Aux Zero Current (nA)	O3 Sensitivity (nA/ppm)	NO2 Sensitivity (nA/ppm)	Response t90 (s)
Min	-8,5	2,8	-493,0	-420,7	11,8
Mean	18,3	9,7	-435,9	-346,4	19,1
Max	47,3	17,3	-343,8	-200,5	28,7
(+/-) 95% Conf.	21,2	5,8	65,1	93,7	6,5
213390055	7,57	14,19	-388,85	-259,31	13,33
213390060	-6,94	9,46	-421,67	-282,96	16,70
213390061	-8,51	12,93	-383,36	-253,64	21,82
213390063	5,67	7,25	-377,63	-229,99	16,42
213390064	-5,36	11,35	-351,94	-200,51	15,09
213390101	34,05	12,30	-456,84	-347,27	17,35
213390103	5,67	11,35	-493,02	-413,48	18,21
213390106	19,23	11,67	-485,82	-417,42	19,02
213390108	14,82	8,51	-392,19	-307,86	16,20
213390109	12,61	17,34	-444,70	-343,49	16,36
213390110	23,96	7,88	-460,74	-355,63	19,95
213390114	20,49	5,04	-463,20	-370,29	15,87
213390115	28,69	15,76	-448,51	-391,25	16,56
213390117	19,86	8,51	-479,09	-411,27	23,05
213390119	1,26	8,20	-475,81	-393,78	27,06
213390120	30,58	10,09	-391,36	-277,76	19,20
213390121	27,11	15,13	-442,99	-340,02	16,94
213390122	12,61	13,56	-430,43	-332,30	11,75
213390124	14,82	10,40	-436,50	-362,72	18,18
213390125	28,37	7,25	-479,09	-382,74	24,17
213390126	5,04	9,14	-430,30	-369,03	17,41
213390127	35,63	8,83	-475,81	-361,82	16,84
213390128	26,80	4,10	-391,36	-336,40	18,91
213390129	47,29	11,67	-442,99	-292,26	22,75
213390131	20,49	2,84	-343,82	-420,73	18,01
213390135	27,11	4,10	-471,55	-390,47	19,72
213390136	21,44	11,35	-443,86	-397,40	28,67
213390137	25,22	5,36	-392,16	-392,52	24,78
213390138	25,22	11,03	-444,50	-368,71	19,72
213390139	18,60	10,09	-467,83	-414,27	23,22
213390140	13,24	7,88	-419,44	-335,45	15,71
213390141	35,94	8,20	-458,04	-348,06	25,76
213390142	10,72	4,10	-451,01	-290,37	15,44
213390143	26,17	13,24	-471,34	-376,91	21,15
213390144	15,76	10,09	-447,31	-358,42	17,58

Figure 5: O3 Gas sensors calibration certification

2.3.4.1 Stability in measurements in open environments

All the sensors used in the trials have, more or less, a long measuring range and an acceptable accuracy. But when this measurements are made in an open, non-controlled environment, this theoretical stability can be affected.

A couple of examples of 2 extreme cases could help to clarify how the environmental conditions can affect to the stability: a temperature measure and noise monitoring.

In normal conditions, between day and night the temperature can change at a maximum of 30°C in 24 hours. Therefore, the changes are always smooth and slow, and it's easy to obtain a stable measurement without needing so many samples, filters and averages.

On the other side, the decibels of noise can change a lot in a very short time, being this measurement affected by multiple, non-related sources, like traffic, industries, construction works, people, among other factors. Therefore, to get a stable measurement of the noise level will be needed a lot of acquisitions, together with filters and processing.

In terms of scalability, this factor has to be taken into account because some measurements may need too much processing to be performed with an acceptable stability and might be not scalable due to energy consumption, data bandwidth or cost. That's why this analysis will take also into account the natural frequencies of each parameter, in section 2.5, in spite that is not feasible to study also the effect of the environment, that could also affect as mentioned here.

2.3.5 Lifetime

When a sensor is produced, it has a lifetime during the measurement keeps reliable. Once the lifetime of a sensor is over, the sensor will decrease the accuracy and the measurement will not be reliable.

Some manufacturers provide this parameter to define the lifetime of the integrated product (in this case, the RD) or the maintenance period.

So, in general, the sensor's lifetime affects mainly to the actual device cost, because the replacement or maintenance costs should be considered accordingly. Its effect on the scalability is clear: a cheap sensor that needs to be replaced or maintained in the short term becomes a non-scalable solution at medium-long term; and vice versa.

In any case, should be take into account that when a sensor reaches the limit of its lifetime, must be changed and we need to either to maintain the RD or replace it. And long lifetime means, in most cases, less maintenance cost. In RERUM, this is the lifetime of the devices that we can get:

Table 9: Lifetime Sensor Table

Sensor type	Part number	Limit	Recommended
Current	i-Snail-VC-50	N/A	2 years
Voltage	CE-VJ03-32MS2	N/A	2 years
Light	TLS2563	N/A	2 years
Temp. & humidity	SHT25	N/A	1 year
Barometric pressure	BMP085	Undefined, but the value can change ±4hPa/12months	10 years(1)
Noise	ZNK14XXX	>2 years	1 year(2)
Weather	Rain Gauge	N/A	5 years(1)
	Anemometer	N/A	5 years(1)
	Wind Vane	N/A	5 years(1)
PM_{10}	GP2Y1010AU	N/A	1 year(2)
SO_2	4-SO2-20	>2 years	
NO	4-NO-250	>2 years	
NO_2	4-NO2-20	>2 years	
O_3	OX-A421	>2 years	
VOC	iAQ-CORE	>10 years	
(1) Based on user's experiences collected through multiple technical forums in Internet. (2) Based on Zolertia's experience using these sensors			

2.3.6 Price

The price is one of the most important parameters to decide which sensor to choose. It is difficult to find a cheap sensor that meets all requirements, especially when considering reliable and/or precise measurements.

Therefore, all sensors have been chosen looking for the best trade-off possible between cost and functionality, something that will be discussed in detail and in terms of scalability in the section 3.

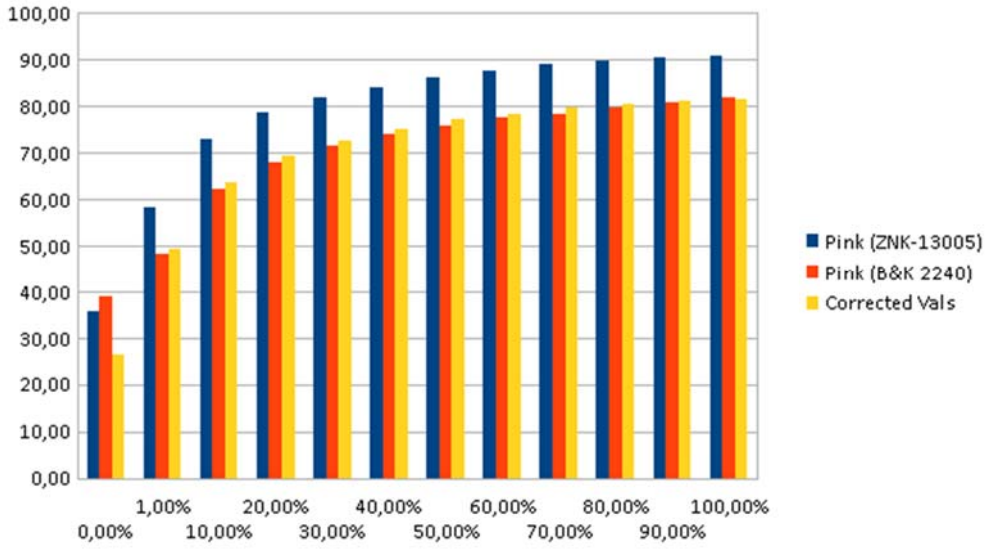
Table 10 shows the prices for the sensors used in RERUM, also including, whenever this information is available, large quantities price to consider potential large deployments.

Table 10: Sensor Prices

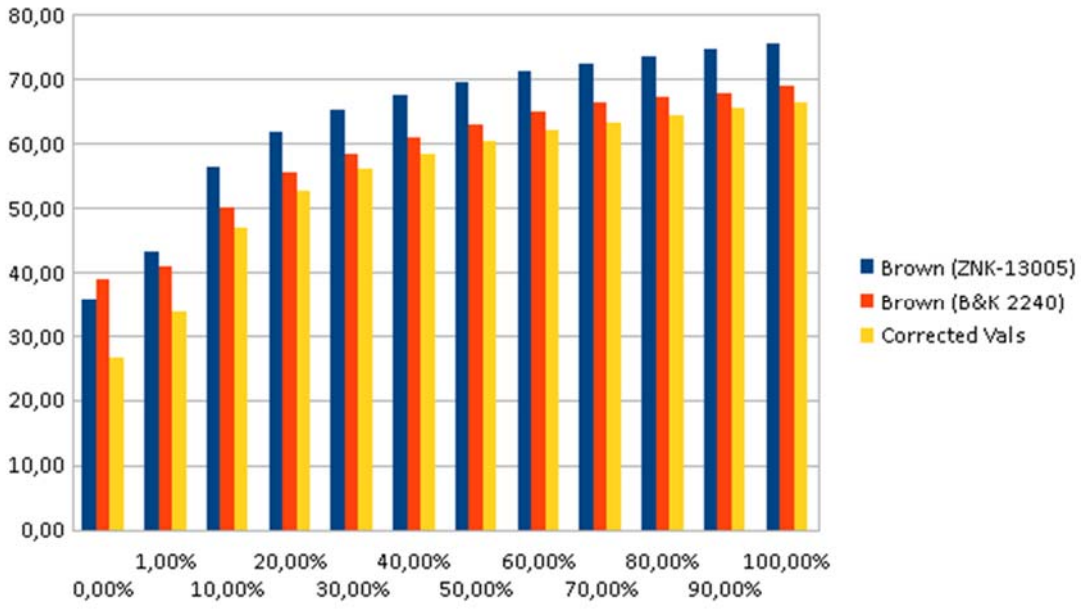
Sensor type	Part number	Price for 1 unit	Price for 1000 units
Current	i-Snail-VC-50	40,-€	34,80€
Voltage	CE-VJ03-32MS2	115,-€	100,-€
Light	TLS2563	2,11€	1,-€
Temp & Humidity	SHT25	12,36€	7,80€
Barometric pressure	BMP085	3,36€	2,03€
Noise meter	ZNK14XXX	35,56€	N/A
Weather meter	Rain Gauge	76,95€	69,26€
	Anemometer		
	Wind Vane		
PM_{10}	GP2Y1010AU	12,50€	N/A
SO_2	4-SO2-20	155,-€	N/A
NO	4-NO-250	150,-€	N/A
NO_2	4-NO2-20	130,-€	N/A
O_3	OX-A421	200,-€	N/A
VOC	iAQ-CORE	37,80€	22,80€

2.4 Lab tests performed to sensors for non-specified features

The noise sensor have tested in the lab in order to know the calibration curve and improve the reliability of the measurement. The measurements has been compared with a reference device, a factory calibrated sound level meter B&K 2240. Table 11 shows the comparison with different noise colours (this means, different spectrum profile's noises) and power levels.



(a)

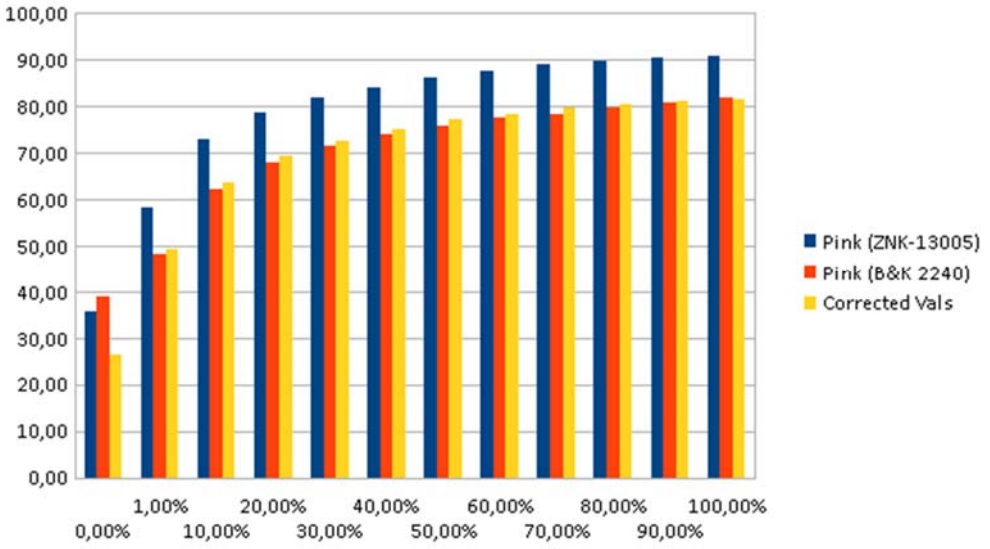


(b)

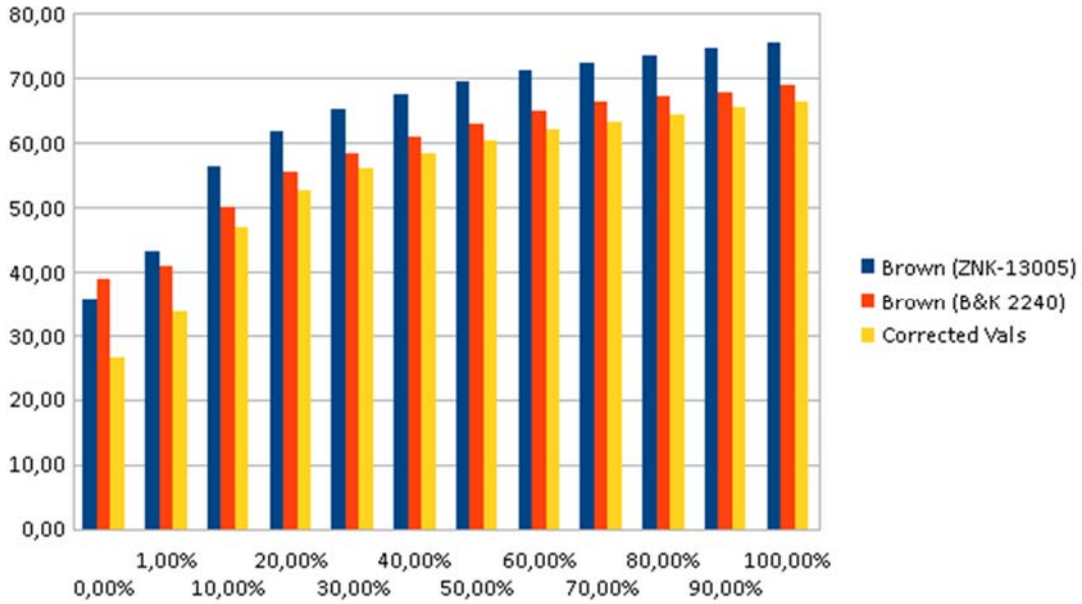
Figure 6 is the graphic result of the calibration process with a pink noise and a brown noise stimulus.

Table 11: Power level comparison between ZNK and B&K 2240 sound level meter results

ABSOLUTE VALUES												
AVERAGE Result	Power Volume of Winamp											
	0,00%	1,00%	10,00%	20,00%	30,00%	40,00%	50,00%	60,00%	70,00%	80,00%	90,00%	100,00%
Pink (ZNK-13005)	35,73	58,33	72,74	78,63	81,75	84,17	86,30	87,58	88,92	89,77	90,31	90,70
Pink (B&K 2240)	38,80	48,20	62,30	68,00	71,60	73,90	75,90	77,50	78,40	79,60	80,70	81,70
Initial Error	3,07	-10,13	-10,44	-10,63	-10,15	-10,27	-10,40	-10,08	-10,52	-10,17	-9,61	-9,00
Corrected Vals	26,53	49,13	63,54	69,43	72,55	74,97	77,10	78,38	79,72	80,57	81,11	81,50
Final Error	12,27	-0,93	-1,24	-1,43	-0,95	-1,07	-1,20	-0,88	-1,32	-0,97	-0,41	0,20
Blue (ZNK-13005)	35,73	53,54	67,99	73,66	76,68	79,05	80,98	82,36	83,69	84,80	85,71	86,56
Blue (B&K 2240)	38,80	41,50	52,50	58,00	61,60	64,00	66,00	67,50	68,80	70,10	71,10	71,80
Initial Error	3,07	-12,04	-15,49	-15,66	-15,08	-15,05	-14,98	-14,86	-14,89	-14,70	-14,61	-14,76
Corrected Vals	26,53	44,34	58,79	64,46	67,48	69,85	71,78	73,16	74,49	75,60	76,51	77,36
Final Error	12,27	-2,84	-6,29	-6,46	-5,87	-5,85	-5,78	-5,66	-5,69	-5,50	-5,41	-5,56
White (ZNK-13005	35,73	53,15	67,37	73,18	76,34	78,84	80,49	82,07	83,43	84,45	85,35	86,31
White (B&K 2240)	38,80	41,90	53,10	58,90	62,10	64,60	66,80	68,20	69,70	70,90	71,80	72,60
Initial Error	3,07	-11,25	-14,27	-14,28	-14,24	-14,24	-13,69	-13,87	-13,73	-13,55	-13,55	-13,71
Corrected Vals	26,53	43,95	58,17	63,98	67,14	69,64	71,29	72,87	74,23	75,25	76,15	77,11
Final Error	12,27	-2,05	-5,07	-5,08	-5,04	-5,04	-4,49	-4,67	-4,53	-4,35	-4,35	-4,51
rown (ZNK-13005	35,73	43,05	56,18	61,77	65,11	67,50	69,40	71,13	72,37	73,49	74,63	75,47
rown (B&K 2240	38,80	40,80	50,00	55,60	58,40	61,00	63,00	64,80	66,40	67,10	67,80	68,90
Initial Error	3,07	-2,25	-6,18	-6,17	-6,71	-6,50	-6,40	-6,33	-5,97	-6,39	-6,83	-6,57
Corrected Vals	26,53	33,85	46,98	52,57	55,91	58,30	60,20	61,93	63,17	64,29	65,43	66,27
Final Error	12,27	6,96	3,02	3,03	2,50	2,70	2,80	2,87	3,23	2,81	2,37	2,63
1Khz (ZNK-13005)	35,73	66,16	81,05	87,16	90,86	92,24	92,30	91,93	91,69	91,55	91,39	91,38
1Khz (B&K 2240)	38,80	60,00	74,40	80,00	85,00	87,10	89,20	90,30	91,90	92,80	94,30	95,00
Initial Error	3,07	-6,16	-6,65	-7,16	-5,86	-5,14	-3,10	-1,63	0,21	1,25	2,91	3,62
Corrected Vals	26,53	56,96	71,85	77,96	81,66	83,04	83,10	82,73	82,49	82,35	82,19	82,18
Final Error	12,27	3,04	2,55	2,04	3,34	4,06	6,10	7,57	9,41	10,45	12,12	12,82
Average Noise	35,73	59,85	74,52	80,41	83,77	85,65	86,64	87,14	87,69	88,17	88,55	88,97
Average Error Di	3,07	-9,10	-11,07	-11,41	-10,47	-10,10	-9,04	-8,24	-7,34	-6,72	-5,85	-5,57
Average Final Err	1,9182142857				-2,0929910714							
1Khz Final Error	7,1566666667				8,0785714286							
Correction Facto			0	-9,2								



(a)



(b)

Figure 6: Power comparison between ZNK and B&K 2240 sound level meter with (a) pink noise and (b) brown noise

2.5 Analysis of the natural frequencies in the selected sensors and measurements

The last critical parameter in terms of the scalability of the system, also something that is almost never shown in the sensors datasheet (because it is not indeed a feature of the sensor but of the combination of sensor, the environment and the type of measurement for which the sensor is used for) is what is known as the natural frequency, defined as the spectrum of the parameter that is being measured.

These frequencies, according to the Nyquist theorem, determine the sampling frequency (and therefore the bandwidth we need to save and/or send the data) for each measurement.

To know which is this optimum sampling frequency for each parameter can be very complex, involving multiple factors such us (also discussed in 2.3.4.1):

- the recommendation from the manufacturers,
- the current legal regulations [ANF001],
- physical factors, such as the environmental conditions,
- the type of measurement
- and the scalability we expect from the system, that can be taken as a consequence of the measurement bandwidth, or as a requirement.

Table 12 shows the minimum and recommended, when apply, sampling interval.

Table 12: Natural measurement frequencies, minimum and recommended.

Measurement type	Part number	Minimum measurement interval	Recommended measurement interval	Comments
Current	i-Snail-VC-50	Limited to ADC, ≥250ns		Depending on data application
Voltage	CE-VJ03-32MS2	200ms		Depending on data application
Light	TLS2563	400ms	*Light is better to be measured as an average over a period (i.e. 1min, 5min, 15min, etc.) due to the possibility of abrupt changes.	Depending on data application
Temp & Humidity	SHT25	As limited by I^2C communication	1 hour	[ANF002]
Pressure	BMP085	1 second	1 hour	[ANF002]
Noise	ZNK14XXX	≥1 second	1 minute	Internal average each second is needed to acquire completely sound bandwidth
Weather	Rain Gauge	Pulsed measure	N/A	Using software interrupts
	Anemometer	Pulsed measure	N/A	Using software interrupts
	Wind Vane	Limited to ADC, ≥250ns	1 hour	[ANF002]
PM_{10}	GP2Y1010AU	0,42ms	1 or 24 hour	According to European normative
SO_2	4-SO2-20	≥45 seconds	1 hour	
NO	4-NO-250	≥30 seconds	1 hour	According to European normative

NO_2	4-NO2-20	≥30 seconds	1 hour	According to European normative
O_3	OX-A421	≥60 seconds	1 hour (maximum daily 8 hour average)	
VOC	iAQ-CORE	1 second (continuous) 11 seconds (pulsed)		Depending on data application

3 Trade-offs

3.1 Trade-offs between system security and energy consumption using DTLS

3.1.1 Introduction

A modified version of Transport Layer Security (TLS), Datagram Transport Layer Security (DTLS) has been introduced to address issues associated with unreliable datagram traffic such as User Datagram Protocol (UDP). DTLS For UC-O2 - Environmental monitoring in particular, it is expected that networks will be formed by a potentially very high number of RDs and therefore scalability is a requirement a mechanism that allows packet retransmissions and reordering whenever it is required, the interested reader is referred to [MR04]. DTLS consist of handshake and record protocol. The main purpose of the DTLS is to authenticate communication parties, negotiate a cipher suite and exchange keys that are later used to protect traffic data. All handshake messages can be divided on six groups of messages (so-called flights) between communication parties; three flights are sent from a client to a server and three from a server to a client. Note that DTLS protocol consists of both secret (a pre-shared key) and public-key (raw public keys and X.509 certificates) options. A detailed description of DTLS operation is provided in D3.1 [RD3.1].

3.1.2 Relevance to RERUM's Use-Cases

For UC-O2 - Environmental Monitoring in particular, it is expected that networks will be formed by a potentially very high number of RDs and therefore scalability is a requirement. By employing DTLS in RERUM, it will allow to scale in the RERUM Use Cases. Furthermore, DTLS can be used transparently by any end applications, which themselves reside at the application layer of the TCP/IP network stack. This transparent feature implies that DTLS could be applied to all four of RERUM's Use-Cases. It directly addresses the following User Requirements, as specified in deliverable D2.1 [RD2.1]:

- UR-7: The user needs to protect his measurements from malicious users;
- UR-25: User requires open solutions for authentication between devices, ensuring the integrity of their data as well as the confidentiality.

In this subsection we focus on the evaluation of DTLS over COOJA, Contiki OS simulator. More specifically, we study the trade-off scenario between energy consumption and network security for DTLS.

3.1.3 Performance evaluation

The goal of this simulation campaign is to analyse the performance of DTLS protocol. To this aim, hereafter, we present simulation performance evaluation results in terms of computation (handshake) latency and energy consumption of a specific implementation of DTLS. More specifically, we study the performance of TinyDTLS (implemented in Contiki) both under Elliptic Curve Cryptography (ECC) with one of the well-known DTLS curves and Pre-Shared Key (PSK) modes.

3.1.3.1 Simulation setup

In this set of simulations over COOJA emulator, we have employed two motes: one was utilized as a client while the other as a server, respectively. We located the two motes close to each other in order to guarantee the reliable and successful wireless communication. We ran our simulations on top of Contiki OS while each set of simulations was executed ten times to derive standard deviation values.

3.1.3.2 Simulation Results: System Security vs Energy Consumption

Overall DTLS handshake latency

The handshake time is calculated from the first *"ClientHello"* message from the client node until the last message (i.e., "*Finished*", Handshake compete) from the server node.

In Table 13 both average values along with standard deviation are provided. Our simulation results show that ECC requires 215.319 seconds in average, while PSK mode only 3.841 seconds. It is worth mentioning that our evaluation took place on top of two constrained devices. We should consider that in most of the applications the gateways are operated over powerful devices, and thus, the computation time for ECC would be halved, since the computation time for the server side will be negligible.

Table 13: Handshake latency for ECC and PSK.

Handshake time (in seconds)			
ECC		**PSK**	
Average	Standard Deviation	Average	Standard Deviation
215.319	6.2126	3.841	0.712

Energy consumption analysis

All energy consumption results were retrieved by utilizing the energest module of Contiki. This energy estimation module maintains a table with entries for all components such as Central Processing Unit (CPU), radio transceiver. Each table entry contains the total time that the corresponding component has been turned on, more specifically, it monitors in real-time the radio and CPU usage (i.e., ECC curve computation) by saving the duration spent in each state (e.g., transmitting, receiving data, awaken, sleeping). This information is then combined with the energy values that are detailed in the component datasheet for each state in order to provide an accurate calculation of energy consumption. In Table 14 the total energy consumption is presented. The results show that the long computation time and multiple packet exchanges of ECC scheme has a straightforward impact on energy dissipation. Indeed, as can be observed, PSK consumes much less energy when compared to ECC heavy in computation scheme.

Table 14: Total energy consumption for ECC and PSK.

Energy consumption (mJ)	
ECC	**PSK**
1161	20.52

3.1.4 Discussion

In this simulation evaluation, our results demonstrate a fundamental trade-off scenario between computation time, energy consumption and system security (i.e., key management). Indeed, from the utility point of view, from the first glance ECC mode presents inefficient performance (i.e., high energy consumption and long handshake time) when compared to PSK. These results can be explained by the fact that in PSK mode, there is little computation, since the nodes basically generate the random values (one single key) during the handshake, and then they exchange them, and thus, presents low-level security. On the other hand, the ECC, which is considered to be secure (since it follows key-management strategy) and efficient approach to public-key cryptography, requires high computations (i.e., of the elliptic curves) both at the client and server side.

We should take into account different type of applications that users may operate. For instance, ECC-based approach could be durable for time driven applications such as temperatures or humidity readings (i.e., a typical RERUM UC-O2 – Environmental Monitoring), since it is necessary for the devices to perform handshakes constantly. Furthermore, in dense scenarios where there are many nodes, ECC mode is more effective in terms of key management, and thus, one time ECC overhead might be suitable, and moreover, the users benefit public key scheme. However, ECC is less efficient in real-time applications (e.g., turning-on the light), where PSK-based schemes are more suitable.

3.2 Trade-offs between network performance, signalling overheads and energy consumption with the BMFA multicast forwarding algorithm

3.2.1 Introduction

Bi-Directional Multicast Forwarding Algorithm (BMFA) for 6LoWPANs was designed and developed in RERUM (during Task 4.2) deliverable D4.1 [RD4.1] to address the needs of RERUM use-cases by achieving very low delay performance and energy consumption. This section focuses on trade-offs between network scalability and its performance (i.e., delay, reliability and energy consumption).

3.2.2 Relevance to RERUM's Use-Cases

In scenarios involving point-to-multipoint network traffic, transmitting to each destination individually with unicast leads to (i) poor utilization of network bandwidth, (ii) excessive energy consumption caused by the high number of packets and (iii) suffers from low scalability as the number of destinations increases.

For UC-O2 - Environmental monitoring in particular, it is expected that networks will be formed by a potentially very high number of RDs and therefore scalability is a requirement. In cases when RDs are powered by batteries, it is impractical or outright untenable to replace batteries very frequently due to high management cost and possibly hard-to-reach installation locations. Thus, long battery life is important. For devices powered from mains, low energy consumption is also important in order to reduce financial cost, but also in order to comply with national and international regulations where applicable.

Through the phase of the experiments we are targeting not only showing comparative results between our algorithm BMFA and its rival Trickle Multicast / Multicast Protocol for Low power and Lossy Networks (TM / MPL), but also to identify their overall behaviour when they are subjected under different configurations. Thus, in this subsection, we present thorough performance evaluation results related with end-to-end delay, reliability and energy consumption.

3.2.3 Performance evaluation

3.2.3.1 Simulation setup

For our investigation, we employ the COOJA emulator [O06]. Our setup consists a number of nodes as discussed previously, with each one of them running either BMFA or TM as multicast forwarding algorithm. Each node was also assigned a role in the network such as being a sink node, a source node or a simple traffic forwarder; the latter is a node that is not subscribed to any multicast groups but that is capable of interpreting and forwarding multicast traffic. For the sake of clarity, we used a radio model based on disks, Unit Disk Graph Medium (UDGM), which models wireless losses based on distance. We configured UDGM so that each node had a TX range of 50m and an interference range of 60m Figure 7. The details of the simulation setup are exposed in Table 15.

3.2.3.2 Simulation Results

Throughout this simulation campaign, we present the network scalability analysis. To this aim, we study the following metrics:

- End-to-End Delay
- Packet Delivery Ratio
- Energy Consumption

More specifically, we here investigate the potential trade-off scenario between network scalability and its performance (i.e., energy consumption and reliability).

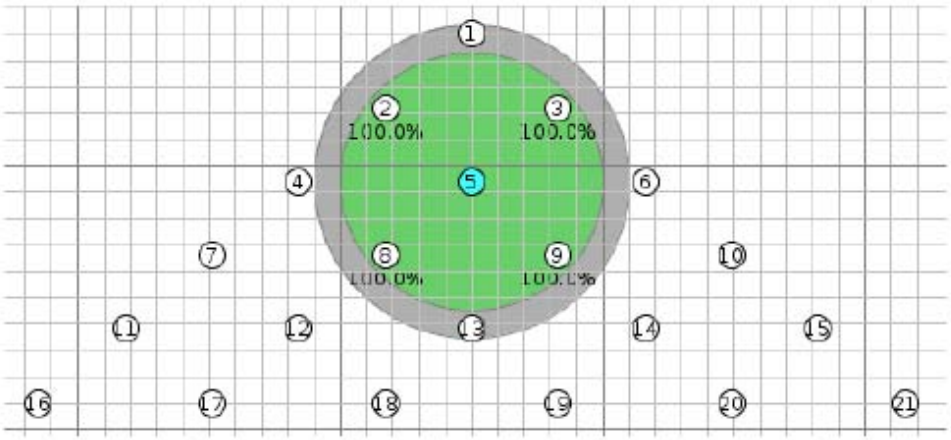


Figure 7: Simulated topology and transmission range within Cooja.

Table 15: Simulation setup.

Simulation Parameter	Value
Number of nodes	21
Number of sources	1 traffic source, 20 sinks
Traffic pattern	CBR and VBR
TM	lmin in [125, 500, 700] *ms*
BMFA	Spread in [2, 4]
MAC	CSMA
Duty cycling	ContikiMAC (CCI 125 *ms*) & NullRDC
PHY	IEEE 802.15.4
Radio medium	Unit Disk Graph Medium (UDGM)
Transmission range	TX: 50 *m*, Interference: 60 *m*

3.2.3.2.1 Scalability Analysis: network density versus performance

End to end delay

As can be observed from Figure 8, TM algorithm does not perform as it was expected based on the configured parameters. For instance, TM configured with Imin = 750 *ms* lead to the lowest end-to-end delay across the board (for different traffic bits rate per density). Moreover, the end-to-end delay supposed to be lower as the lmin decreases, while our simulation results present the opposite. Finally, high delay performances were expected for TM scheme since ContikiMAC, due to packet trains, is inefficient for broadcast transmissions.

On the other hand, under BMFA scheme, for low densities (i.e., up to 0.35) the end-to-end delay declines slightly as the inter-packet delay increases, while when the bit rate gets lower (Variable Bit Rate with 1-2 *s* inter-packet delay) the delay reaches its maximum. This phenomenon is due to that as the bit rate increases the more a packet waits into the cache; it does not get transmitted until all packets preceding it are forwarded first. Furthermore, an inter-packet delay of more than 1 *s* leads to the opposite results since all existing packets in the cache are getting forwarded before the next one arrives. On the other hand, for high densities (i.e., 0.71) the delay continues its descending trend as the inter-packet delay increases.

Based on the fact that for high network densities a node is expected to be selected as preferred parent from a greater number of nodes (RPL's DODAG becomes shallow and wide), more packets are expected to wait into its cache until they get forwarded; in this case packets transmitted with VBR do not arrive neither too soon nor too late to the recipient nodes resulting to even better results. To summarize the performance of the two algorithms, BMFA outperforms TM for at least five times in under any configuration.

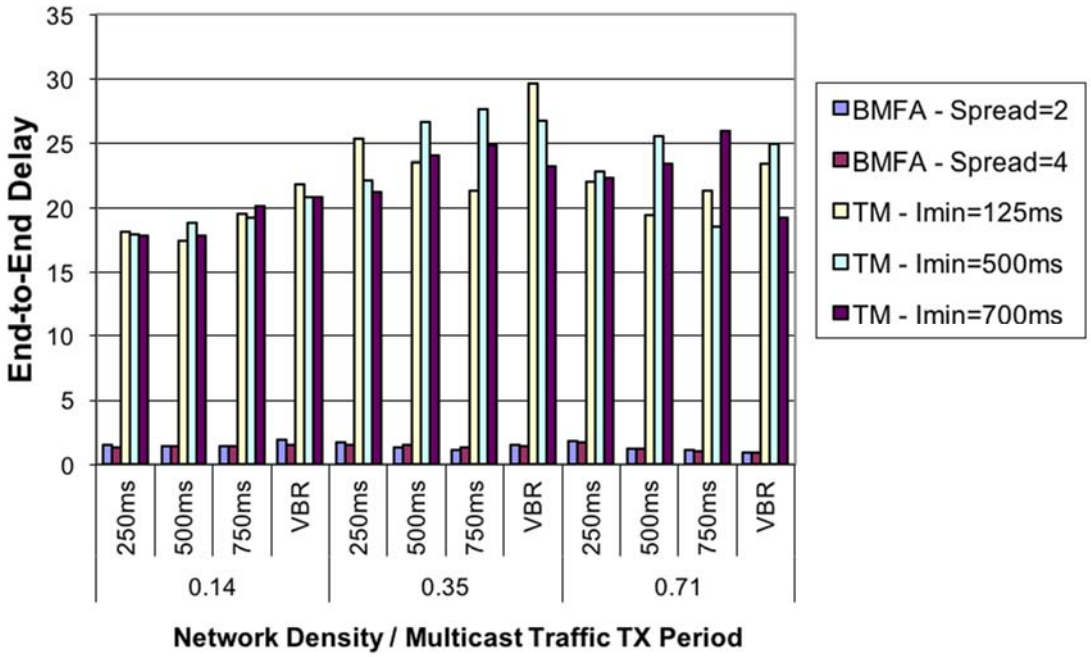


Figure 8: End to end delay performance.

Packet Delivery Ratio

Regarding the reliability, as can be observed from Figure 9, TM performs better in dense networks, due to its path redundancy reliance. While there is a slight improvement moving from network density 0.35 to 0.71, TM achieves its highest results in low densities. Despite the fact that higher densities can

boost its performance, the signal collisions that occur can reduce its efficiency significantly. The more nodes are in range, higher the probability of the noise (i.e., interference) in the network; especially in the case of ContikiMAC which uses data packet trains. It is worth to mention that the ICMPv6 messages can be dropped for the same reason. Moreover, irrespective of network density and Imin, TM achieves lower packet loss as the transmitting bit rate decreases. This phenomenon is due to the bit rate increases, the probability of having packets overwritten in the local cache before getting forwarded gets higher. On the other hand, under BMFA scheme, the higher is the density, the larger is the set of parents, and consequently, better choice of preferred parent can be made. BMFA presents better results both for low and high density when compared against the intermediate. Moreover, regardless of network density and forwarding delays, reducing the traffic bit rate the packet delivery ratio increases.

By comparing the performance of the two algorithms, according to our simulation performance evaluation BMFA performs better than TM. However, TM may outperform BMFA under different configuration, as it has been presented in [OP2013], [OPT2013], where TM achieves better results.

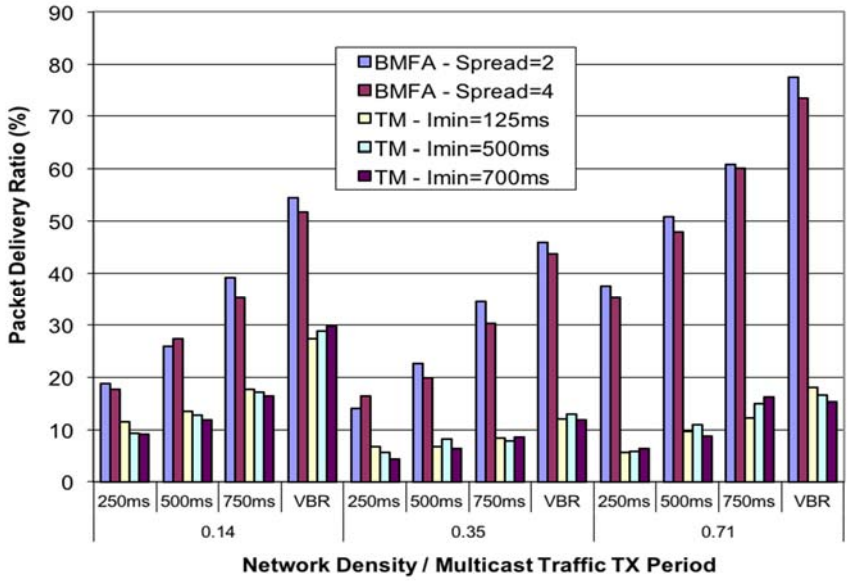


Figure 9: Network reliability in terms of packet delivery ratio

3.2.3.2.2 Energy Consumption

Through the facilities provided by Contiki's energy consumption estimation module (energest) [DOTH07a, DOTH07b], we measured the time each node spent in each of the following three states over the duration of each experiment: i) MCU active, ii) RF listening / receiving, iii) RF transmitting. Since we are simulating the exp5438 (Texas Instruments (TI) MSP430F5438 experimenter board) [TI14, TI07], we then converted these time values to estimated energy consumption based on typical datasheet power levels at an operating voltage of 3.0V. This includes the consumption of the Micro-Controller Unit (MCU), a Texas Instruments (TI) MSP430F5438 [TI14], as well as the consumption of the TI CC2520 radio transceiver [TI07]. The values are summarised in Table 16.

Table 16: Typical exp5438 current draw with an operating voltage of 3.0 V at 25°C.

Mode	Current consumption
MCU active @ 8 *MHz*, Code execution from Flash	2.50 *mA*
MCU in deep sleep (LPM3, XT1LF TI14)	2.60 *μA*
CC2520 Frame reception at an input level of -50 *dBm*	18.5 *mA*
CC2520 Frame Transmission, 0 *dBm* output power	25.8 *mA*

NullRDC keeps radio transceivers always on (no duty cycling). As a result, the majority of energy is consumed during idle listening or packet reception, with other components contributing insignificantly. For this reason, we only consider ContikiMAC for the evaluation of the two algorithms in terms of energy consumption. Generally for TM it can be observed (Figure 10) that for higher densities less energy is consumed due to the fact that agreement between all nodes can be achieved with fewer ICMPv6 control message exchanges; in other words inconsistencies can be solved with fewer hop-by-hop transmissions. This can be also observed by the fact that as the density increases less energy is required for transmitting than for listening.

For BMFA, irrespective of network density, as the inter-packet delay between the transmitted packets increases, the energy consumption decreases since fewer packets are forwarded during the experiment. In the case of the highest density (0.71) and for high bit rate we can see that the energy consumption of BMFA approaches the one of TM's. This happens because nodes are consuming too much energy by keeping the radio on as a result of picking up transmissions from their large number of neighbours, despite the fact that they only forward packets received only from their children or preferred parent. By comparing the two algorithms we can see that BMFA is more energy efficient than TM since it forwards each packet only once and there is no ICMPv6 message exchange. Moreover, we must highlight that the energy consumption for CPU indicates the complexity of the two algorithms and it becomes noteworthy that TM's complexity is much higher than BMFA's.

Generally, BMFA can outperform TM in energy efficiency especially in low density networks where TM consumes four times more energy; and assuming that TM can be configured to achieve higher PDR, its energy consumption is expected to be even higher.

3.2.4 Discussion

In the context of this deliverable, we have evaluated the RPL-based Bi-Directional Multicast Forwarding Algorithm (BMFA) for 6LoWPANs. BMFA was developed during Task 4.2, while it was introduced in RERUM deliverable D4.1 [RD4.1] for smart city applications and RERUM use cases, UC-O2 – Environmental Monitoring in particular. Through the phase of the simulations, we targeted not only presenting comparative results between proposed BMFA and its rival Trickle Multicast, but also to identify their overall behaviour when they are subjected under different configurations.

To this aim, we investigated the scalability issue and its impact in the network performance. Our thorough simulation performance evaluation, conducted with COOJA emulator on top of Contiki OS, demonstrates that BMFA can outperform TM in energy efficiency especially in low-density networks where TM consumes four times more energy; and considering that TM can be configured to achieve higher PDR, its energy consumption is expected to be even higher.

Furthermore, we observed a typical trade-off situation between network performance, energy consumption and reliability. More specifically, our results show that BMFA outperforms TM, by terms of reducing the end-to-end delay, design complexity and energy consumption. On the other hand, TM severely outperforms BMFA in reliability, since TM was designed to be reliable scheme.

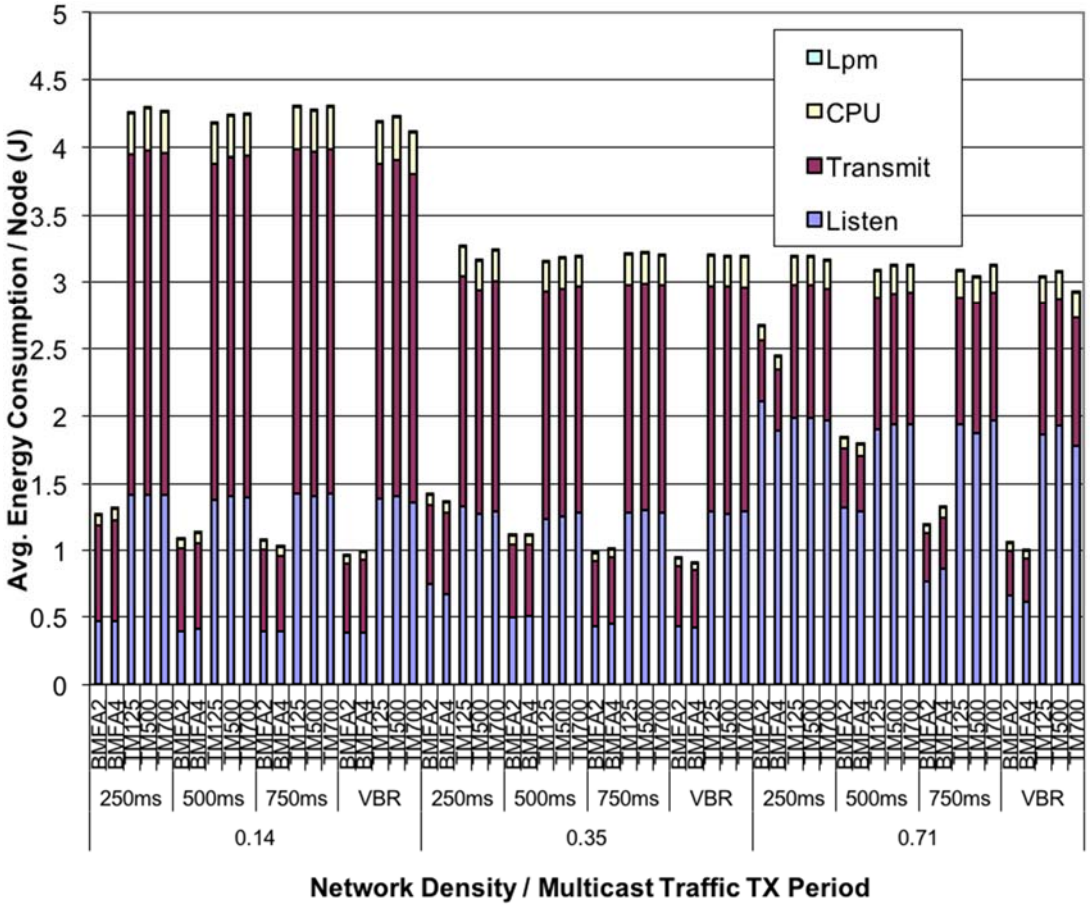


Figure 10: Average node energy consumption

3.3 Trade-offs between network performance and energy consumption of the CADC congestion-aware duty-cycling algorithm.

3.3.1 Introduction

In sensor applications such as surveillance, fire detection and object-tracking systems, a sudden event can lead to simultaneous generation of data by multiple sources. This will cause network congestion as the paths approach the sink node [PPGNA2015]. Congestion leads to packet losses, which in turn lead to re-transmissions thus causing energy consumption that could have been avoided by proactively preventing the network congestion. Furthermore, while the network is in a congested state, RDs will stay awake for a period of time in order to resolve it, and this leads to additional energy consumption.

If a Medium Access Control (MAC) protocol is tuned to a specific achievable throughput, then the network will suffer from data loss due to its incapability to adapt to the traffic needs [YHE2002], [PGST2015], [BYAH2006], [PBGN2014], [MLTK2008] and [D2011]. To reduce packet loss, wake up frequency must be increased. However, this situation leads to a trade-off, since it consumes more

energy. Ideally, the period between wake ups for the Radio Duty Cycle (RDC) should adapt to traffic loads increasing during heavy traffic and decreasing when the nodes are idle.

The main objective of this section is to study the trade-offs between the duty-cycling and 6LoWPAN performance, since increased channel samplings significantly increase the performance of the network in terms of goodput, delay and packet loss. However, increased duty cycles have significantly higher energy consumption when the network is idle.

The work discussed in this subsection is joint work with Loughborough University, Computer Science and has been published in [MOGP2014]. Furthermore, the implementation details were introduced in RERUM deliverable D4.2 [RD4.2].

3.3.2 Relevance to RERUM's Use-Cases

In scenarios involving event-based deployments, network traffic is of a very bursty nature: The network is idle for most of the time with the only traffic being occasional network control packets. When an event occurs to a parameter under observation, multiple RDs may attempt to transmit simultaneously. An example of this is the simultaneous transmission of ambient temperature readings when the measured temperature exceeds a high threshold.

The scheme proposed here is relevant to UC-O2 - Environmental monitoring, whereby deployments will be formed by a large number of devices. Topologies are potentially going to be very dense and this makes them particularly susceptible to congestion due to a large number of devices occupying the wireless medium in the small geographic area. In cases of battery-powered RDs, it is impractical or outright untenable to replace batteries very frequently due to high management cost and possibly hard-to-reach installation locations. Thus, long battery life is important and the proposed scheme perfectly fits the requirement for prolonging the network lifetime. For devices powered from mains, low energy consumption is also important in order to reduce financial cost, but also in order to comply with national and international regulations where applicable.

For the indoor use cases, the devices can be plugged in the power outlets, so energy efficiency may not be so critical. However, frequent battery replacement is a costly nuisance for the end user, while regulations also need to be adhered to.

3.3.3 Performance Evaluation

3.3.3.1 Simulation Setup

To evaluate the performance of Congestion Aware Duty Cycle (CADC) scheme, we performed a thorough simulation campaign [PKGCN2016] on top of Cooja [O06] emulator. We compare CADC against ContikiMAC, X-MAC and BEAM in a simulated environment. We conducted in excess of 10.000 simulation executions under various network conditions.

ContikiMAC serves as a base-line for our comparisons. We chose to compare our scheme with X-MAC because of its high research impact and because an implementation of X-MAC with Contiki OS already exists, making it demonstrably deployable. Lastly, we chose BEAM because its dynamic RDC approach makes it very relevant to the research presented here and because it constitutes an improvement over X-MAC. Additionally, this choice of algorithms offers an insight into the differences between dynamic RDC techniques and various configurations of their static equivalents.

For our tests we use Constant Bit-Rate (CBR) traffic with a 24-byte application-layer payload, and Variable Bit-Rate (VBR) traffic. For VBR traffic, sources were randomly setting their inter packet transmission interval to a value in the range 62.5 *ms* to 250 *ms*. The packet transmission interval was randomly changing every few (between 20 - 70 packet transmissions) packet transmissions and each

VBR experiment was repeated 20 times. MAC layer packet queues were set to 10 packets and were used for buffering outbound traffic.

In Contiki, the default configuration of Channel Check Rate (CCR) is 8 (each node will wake-up 8 times per second). We conducted comparisons with multiple different static CCR configurations (i.e., 4, 8, 16, 32 and 64) to get representative results.

The packet transmission intervals used during the simulations were: 500 *ms*, 250 *ms,* 125 *ms* and 62.5 *ms*. This range of CCR values was sufficient for understanding how a static Radio Duty Cycle (RDC) would behave in each case (low, medium or high CCR) and consequently how CCR can affect the performance of static RDCs. The details of the simulation setup are exposed in Table 17.

Table 17: Simulation configuration permutations

	ContikiMAC	X-MAC	BEAM	CADC
MAC layer	CSMA	CSMA	BEAM	CADC
RDC layer	ContikiMAC	X-MAC	BEAM	CADC
CCR (*Hz*)	4, 8, 16, 32, 64	*dynamic*	4, 8, 16, 32, 64	*dynamic*
min_{ccr} (*Hz*)	config-dependent	config-dependent	4	4
max_{ccr} (*Hz*)	config-dependent	config-dependent	64	64
N	N/A	N/A	N/A	10
Q_h	N/A	N/A	100%	90%
Q_l	N/A	N/A	60%	60%
F_tT	N/A	N/A	N/A	20%

For simulations we used emulated Tmote Sky nodes and the Unit Disk Graph Medium (UDGM) to emulate the lossy nature of the radio environment. At the routing layer, we rely on a broadly used and scalable Routing Protocol for Low power and Lossy Networks (RPL) protocol [TED2010]. Furthermore, during all simulations there was constant background control plane traffic, such as IPv6 Neighbour Discovery Solicitations and Advertisement messages, as well as RPL control packets.

Two different topologies were used for the simulation: i) a line topology of 9 nodes (1 source, 1 sink and 7 intermediate nodes, with varying distance in hops between source and sink across different runs) and ii) a 25 node random topology with 6 sources, 1 sink and 18 intermediate nodes (Figure 11).

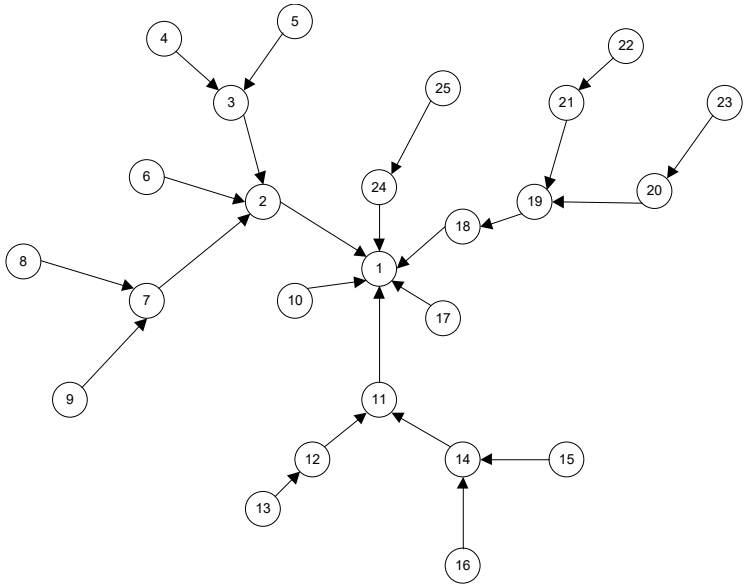


Figure 11: Indicative routing topology during random topology simulations.

In terms of node degree, random topologies had an average node degree of between 4 and 5. Degree in the linear topology was 2 for all intermediate nodes and 1 for the two edge nodes. Node degree is the number of a node's single-hop neighbors (can directly communicate with one another) at the link layer.

In VBR simulations, the distance between the source and the destination was 4 hops. The first topology aimed at presenting each mechanism's performance in a low density network as well as how the hop count affects the performance of each algorithm. The aim of the simulations in a random topology was to investigate the impact of node density on algorithm performance.

Each simulation was repeated 15 times with a new random seed per iteration, while the following metrics were recorded:

- Goodput as the total number of unique packets received by the sink.
- Packet loss.
- Delay: in simulations we measured end to end delay (from source to sink).
- Energy consumption: measured as the sum of four components: i) Energy consumed by the CPU under normal operation, ii) energy consumed when the CPU is in sleep mode, iii) energy consumed by the RF transceiver while transmitting and iv) energy consumed by the RF while listening or receiving.
- Code footprint and memory requirements.

3.3.3.2 Goodput

Figure 12 shows the percentage of packets received successfully for different scenarios of channel check rates (i.e., 4, 8, 16, 32 and 64) and packet transmission intervals. Figure 13 shows the impact of the CCR configurations in goodput under heavily congested scenarios. Figure 14 highlights each algorithm's best goodput configuration for various packet transmission intervals.

Our performance evaluation results demonstrate that CADC and ContikiMAC outperformed X-MAC and BEAM. CADC's goodput is similar to the highest CCR configuration of ContikiMAC.

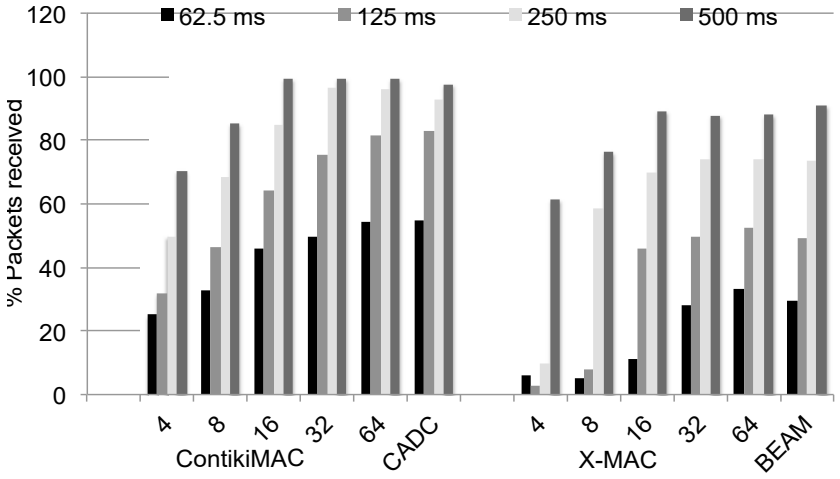


Figure 12: Percentage of packets received successfully under different CCRs and inter-packet transmission intervals (simulations, random topology).

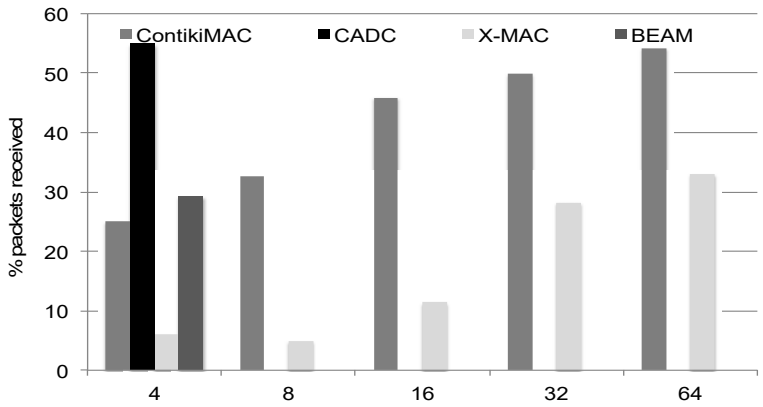


Figure 13: ContikiMAC and X-MAC best performances. BEAM and CADC in normal operation mode (simulations, random topology): Percentage of packets received successfully under different CCRs (transmission interval = 62.5 *ms*).

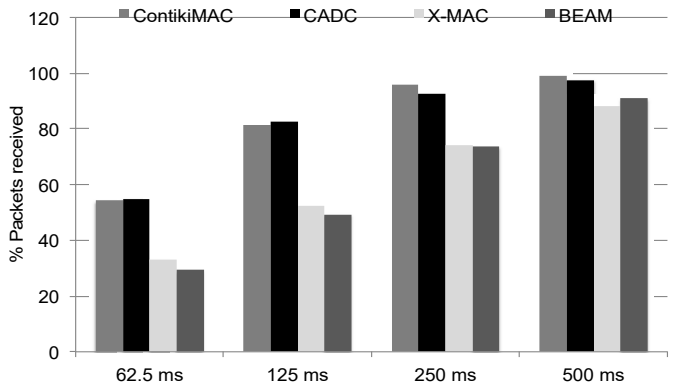


Figure 14: ContikiMAC and X-MAC best performances. BEAM and CADC in normal operation mode (simulations, random topology): Percentage of packets received successfully under different packet transmission intervals (CCR = 64).

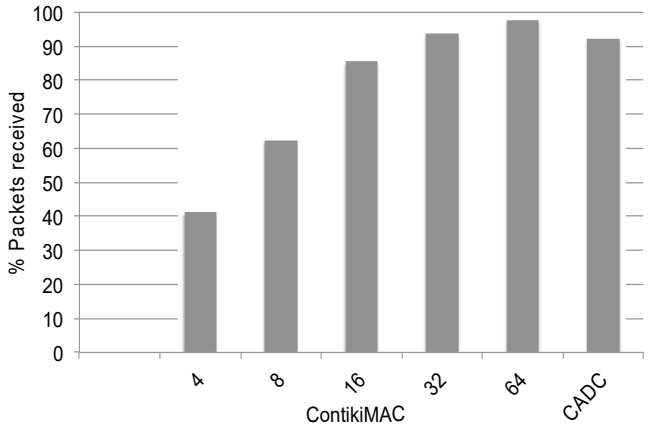


Figure 15: Percentage of packets received successfully under VBR traffic (simulations).

Figure 15 demonstrates the percentage of packets received successfully by CADC under different transmission intervals. As can be observed, the percentage of successfully received packets was slightly higher when the number of transmitted packets increased. The main reason for this is that some packets were dropped during CADC's initial channel check adaptation phase. Therefore, the results demonstrate how dynamic RDCs such as CADC may suffer from slight decreases in goodput when the amount of traffic fluctuates.

During VBR simulations, traffic sources frequently reconfigured their packet transmission intervals. CADC's goodput was similar to ContikiMAC's best CCR configurations (i.e., channel check rates configured at 32 and 64). CADC's goodput performance was slightly worse than the one demonstrated under CBR traffic. This is expected since random packet transmission intervals will result in very frequent CCR adaptations, which cause slightly higher packet losses.

Under both CBR and VBR traffic sources, CADC significantly outperformed ContikiMACs default CCR configuration by between 30% and 60%.

Figure 16 illustrates results comparing the performance of the different mechanisms over longer paths. Here we only include the best-performing CCR configurations of ContikiMAC and X-MAC. We observe that when the hop count between the source and the sink increased, CADC achieved the most stable behaviour. CADC's goodput dropped by 24% when we increased the distance by 8 hops. On the other hand, ContikiMAC's goodput dropped by 30%, X-MAC's by 55% and BEAM's by 43%.

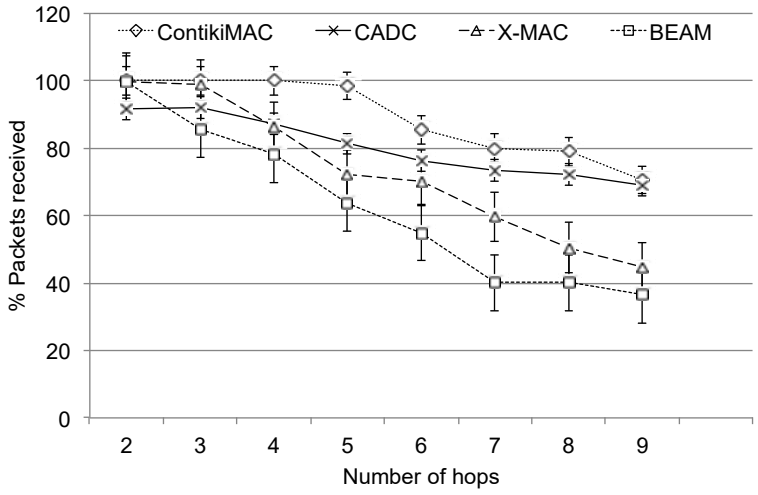


Figure 16: Percentage of packets received successfully over hop count (simulations).

3.3.3.3 Packet Loss

Figure 17 presents a detailed analysis of packet loss under different transmission rates during random topology simulations. CADC and ContikiMAC achieved lower loss under all simulated topologies.

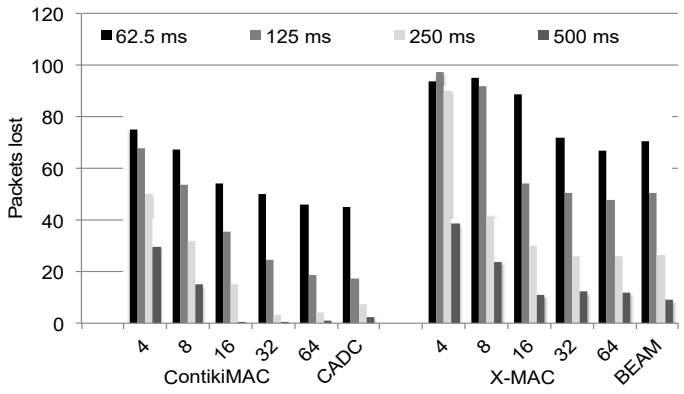


Figure 17: Packets lost under different CCRs and inter-packet intervals (simulations, random topology).

When ContikiMAC is configured with a CCR of 64, it's delay times were lower than CADC. However, this observation is expected since CADC takes time to converge from a lower to a higher CCR. Overall, CADC achieved lower delay times than ContikiMAC's configurations of 8, 16, 32. In addition to measuring packet loss, this experiment serves as a metric to evaluate bi-directional traffic goodput, due to the fact that in this experiment all received packets were unique. Therefore, packet loss of 30% corresponds to a 70% goodput.

3.3.3.4 Packet Delay

Figure 18 shows the simulation results for packet delay against hop count in a line topology and Figure 19 for random topologies. Simulations indicate that X-MAC and BEAM achieved lower delay times than CADC and ContikiMAC under low traffic rate configurations. However, when the network was operating with high traffic rates (62.5% inter-packet interval), CADC and ContikiMAC achieved similar delay times as X-MAC and BEAM. Moreover, it can be observed that ContikiMAC's performance under low CCR configurations was very poor, with approximately a ten-fold increase in delay between the corner CCR values of 4 and 64.

Overall, CADC achieved better delay times than most of the configurations of ContikiMAC and X-MAC. During simulations of low- to medium-congestion scenarios, CADC's delay was higher than BEAM's. On the contrary, when the network was heavily congested (packet interval 62.5 *ms*), CADC achieved lower delay times than BEAM.

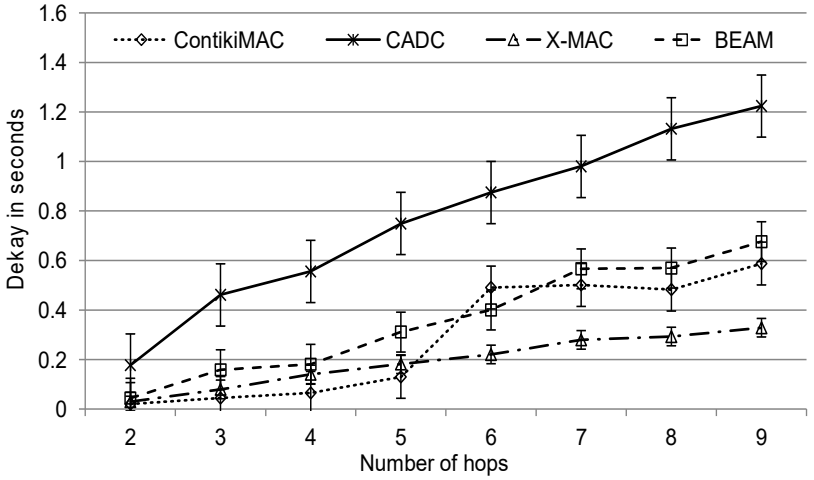


Figure 18: Packet delay over distance in hops (simulations).

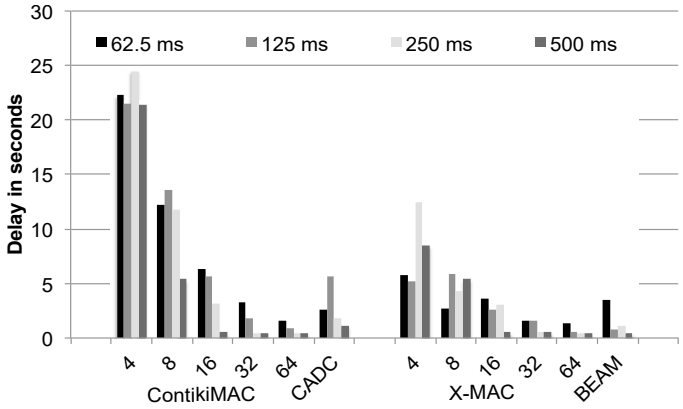


Figure 19: Average packet delay for different CCRs and packet transmission intervals (simulations, random topology).

3.3.3.5 Energy Consumption

Figure 20 shows the average energy consumption per second per node when the network is idle. Hence, we observe that energy consumption values for various device states is different. When the network was idle, energy consumption was significantly increased under higher CCR values. If we further analyse energy consumption due to radio TX/RX, it is noticeable that it approximately doubles for each increase in the value of CCR. With ContikiMAC and X-MAC, the CCR is pre-configured with the default CCR value in Contiki (for the majority of platforms this is 8). In order to increase the bandwidth, a WSN must be configured with a higher CCR. However, configuring a WSN with a CCR of 64 will result in approximately 4 times higher RX/TX energy consumption compared to the default CCR. Therefore, in real application scenarios it is unrealistic to assume that a deployment of battery-powered nodes will be configured with a CCR of 64.

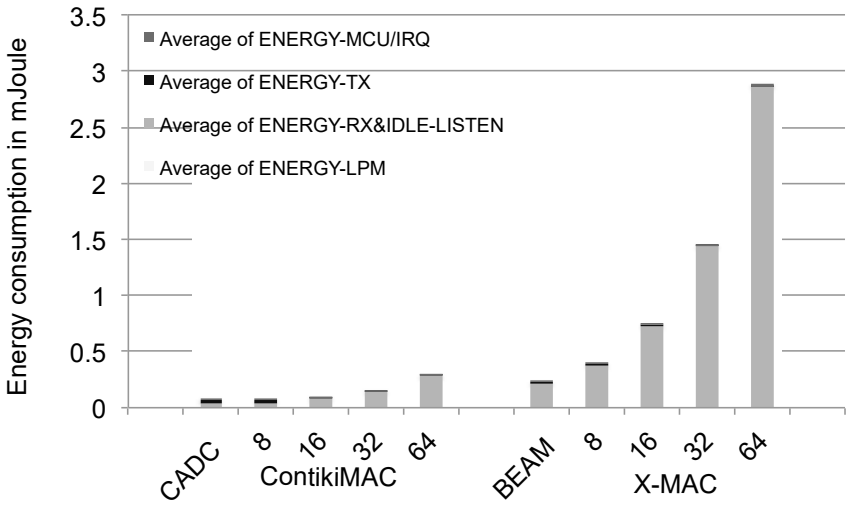


Figure 20: Average idle network energy consumption/sec per node (simulations).

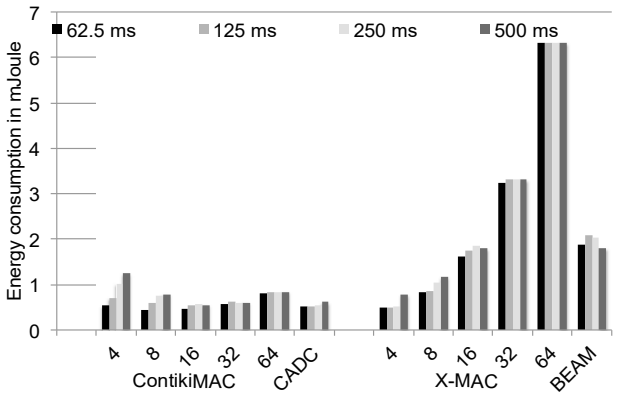


Figure 21: Energy consumption (per node) for each transmitted packet under different inter-packet intervals and CCRs (simulations).

Figure 21 illustrates the per node average energy consumption in the network for each packet transmitted by the sources. In contrast to what we observed under the idle network state, ContikiMAC demonstrated varied energy consumption across its various CCR configurations. Energy consumption with X-MAC significantly increases with CCR. Since ContikiMAC has lower idle energy consumption, it does not always follow X-MAC's energy consumption patterns. When significant traffic is present in the network, it can be observed that ContikiMAC's CCR configurations of 16 and 32 achieved lower energy consumption than configurations of 4 and 8. This can be attributed to the unique design of ContikiMAC's packet transmission scheme, whereby the duration of each packet train gets longer as CCR decreases. Therefore, energy saved when the network is idle during low CCR configurations can be overshadowed by the energy consumed due to transmissions.

Even though the same principle applies to X-MAC, the above phenomenon is not observed with that scheme due to X-MAC's significantly higher energy consumption when the network is idle. Therefore, the amount of extra energy consumed per transmission is too small to be observed when compared to X-MAC's overall energy consumption. Furthermore, CADC and BEAM outperformed most configurations of their static CCRequivalents.

In order to further analyse energy consumption, a comparison between the energy efficiency of the compared protocols (energy over goodput) is presented in Figure 22. In this case, CADC and ContikiMAC significantly outperformed BEAM and X-MAC. Moreover, X-MAC's energy efficiency

decreased significantly with decreased inter-packet intervals. This is attributed to its higher packet loss. When we compare configurations of ContikiMAC and X-MAC that can achieve similar to CADC's and BEAM's goodput, BEAM consumed a quarter of the energy than X-MAC, while CADC consumed approximately half of what ContikiMAC did.

When compared with various configurations of ContikiMAC, X-MAC and BEAM, CADC achieved both the best energy efficiency and the best overall energy consumption.

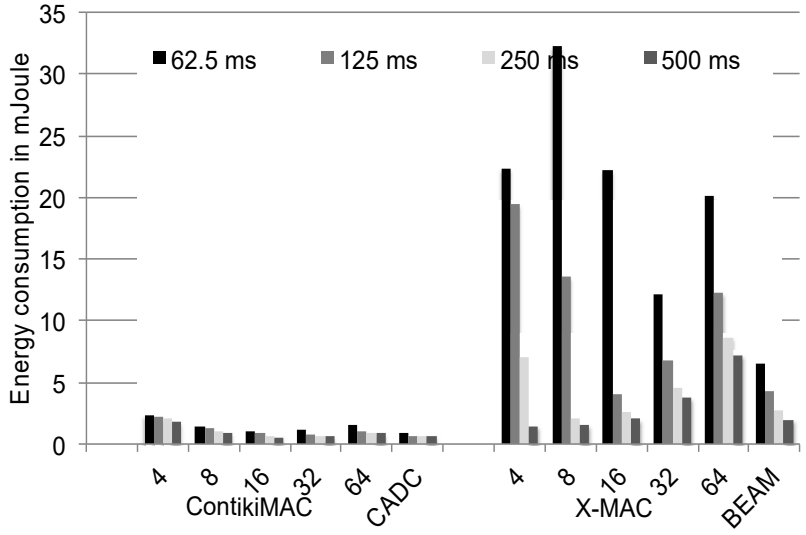


Figure 22: Energy consumption (per node) for each successfully received packet under different inter-packet intervals and CCRs (simulations).

3.3.3.6 Code Footprint and Memory Requirements

Table 18 and Table 19 show the memory requirements of the protocols under comparison when built for different hardware architectures. The numbers represent algorithm sizes at both MAC and RDC layers. X-MAC has the lowest memory requirements while CADC requires approximately 1 kB to 2 kB additional memory, with the exact difference depending on the target architecture and toolchain. This was anticipated, since CADC incorporates more functionality in order to precisely calculate the desired cycle at each given point in time.

Table 18: Algorithm footprints in bytes (MSP430 GCC toolchain).

	ContikiMAC / CSMA	X-MAC CSMA	BEAM	CADC
text	2378	2030	2544	3236
data	24	38	38	28
bss	366	250	274	420
dec	2768	2318	2856	3684

Table 19: Algorithm footprints in bytes (SDCC toolchain).

		ContikiMAC / CSMA	CADC
In Flash (bytes)	CODE	7576	9426
	CONST	58	58
In RAM (bytes)	XRAM	433	508
	DATA	0	0

3.3.4 Discussion

In this subsection, we focused on duty-cycled radio operation with congestion awareness. Congestion is likely to occur on event-based deployments, whereby changes to an observable parameter (e.g., ambient temperature) can lead to multiple RDs attempting to transmit data simultaneously. As a result, we may observe dynamic or even bursty traffic in the network. By tackling the congestion on a low-power wireless network, we decrease packet retransmissions, and we reduce the amount of time RDs have to spend awake before it can exit its congested state. The scheme proposed in this subsection is relevant to RERUM UC-O2 – Environmental Monitoring, whereby deployments will be formed by a large number of RDs.

The main objective of this research study was to study the trade-off scenario between the duty-cycling and 6LoWPAN performance. As shown in previous sections, increased CCR values significantly increase the performance of the network in terms of goodput, delay and packet loss. Furthermore, increased CCR does not necessarily lead to higher energy consumption when the network is active. On the contrary, increased duty cycles have significantly higher energy consumption when the network is idle. Assuming that a WSN is configured with a lower CCR, its lifespan can be extended significantly. Therefore, it is unrealistic to assume that a battery powered WSN will be configured with a high CCR.

Our results have shown that under different sensor network conditions (traffic load and distance), different CCR configurations achieved the optimal performance in terms of energy consumption, goodput, delay and packet loss. To be more precise, ContikiMAC's energy consumption per successfully received packet varied between the different packet transmission intervals and channel check configurations. We conclude that there is no single best CCR, and setting the correct value is a challenging task that should take into account network configuration, topology as well as anticipated traffic patterns. The task is even more complex in a multi-application 6LoWPAN deployment with arbitrary traffic patterns. This problem can be solved with the use of protocols such as CADC and BEAM which dynamically adjust duty cycling to adapt to changing network conditions. Simulation results have shown that the proposed CADC mechanism successfully copes with the different network parameters used during our simulation campaign. Furthermore, CADC achieved comparably high goodput, the lowest possible packet losses, the lowest energy consumption and very competitive delay times.

3.4 Trustworthiness Problems in sensing – a game theoretic approach

3.4.1 Introduction

Due to the extensive deployment of sensor-based technologies in many sensitive fields such as military and medical applications, their security is of topmost importance to their end-users. As with every other sensitive networking technology, RERUM, in order to maintain its security, requires a set of policies effectively implemented in an automated fashion to be in place. By avoiding human

intervention, faster and more objective safety-related decisions can be made, which could reduce errors, and therefore making it more difficult for a potential attacker to achieve a successful attack.

In this section, we present two game-theoretic models of Intrusion Detection and Intrusion Prevention System for Wireless Sensor Networks. The latter is an extension of the former one, with the additional feature that the defender has the ability to recover RDs that have been previously compromised.

To validate the theoretic findings, we run two sets of simulations. We employ Sensomax [HC2014, HC2013], an agent-based sensor network middleware, which supports executing multiple applications with regards to their operational paradigms. Subsequently, we evaluate the model in a simulated IPv6-based network, on top of COOJA simulator.

3.4.2 Relevance to RERUM's Use-Cases

RERUM had put a significant emphasis on improving security and network reliability of IoT systems. Due to IoT-based constraint devices (consequently RDs), and long unsupervised operations, the key challenges remain to be the development of lightweight methods that able to efficiently detect attacks under constrained computational resources.

The scheme proposed here is relevant to UC-O2 - Environmental monitoring, whereby deployments will be formed by a large number of devices. Moreover, due to the large number real-world deployments of IoT in many sensitive fields such as military and medical applications, their security is essential to their end-users.

3.4.3 Application of a Game Theoretic Approach in Smart Sensor Data Trustworthiness Problems

In this section, both models are presented, the ID model as "Modification Detection Model" and the IP under the name "Modification Correction Model".

3.4.3.1 Attacker Model

In our models, the attacker modifies compromised nodes in order to make them report erroneous values. We make the following assumptions about the deployment:

- All network traffic is encrypted.
- All sensor measurements are signed.
- The deployment's topology is not publically available. This is a reasonable assumption, since the logical network topology (e.g. routing topology or cluster membership) is created and maintained at runtime by an algorithm that relies on criteria which can change over time and which are not known a-priori, such as the quality of radio links.

For our attacker, we make the following assumptions:

- He is highly motivated
- He is external to the system
- He can actively initiate attacks against nodes. The firmware running on nodes is susceptible to a bug, and the attacker has discovered it. Therefore, an attack against a node is always successful.
- He has high availability of time, but not enough to break cryptography and signature schemes.
- Because of his inability to break signatures, the attacker can actively introduce neither his own traffic nor his own malicious nodes. Therefore, his only option is to compromise an existing, legitimate node.

- Because of his inability to break cryptography, the attacker can passively overhear traffic but cannot understand the contents of network packets. As such, he cannot synthesise the deployment's topology from passive eavesdropping.
- The attacker has high, but not unlimited financial resources. Therefore, he can choose to attack the entire network if he wishes to do so, but his criterion is to optimise the financial benefit of an attack.

Thus, the attacker can choose how many nodes he wishes to attack, but being oblivious about the network topology he has no way of identifying which nodes would maximise damage to the network.

3.4.3.2 Modification Detection Model

In this model a game between the defender who could be the security team responsible for the seamless operation of an IoT system that gathers data about a measured characteristic (e.g. temperature) of the area under monitoring and the attacker who randomly chooses which sensors to attack and tries to make the network transmit as much incorrect information as possible, is replicated.

According to this scenario, the defender needs to monitor a specific, predefined area. In order for this to be accomplished, sensors have to be deployed throughout the whole area of interest. The question that rises is what should be the *density* (i.e. number of sensors per area unit) that the defender must choose. Since the area under investigation is predefined, it is only the number of sensors that can affect the density. Hence, the number of sensors is part of the strategy of the defender and throughout the game the player should try to find the most beneficial value from within a set of realistic choices.

In addition, there is a significance coefficient for every sensor. This coefficient is proportional to the level of trust that is related to the information transmitted by this particular sensor and echoes the probability that the measurements provided by the sensor are indeed true. The reasons why this coefficient differs from sensor to sensor varies from the kind of measurements that are taken to structural features of the sensing elements.

Apart from this parameter, *tolerance* is also part of the defender's strategy and it is a property of the whole network. Having defined compromised / uncompromised / total information as the sum of significance coefficients of compromised / uncompromised / all sensors respectively, tolerance denotes the minimum portion of the total information that the compromised information should be, in order for the latter to be believed by the defender. In other words, it denotes the minimum value that the fraction (1) can have in order for the incorrect information that has been injected in the network to be treated as correct. We call this fraction *Attack Coefficient* (AC):

$$AC = \frac{compromised\ information}{total\ information}$$

At this point and before continuing it is essential that some basic assumptions of the model are presented.

Assumptions

- Players are rational.
- Full area coverage is desired.
- Two sensors of the same network with identical specifications, operating under identical conditions can still report slightly different values.
- A compromised sensor cannot affect the information that other sensors transmit.
- The attacker's goal is to make a sensor transmit faulty data that demonstrate noteworthy deviation from the data that uncompromised sensors transmit. Otherwise the attack is pointless.

- Compromised network is the network into which the injected faulty information is believed by the defender.

Under those assumptions, the network operators that handle the data generated from all sensors try to only take into account the non-compromised data. Therefore, if fraction (1) is greater than tolerance then the incorrect information is considered to be accurate, correct data is disposed of and the attempt for compromise of the network is considered successful, with maximisation of attacker's payoff. Otherwise, the network is not considered compromised and that would imply a lower payoff for the attacker. Intuitively, tolerance should only be a value greater than 0.5 and of course less or equal to 1 as it denotes a percentage. In this way, the weighted information that will be ultimately 'believed' will correspond to at least the half of the total weight, which is realistic.

The attacker can only affect the number of sensors to attack considering that every attacks bears a cost. Therefore, the optimal strategies are not obvious and a game theoretic approach would be suitable to adopt.

The payoff function with the help of which, a payoff matrix will be filled, is function (2). As expected it is affected by the aforementioned parameters.

$$Attacker's\ Payoff =$$
$$\left(\frac{incorrect\ sum}{total\ sum} \geq tolerance\right) * reward\ for\ compromising\ the\ network$$
$$+sensor * cost\ per\ sensor - attacks * cost\ per\ attack \quad (1)$$
$$+tolerance * tolerance\ cost$$

where,

$$\left(\frac{incorrect\ sum}{total\ sum} \geq tolerance\right) = \begin{cases} 1, & if\ inequality\ holds \\ 0, & if\ inequality\ does\ not\ hold \end{cases} \quad (2)$$

Since it is a zero-sum game, the attacker takes advantage of the defender's expenses. This is why everything that has a cost for the defender, like the total cost of sensors, counts in favor of the attacker in formula (2). According to the definition of the compromised network given earlier, if:

$$\left(\frac{incorrect\ sum}{total\ sum} \geq tolerance\right) = 1$$

then the attack is considered successful and the corresponding reward is given to the attacker. The necessity of *tolerance cost* lies in the fact that the greater the tolerance is, the greater part of the whole information, should be faulty in order for it to be "believed". That motivates the attacker for a more comprehensive attack and therefore for a less possible recovery by the operators or engineers of the network. Through under this perspective, it could be preferable for the network to suffer a mild assault that will compromise the network temporarily, than risk suffering a massive one that will make it totally useless or unaffordable to be fixed. It should be noted that the payoff function has no units of measurement. It is just a necessary quantification of the advantage derived for each player due to the actions taken so that the problem can be solved and resembles the role of a utility function.

It is worth noting that although defender is not aware of which piece of information is compromised or uncompromised, it is still possible to use the outcome of formula (2). In other words, although defender cannot distinguish between correct and faulty data, they are aware of the payoff that he receives when both defender and attacker choose specific strategies.

In addition, there is a chance that the following inequality holds:

$$\frac{correct\ sum}{total\ sum} < \frac{incorrect\ sum}{total\ sum} < tolerance$$

In this case, the compromised information will not be believed although it is greater portion of the total information than the correct information and therefore no reward for compromised network is given to the attacker. This is only possible for tolerance greater than 0.5 (50%).

Since every strategy of the defender consists of a pair (*m*, *n*) where *m* is the number of sensors used and *n* is the tolerance employed, we have two-dimensional strategy sets. One way for this to be tackled and thus for the optimal strategies to be found, is the procedure we outline here. This procedure has been previously applied in other domains [SKTO2013] [WSRED2010]. The algorithm is described by the following piece of pseudo-code and Figure 23, in which green denotes defender's strategies and orange the ones of attacker. Sensor weights are the aforementioned significance coefficients which can shape defender's strategy but their value cannot be affected by the defender and therefore it is in grey colour.

```
for all examined number of sensors
  *set all the significance
  coefficients equal to 1
  *given the sets of strategies for
  number of attacks and tolerance
  level find Nash Equilibrium (most
  beneficial strategies) and the
  defender's rewards they lead to
end for
consider as Nash Equilibrium of the
whole scenario the strategies that
lead to the least reward for the
attacker out of all that were found
from all those sub-games
```

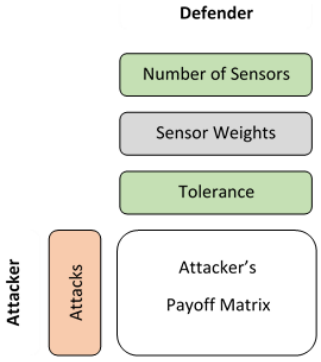


Figure 23: Schematic description of the ID model**.**

3.4.3.3 Modification Correction Model

In the use case for this model there is an attacker who attacks sensors and a defender that protects them, but there are two major differences from the detection instance. Firstly, the defender in this game knows which sensors are attacked, as opposed to not even knowing that there were sensors under attack, trying to figure that out by the collected data. Secondly, the game now is repeated for many rounds. However, all decisions are made at the beginning and remain unchanged for the whole game which makes the game static, although in a repeated form. The goal of attacker is once more to compromise the network with the least possible cost while defender's is to keep the network uncompromised with the least possible cost.

The payoff function this time is:

$$\begin{aligned} Attacker'sPayoff \\ &= total\ number\ of\ attacks * (reward\ for\ compromised\ sensor - attack\ cost) \\ &+ total\ recoveries \\ &* (recovery\ cost\ per\ sensor - reward\ for\ compromised\ sensor) \\ &+ number\ of\ sensors \end{aligned}$$

$$* sensor\ cost + \left(\frac{compromised\ sensors\ at\ the\ end}{total\ number\ of\ sensors} > tolerance\right) \qquad (3)$$

$$* reward\ for\ compromised\ network$$

where,

$total\ number\ of\ attacks = \sum_{i=1}^{n} attacks\ at\ round\ i\,,\ for\ n\ number\ of\ rounds$

$total\ recoveries = \sum_{i=1}^{n} recoveries\ at\ round\ i\,,\ for\ n\ number\ of\ rounds$

As usual,

$$\left(\frac{compromised\ sensors\ at\ the\ end}{total\ number\ of\ sensors} > tolerance\right) = \begin{cases} 1, & if\ inequality\ holds \\ 0, & if\ inequality\ does\ not\ hold \end{cases}$$

Additional assumptions in this model are the following:

$$reward\ for\ compromised\ sensor < attack\ cost \quad (4)$$

$$recovery\ cost\ per\ sensor > reward\ for\ compromised\ sensor \quad (5)$$

because if (5) does not hold then the attacker would not seek the additional reward for compromising the network and would attack as many sensors as possible and similarly if (6) does not hold then the defender could overspend his resources by protecting more sensors than necessary. Of course, since it is a zero-sum game again, everything that has a cost for the defender rewards the attacker. For example, an attack has a cost for the attacker and as a result a compromised sensor occurs which rewards the attacker since it is opposed to defender's interest. Similarly to the previous model, if:

$$\left(\frac{compromised\ sensors\ at\ the\ end}{total\ number\ of\ sensors} > tolerance\right) = 1$$

then the network is considered compromised and the corresponding reward is given to the attacker. The pseudo-code for this model is:

```
for all examined number of sensors
  for all examined distributions
    for every examined mean value generate 5 numbers that follow the
    current distribution with this mean value. Let those 5 numbers be
    possible numbers of attacks and let attacker's strategies set be the
    set of those 5 values. For every pair of values (attacks,
    recoveries) set the corresponding element of payoff sub-matrix equal
    to Payoff from the function previously given and find Nash
    Equilibrium of every matrix (Figure 24)
    end for
  end for
end for
```

Consider as Nash Equilibrium of the whole scenario the strategies that lead to the least reward for the attacker out of all rewards that correspond to all Nash Equilibria that were previously found for all the matrices.

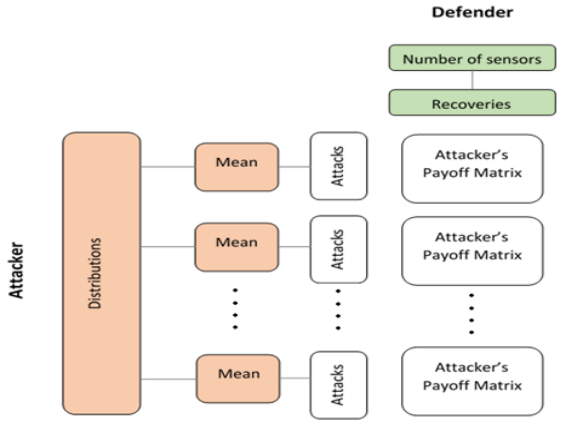


Figure 24: Schematic description of the IP model.

The schematic representation of the model is as follows in Figure 24, according to which, attacker's strategies are defined by the distribution that the number of attacks follow and their mean while defender's strategies are defined by the number of sensors adopted and the maximum number of recoveries performed in each round which is the same for all rounds. At any round of the game, the attacker can only make as many attacks as the uncompromised sensors in the network and the defender can only make as many recoveries as the compromised sensors in the network.

3.4.4 Performance Evaluation

In this section we present and discuss the outcomes of the theoretical models. Results are visualized as a triple graph as explained below. The distributions used to describe attacker's behaviour throughout this work and also significance coefficients, are distributions that are commonly used to describe various elements of network activity [H2005], [C2009], [BRS2011].

As far as the ID model is concerned, the sample values that were used for formula (2) are: Number of sensors: [500, 600]; Tolerance: [0.55, 0.9]; Number of attacks: [500, 600]; Significance coefficients: All equal to 1, following Uniform(1,4), Normal(2.5, 0.25); Reward for compromising the network = 10; Cost per attack = 1.2; Cost per sensor = 2.3; Tolerance cost = 10. Conclusions can be extracted by Figure 25. In order to interpret the figure, bear in mind that we help defender to take the best possible decision regarding the maximization of their payoff. In Figure 25, we can see the Nash Equilibria of all the sub-games that occurred. The horizontal axis in all sub-graphs of a figure is the number of Sensors. A Nash Equilibrium can be seen, as a vertical line that goes through all three sub-graphs. If (x, y_1), (x, y_2) and (x, y_3) are the points that this line cuts the blue lines of sub-graphs 1, 2 and 3 (counting from the upper to the lower one) respectively, that would mean that when the attacker decides to make y_2 attacks, the best response by the attacker would be to employ x sensors and tolerance equal to y_3. That strategy would lead to a payoff for the attacker equal to y_1. No matter what x, y_1, y_2 and y_3 can be, the vector (x, y_1) represents the best strategy that the defender can adopt in order to oppose to attacker's y_2 strategy and vice versa. In other words, if the defender decides to employ x sensors with tolerance equal to y_3, then the strategy that would lead the attacker to the greatest individual payoff (which is y_1 in this case) is a number of attacks equal to y_2. Any other number of attacks would lead to a payoff less than y_1. On the other hand, if the attacker decides to perform y_2 attacks, the defender's strategy that would lead the attacker to the least possible individual payoff (which is of course equal to y_1 again) is the employment of x sensors and tolerance equal to y_3. For every fixed price of Sensors, we firstly plot the corresponding Value price of the first sub-graph. This is the attacker's payoff, if both players choose the strategies that form the Nash Equilibrium of this sub-game. Since this is the attacker's payoff and the number of sensors is chosen by defender, the optimal number of Sensors is the one that will lead to the smallest value of the first sub-graph. This number is 511 sensors here, which will lead to 703.3 reward for the attacker (the smallest achievable as seen in the same sub-graph).

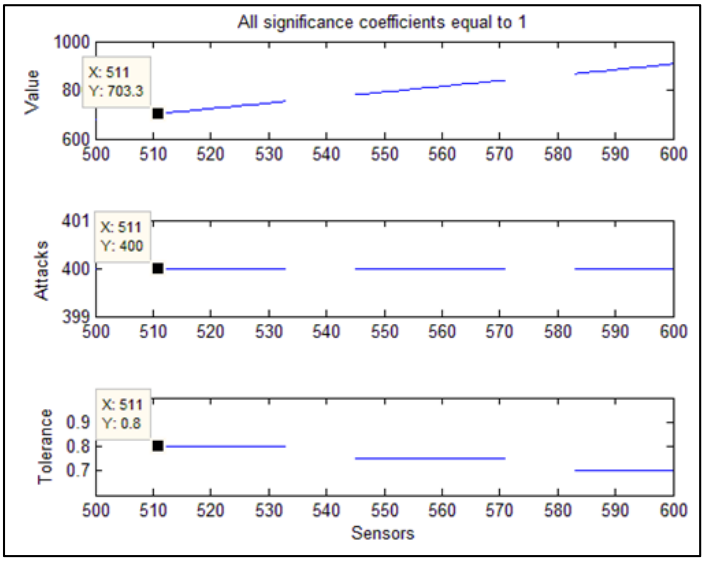


Figure 25: Attacker's Payoff (Value), Number of Attacks and Tolerance for the NE that occurs for every different num. of Sensors when all significance coefficients are equal to 1.

Given that the defender chooses to utilize 511 sensors, there is now a game that only the optimal Tolerance and the optimal number of Attacks remain to be found. These are calculated by game theoretic methods and are found to be 400 Attacks and 0.8 Tolerance. We can see that the graph does not consist of a continuous line because there is not a Nash Equilibrium for every price of Sensors. The gaps are observed close to the values that correspond to combinations of strategies that would make the following equality to hold:

$$\frac{incorrect\ sum}{total\ sum} = tolerance$$

Graphs for the two remaining scenarios of Significance Coefficients distribution, Uniform (1,4) and Normal (2.5, 0.25), are not demonstrated because they were found to be very similar.

For the IP model, results from which are displayed in Figure 26, values for the involved parameters in formula (4) are: Number of sensors: [200, 400]; Number of recoveries: [1, 70]; Distribution of number of attacks: Normal, Poisson, Exponential; Reward for compromised sensor = 1.5; Attack cost = 3; Mean values created per distribution = 5; Recovery cost per sensor = 5; Cost per sensor = 4; Reward for compromised network = 2000; Tolerance = 0.5; Number of attacks: [10, 120].

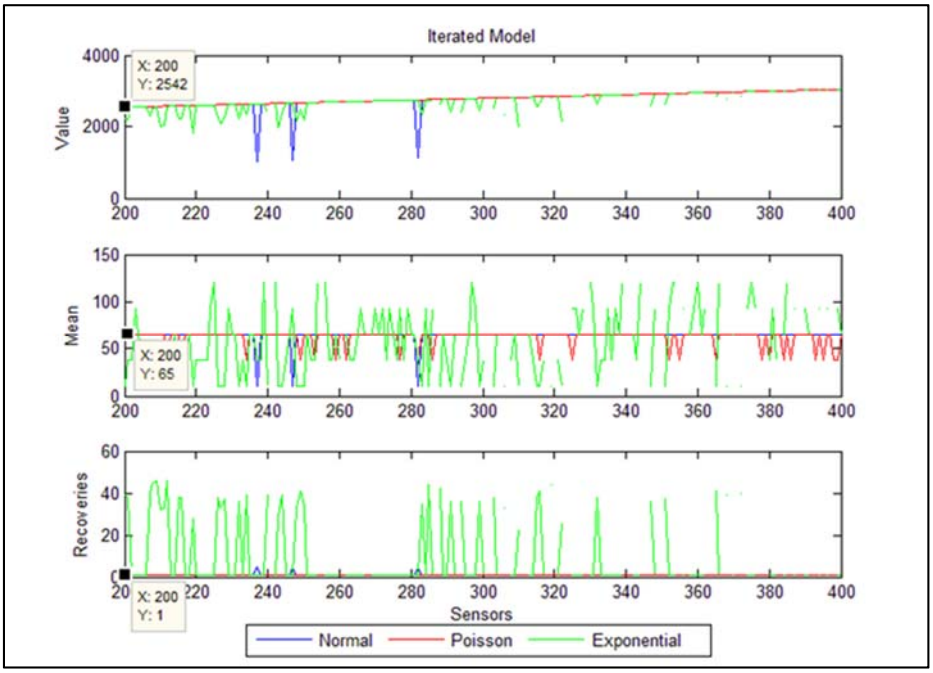


Figure 26: Attacker's Payoff (Value), Mean values and Number of Recoveries of the Nash Equilibria found in the Iterated model.

In this version of the game defender chooses firstly the optimal number of Sensors. This time there are many distributions available so, when the defender chooses a figure for Sensors that will correspond to three ones for Value, one for every available distribution (first sub-graph). As the attacker will later choose distribution in an attempt to maximize the Value, he/she will choose the one that corresponds to the maximum among the three aforementioned Values. Thus, the best practice for defender at this stage is to choose the number of Sensors that corresponds to a triplet of Values such that their maximum is the least among all the maximums of Value for all the possible triplets (i.e. the triplets that occur for all possible prices of Sensors). In formal terms, if on the horizontal axis for Sensors = t we have Normal(t), Poisson(t) and Exponential(t) being the corresponding Values of the first sub-graph, defender should choose t, such that:

$$max\{Normal(t), Poisson(t), Exponential(t)\} = \min\{\max\{Normal(s), Poisson(s), Exponential(s)\}\}, \forall\ possible\ s$$

In this formula, t and s are not parameters of the distributions but the number of sensors and Distribution(t) or Distribution(s) denote the spots on the corresponding sub-graph for the given values of sensors. In this illustration t=200 sensors, which corresponds to Poisson distribution (Figure 26). Using the same rationale as the IP game instance, for Sensors = 200, the parameters that remain to be determined are the Mean of the Poisson distribution and the number of Recoveries. For fixed number of Sensors (200) and Distribution (Poisson), the Mean can be no other than 65 (second sub-graph) and the Recoveries can be no other than 1 (third sub-graph).

After demonstrating the method according to which the optimal strategies for both players were found, it can be verified that the Nash Equilibria described in both models follow the definition of Nash Equilibria, which is that none of the players would be tempted to unilaterally deviate from their strategy because that would lead to a worse reward for the player that made the change.

3.4.5 Validation in a Cluster-based Deployment

In this section we conduct a number of experiments to validate the IP and ID models. We exploit the clustering facilities offered by Sensomax, in order to increase network scalability, reduce packet loss and increase node mobility.

The energy consumption for RERUM is a crucial factor to consider, whilst designing their security protocols. The proposed models take advantage of relatively high-level mathematical operations in order to detect and prevent attacks. For a security model to be realistic in this context however, it needs to be shown that in practice it is feasible to implement within constrains imposed by implementation technologies and platforms. Therefore, in addition to evaluating model effectiveness against attacks, we also examine how their adoption contributes to energy expenditure on a per-node basis as well as network wide. In the final phase of our experiment, we will measure and report the potential latency in the data transmission, which could be imposed by the utilization of these models in middleware.

In all our experiments, both were programmed as two separate applications in every sensor node. Those two applications can be executed concurrently in order to detect and prevent attacks, whilst sensor nodes are carrying out their normal operation and meeting the requirements of their given task. The application itself resides in a single node, known as the cluster-head, where all the top-level executions happen. The IDM and IPM applications (i.e. model logic) are present in every sensor node, whilst being executed only in the cluster-heads.

For the first phase of our experiment a network of 600 virtual nodes (hereafter referred to as the "Defender") was created in SensomaX Companion Simulator (SXCS) [H2013], incorporating 30 clusters, each containing 20 nodes. All nodes were programmed to constantly report Temperature readings at 1-second intervals. A second network (hereafter referred to as the 'Attacker'), containing 600 nodes without any clustering mechanism was also created to report false temperature readings. Each experiment reported in this section was repeated 100 times to gain the average values. Figure 27 shows the average number of attacks required before detection.

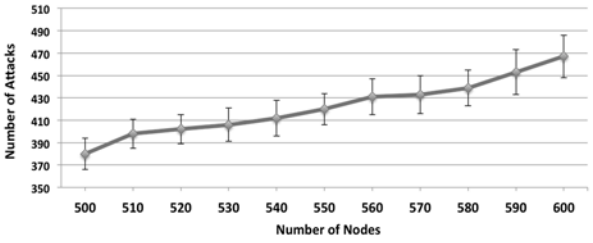


Figure 27: IDM's required number of nodes vs. number of attacks.

For a 510-node network, the average number of attacks is 398. This result is on par with the results reported in Figure 25, Channel Check Rate (the standard deviation, which covers the 400 attacks reported earlier. Figure 28 shows the number of nodes required for the IPM to operate successfully based on a variable number of attacks. The results reported in this figure are also relatively on par with the results reported in Figure 28, given the standard deviation around the mean values.

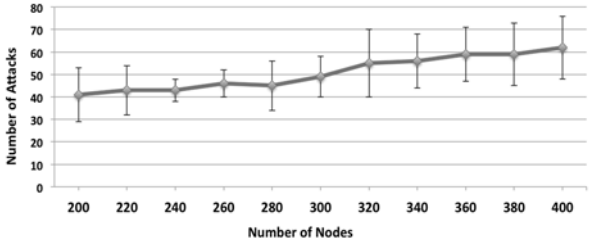


Figure 28: IPM's required number of nodes vs. number of attacks.

In the next phase we will investigate the extra energy consumption required by the two models in comparison with the normal operation of the network without them. The results reported in this

section have been recorded whilst running the above-mentioned experiments. The red line in Figure 29 shows the average increase in the energy consumption of each node running the IPM, whereas the blue line represents the average energy increase required for running the IDM, with regards to the number of nodes involved in the network.

As Figure 29 shows the increase in the energy consumption for each model is less than 10%, which is a reasonable and acceptable result, given the intensive operation involved in game theoretic computation.

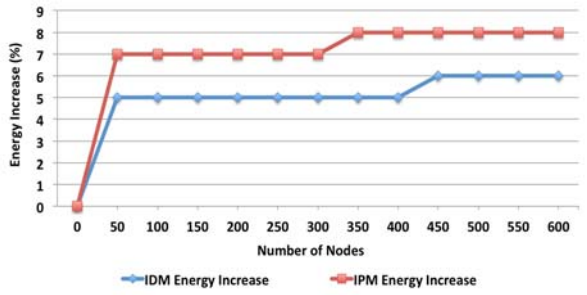


Figure 29: Node's energy increase by the utilization of IDM and IPM.

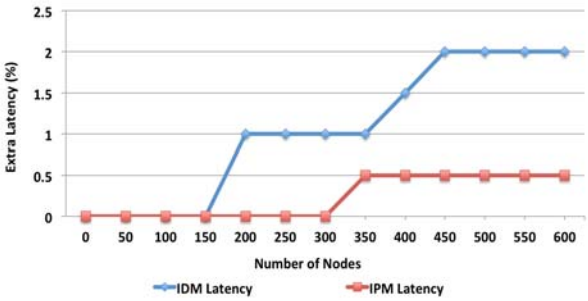


Figure 30: Latencies induced on the data transmission by IPM and IDM.

Lastly, Figure 30 represents the latency imposed by the IDM (blue line) and IPM (red line) on the data transmission compared with the normal operation of the network without the two models. Both models induce less than 2% latency, which is negligible given the number of nodes involved in the network interaction.

3.4.5.1 Validation in an IPv6-based Deployment

We use Cooja [O2006], the network simulator distributed with the Contiki Operating System for the Internet of Things. Within COOJA, we simulate an IPv6-based wireless sensor network. Network nodes use 6LoWPAN [MKHC2007] and the RPL [W++2012]. We simulate a network with 1 traffic sink and 40 traffic sources, distributed in a 200x200 grid. Node distribution is entirely random, with the only limitation being that all sources must have a network path to the sink. We choose to simulate a network of 40 nodes in order to achieve full area coverage, as is the assumption in the model. We use 10 different random topologies and for each topology we repeat the experiment 10 times using a new random seed for each iteration. A sample topology is illustrated in Figure 31. The green circle corresponds to node 12's communication range, whereas nodes in the grey area are subject to RF interference due to node 12's transmissions, but the received signal strength is not sufficient for correct frame reception.

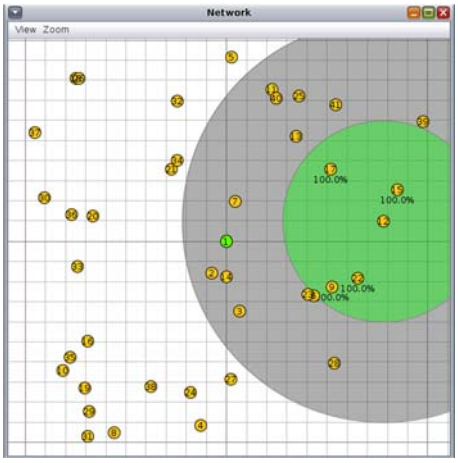


Figure 31: Sample simulated topology within Cooja.

In the remainder of the section, we use the following notation:

- n is the index of a node
- $N = \{n: n \in \mathbb{Z}^+ \wedge n \leq 40\}$ a set containing all network nodes
- $C = \{n: n \in N \wedge node\ n\ is\ compromised\}$ the compromised nodes set
- T is the defender's chosen tolerance
- $D_n: n \in N$ is the degree of node n, discussed below
- $S_n: n \in N$ is the significance of node n, also discussed below

In the model, the choice of node significance is based on a random distribution. In our simulations we model node significance as a function of network density. We first calculate the node degree D_n for each network device, which is calculated as the number of other network nodes within communication range (green area in Figure 31). The significance S_n for node n is subsequently calculated as follows:

$$S_n = \frac{\max(\{D_i: i \in N\})}{D_n}$$

Thus, S_n corresponds to the maximum node degree observed in the network, divided by the node's own degree. As discussed above, all nodes in the network have a path to the sink, therefore they have at least one other node within communication range. Thus, $D_n > 0$ and therefore the significance calculation's denominator is always non-zero. With this calculation, nodes in dense areas will have lower significance, while nodes in sparse areas will have a high one. The reasoning behind this calculation is that the network is used to gather sensory information about an environmental parameter in a geographical region. Even when two identical devices measure the same parameter, measurements are likely to be slightly different due to sensing element manufacturing inaccuracies and due to slight fluctuations of environmental parameters even within the same geographical area. Therefore, in an area where multiple nodes are reporting, each node's measurement will be of lower significance, whereas in a sparse area where only a few nodes are reporting, each nodes measurement will bear more weight. This is captured by our calculation.

The model has found that the optimal attacker strategy is to compromise approximately 78.27% of the total number of nodes in the network (400 out of 511). With this in mind, in each experiment the attacker compromises a random set of 31 nodes ($|C| = 31$). Furthermore, the model has found that the defender's optimal strategy is to select tolerance level $T = 0.85$.

An attack is successful if the defender believes the erroneous value to be accurate and this is only try if the Attack's Coefficient AC (section 3.4.3.2) is greater than the defender's tolerance T:

$$AC = \frac{\sum_{j \in C} S_j}{\sum_{i \in N} S_i} > T$$

3.4.5.1.1 Results and Analysis

Figure 32 illustrates the densities of the ten network deployments under investigation in our simulations. For all deployments, the minimum node degree D_n was between 1 and 3, whereas maximum node degree was between 7 (topology 1) and 13 (topologies 3 and 5).

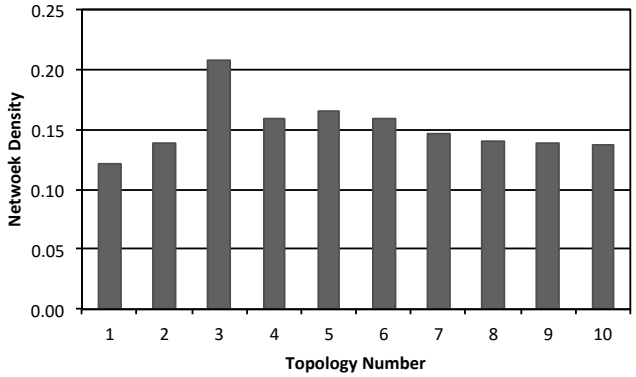


Figure 32: Topology densities.

Figure 33 illustrates attack coefficients for each iteration. Across the entire experiment set of, the attacker was successful only three times. For all other iterations detection was successful. The three successful attacks were observed in topologies 3 and 5, which were the ones with the highest network density. This suggests there may be a correlation between the model's effectiveness and the network density. We shall investigate this further as part of our future work.

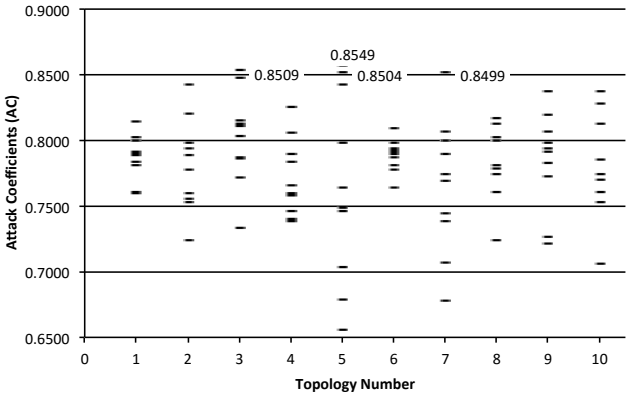


Figure 33: Attack Coefficients per experiment.

3.4.6 Discussion

RERUM had put significant emphasis on improving the security and network reliability of IoT systems. Due to the severe resource constraints in IoT-based devices, and consequently with RDs, and long unsupervised operations, the key challenges remain to be the development of lightweight methods that able to efficiently detect attacks under constrained computational resources. In this section, we have shown how Game Theory can be used as a useful approach in order to detect and prevent intrusions in RERUM. More specifically, two game-theoretic models were presented, namely, Intrusion Detection and Intrusion Prevention System for IoT networks. The scheme proposed in this subsection

is relevant to RERUM UC-O2 – Environmental Monitoring, whereby deployments will be formed by a large number of RDs.

We have demonstrated the effectiveness of the detection and prevention models, by two methods of validation. Firstly, with the help of Sensomax the adoption of which not only had insignificant effect on the network's overall power consumption and data transmission latency but also its results matched the ones of the analytical models. Secondly, by using the COOJA simulator for networks of embedded objects, we investigated the effectiveness of the detection model in an IPv6-connected network of smart RDs.

We studied the trade-off scenario between network performance and its density (i.e., scalability). Our performance evaluation results show that higher the network density, more successful attacks (i.e., three) were observed in topologies, which implies a potential correlation between the model's effectiveness and the network density.

3.5 Trade-offs in sensor's space

With part of the work described in section 2 we will be able to take some conclusions about the trade-offs that will affect to the sensors.

For the moment we have identified two of them that are listed below and will be detailed in this deliverable. Other ones might appear while finishing the section 2 and will be added later on.

3.5.1 Sensor cost vs. data reliability

In the following table we'll find a review of the cost and a cross analysis with the reliability of the data for each sensor, using colours to reinforce the cheaper (light green) and more expensive (light right) prices and the less (light green) and most (light red) reliable gathered data:

Table 20: Sensor cost vs. reliability comparison

Measurement type	Sensor	Unit Price	Data reliability analysis
Current	i-Snail-VC-50	40,-€	The accuracy is pretty good, but the extended range we are expecting for the measurement, provoke a poor resolution that compromise the ratio cost/reliability. Other options might be studied.
Voltage	CE-VJ03-32MS2	115,-€	Once again, the extended range for an analogue sensor, decrease the resolution. The accuracy is pretty poor. But in any case, the reliability is enough for the type of measure it will perform.
Light	TLS2563	2,11€	Digital and reliable sensor. The only recommendation is to properly adjust the range of measurement in order to maximize the precision for the measured environment.
Temp & Humidity	SHT25	12,36€	Good accuracy, right range for the RERUM UCs, great precision due to the calibration in fab, and excellent

			stability. We can only complain on the guaranteed lifetime and a bit of the price, pretty higher than similar, but less reliable, options.
Barometric	BMP085	3,36€	Very good trade-off between accuracy, resolution and range, with good precision and measurement stability and long lifetime; it's undoubtedly a "best value" deal.
Noise	ZNK14XXX	35,56€	Resolution and range fit on the UCs demands. However, the accuracy is almost the minimum acceptable one and it needs to be re-calibrated often if we want to keep that, incurring in big maintenance costs.
Weather meter	Rain Gauge Anemometer Wind Vane	76,95€	Appropriate range for the measurements but very poor resolution in a not good enough tested and documented sensor kit: this kit has been created for makers and technology enthusiasts, not for professional use.
PM_{10}	GP2Y1010AU	12,50€	There is not a lot of documentation about this sensor which indicates that is not for professional use. Nonetheless, the range and the resolution really fits on the requirements. Physically looks like the sensor require maintenance (namely cleaning) to keep its accuracy and precision.
SO_2	4-SO2-20	155,-€	Excellent features for this set of gas sensors from the same electrochemical technology, all of them are very accurate and with excellent resolution, theoretical stability and precision. The technology they are based on require a 2-years replacement, an acceptable time.
NO	4-NO-250	150,-€	
NO_2	4-NO2-20	130,-€	
O_3	OX-A421	200,-€	
VOC	IAQ-CORE	37,80€	Very good option for this type of measurement with this digital sensors that provides the right range in a very good resolution and really stable sensor, also with extended lifetime.

After this analysis, we can conclude that within the project different trade-off options can be found in terms of cost vs. measurement reliability:

- We have options with a very good trade-off, cheap and reliable, that can provide a lot of scalability on the system: they can provide a good quality on the measurement, allow to deploy a

big amount of devices, and just have a small impact on the cost. This is the case, for instance, of the light, the temperature and humidity and the pressure sensors.

- We can find very reliable sensors that have a medium or large cost, suitable to perform a good measure but not scalable to be installed in large amount due to its impact on the budget. This is the case, for instance, of the gas sensors.
- We have cheap sensors that, however, could not be deployed in large scale because their accuracy is poor, unless a big amount of them are used in a join way, with a process that unifies their measurement to make them more reliable. This is the case, for instance, for the PM_{10} sensor.
- We found sensors which are pretty expensive and with a poor performance, like the noise one, that can't scale from both points of view and the recommendation is to find some way to improve at least one of this vectors or replace them.
- And we have also sensors that are not good for anything, neither cost or reliability, turning on a low scalability in this sense. The recommendation in that case for future modifications of this system is to replace for better (and likely more expensive) more reliable alternatives, to really make the system scalable. This is the case, for instance, for the weather sensors.

In terms of scalability, the results can be placed onto the following matrix:

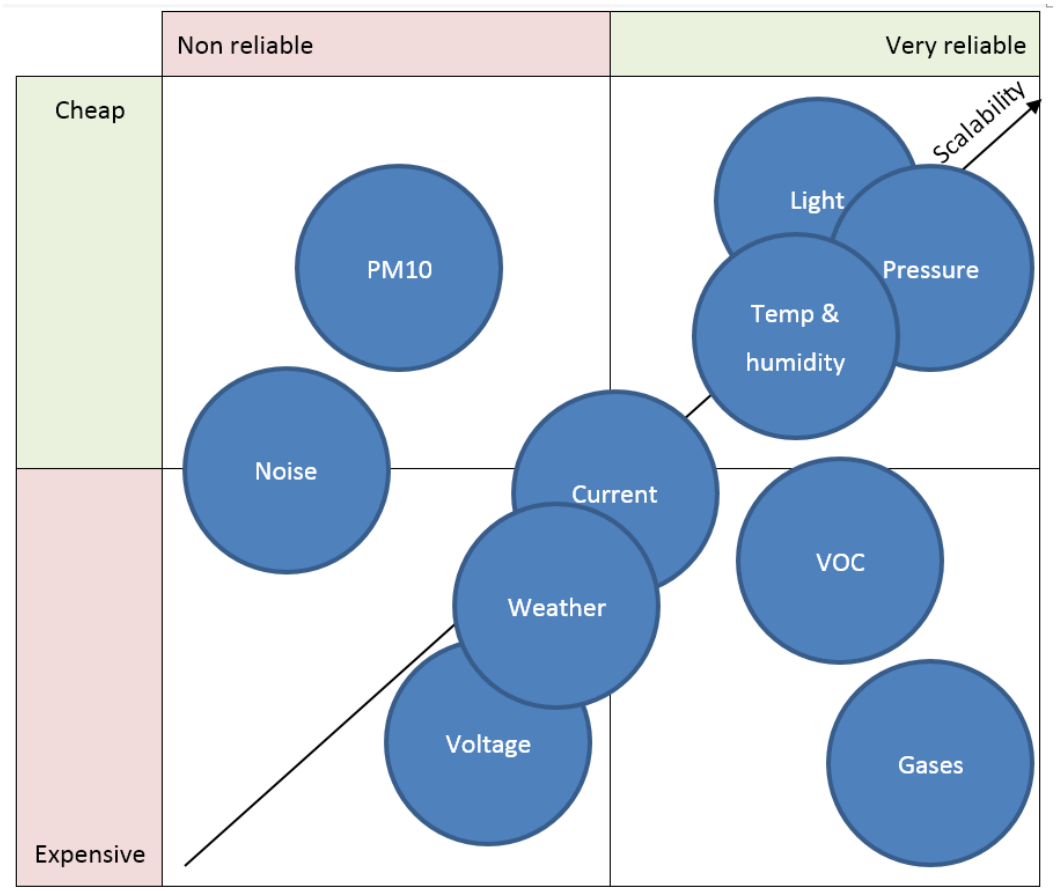


Figure 34: Cost vs. measurement reliability design analysis

3.5.2 Measurement sampling rate vs. transmitted data

Another important trade-off for the measurements of a system pretending to scale in the type of scenarios that are contemplated within the RERUM project, is the relation between the amount of data that is being transmitted in a limited spectrum wireless system and the amount of data needed for a proper processing of the data gathered.

That amount depends a lot on the type of measurement we are performing and the quality we expect from that measurement. So let's have a deep look to the natural frequencies of the performed measures and the expected resolution for them, and extract some conclusions about this:

As seen in Figure 34, most of the measurements neither require real time communication nor do they consume a lot of bandwidth.

For instance, on a sort of worse case where there is a channel with an effective bandwidth of just 1200kbps, that will need to support devices sensing weather information (4B) + Temperature & humidity (3B) + Light (2B) + Noise (2B), every minute, would mean that every channel will support up to **818** different devices.

Therefore, in its worst case, a network with a cell topology (without a mesh) and only using one frequency channel, clearly demonstrates the scalability of the system. It allows the system to grow in different dimensions, such as in the number of devices connected, in the number of sensors per device and in the sampling rate, making also the whole system more robust. For instance, this over dimension of the system might enable to create an adaptive system which, depending on the context and in certain situations, such as an emergency, can increase temporally the data rate to get more accurate or reliable measurements.

Table 21: Measurement rate cost vs. reliability comparison

Measurement type	Recommended sampling rate	Resolution | Min. payload (in bytes [B])	Comments
Current	1 minute averages	12,2mA | 2B	Can change every second but, for the applications, average consumption is enough
Voltage	1 minute averages	61mV | 2B	
Light	1 minute averages	1lux | 2B	The averages are used to avoid false information like a shadow or a cloud passing in front of the sun
Temp. & Humidity	1 minute indoor; 5 minute outdoor	±0,1°C | 1B ±0,1% RH | 2B	The air smooths the change of the temperature according to its mass. In a closed environment the change can be faster. No need for averages because the changes are smoothed by nature
Pressure	5 minutes	0,03hPa | 2B	Measurement very affected by the atmosphere, slowly changing during the day
Noise	1 sec. integration windows; 1min., 15min. and 1 hour averages	0,12dB_A | 2B	Are regulated by EC, requesting averages minimum every hour. Nonetheless, might be interesting to monitor more

			often to detect short periods of noise.
Weather	1 min. averages and peaks	Rain: 0.2794mm/pulse | 1B Wind speed and gusts: 2,4km/h/pulse | 2B Wind direction: 22,5° | 1B	Rain and wind need constantly to be measured to make averages. In general, averages are enough excepting for the wind gust that should be measured as peaks during a certain period.
PM_{10}	1 hour averages	0,1mg/m^3 | 2B	Partially regulated by EC rules, requesting averages minimum every hour.
SO_2	1 hour averages	0,014ppm | 2B	
NO	1 hour averages	0,013ppm | 2B	
NO_2	1 hour averages	0,013ppm | 2B	
O_3	1 hour averages	0,015ppm | 2B	
VOC	1 min. indoor; 5 min. outdoor	CO_2: 1ppm | 1B TVOC: 1ppb | 1B	The air dissolves and distributes the VOC across the air. In a closed environment the change can be faster. No need for averages because the changes are smoothed by nature

3.6 Compressive sensing encryption

3.6.1 Introduction

Compressive sensing (CS) theory has received much attention in recent years, both as a compression and an encryption mechanism. Leveraging the fact that a natural signal x is sparse or compressible, it enables its faithful recovery from a small set of linear and non-adaptive measurements: $y=\Phi x$. Typically, the measurement matrix Φ is considered as the encryption key, describing the subspace on which the initial signal is projected, and this is usually generated from a shared secret between a transmitter and a receiver. We have presented a thorough analysis of the CS encryption/compression characteristics in RERUM Deliverable D3.2 and RERUM Deliverable D3.3.

In this section, we present results regarding the security strength against three common attacks. Furthermore, we measure the energy consumption and the execution time required to run the CS encryption in a RD.

3.6.2 Relation to RERUM UCs

The CS mechanism is part of the Data Encrypter/Decrypter security component of the RERUM architecture. More specifically, the CS encryption/compression takes place in the RD, while the corresponding decryption/decompression is performed by the RERUM GW. This mechanism will be mainly used in all RERUM use-cases apart from the smart transportation. The main advantage of the CS mechanism that was deemed very important for RERUM's deployments is that it can perform simultaneously encryption and compression on the data that are to be transmitted from a RERUM Device (RD). This is achieved on the time domain, which means that each RD gathers some measurements and since it is not required to transmit them immediately, it stores them and after a block of measurements it compresses and encrypts them and then transmits much less measurements.

Then, at the receiver (RERUM Gateway) the measurements are decrypted and a very good estimation of the initial measurements is being computed.

CS is very important for the RDs because it is very lightweight and can run without issues to constrained devices. The accuracy error at the decryption at the RERUM Gateway depends on many factors, but mainly on the smoothness of the changes of the signal or the "sparsity" of the measurement signal on some domain (i.e. FFT). Our experiments have shown that for smooth signals (signals that do not change abruptly, i.e. temperature, humidity, gases, voltage, etc.) the error at the receiver can be very low even if the compression is very high. Other signals that can show abrupt changes (i.e. noise, light, current, rain, wind) can also be compressed and encrypted with the CS technique, either using a different domain or by compressing them less. To address this problem, in RERUM Deliverable D4.2 we presented the Adaptive CS framework that can adapt the compression according to the sparseness of the signal. In any case, the number of packets that will be transmitted will be much less than without CS due to the inherent compression capabilities.

Thus, using the CS encryption technique, the RERUM Devices can save a large amount of energy and maximize their lifetime. The RERUM network can face lower amounts of traffic since a percentage of the original number of packets will be transmitted, which contributes to the **scalability** of the network, because larger number of devices can be supported without congestion problems.

3.6.3 CS security strength

Here, we investigate CS security strength against three types of attacks:

- *Ciphertext-only attacks* (COAs), when an attacker is able to capture encrypted data transmitted by the legitimate RDs,
- *Known plaintext attacks* (KPAs), when the attacker has knowledge about the data prior to encryption,
- *Chosen plaintext attacks (CPAs)*, when an attacker has the ability to provide specific plaintext to a crypto-system, aiming to reveal potential weaknesses.

3.6.3.1 Ciphertext-only attacks

In this case, the adversary has full knowledge of the encrypted signal (ciphertext) y, and tries to guess matrix $\mathbf{\Phi}$. When a brute force attack is applied, the attacker performs an exhaustive search over a grid of values, until the correct key is found. In the following, we explore the feasibility of decrypting $\mathbf{y}$ when a wrong key (measurement matrix) is used. We compare the performance of three different encryption matrices, namely Gaussian, Bernoulli, and Structurally Random Matrix (SRM), for block size N varying in $\{32, 64, 128, 256\}$, and data sparsity level S in {5%-15%}. The encryption strength is expressed by the ability of the attacker to achieve a low reconstruction error. The memory requirements for all matrices depend on the block size N and the compression rate M, as the actual size of each matrix is MxN. Regarding the memory requirements, for the Gaussian and SRM matrices, it is required $MxNx4$ bytes of memory, while for the Bernulli ones it is required $MxNx1$ bytes as these consist of {-1, 1} entries with 50% probability.

In particular, we generate an ensemble of 100 encryption matrices and 100 blocks of data samples of size N that are sparse in the DCT (Discrete Cosine Transform) domain, with non-zero DCT coefficients independently drawn from a normal distribution $\mathcal{N}(0,1)$. The compression rate is selected as $CR = 1 - 5S$ in order to guarantee signal reconstruction, which is performed using the OMP algorithm. The SRM used is based on the Block Walsh-Hadamard Transform (BWHT) with a local pre-randomizer and a block size equal to 32.

For each experiment, we encrypt each data block with one out of the 100 encryption matrices, and use the remaining 99 for decryption. The evaluation is based on the reconstruction error defined as: $e = \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 / \|\mathbf{x}\|_2^2$, where $\hat{\mathbf{x}}$ is the decrypted signal, obtained by using OMP, and $\|\cdot\|_2$ stands for the ℓ_2 norm.

In Figure 35, Figure 36, and **Figure 37**, the reconstruction error versus the block size is reported for each encryption matrix, and across all sparsity levels. The reconstruction error, when using the wrong secret key, increases as the block size increases. Interestingly, we observe that the performance of Gaussian and Bernoulli matrices is almost the same, while the SRM exhibits clearly a lower reconstruction error. This is because the SRM exhibits a stronger deterministic part in its construction procedure, compared to the fully random Gaussian and Bernoulli matrices. Additionally, we observe that the increase in sparsity level also causes an increase in the reconstruction error across all matrix types and block sizes.

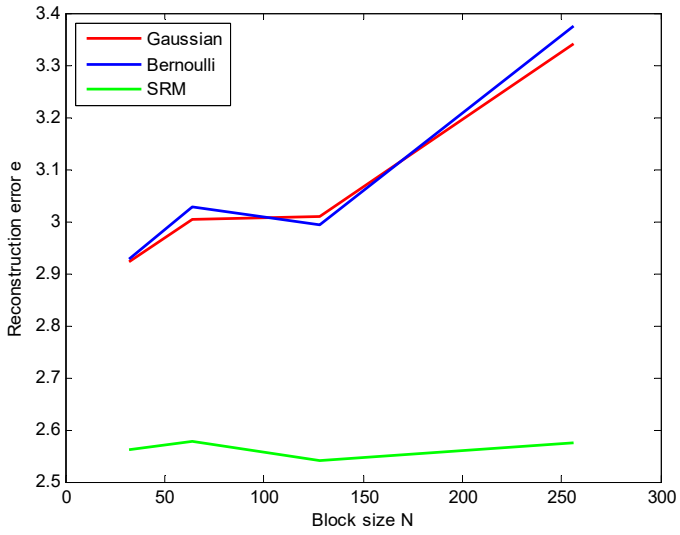


Figure 35: Reconstruction error vs. block size for S = 0.05

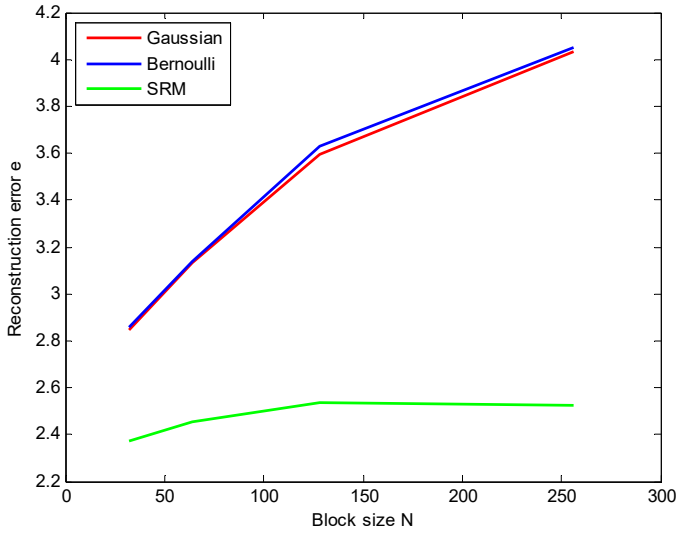


Figure 36: Reconstruction error vs. block size for S = 0.1

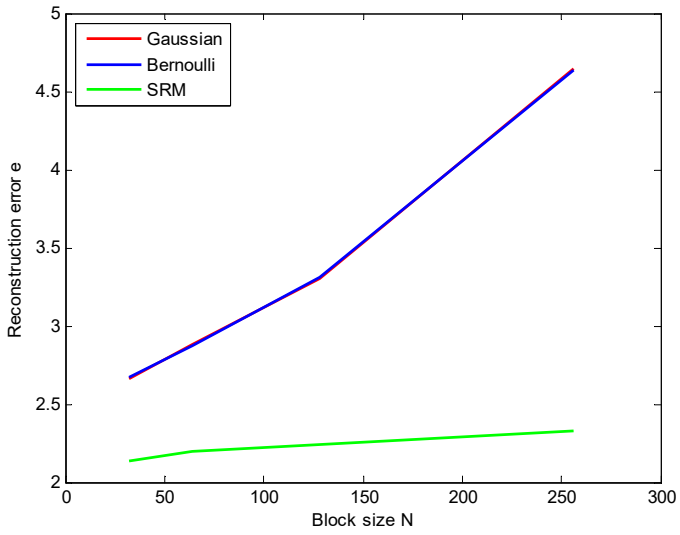


Figure 37: **Reconstruction error vs. block size for S = 0.15**

3.6.3.2 Known plaintext attacks

In this case, a malicious eavesdropper has gained access to an instance of the signal $\mathbf{x}$ (plaintext), and its corresponding random measurements $\mathbf{y}$ (ciphertext), and from this information tries to infer the corresponding encryption matrix. We follow the framework that is described in [CMPRS15] in order to study this type of attack, where the attacker attempts to recover the encryption matrix by formulating an appropriate subset-sum problem (SSP) [CMPRS15]. The resistance of the CS scheme against KPA is not based on the hardness of the corresponding SSP, but on the large number of candidate solutions among which an attacker should find the only true solution to guess each row of $\mathbf{\Phi}$.

The procedure we follow for evaluating the signal recovery quality achieved through a KPA is as follows:

1. The attacker, who gains access to a single plaintext-ciphertext pair $(\mathbf{x}, \mathbf{y})$, attacks row-by-row the corresponding true encoding matrix $\mathbf{\Phi}$ (whose elements in this case are i.i.d. Bernoulli variables) by executing Monte Carlo searches for each matrix row, until a large number of candidate solutions $\hat{\mathbf{\Phi}}$, that verify $\mathbf{y} = \hat{\mathbf{\Phi}}\mathbf{x}$, is found.

2. Furthermore, he tests the quality of each solution by getting a reconstruction $\hat{\mathbf{x}}$ of the original signal using the corresponding candidate matrix, and calculating reconstruction error $e = \|\mathbf{x} - \hat{\mathbf{x}}\|_2 / \|\mathbf{x}\|_2$. Then, he keeps those candidate solutions that achieve a low value of e.

3. Finally, he uses the chosen solutions to recover a new unknown plaintext $\mathbf{x}'$ from its ciphertext $\mathbf{y}' = \mathbf{\Phi}\mathbf{x}'$. Now, the reconstruction error, which is unknown to the attacker in this case, is $e' = \|\mathbf{x}' - \hat{\mathbf{x}}'\|_2 / \|\mathbf{x}'\|_2$, where $\hat{\mathbf{x}}'$ is the plaintext as recovered by the attacker when using a candidate solution $\hat{\mathbf{\Phi}}$.

Next, we show that although the attacker is able to recover a large number of candidate solutions that achieve a low reconstruction error e, he can never recover $\mathbf{x}'$ such that $e' \approx e$ thus his attack is unsuccessful. For this reason, we generate data that are sparse in DCT domain, with block size N varying in $\{32, 64, 128, 256\}$ and sparsity level S in {5%-15%}. Non-zero DCT coefficients are drawn uniformly at random from $\{-L, \ldots, -1, 0, 1, \ldots, L\}$, with $L = 10^4$, and the compression rate is chosen as

$CR = 1 - 5S$, in order to guarantee signal reconstruction, performed with OMP algorithm. We report empirical cumulative density functions (CDFs) of the reconstruction error e and corresponding e' for 2000 candidate solutions for the encryption matrix Φ. In Figure 38, **Figure 39**, Figure 40, Figure 41, Figure 42, and Figure 43, we show the errors e and e' for the different sparsity levels.

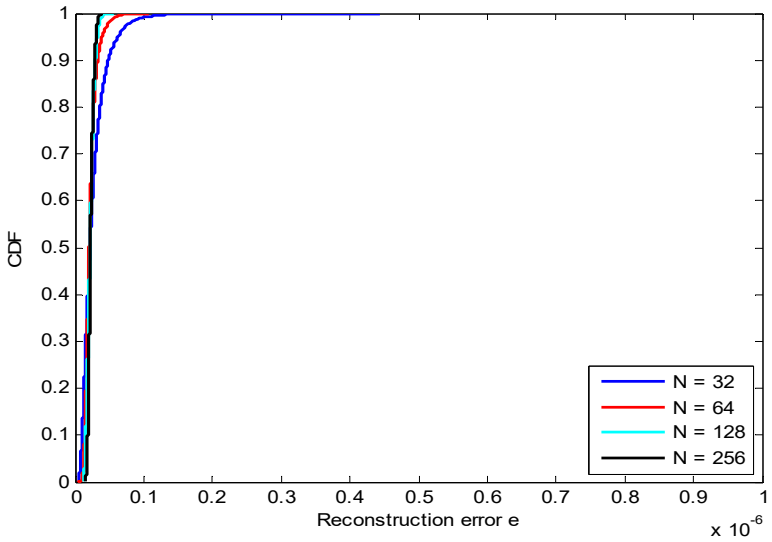


Figure 38: Reconstruction error ***e*** for *S* = 0.05

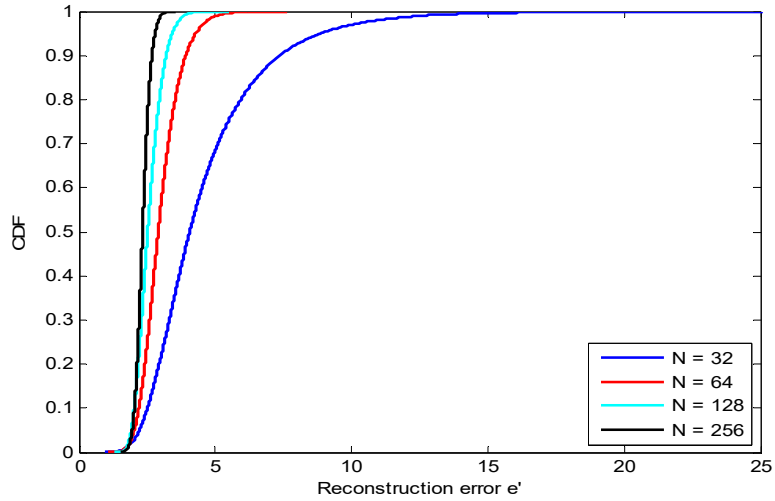


Figure 39: **Reconstruction error e' for *S* = 0.05**

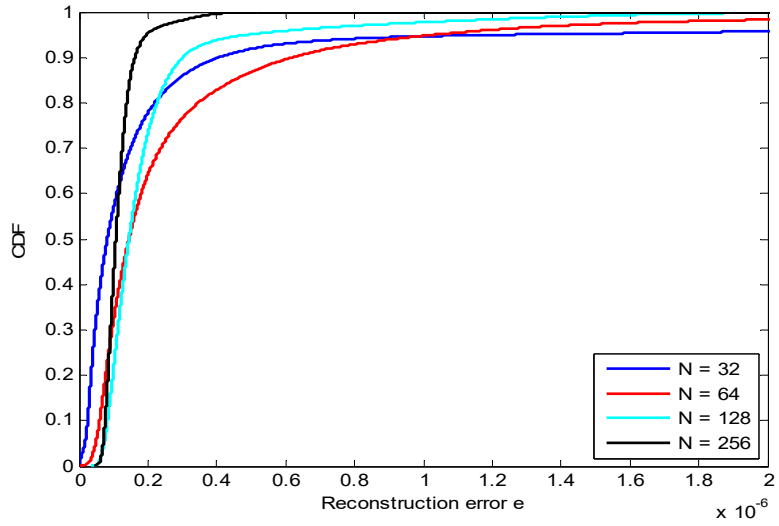


Figure 40: Reconstruction error $\boldsymbol{e}$ for S = 0.1

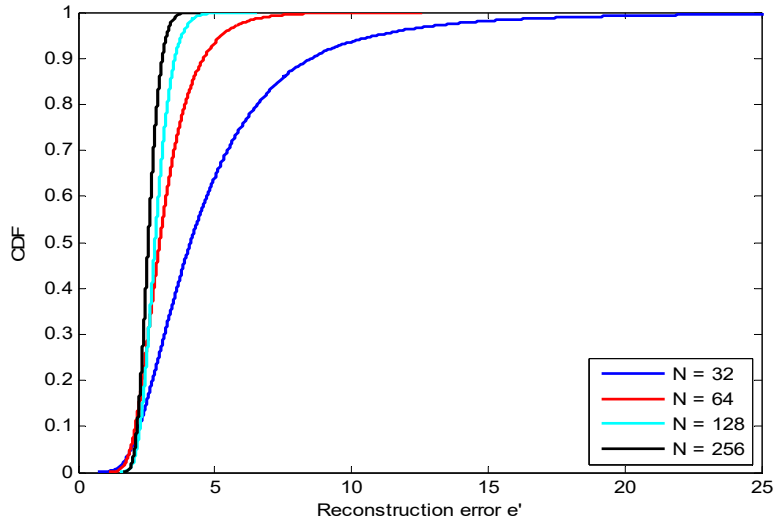


Figure 41: Reconstruction error $\boldsymbol{e}'$ for S = 0.1

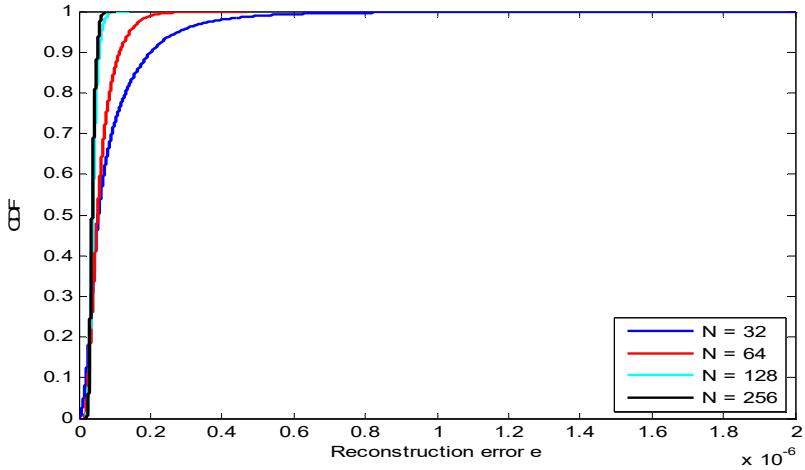


Figure 42: Reconstruction error $\boldsymbol{e}$ for S = 0.15

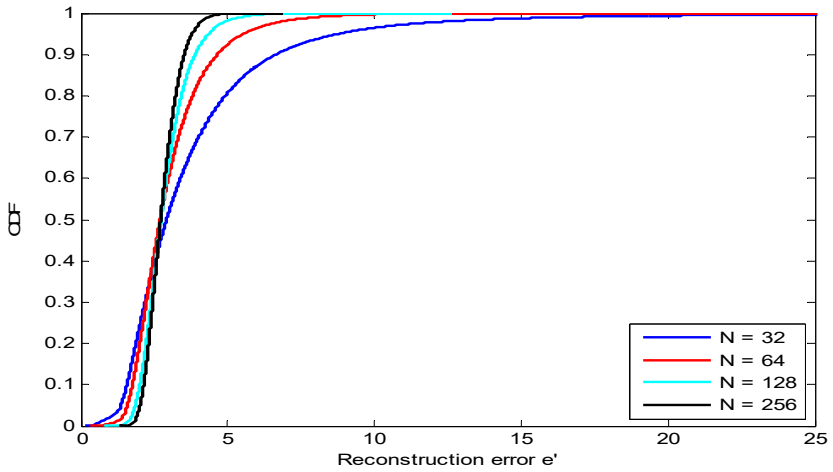


Figure 43: Reconstruction error e' for $S = 0.15$

3.6.3.3 Chosen plaintext attacks

In this scenario, an attacker has gained access to the crypto-system, and he is able to provide specific plaintexts in order to reveal potential weaknesses. Recall that ciphertext is derived after a multiplication between matrix $\mathbf{\Phi}$ and plaintext $\mathbf{x}$ takes place. Supposing an attacker has the ability to launch a CPA, he can provide plaintext $\mathbf{x}$, where all of its values, except in a specific location (index) j, are equal to zero. This results in revealing column j of matrix $\mathbf{\Phi}$, and by repeating this procedure N times, the attacker can reveal all columns of $\mathbf{\Phi}$; hence, CS encryption is highly vulnerable to CPAs.

In RERUM Deliverable 3.2, we proposed a method based on chaos sequences, which makes CS immune to this attack. Using a chaos sequence, we created a secret sparsifying basis that was used to construct matrix $\mathbf{\Phi}$. For investigating the encryption strength of this scheme, we define two types of attackers: (i) oblivious, and (ii) non-oblivious.

We assume that the oblivious attacker can successfully launch a CPA, and he is also able to capture all ciphertexts transmitted by a legitimate user. Nevertheless, he is not aware of the chaos-based encryption scheme, so he tries to decrypt the captured data using the typical sparsifying basis Ψ. Figure 44 shows that the reconstruction error the attacker experiences, is too high for proper decryption. For comparison, Figure 45 shows the corresponding error for a legitimate receiver (note that the x-axis scale is different).

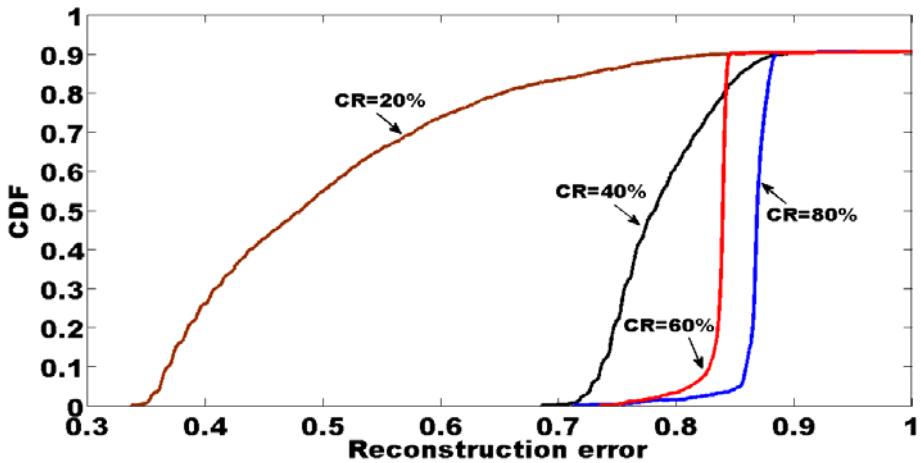


Figure 44: Error for various compression rates for the oblivious attacker

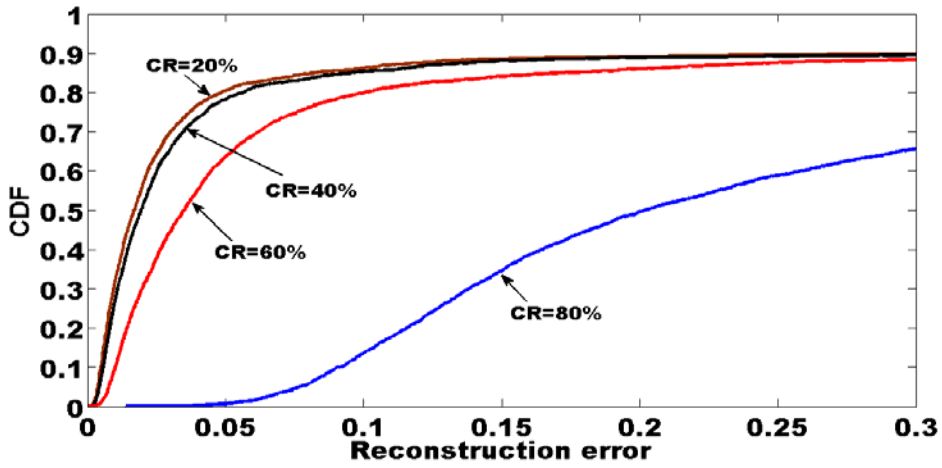


Figure 45: Error for various compression rates for the legitimate receiver

The non-oblivious attacker, is also capable of performing a successful CPA, and capture the corresponding ciphertexts. Additionally, he is fully aware of the chaos-based encryption scheme used. Regarding the chaotic sequence considered by a legitimate transmitter, C(d, k, c1), where d is the sampling distance, k its total length, and c1 the initial value, we assume that the non-oblivious attacker knows d and k, but not c1. Nevertheless, he attempts to decrypt the captured ciphertext by guessing c1. For empirically evaluating the reconstruction error the attacker experiences, we select c1'=(1-d)*c1 as the guessed value, where d=[5%, 10%, 15%, 20%, 25%] is the deviation from the real initial value.

As shown in the following figures, for high compression rates (CRs), the attacker experiences a very high error, regardless of the deviation from the correct c1. For the lower rates, only when the deviation is very small (5%), he can decrypt data with a low error.

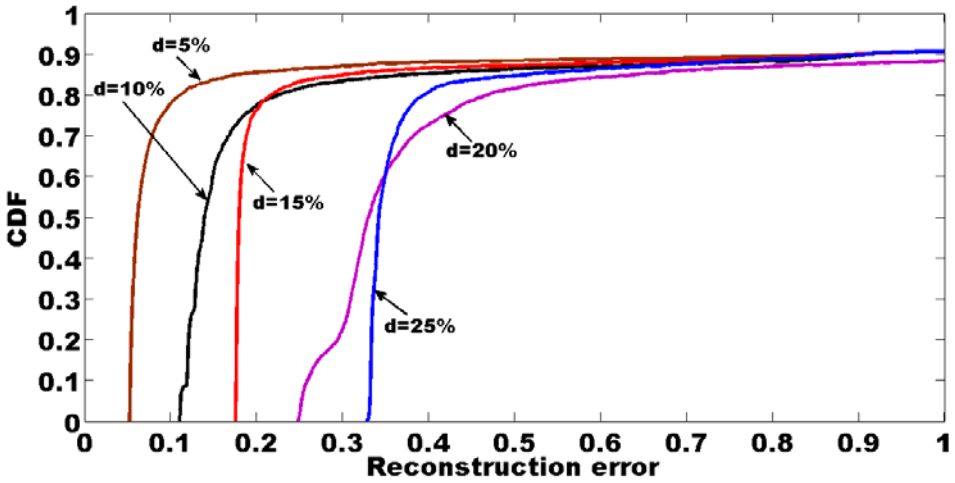


Figure 46: Error for the non-oblivious attacker when CR=20%

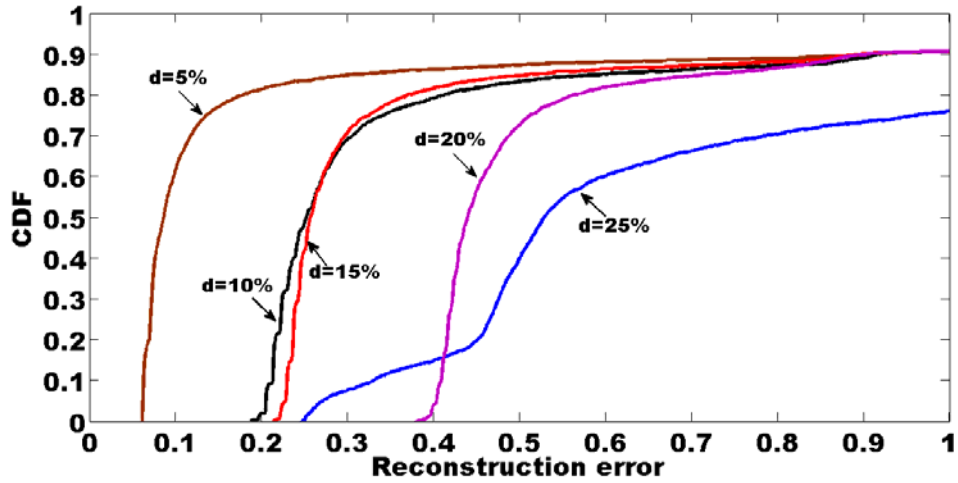


Figure 47: Error for the non-oblivious attacker when CR=40%

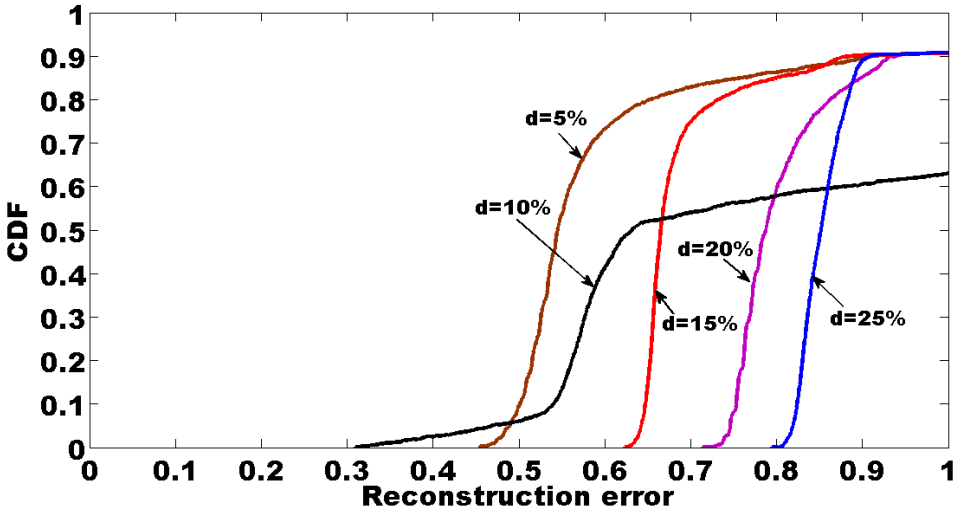


Figure 48: Error for the non-oblivious attacker when CR=60%

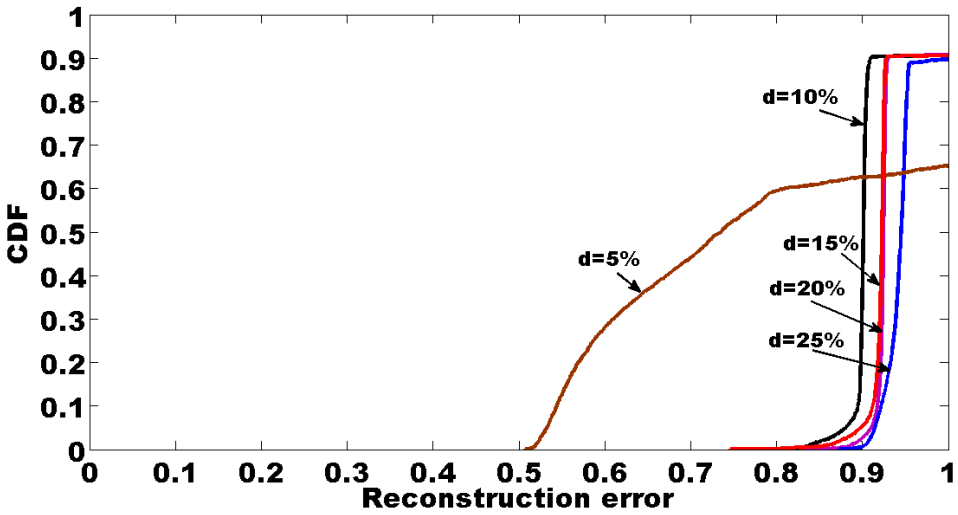


Figure 49: Error for the non-oblivious attacker when CR=80%

3.6.4 Analysis of energy consumption for the CS encryption

We implemented the CS encryption algorithm, along with the chaos sequence generation module, in a RE-Mote, based on the Contiki OS. The algorithm is used on an experimental implementation of the use case UC-I2 for indoor comfort quality monitoring, encrypting sensory data (ambient temperature, ambient light, humidity, etc.). The encrypted data are provided through a built-in COAP server.

CS encryption mainly involves the multiplication of a data vector with the measurement matrix (encryption key). The lower the compression rate, the higher the size of the encryption key; therefore, the higher the number of the required multiplications. In Figure 50 we show the energy consumed for **encrypting** a single block of plaintext for various compression rates.

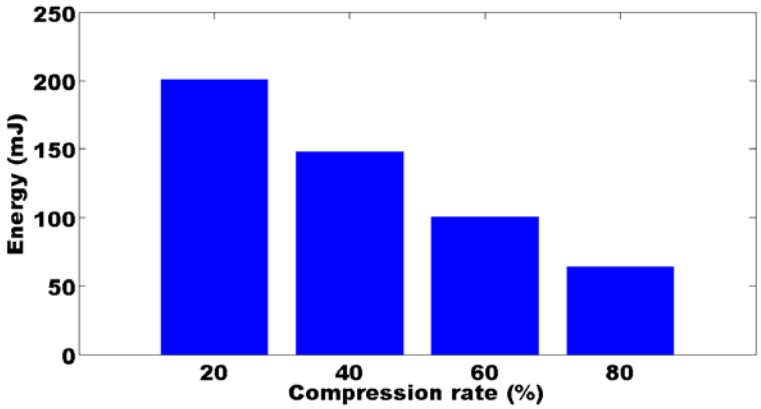


Figure 50: Energy consumption of CS encryption for various compression rates

Figure 51 shows the time required for the multiplications to complete. As expected, the higher the compression rate, the smaller the required time because fewer multiplication are performed. Figure 52 and Figure 53 show the energy consumption and execution time of the chaos sequence generation module of various sizes, respectively.

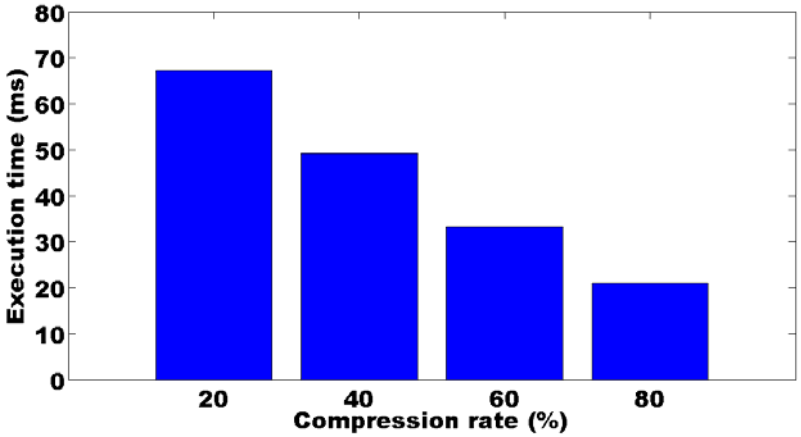


Figure 51: Execution time of CS encryption for various compression rates

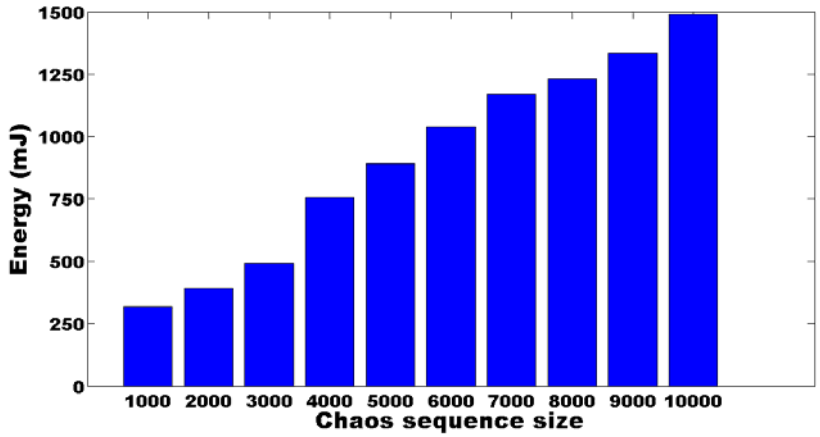


Figure 52: Energy consumption of the chaos sequence generation module for various sizes

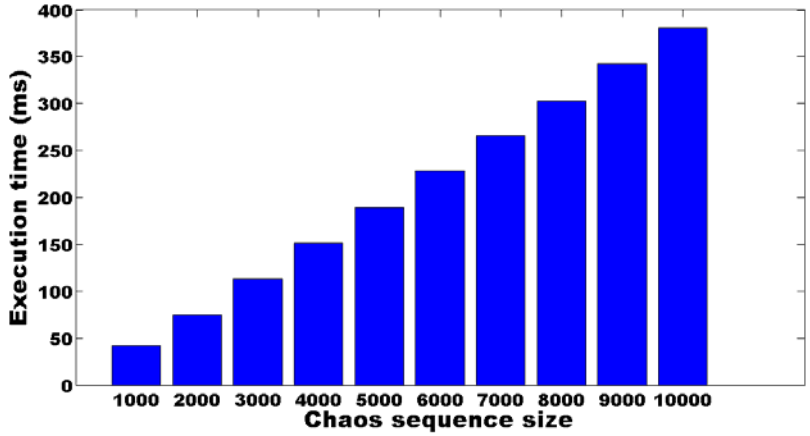


Figure 53: Execution time of the chaos sequence generation module for various sizes

Next, we demonstrate how CS minimises the energy consumption, as data are compressed while encrypted. We collect encrypted ambient light measurements from a single RE-Mote loaded with the firmware that contains our CS crypto-system. The light sensor collects measurements every 2 seconds, while data are transmitted (over the COAP protocol) every 3 seconds to a remote server. At the same time, we measure the energy consumption of the mote using powertrace [DEFT11]. The consumption refers to the CPU, LPM (low power mode), Transmit, and Listen operations. We run the experiment without CS, until 640 measurements are collected. We repeat the experiment, this time enabling CS with a compression rate of 50%, till the same amount of information is collected. In the second case, fewer packets are transmitted, however, after the reconstruction ends, 640 measurements are available (each packet carries a single measurement).

Figure 54 and Figure 55 show the cumulative sum of the power consumed due to the CPU operation. Observe that in the case where no CS is used, the consumption is higher. This is because more packets have to be transmitted; hence, all related operations (COAP handling, etc.) increase. For the same amount of information, 99 packets have to be transmitted in the no CS case, while 75 packets when CS is used. Regarding the energy due to the transmit operation (Figure 56, Figure 57), when CS is used is reduced, as less packets are transmitted, so the RF transceiver of the mote consumes less energy.

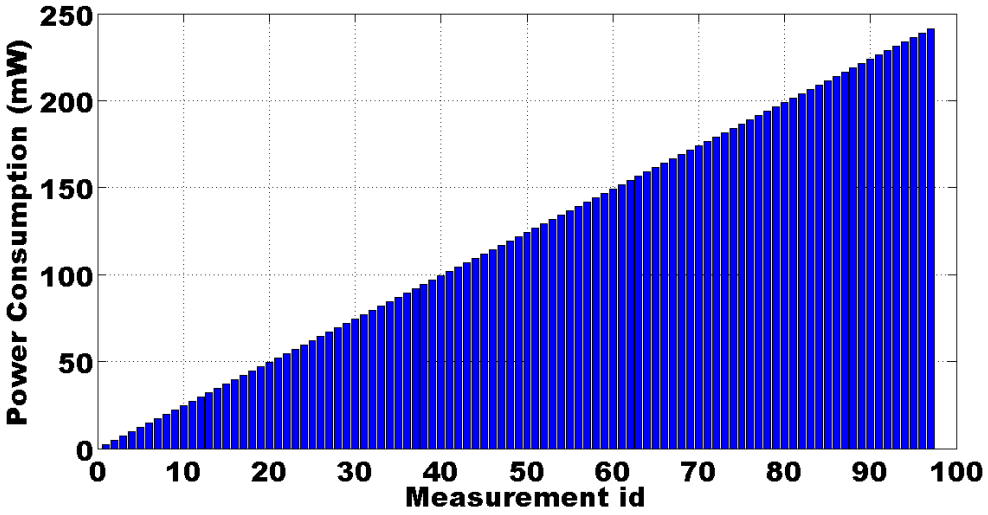


Figure 54: Power consumption due to CPU operations when no CS is used

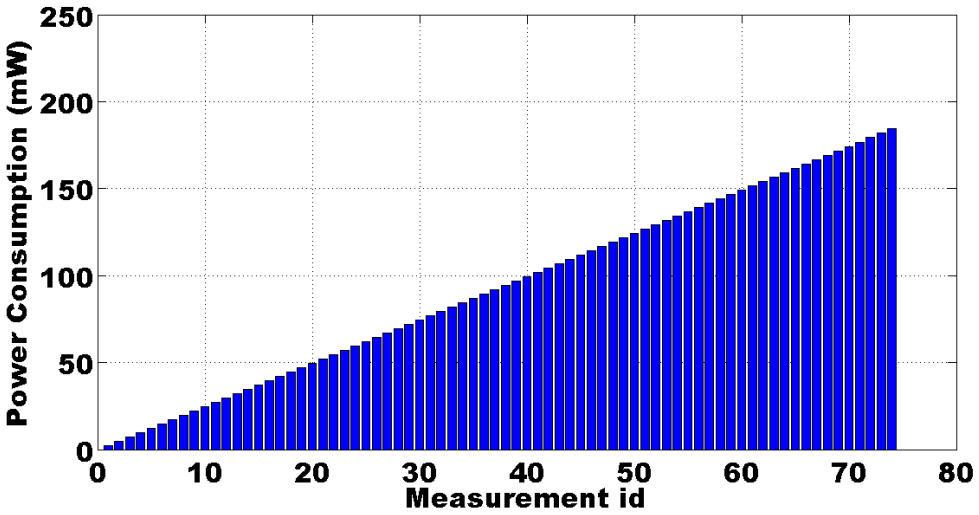


Figure 55: Power consumption due to CPU operations when CS is used

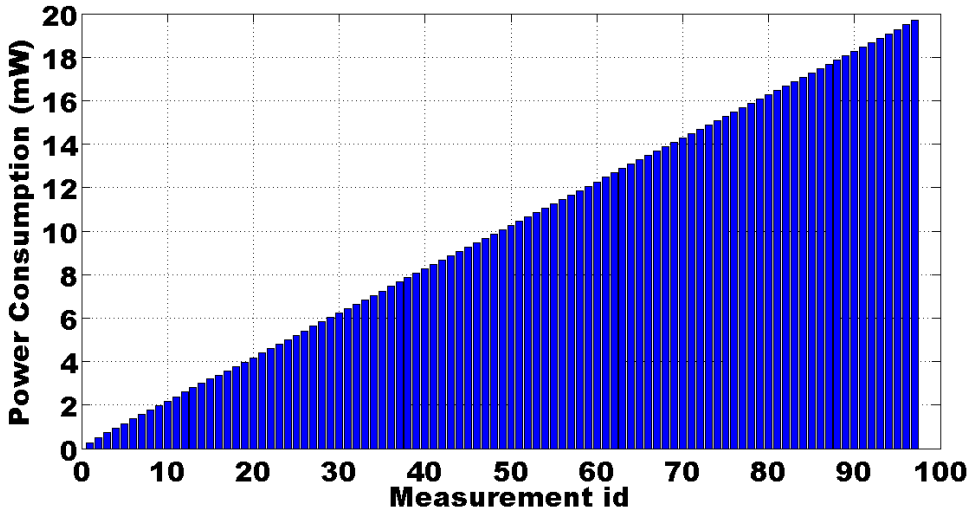


Figure 56: Power consumption due to Transmit operations when no CS is used

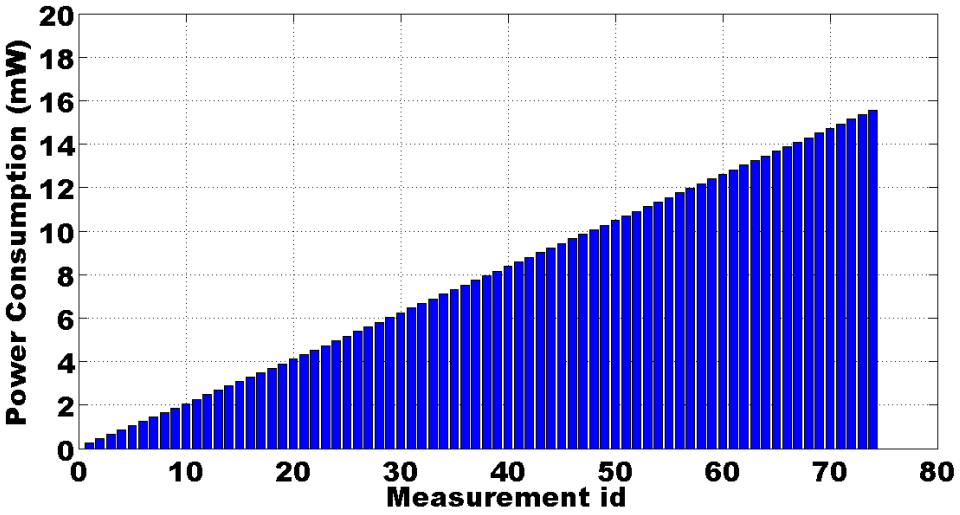


Figure 57: Power consumption due to Transmit operations when CS is used

3.6.5 Discussion

In this subsection, we presented a detailed performance analysis of the security strength of CS against three types of common attacks. From the presented results, it is clear that the proposed scheme can really contribute to the mitigation of the three presented attacks, which can be a huge advantage for improving significantly the performance of the overall network, minimizing packet losses, minimizing traffic congestion and maximizing efficiency. Furthermore, the overall security of the network is increased with what has been clearly shown to be a very lightweight and energy efficient technique. This can be very useful in any IoT system, such as RERUM, where **large scale deployments** can be susceptible to (i) attacks, because they can't be easily monitored and (ii) network congestion, because large numbers of devices can create large volumes of traffic.

More specifically, we showed above that with our proposed CS technique, in the ciphertext-only attacks, an attacker experiences a higher reconstruction error when Gaussian or Bernulli matrices are used for the encryption. When an SRM matrix is used, the reconstruction error reduces it exhibits a stronger deterministic part in its construction procedure . The lower the reconstruction error for the attacker, the lower the encryption strength of CS when this matrix is used. Regarding the memory requirements, the Bernulli matrix requires less memory, compared to the SRM and Gaussian matrices, as it consists of {-1, 1} entries that can be stored as 1-byte values.

For the known plaintext attack, an attacker although he has knowledge of plaintext-ciphertext pairs, he is not able to correctly guess the encryption matrix due to the large number of possible solutions. For the chosen plaintext attack, our chaos-based proposed solution make CS immune to this attack. The evaluation results show that an attacker experiences a very high error. Furthermore, the implementation of the chaos-based crypto-system, and the presented results regarding the execution time and the consumed energy, prove its feasibility for use in constrained devices.

Furthermore, the results above showed clearly that when CS is used both the execution time and the energy consumed within an RD are reduced. This happens because the data compression leads to the transmission of fewer packets. Thus it is evident that CS is very lightweight, and can really be of significant advantage to maximizing (i) the energy efficiency of large IoT deployments, (ii) the performance (in terms of congestion avoidance) and (iii) the security of IoT networks.

3.7 Fairness vs Throughput

3.7.1 Introduction

IoT is expected to benefit from the deployment of 5G networks, i.e., networks that can use multiple access technologies (access networks – AN), offering both more reliable connectivity and energy efficiency. Furthermore, multiple access technology provides improved handling of radio resources, since more degrees of freedom are added. As a result multiple access technologies offer improved fairness among devices, at the cost of additional computational and networking complexity.

Within the context of RERUM, multiple access technologies will provide an optimum trade-off between network throughput and individual RD throughout, in order to avoid common situations in dense deployments (e.g., large number of sensors) where many devices lost connectivity due to severe interference conditions, or due to the lack of those mechanisms that can allocate resources from different access technologies to devices that have dual access (e.g., LTE and WiFi).

In other words, RERUM recommends these specific multi-RAT cooperative mechanisms for those Network Providers (Telco Operators) that wish to provide also IoT Services. Such providers that support the new cellular/WLAN interworking standards can benefit from these mechanisms, since the will be able to offer better resources utilization and better QoS to their customers.

This mechanism can be part of the RERUM communication manager functional component and more specifically, it can be implemented within the routing and scheduling modules. These modules will be responsible for the optimum allocation of the resources to the RDs and the routing of the packets through the optimum route and the best RAT available at the moment of the transmission. The mode of operation is depicted in Figure 58.

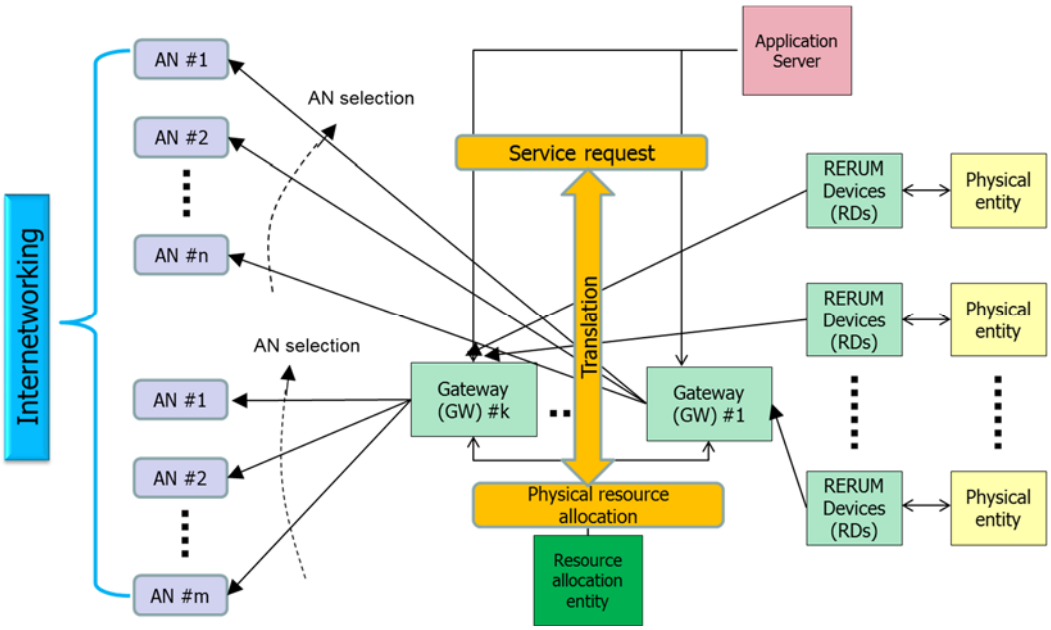


Figure 58: The overall architecture and its relation to the RERUM architecture.

3.7.2 Relation to RERUM UCs

The proposed network access mechanisms could be directly applied to all RERUM deployments but is recommended that deployments that involve high mobility and hence increase the possibility for frequent hand-overs between different RATs, should be avoided. In this sense, the proposed mechanism would ideally fit to the following RERUM use cases:

- Environmental monitoring (UC-O2)
- Home energy monitoring (UC-I1)
- Comfort quality monitoring (UC-I2)

In these deployments, assuming that the network provider owns both the WAN (e.g., LTE) and WLAN (e.g., WiFi) networks and that the RDs have both a WAN and WLAN radio interface, we can achieve considerably higher spectrum efficiency and fairness in the resource utilization by the RDs.

Practically speaking, let us consider for example the UC-O2, where hundreds of RDs are deployed in a specific geographical area and each RD is connected to the internet (i.e., to the application server through the MW) either using a GW or directly, while the access technology can be either a WLAN or a WAN. Assuming that the WAN and WLAN is handled by the same network operator, then by employing this network mechanisms, the operator can increase both the number of connected devices to the internet, since it can optimally allocate resources from both RATs to the RDs (or GWs), and improve the QoS for each RD, by maximizing the individual SINRs. Otherwise, trying to deploy hundreds of RDs without exploiting different access technologies, may lead with high probability to network congestions ad bad QoS.

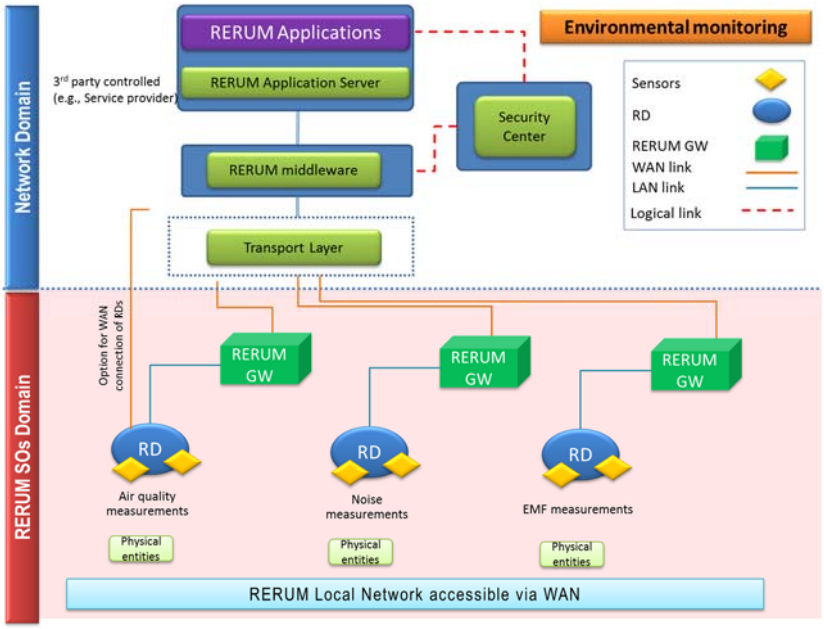


Figure 59: An example where the proposed example can be applied as part of the RERUM deployment

3.7.3 Analysis of the fairness-throughput trade-off

In this Section, important statistical metrics of the instantaneous SINR, namely the probability density function (PDF) and the cumulative distribution function (CDF) are presented for the conventional cellular and the hybrid Cellular/WLAN networks. Details about the derivation of these formulas can be found in [BLKS2014]. In short, the performance metrics are derived after taking into account the following cases:

- The RDs are connected to the internet through a GW while the GW has the capability for a WAN and a WLAN network access interface. In that case the operator optimally allocates the resources to the GWs, depending on the requested Quality Channel Indicators (indoor use cases).
- The RDs are connected to the internet through a WLAN GW or directly to the WAN (outdoor use cases).

Case 1: SINR for RDs that connect to the internet via the eNB

For the hybrid Cellular/WLAN network, the following two complementary communication cases are investigated, namely one-phase direct eNB→RD communication and two-phase indirect eNB→AP→RD communication. Considering the case where the target RD is directly connected with the, the SINR at the target RD is given by

$$\gamma_{eu,k} = \frac{\gamma_{eu_{\mathrm{k}}}}{1 + \underbrace{\sum_{i=1}^{|\mathcal{I}_{AP,k}^{C}|} \gamma_{I_{ea,i}}}_{\mathrm{eNB}\overset{\smile}{\rightarrow}\mathrm{AP}} + \underbrace{\sum_{i=1}^{|\mathcal{I}_{UE,k}^{C}|} \gamma_{I_{eu,i}}}_{\mathrm{eNB}\overset{\smile}{\rightarrow}\mathrm{RD}}}$$

where $\gamma_{I_{ea,i}}$ is the INR due to the cellular co-channel interferers, which follows the exponential distribution with mean values $\overline{\gamma}_{I_{ea,i}}$.

Case 2: SINR for RDs that connect to the internet via the GW

Considering the case where the target RD is connected to an eNB via a WLAN AP, in the first phase of the transmission the AP will experience interference from the eNB serving other UEs and the APs (e.g., IoT GW) at the same cellular frequencies. In the second phase the device experiences the interference from the APs that serve RDs at the same WLAN frequencies. In this sense, during the first phase, the instantaneous SINR at the m*th* AP, directly connected with the nth eNB, will be

$$\gamma_{ea,m} = \frac{\gamma_{ea_{\mathrm{n,m}}}}{1 + \underbrace{\sum_{i=1}^{|\mathcal{I}_{AP,m}^{C}|} \gamma_{I_{ea,i}}}_{\mathrm{eNB}\overset{\smile}{\rightarrow}AP} + \underbrace{\sum_{i=1}^{|\mathcal{I}_{UE,m}^{C}|} \gamma_{I_{eu,i}}}_{\mathrm{eNB}\overset{\smile}{\rightarrow}UE}}$$

where $\gamma_{ea_{\mathrm{n,m}}}$ the instantaneous SNR at the m*th* AP, following the exponential distribution with mean value $\overline{\gamma}_{ea_{\mathrm{n,m}}}$. In the second phase of the communication, where the RD receives the desired signal from one AP and interfering signals coming from the m*th* APs, the instantaneous SINR at the kth UE can be expressed as

$$\gamma_{au,k} = \frac{\gamma_{au_{\mathrm{m,k}}}}{1 + \underbrace{\sum_{i=1}^{|\mathcal{I}_{UE,k}^{W}|} \gamma_{I_{au,i}}}_{\mathrm{AP}\overset{\smile}{\rightarrow}UE}}$$

where $|\mathcal{I}_{UE,m}^{C}| + |\mathcal{I}_{UE,k}^{W}| = N$. $\gamma_{au_{\mathrm{m,k}}}$ is the instantaneous SNR at the k*th* RD and $\gamma_{I_{au,i}}$ is the INR due to WLAN co-channel interferers.

Derivation of the probability density and cumulative distribution functions for SINRs (Case 1, Case 2 and Total SINR)

The PDFs of the total instantaneous INR caused o the k*th* device by the eNB serving RDs at the same frequency, i.e.,

$$\gamma_{I_{eu,k}} = \sum_{i=1}^{|\mathcal{I}_{UE,k}^{C}|} \gamma_{I_{eu,i}}$$

and that of the total instantaneous INR caused to the k*th* device by the eNB serving APs at the same frequency, i.e.,

$$\gamma_{I_{ea,k}} = \sum_{i=1}^{|\mathcal{I}_{AP,k}^{C}|} \gamma_{I_{ea,i}}$$

Which follow the chi-square distribution. Assuming i.n.d. fading conditions, the PDF and the CDF of $\gamma_{eu,k}$ has the form of the equations in [BLKS2014, Table II, Form B]

For i.i.d. fading conditions, the PDF and the CDF has the form of the equations [BLKS2014, Table III, Form B]

$$\gamma_{ea,m} = \frac{\gamma_{ea_{\mathrm{n,m}}}}{1 + \underbrace{\sum_{i=1}^{|\mathcal{I}_{AP,m}^{C}|} \gamma_{I_{ea,i}}}_{\mathrm{eNB}\overset{\smile}{\rightarrow}AP} + \underbrace{\sum_{i=1}^{|\mathcal{I}_{UE,m}^{C}|} \gamma_{I_{eu,i}}}_{\mathrm{eNB}\overset{\smile}{\rightarrow}UE}}$$

In the case of the indirect connection of the RD to the eNB, there are two communication phases. During the first one (i.e., eNB → AP), the SINR at the target AP is given by

where $\gamma_{ea_{\mathrm{n,m}}}$ is the INR due to the cellular co-channel interferers, which follows the exponential distribution with mean values $\overline{\gamma}_{ea_{\mathrm{n,m}}}$.

In the second phase one (i.e, AP → RD), the instantaneous SINR at the target RD is expressed as

$$\gamma_{au,k} = \frac{\gamma_{au_{\mathrm{m,k}}}}{1 + \underbrace{\sum_{i=1}^{|\mathcal{I}_{UE,k}^{W}|} \gamma_{I_{au,i}}}_{\mathrm{AP}\overset{\smile}{\rightarrow}UE}}$$

where $\left|\mathcal{I}_{UE,m}^{C}\right| + \left|\mathcal{I}_{UE,k}^{W}\right| = N, \gamma_{au_{\mathrm{m,k}}}$ is the instantaneous SNR at the k*th* RD and $\gamma_{I_{au,i}}$ is the INR due to WLAN co-channel interferers. Both $\gamma_{au_{\mathrm{m,k}}}$ and $\gamma_{I_{au,i}}$ are exponentially distributed with mean values $\overline{\gamma}_{au_{\mathrm{m,k}}}$ and $\overline{\gamma}_{I_{au,i}}$.

Assuming i.n.d. fading conditions, for the first phase, the PDF and the CDF of $\gamma_{ea,m}$ is of the same form as that of $\gamma_{eu,k}$ by substituting $C = \gamma_{ea,m}, X_1 = \left|\mathcal{I}_{UE,m}^{C}\right|, X_2 = \left|\mathcal{I}_{AP,m}^{C}\right| Y = \overline{\gamma}_{ea_{\mathrm{n,m}}}, Z_{1,i} = \overline{\gamma}_{I_{eu,i}}, Z_{2,i} = \overline{\gamma}_{I_{ea,i}}$.

For the second phase, the PDF and the CDF of $\gamma_{au,k}$ is of the forms [BLKS2014, Table II, Form A]. Assuming i.i.d. fading conditions, the PDF and the CDF can be found in [BLKS2014, Table III, Form B].

The goal of the derivation of the mathematical formulas is to be able to use them in optimization problems, so that the calculations can be performed instantly and in real time. This is of critical importance, because non-real time calculations would affect the performance of the system and would in practise result in wrong results, since the calculations would be not applicable to the network conditions at the time of decision making. In other words, the resource allocation to the RDs would be out-dated each time the calculations were performed.

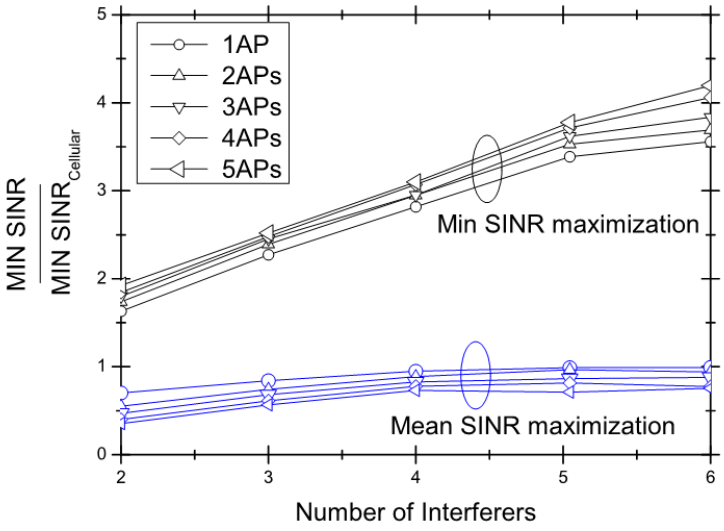


Figure 60: The minimum UE SINR gain over the cellular scheme, considering the two optimization problems, i.e., the maximization

3.7.4 Discussion

Since the SINR of the RDs is known in closed form, they can be used directly aiming either at maximizing the total SINR (i.e., the network throughput) or the individual SINR of the RDs, so that all RDs attain almost the same performance. In Figure 60, it is shown that maximizing the cell's overall SINR greatly affects the individual RDs SINRs. Having the minimum RDs SINR of the cellular scheme as a benchmark, we observe that the greedy solution results in a performance which is worse than that of the cellular scheme, in terms of the maximum minimum RD SINR.

Concluding, the utilization of the proposed mechanism will provide the following benefits to RERUM (and IoT in general) deployments, where the devices can use multiple access technologies:

- Increased number of served devices thanks to the coordinated resource management
- Fairness to the RDs, i.e., provide the minimum required QoS to all devices
- Increased reliability in terms of service availability

The increased number of devices is illustrated by the upper group of curves, where the gain of the MIN SINR over the cellular MIN SINR (e.g. the case where no cooperative RATs are used) is higher than 2. In other words, by employing cooperation between different Rats, we can increase the achievable SINR for each RD. Furthermore, by definition, since the MIN SINR is higher (maximized) than that in the conventional case, this means that all RDs achieve a minimum required QOS, in contrast to the conventional case, where some RDs may not be served at all. This also means that higher reliability is attained, since more RDs get access to the network.

4 Heterogeneous Networking for RDs

4.1 Cooperative heterogeneous network for RDs

4.1.1 Introduction

Resource allocation plays a fundamental role in wireless networks, especially in dense heterogeneous ones, where multiple problems co-exist, such as the congestion of radio resources, the co-existence of multiple radio access technologies (RAT)s and the need to serve all devices with fairness. To this end a hybrid cellular/WLAN communication was proposed in [RERUM-D4.1] where the mobile users can be served by either the eNB or a WLAN access point (AP), depending on the selection strategy (e.g., the signal to interference plus noise ratio (SINR)). According to this scheme, the WLAN APs are wirelessly connected to the eNB and share this broadband connection with specific users over WLAN frequencies. The users select their serving node, i.e., the macro-cell eNB or a WLAN AP, based on a performance criterion. The aim of this architecture is to reduce the transmission power from the eNB to users with bad channel conditions (e.g., cell edge users) and thus the interference at both the cellular and WLAN part of the hybrid network, while avoiding modifications to the existing cellular network. For the completeness of this Section, the overview of the resource allocation scheme that was introduced in [RERUM-D4.1] is depicted in Figure 61.

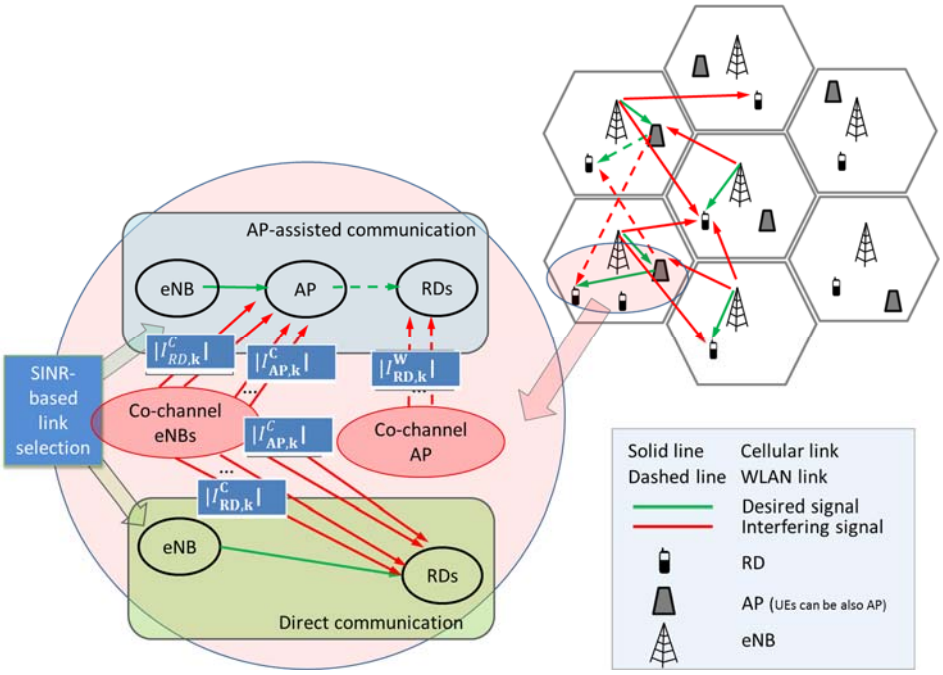


Figure 61: An overview of the proposed multi radio access scheme [RD4.1].

As an offload technology, this approach has the following benefits as compared to other technologies. Regarding RERUM, these benefits can be obtained by enabling the RDs and GWs to be compatible with multi-RAT access networking, i.e., implement these mechanisms.

- Cost-effective method to access mobile broadband
- Widespread existing deployments at home, in many public places, and most importantly, offered on most of the mobile devices
- WiFi technology is widely used by many consumers to access the internet, and is also available on the vast majority of mobile devices. Thus, offloading using available an WiFi network is a route which can lead to minimized data costs for providers and consumers.

- No significant modifications to the existing cellular network is required
- Offloading data between licensed and unlicensed spectra greatly helps in efficient spectrum utilization [WIPRO]
- Capability to address new users and devices without mobile subscription [CISCO]
- Standards availability for integration into mobile core networks
- Reduces energy consumed by the mobile devices as short-range communication (e.g., WiFi) typically consumes less energy-per-bit than long-range communication (e.g., 3G/4G)
- WiFi uses unlicensed spectra and developing new applications and technologies in an unlicensed spectrum environment does not require approval from government authorities. However, there could be possible future regulations pertaining to these unlicensed bands.
- Automatically and seamlessly provides the best option available to the mobile device without involving it
- WiFi offloading approach uses a SINR maximization criterion, which increases the overall network throughput and reduces intracell and intercell interferences.

That optimization scheme aims at maximizing the overall cell's signal-to-interference-plus-noise (SINR), without taking into account the fairness among the devices. In other words, maximizing the overall cell SINR without considering the individual SINR requirements of each user/device may result with high probability that each time only a limited number of user/device are served, while other devices may be unable to meet their quality-of-service (QoS) requirements. In order to avoid this situation and attain the fairness among the devices, the resource allocation algorithm in [RERUM-D4.1] has been enhanced towards satisfying the QoS requirements of all the devices. Obviously, the introduction of the fairness results in a trade-off between the overall cell's SINR performance and the individual SINR performance.

4.1.2 Relation to RERUM UCs

The proposed scheme can be part of the RERUM communication manager functional components and more specifically the routing and scheduling module and operates complementary to the mechanism discussed in 3.7. Thus, the relation to the RERUM UCs is as discussed in in 3.7.

4.1.3 Achieving fairness-throughput trade-off

The introduction of the fairness requirement in the algorithm [RERUM-D4.1] modifies the optimization problem that has to be solved. Instead of maximizing the overall cell's SINR, i.e.,

$$SINR_{opt} = \underset{\{All\ interference\ combinations\ among\ devices\ and\ APs\}}{\arg\max\{SINR_{total}|_{n=1}\}}$$

now the optimization problem becomes

$$SINR_{opt} = \underset{\{All\ interference\ combinations\ among\ devices\ and\ APs\}}{\arg\max\{\min(SINR)_k\}, \forall kth\ device}$$

Obviously, an exhaustive search among all the combinations would be a quite complex process. For example for a relatively small number of APs, e.g., 7 and devices, e.g., 5, the total number of searches would be $E = (M+1)^N = 279.936$. To this end, a greedy algorithm has been developed, which aims at a very fast suboptimal but near-optimal solution, which considerably reduces the number of searches among all possible network combinations.

The mode of operation is described in detail in Table 22and is summarized as follows. For a network setup with *N* devices and *M* APs, the algorithm initially considers only the *Q* out of the *N* devices and

finds among all the $(M + 1)^Q$ combinations the optimal network topology that maximizes the minimum SINR among the devices. Then continuing to the *Q + 1* device, the algorithm looks for the *AP* that this device will be connected to, searching among the M APs, in order to maximize the overall SINR given that the previous Q UEs cannot change their AP. Similarly, considering that the previous *Q + 1* devices cannot select a different AP, the algorithm continues to the *Q + 2* UE and searches for that AP that the device will be connected to, in order to maximize the overall SINR. In this way the total number of searches is given by $G = (M + 1)Q + (N - Q) * M$, (e.g., 37 for 7 APs and 5 devices and an initial exhaustive search for *Q*=2 devices).

Table 22: SINR optimization with user/device fairness

Input:

- The number of total devices in the cell, *N*, and the number of total WLAN APs, *M*.
- The number of devices, *Q*, for the optimal topology in the initialization stage.
- The channel gains between the eNB and the devices.
- The channel gains between the eNB and the APs.
- The channel gains between the APs and the UEs

Initialization stage:

Considering *Q* devices within the cell, find the optimal topology

that maximizes the minimum individual SINR among all devices

Save the optimal topology in the vector *S = [t(1) t(2) ... t(Q)],*

where *t(·)* = 1 if the AP is the eNB

and *t(·) = M + 1* if the AP is the *M*th WLAN AP.

Greedy stage:

for *j := Q + 1 to N*

Considering the topology *S(1) = [t(1) t(2) ... t(j − 1) t(j)]*

for *a := 1 to M + 1*

Calculate the individual SINR of each device for the topology

S(a) = [t(1) t(2) ... t(j − 1) t(j) t(a)]

End for

Among the *S(a) = [t(1) t(2) ... t(j − 1) t(j) t(a)]*

Find the topology that maximizes the minimum individual SINR

Set *S(j) = [t(1) t(2) ... t(j − 1) t(j) t(a)]*

End for

As for an example let's consider a macro-cell with one eNB, *N*=5 d**e**vices and M=3 WLAN APs and assume that the initial optimum network topology is calculated for *Q*=2 devices. The optimum topology is the one that satisfies the optimization problem (eq. 2), i.e., the minimum SINR of each device is maximized. Mathematically speaking, the initial topology is represented by the vector $S = [2\ 1]$, where the *k*th element of the vector denotes the serving point (for example 1 for the eNB and 2, …, M+1 the rest of the WLAN APs) where the *k*th device is connected to. In this specific example, the vector $S = [2\ 1]$ denotes that the 1st device (1st element of the vector) is connected to the 3rd (*M+1 = 2+1*) AP and the 2nd device (2nd element of the vector) is connected to the eNB (*M=1)*. Then, the algorithm continues to the 3rd device and searches among the *M*=3 APs those AP where the 3rd device will connect to, in order to maximize the minimum individual SINR of this device. More specifically, the algorithm will calculate the individual SINR for the 3rd device for the following network topologies: $S = [2\ 1\ 1], S = [2\ 1\ 2]. s = [2\ 1\ 3], S = [2\ 1\ 4]$ and will select the topology that maximizes the minimum SINR of the 3rd device. Note that the first two APs are kept fixed. Assuming the algorithm selects the topology$S = [2\ 1\ 3]$, it will continue to the 4th device by considering the following topologies: $S = [2\ 1\ 3\ 1], S = [2\ 1\ 3\ 2], S = [2\ 1\ 3\ 3], S = [2\ 1\ 3\ 4]$. The algorithm will continue until all the devices have connected to an AP (eNB or WLAN).

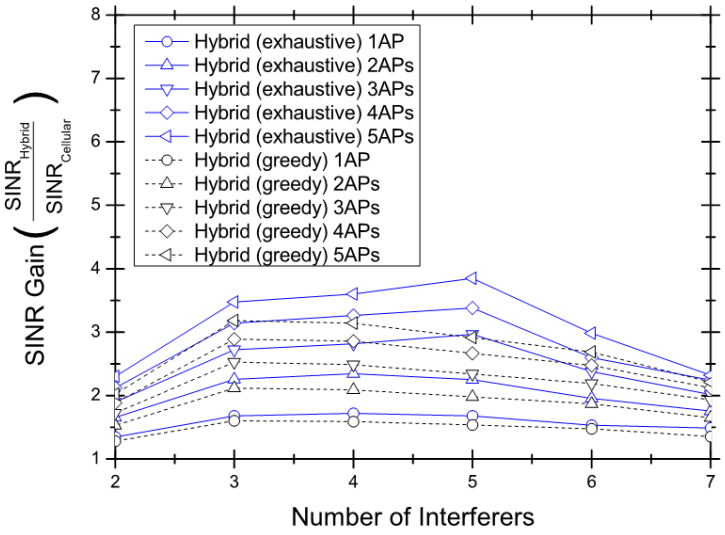


Figure 62: The SINR gain of the proposed Cellular/WLAN scheme over the cellular (Single cell scenario).

4.1.4 Discussion

5G network deployments are expected to be the core component for the IoT revolution, since they will satisfy crucial requirements such as high network availability (through the ability to connect to multiple RATs), energy efficiency and higher spectrum efficiency. RERUM aspires to make those networking recommendations for IoT deployments, which will boost their performance in terms of spectrum utilization and fairness between the numerous IoT devices. The proposed mechanism targets the Telco Operators that plan to provide IoT services as well, minimizing the risk for network congestions and waste of cellular resources.

With respect to the RERUM UCs, i.e., UC-O2, UC-I1/2, where this scheme can be applied to, we can expect considerably higher spectrum efficiency and fairness in the resource utilization. Practically, this would mean a larger number of RDs served in a specific geographical area (e.g., buildings, public squares) and a better QoS for each of those RDs. This is evident from Figure 62, where the SINR gain of the proposed scheme over the conventional one (e.g, no multi-RAT access) ranges from 1 to 4 depending on the available WLAN Aps that are available in the area. RERUM aims at demonstrating

how network operators can deploy IoT use cases without wasting valuable resources, especially those that are quite costly, like mobile radio resources, while on the other hand avoid overcrowding WLAN.

Furthermore, we can observe that deploying hundreds of RDs in a specific geographical area without any provisioning on the access network, it will cause considerable interference, network congestion and bad QoS. IoT services is different than conventional telecom services and these differences must be taken into account when designing the access network, e.g., by applying smart network cooperative techniques and resource allocation among different RATs.

4.2 Overhead

4.2.1 Introduction

Over the last decade, cellular networks have evolved from providers of ubiquitous coverage for voice-communication services to "anywhere-anytime-available" access ports for high data rate, Internet-based data services. A 3-order of magnitude increase in the supported data rates has been achieved (from several kbps in 2G-GPRS up to tens of Mbps in the latest LTE systems). Nevertheless, even these 4G data rates may soon prove inadequate, since the need for mobile data capacity is growing at an unprecedented extremely fast pace. Recent market studies, conducted by global organizations, wireless fora telecom companies, and operators, have indicated that mobile data traffic is (at least) doubled every year. Projecting this demand a decade ahead, we are faced with the so-called "1000x data challenge" or "capacity crunch", which should be efficiently dealt with by future service providers. Based on shorter-term forecasts, the study of predicts a factor of 13 growth in mobile data for the 2012-2017 period. This increase is justified by several trends, such as:

- the increase in the number of mobile devices, as by 2017 there will be 1.4 devices/holder;
- the increased penetration of machine-to-machine (M2M) devices, as billions of low data-rate devices with cellular connectivity are expected to be deployed and operate in the foreseen future;
- the increase in the usage of high-end portable devices like tablets and smartphones, as each smartphone is expected to generate more than 2.7 GB per month (contrary to today's 350 MB/month figure) by 2017;
- the shift to data-hungry mobile video services, as currently half of the mobile traffic is video and in five years, it will have dominated the total load, possessing approximately the two thirds of it.

Distributed Multiple Input Multiple Output (D-MIMO) networks have attracted great research interest for their potential to satisfy the very high data rates demands of future wireless networks and exploit the large number of deployed devices. Depending on the ratio between the number of the access nodes and terminals, as well as the kind of information that is shared among the network elements, several different techniques have been proposed, which aim at the interference mitigation or the sum-rate scaling, such as interference alignment (IA), joint multiuser beamforming (JMB) and dirty paper coding (DPC). A substantial amount of information must be shared among the network elements, for performing various operations (such as optimal spatial processing, beamforming, power allocation, etc.). This information is required for CSI estimation, time and frequency synchronization, data sharing (cooperative transmissions) etc., and the required overhead signaling significantly increases with the number of access nodes and terminals. Using the assumptions presented in [PH2012, Section II, System model] Figure 63 illustrates a simple example for the overhead bits that are required for exchanging channel information among nodes in a conventional distributed MIMO system. As seen, there is an optimum number of the MIMO order where the effective throughput is maximized. From that point, adding a new node to the distributed MIMO system results in the reduction of the effective throughput, i.e., the signalling information is greater than the actual information bits.

In this sense, when all the transmitters are communicating during the data portion of a frame, the effective sum-rate is reduced by a non-negligible factor when compared to the information-theoretic sum-rate. In this context, the overhead signaling reduction of the D-MIMO gained increased interest and several techniques have been proposed towards this aim. In [PH12], a novel concept was introduced, where a D-MIMO network employing IA, is partitioned into orthogonal groups (e.g., in a time division multiple access (TDMA) fashion), eliminating in this way any kind of interference in the network. Additionally, the overhead penalty on the sum-rate has been investigated, showing that the effective sum-rate goes to zero as the number of users increases, whilst the orthogonal partitioning is shown to increase the effective sum-rate. The partitioning of a D-MIMO network into orthogonal groups is a very promising concept for maximizing the effective sum-rate.

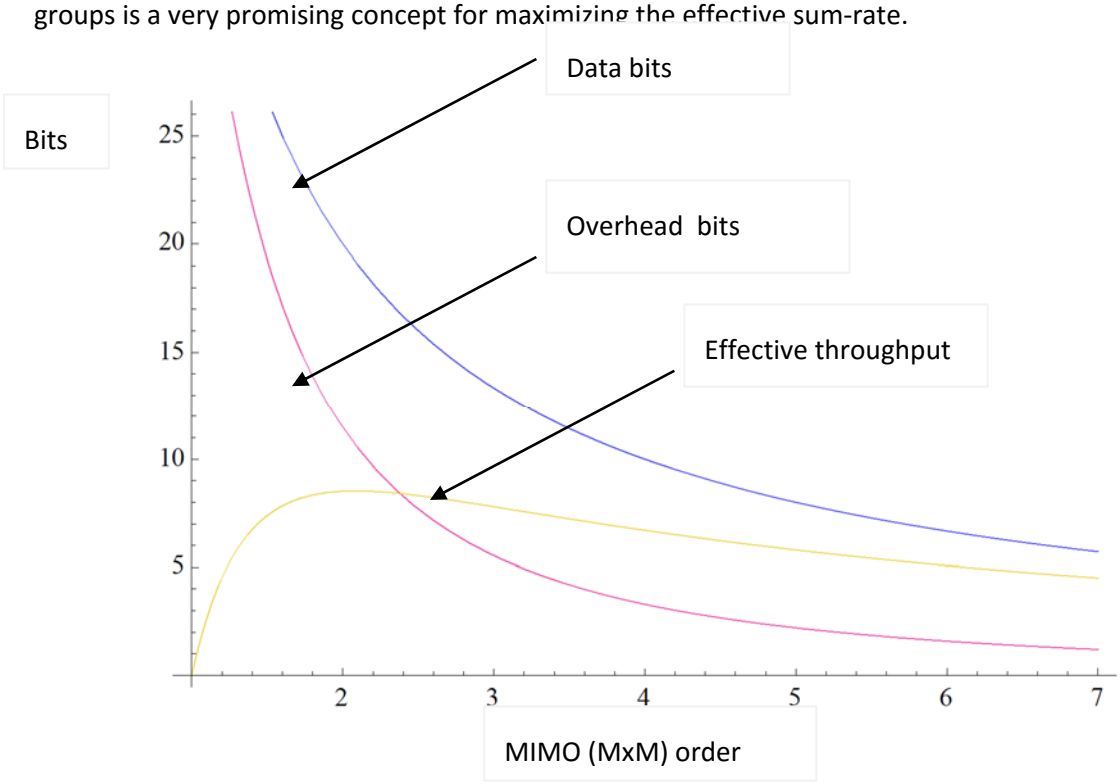


Figure 63: Example of overhead for a simple MIMO system

In this section we will investigate the impact of those D-MIMO based techniques on the overhead signalling within the network. Furthermore, we examine the optimal partitioning in terms of maximum effective sum-rate and based on the maximum allowed portion of the frame that is available for overhead.

4.2.2 Relation to RERUM UCs

The scope of the proposed scheme is to reduce the overhead in distributed MIMO deployments. Practically, this means that it can be used in use cases where hundreds of sensor devices are scattered in an area and transmit the sensed data to an aggregation point, while these devices can communicate also with each other (device-to-device) in order to share data. Then, these devices can send the sensed information to the aggregation point using D-MIMO transmission techniques, in order to minimize the transmission data and hence increase their energy efficiency.

In this sense, the proposed mechanism could be applied in the RERUM UC-O2 (Environmental monitoring), where the number of deployed devices can be very large, while the battery lifetime is something crucial. Furthermore, such static deployments are characterized by almost steady channel conditions, a feature that makes this mechanism the appropriate way to achieve high spectrum/energy efficiency. In terms of RERUM architecture, this scheme is part of the Communication manager functional component.

4.2.3 System model

The downlink of a D-MIMO network is considered, where K distributed GWs (transmitters) communicate with M = K spatially distributed single-antenna RDs through a time-varying fading channel (Figure 64).

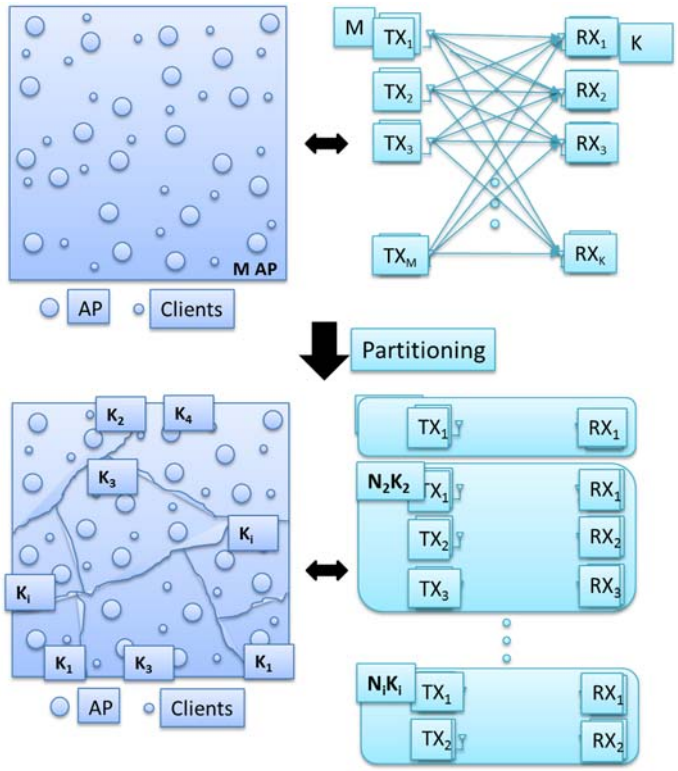


Figure 64: Partitioning a hybrid access network for reducing signalling overhead. The RDs form partitions and then use D-MIMO techniques to transmit data

The following assumptions have been made:

- The transmit antennas are connected through a high throughput backhaul.
- A time-varying fading channel is considered, with the fading amplitudes at each path being Rayleigh distributed.

Transmissions and receptions are synchronous

- Perfect channel state information is available at the transmitters
- We assume an error-free, zero delay feedback link from each node to the GW
- The transmitter has an average power constraint.

The GWs employ the zero forcing beamforming technique, where the beamforming weights (Wk) are appropriate selected in order to satisfy the zero-interference condition. The zero-interference can be obtained using the pseudoinverse of the channel gain matrix as weights.

Regarding the overhead model, i.e., the information that has to be exchanged between the nodes in order to realize the network MIMO, we follow the model of [PH12], where the communication is divided into frames of T symbols duration, with each frame consisting of two parts. The first part is devoted to overhead, which includes symbols required for training, feedback, synchronization etc, while the second part is utilized for data transmissions (Figure 65).

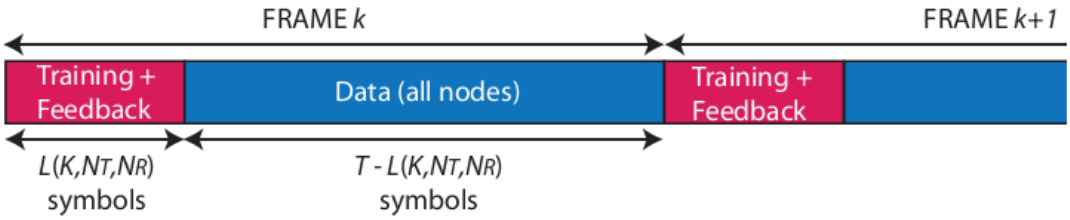


Figure 65: Frame structure - overhead and data

4.2.4 Reducing the overhead

In this section, we consider the practical important case, where a maximum allowed overhead size is taken into consideration. Thus, the overhead portion of the frame is less or equal to a predefined threshold α_{th}, *i.e.* $\alpha < \alpha_{th}$. In that case the bounded Knapsack problem (BKP) can be applied as follows:

Given n item types and a knapsack with p_j, profit of an item j, w_j, weight of an item of type j, b_j, upperbound on the availability of items of type j, c, capacity ofthe knapsack, select a number $x_j (j = 1, \ldots, n)$ of items of each type so as to maximize $z = sum(p_j x_j)$ subject to $sum(w_j x_j) \leq c$. It is noted that the BKP is a generalized of the zero-one knapsack problem and can be simplified to the latter one using the algorithm presented in Figure 66. In our case our objective is to obtain the optimal network partitioning that maximizes the effective throughput subject to a predefined threshold for the maximum allowed overhead. Hence, after the BKP has been transformed to a zero-on Knapsack problem, the exact optimal solution coincides with the one provided via the Greedy-Split algorithm.

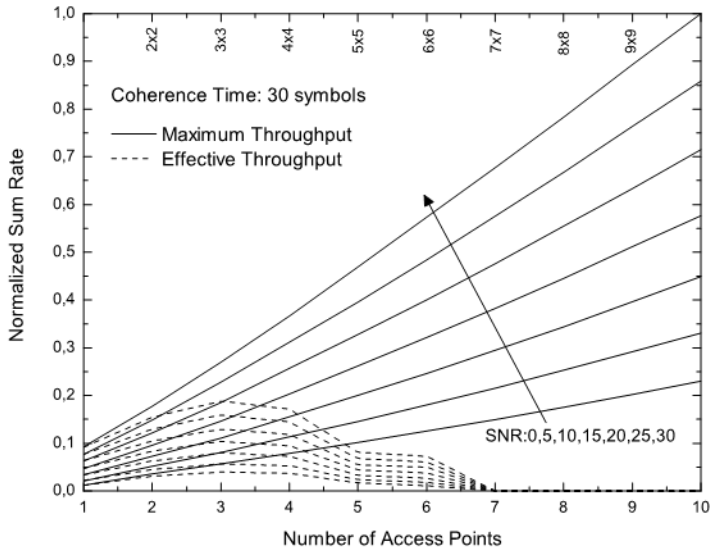


Figure 66: The impact of overhead as the number of access points (e.g., RERUM GWs) increases

Table 23: SINR optimization with RDs fairness

input: n, p_j, w_j, b_j
output: $\hat{n}, \hat{p}_j, \hat{w}_j$
begin
$\hat{n} := 0$
for $j := 1$ to n **do**
begin
$\beta := 0$
$k := 1$
repeat
if $\beta + k > b_j$ then $k := b_j - \beta$
$\hat{n} := \hat{n} + 1$
$\hat{p}_{\hat{n}} := kp_j$
$\hat{w}_{\hat{n}} := kw_j$
$\beta := \beta + k$
$k := 2k$
until $\beta = b_j$
end
end

4.2.5 Discussion

In this section, selected numerical performance evaluation results are presented and discussed in order to demonstrate the scalability performance of cooperative network access techniques. These results include performance comparisons, in terms of effective sum rate, of various communication scenarios. The parameters consider in all cases are: single antenna and randomly deployed GWs and RDs and independent and identical distributed Rayleigh fading conditions. In Figure 66 the normalized sum rate is plotted as a function of the number of the GWs for various values of the SNR and the CCT. The normalized sum rate has been evaluated for both cases of maximum achievable sum rate (ideal case without overhead) and effective sum rate (with optimum partitioning). In all cases in this figure, the performance improves as the SNR increases, while for the maximum NSR the performance increases as the number of GWs increases. Considering the effective normalized sum rate it is important to note that the linear scaling of the sum-rate is not maintained as the number of GWs increases, due to the overhead, whilst the performance considerable improves as CCT increases.

RERUM (and in general IoT) deployments, where hundreds of static devices are scattered within a geographical area are expected to benefit from this overhead reduction mechanism, since they can exploit the advantages that D-MIMO offers, while minimizing the overhead that is needed. Practically, for uses cases like RERUM UC-O2, this means that the RDs can exploit their geographical distribution in order to cooperate and improve the spectral efficiency/energy efficiency.

Furthermore, we have observed that while cooperative network access may result in important throughput improvements, there is a limitation in practical deployments. Deploying hundreds of RDs without any provisioning on the access network, it will cause considerable interference, network congestion and bad QoS. Thus, these smart access network techniques, as well as their limitations must be taken into account when designing the deployment of IoT services, in order to avoid bad quality of services.

4.3 Timeliness of information in dense networks

The concept of the Age of Information (AoI) was recently introduced by Yates et al. in [KGRK11] and then formalized in [KYG12]. This novel metric answers how fresh is a piece of information arriving at the receiver? It is different from the delay, since it includes the time from when a destination has received the last update about a particular piece of information (e.g. the temperature, the water flow/level etc.) from a source, which in our case is an RD. It also has a broader scope than the delay, since it measures a quality of a particular piece of information not a quality of the individual packets themselves. In certain UCs only the most updated measurement of a particular piece of information is relevant, e.g. a fire alarm triggered in an RD of the UC-I2 is of critical importance to have minimal delay from the moment it is captured at the RD to the moment that it will arrive at the MW server(s).

Here[1] we study a scenario where an RD node is immersed in a dense IEEE 802.11 WLAN, where a number of non-RERUM devices are associated. The RD tries to send information to the RERUM Middleware (MW). Dense WLANs are a specifical scenario that will be covered in the forthcoming IEEE 802.11ax HEW (High Efficiency WiFi) standard. The IEEE 802.11ah standard is also specifically designed for the IoT, hence of high relevance here. In this standard, an Access Point (AP) can cover up to 1 km in range, and it is fair to assume that overlapping networks with hundreds of devices would not be uncommon. Devices will have to compete for the channel with possibly hundreds of other devices (RDs and user access devices), with a very hetereogeneous population of traffic patterns. For example, there could be devices trying to offload traffic from the existing cellular infrastructure (see, for example, Section 4.4), futher congesting existing IEEE 802.11 WLANs, as in the 5G HETereogeneous NETwork (HETNET) paradigm. Competing with numerous devices degrades both throughput and delay performance, due to the increasing number of collisions, and in case of traffic burstiness, increases the idle time.

4.3.1 Relevance to RERUM's UCs

In this work, we let the MAC know about the "freshness" of a packet received from the application layer, along with the particular application that generated it, in order to develop a strategy to minimize the AoI at the receiver. Briefly, it will always try to send the packets carrying the freshest update of that particular information, without trying to transmit (or re-transmit) older packets. LUPMAC is therefore not for general purpose use in monitoring. However, such a policy, which is applicable at the network manager of the RERUM Architecture can fit all use cases in accomodating critical importance messages such as, for example, in UC-I2 a message that can indicate a fire alert due to rapid temperature increase, or in UC-O1 a set of messages that might indicate a traffic accident.

4.3.2 Age of information as a measure of timeliness

Assume a packet with the desired information I is generated at time t_{i-1} from a a source sensing that information. An example curve of the age of information I over time is depicted in Fig. x above, with

The receiver receives it at time t'_{i-1} s. The packet will then have an age of $\varepsilon_{i-1} = t'_{i-1} - t_{i-1}$ s, so the age of the information I will be at that time ε_{i-1} s. Then, if it is not receiving new packets, the AoI will increase over time with slope of one unit per time unit. The next packet carrying the updated information I is generated from the transmitter at time t_i s. It is received at time t'_i s. The age of that packet would then be $\varepsilon_i = t'_i - t_i$ s. If this packet is fresher than the current AoI (i.e. $\varepsilon_i < t'_i - t'_{i-1} + \varepsilon_{i-1}$) then the AoI will jump down to ε_i seconds, otherwise it will continue increasing. The AoI will continue to have this characteristic sawtooth behaviour, and it is possible to reconstruct its curve by interpolating between the various samples when packets are received. Then it is possible to

[1] This work has been accepted for presentation at the ICT 2016 [F++16], and has been done in collaboration with the University of Lund, through the ELLIIT excellence center.

reconstruct various metrics; for example, it is possible to reconstruct the average AoI by calculating the integral over time of the curve as a sum of trapezoids and dividing over the elapsed time.

4.3.3 LUPMAC: Latest UPdate MAC

In LUPMAC (i.e. the scheduler of the Communication Manager –see Fig. 19 D.2.5 we consider that the MAC is aware of the time a packet is generated in the upper layer (IoT Resource for RDs / or Service Manager for the MW backend side). If we assume the sources sending the respective pieces of information do not scramble the order of the generated packets, LUPMAC can simply assume the newest packets from the source are also the freshest. The applications running in an RD all map to an physical entity observed/monitored, and have an ID based on the VE . The ID thus identifies one information stream. This ID is stamped into the packet at generation time, for example in a field in the header of the network packet. When a new packet p' arrives from the upper layer, the MAC inspects the packets in the transmission buffer $\mathcal{P}$, including the packet in backoff (i.e. the one at the front of the buffer queue), to check if there is one that has the same ID as the newly arrived packet. We call this subset $\mathcal{P}_i$, where i is the source ID. Then, the MAC checks each packet $p \in \mathcal{P}_i$; if p is older than p', it is substituted with a copy of p'.

In the IEEE 802.11 standards the access mechanism is the so-called Distributed Coordination Function (DCF). A frame (that encapsulates a packet) waits a random time before being transmitted. A frame in this state is said to be in "backoff". If a collision occurs after a backoff period, the frame goes again into the backoff state, with a longer period to wait (on average). After a number of retransmissions[2] the frame is dropped. In case of an heavily loaded network, there is always a chance that the packet in front of $\mathcal{P}$ has already been into several retransmissions. So the chance for this particular packet to be dropped is higher, with negative effects on the AoI at the receiver end. In order not to have a newer packet at the last stage of the backoff be thus penalized, if the only substituted packet is the one currently in backoff, a copy is appended at the end of $\mathcal{P}$. Also, in order to not have too many packets of a particular PE in $\mathcal{P}$, only two copies of a packet from a particular PE are allowed in the buffer. If there are no packets substituted, p' is appended at the end of $\mathcal{P}$.

In order not to transmit multiple copies of the same piece of information, upon the reception of an ACK for p' (i.e. p' is succesfully transmitted), LUPMAC will delete every packet in $\mathcal{P}$ having the same ID as p'.

It is important to point out that LUPMAC is not doing deep packet inspection in order to substitute or remove packets in the MAC buffer. The ID could be inserted in the packet header at the application layer, and then propagated all the way to the MAC layer in a field in the header. It is also unreasonable for applications in the RD node to scramble *the order* of the generated packets, so LUPMAC will just infer the freshness of the piece of information contained in the packet by the time it is received from the upper layer, i.e. the latest received packet is the freshest.

4.3.4 Setup and Results

We consider a scenario where an RD node is immersed in a dense IEEE 802.11 WLAN with no hidden nodes, in order to better inspect the effects of LUPMAC on the average AoI. Such scenarios could occur, for example, in UC-O2, where a public hotspot serves a large number of users, and RDs, using the existing infrastracture to send information remotely.

Here, an RD node is sending multiple information streams to a remote MW server. Consider that a packet from the RD has to be sent first via the wireless channel to the GW, then routed via the internet, to a remote MW server. An RD serves a number of applications, with each one of those measures a particular piece of information mapping to an PE, and sends updates about their own information to th MW server. In case LUPMAC is used, the applications running in the application layer insert their

[2] Seven retransmission, in the current basic access mechanism,

unique ID in a field in the packet header, that is propagated all the way to the MAC layer, in order to let LUPMAC know which application generated that particular packet. Then, there is a network layer, and then an IEEE 802.11 MAC, that holds the packets generated by the various sources in its buffer. Next there is an IEEE 802.11 PHY to access the channel.

Table 24: The LUPMAC algorithm

```
1: on event p' comes from the network layer do
2:     n ← 0
3:     for all p ∈ P do
4:         if p.id == p'.id ⋀ p'.age < p.age ⋀ n < 2 then
5:             Substitute p with a copy of p'
6:             n ← n + 1
7:     if n == 0 then
8:         Append p' at the end of P
9:     else if n==1 ⋀ p' is at the front of P then
10:        Append p' at the end of P
11: on event ACK received upon transmission of p' do
12:    for all p ∈ P do
13:        if p.id == p'.id then
14:            remove p from P
```

The RD node is competing for the channel with a number of contenders, each one having some service demand to be served by the GW. The contenders send requests to the GW, which in turn fullfills their requests by sending back content to them. They send relatively small packets for the request, and receive packets of various sizes back.

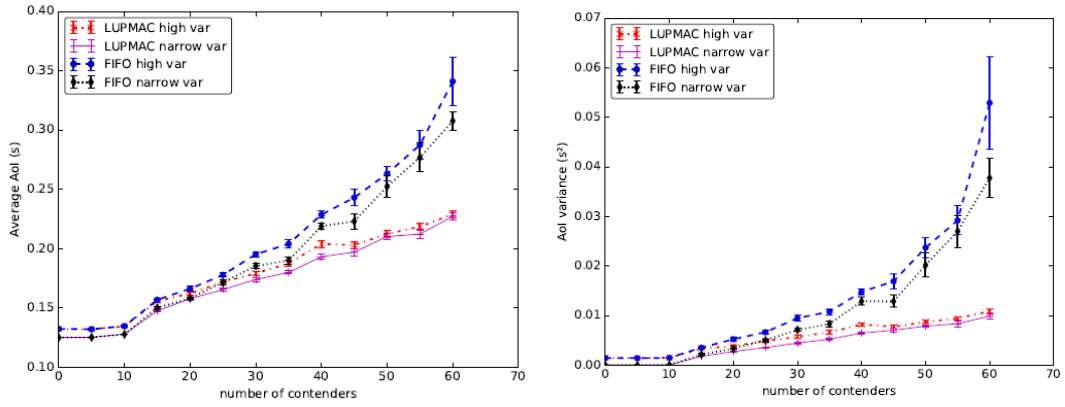
The remote internet link between the GW and the MW server introduces a delay according to a random distribution. Since the metro (or backbone) part of the network is usually reliable, at least in big cities, we will assume the remote link to be reliable, so no packet is dropped there. This models, for example, a routed path to a remote destination via the internet. On the other end, we tested the reliability of LUPMAC by measuring the average AoI both with high and low variance in the network part of the simulation.

For the results presented below contains the simulation parameters used. All plots are presented with 95% confidence, allowing for a sufficient warm-up period before taking measurements. There is an RD node uploading data to the MW server. The RD node is using an IEEE 802.11g WLAN with a number of contenders varying from 0 to 60. It is uploading small packets deterministically at a fairly slow rate (10 pk/s). The contenders are issuing requests to a remote server with exponentially distributed interarrival times, with an average rate of 100 pk/s, in order to increase the traffic load on the WLAN and congest it. The request packets are small (10 bytes on average, exponentially distributed), while the reply packets are uniformly distributed from small packets (14 bytes, a control frame) to big packets (1000 bytes). The delay on the wire connecting the access point to the remote server is considered to be a reliable metro/backbone connection. The average roundtrip time is however considered to be challenging with respect to VoIP traffic (150ms).

The average AoI and its variance are measured with an increasing number of contenders in the case that the delay has narrow variance, i.e. the one-way delay is uniformly distributed between 74ms and 76ms (so as to have an average roundtrip time of 150ms) with LUPMAC or the standard IEEE 802.11 FIFO approach. Then it is tested in the case it has a large variance, i.e. the one-way delay is uniformly distributed between 0s and 150ms (still an average round trip time of 150ms) with LUPMAC or the standard IEEE 802.11 FIFO approach. the AoI for all

Table 25: Simulation Parameters

	Parameter	Value
Physical	Frequency	2.4 GHz
	Noise Power	-110 dBm
	SINR Threshold	4 dB
	Transmission Power	20 mW
	Reception Threshold	-85 dBm
	Data Rate	54 Mbps
	Slot Time (σ)	9 μs
Scenario	Scenario dimensions	600 x 400 m
	Channel model	Free space
	Free space exponent	2
App	number of sensor nodes	1
	number of contenders	variable
	information generation (sensors only)	every 0.1 s
	request generation (contenders only)	$\sim$ exp{0.01} s
	Packet length (sensors)	10 bytes
	Packet length (contenders)	$\sim$ exp{10} bytes
	Requested packet length (contenders)	$\sim U(14, 1000)$ bytes
MAC	type	802.11g (AC1)
	buffer length (packets)	100



(a) (b)

Figure 67: Average Age of Information (a) and variance (b) measured at the destination with narrow variance on the wire delay (U(74ms; 76ms)) and high variance (U(0s; 150ms)) with LUPMAC or the standard IEEE 802.11 FIFO approach.

As we can see in Figure 67 the difference between high and narrow variance in the standard case (i.e. IEEE 802.11 FIFO) is quite small, only a fraction of the average AoI even with a totally saturated network with 60 contenders. In both cases the average AoI grows almost two tenths of a second from 10 to 60 contenders. This is quite a high increase, considering that the source on the sensor node is generating one packet every tenth of a second. As we can see, LUPMAC significantly improves the AoI in case of a highly saturated scenario (when the number of contending devices grows over 30), with an improvement of almost a tenth of a second with 60 contending devices on the average AoI. Also, the AoI appears more stable, as the variance grows much more slowly when LUPMAC is utilized at the RD communication manager. The improvement over the average AoI is extremely good, considering that the RD node generates one packet each tenth of a second. The improvement can be explained by the number of replaced packets in the MAC buffer when LUPMAC is used. In Figure 68 the percentage of

the replaced packets according to Algorithm 4 over the totality of packets sent by the application layer in the sensor is presented. As we can see, LUPMAC starts to replace packets in the MAC buffer as soon as we have a sufficiently high number of contenders in the WLAN (in this case ≥ 15), exactly when the average AoI starts to diverge from the one measured in the standard case (i.e. IEEE 802.11 FIFO).

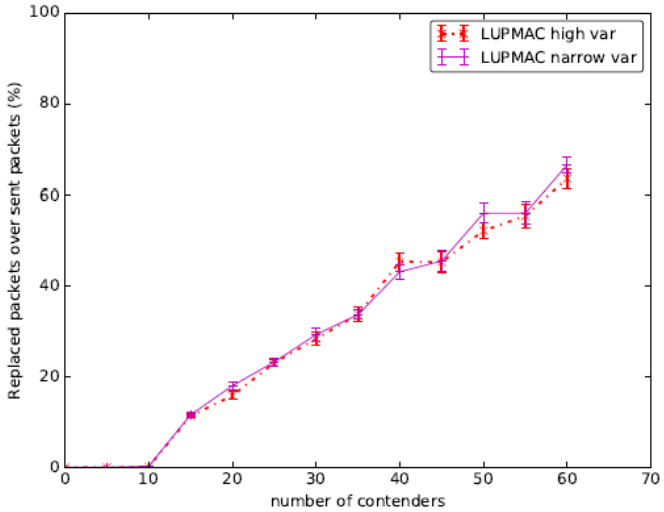


Figure 68: Percentage of the replaced packets according to Algorithm 1 over the totality of packets sent by the application layer in the sensor.

If we allow for a faster update generation, the benefits are overwhelming. In Fig. 5, the source on the sensor node is allowed to generate up to 100 pk/s, i.e. one packet every hundreth of a second with 30 contenders and narrow variance on the one-way network delay. Notice that the y-axis is in log-scale. When LUPMAC is enabled, the average AoI is improved by up to an order of magnitude compared with when the sensor is simply relying on an unmodified IEEE 802.11 MAC. In addition, when LUPMAC is used, the average AoI is fairly stable, and its variance limited.

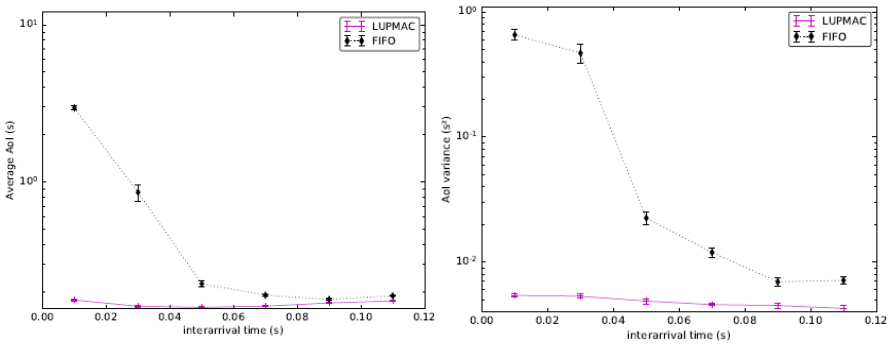


Figure 69: AOI benefits of LUPMAC vs. unmodified 802.11.

4.3.5 Discussion

We have extended previous AoI works with a more practical implementation for RERUM, towads investigating the performance of 802.11-based Access Points acting at the same time as RERUM Gateways. This means that we consider devices that the same time support general use networking devices while able to provide good quality of service to the RDs, under AoI criteria. We have done so by introducing a new cross layer approach between the application layer and the MAC layer, called Latest UPdate MAC (LUPMAC), aimed at modifying the existing IEEE 802.11 in order to minimize the average AoI at the receiver end.

Our results indicate that such an option is feasible for the REURM use cases considered. For the indoors the user density decsribed is in the acceptable region backed by our results (Figure 67 and Figure 68)

for even considering large installations – given the load incurred by our contending nodes. Sinilarly is the case for the outdoors UCs. When LUPMAC is enabled, the average AoI is improved by up to an order of magnitude compared with when the sensor is simply relying on an unmodified IEEE 802.11 MAC. In addition, when LUPMAC is used, the average AoI is fairly stable, and its variance limited

4.4 Load coupling Characterization

4.4.1 Offloading under cell load coupling

In a cellular-based IoT supporting network, frequency reuse is employed, and thus base stations or gateways using the same frequency band interfere with one another. The expected level of resource usage in the time-frequency domain of such a cell is its *load*.

To optimize system performance, load balancing has to be performed across the macro network and potential supporting infrastructure (small cells or WiFi), in the context of offloading. However, since the load of a cell depends on the load of other cells because of the coupling of interference and the requirement to serve a given per-cell demand to tune performance one has to balance loads and demands. The core problem is that a non-linear coupling relation of cells loads arises, making analytical characterization of the load hard.

4.4.2 Relevance to RERUM's UCs

Driven by UC-O1, we consider a separate scenario in which the cellular network offloads to a supporting secondary network of gateways. The suppporting network is either based on small cell, or more typically a WiFi network (for example in UC-O1 this could be WiFi-enabled gateways). Notice that in the case of a small-cell the network resources are shared with the cellular network, while for WiFi orthogonal network are in play, however a dense WiFi infrastructure also will suffer from the same load coupling.

4.4.3 Overall System and Offloading model

Here we denote a vector using bold lower case letters, say $\boldsymbol{x}$, a matrix by bold capital letter, $\mathbf{Y}$, and its $(i,j)^{\text{th}}$ element by its lower case y_{ij}. A *positive* matrix, denonted by $\mathbf{A} > 0$ iff $a_{ij} > 0$ for all $(i,j)_s$. Similarly, a *non-negative* matrix $\mathbf{A} \geq 0$ if $a_{ij} \geq 0$ for all $(i,j)_s$. Similar definitions apply to our vectors.

Consider a cellular network of n base stations that interfere with each other. We focus on the downlink communication scenarios, considering the feedback the users RDs receive in the UC-O1 case, comprising of heavy multimedia (maps with real time traffic estimation) where base station $i \in \mathcal{N} \triangleq \{1, \cdots, n\}$ transmits with power $p_i \geq 0$. We use the term "cell i" interchangeably with "base station i" and "user" with "RD", since each user is assumed to carry one RD. All transmit powers are in a vector $\mathbf{p} > 0$. BS i serves only one group of RDs $\mathcal{J}_i$, with $|\mathcal{J}_i| \geq 1$and each RD can get a rate of up to R_{ij} from its serving cell. Thus, the data to the RDs can be interpreted as best-effort, served subject to network conditions.

Data offloading is assumed, in the sense that the demand of each RD can be served by a "supporting" network (of say gateways WiFi gateways or Small Cell). We assume a total of such n' complementary cells in the supporting network, denoted by the set $\mathcal{N}' = \{1, \cdots, n'\}$. Each supporting cell i transmits with power $p_{Si} \geq 0$. We consider that every regular-cell user $j \in \mathcal{J}_i$ in the regular cell $i \in \mathcal{N}$ corresponds to a unique virtual complementary-cell user $b \in \mathcal{J}'_a$ in the complementary cell $a \in \mathcal{N}'$, using a mapping function μ: $(a,b) = \mu(i,j)$. We take the demands d_{ij} and $d_{\mu(i,j)}{}'$ served in the regular and complementary cells, respectively, to be the variables we want to optimized, subject to a total demand constraint: $d_{ij} + d_{\mu(i,j)} \leq D_{ij}$ ensuring that the total demand served is not more than the demand D_{ij} requested. All demands $\{d_{ij}\}$ and $\{d_{\mu(i,j)}{}'\}$ are in vectors $\mathbf{d} \geq 0$ and $\mathbf{d}' \geq 0$, respectively.

We assume there is at least one user j in cell i with $d_{ij} > 0$, otherwise $p_i = 0$ and so base station i could be omitted from the solution; likewise for the supporting cells. Thus without loss of generality, we have $p_i, p_{Si} > 0$.

We consider two types of supporting networks, either only small cells or WiFi. For small-cell offloading, both the regular cellular network and small-cell network use the same 3GPP frequency band, so, they interfere with each other. For WiFi offloading, the frequency band used in the WiFi network is orthogonal to that of the cellular network, hence there is no mutual interference. The hybrid case of using both can be a simple extension to the work presented here.

4.4.4 The Load Coupling Equation

We begin the analysis presenting the coupling model for a cellular network with no any supporting network to introduce the reader to the basic notions. Consider $\boldsymbol{l} = [l_1, \cdots, l_n]$ be the load of the cellular network, where $0 \leq \boldsymbol{l} \leq 1$. Load l_i is the usage faction of the shared resource in cell i. For example in LTE, this load could be interpreted as the expected fraction of time-frequency resources (resource blocks, or subcarriers in timeslots) that are scheduled to deliver data. Take the SINR of user j in cell i as $SINR_{(i,j)}(\boldsymbol{l}) = \frac{p_i g_{ij}}{\sum_{k \in \mathcal{N}\{i\}} p_\kappa g_{kj} l_k + \sigma^2}$, where σ^2 represents the noise power and g_{ij} is the channel gain from base station i to RD j. This SINR provides a decent approximation of more complicated cellular models. Intuitively, l_k can be interpreted as a probability of having active interference from cell k on all the sub-carriers of a resource block.

An achievable rate can be considered to be the one given by the Shannon equation, with: $r_{ij} = B\log\left(1 + \mathrm{SINR}_{(i,j)}(\boldsymbol{l})\right)$, where B is the bandwidth for one resource unit. To deliver a demand of d_{ij} nat for user j, the i^{th} base station therefore needs to use: $l_{ij} = \frac{d_{ij}}{r_{ij}}$ resource units. We assume that at total S (time and frequency) resource units are available. Summing over all users in cell i, we get the total load for the cell:

$$l_i = \frac{\sum_{j \in \mathcal{J}_i} l_{ij}}{S} = \frac{\sum_{j \in \mathcal{J}_i} \left(\frac{d_{ij}}{r_{ij}}\right)}{S} = \frac{1}{S \cdot B} \sum_{j \in \mathcal{J}_i} \left(\frac{d_{ij}}{\log\left(1 + \mathrm{SINR}_{(i,j)}(\boldsymbol{l})\right)}\right),$$

which is dependent on the entire load vector as one can directly see. To simplify notation, we can normalize d_{ij} and r_{ij} by the total resource units MB. So, without loss of generality we imply that $SB = 1$ henceforth.

Let $f_i(\boldsymbol{l}) \equiv l_i = \frac{1}{S \cdot B} \sum_{j \in \mathcal{J}_i} \left(\frac{d_{ij}}{\log\left(1 + \mathrm{SINR}_{(i,j)}(\boldsymbol{l})\right)}\right)$. Then in vector form $\boldsymbol{f}(\boldsymbol{l}) = [f_1(\boldsymbol{l}), \cdots, f_n(\boldsymbol{l})]^T$, which means $\boldsymbol{l} = f(\boldsymbol{l}, \boldsymbol{d}, \boldsymbol{p})$ showing clearly now the dependence of the the demand $\boldsymbol{d}$ and power $\boldsymbol{p}$ on the load and vice versa. Observe that in our **load coupling equation** load $\boldsymbol{l}$ appears in both sides and thus cannot be readily solved in closed-form. To indicate that a load vector is a solution of the load coupling equation, we shall denote it as $\boldsymbol{l}^*$ and we shall call it *feasible* if it is also $\boldsymbol{l}^* \geq 0$.

4.4.5 Load Coupling with Supporting Network

For a cellular network with small-cells, we can assume the two (macro/small) layers operating in the same frequency band and treat them as one *hetnet*: Simply assume that the set of n base stations in the macro network is combined with the set of n' base stations in the small-cell network to form a new set of base stations of size $n + n'$. All these will interfere with one another.

On the other hand is the supporting network is WiFi, the two networks operate in different bands. The WiFi network however also follows the load-coupling system relation for itself, separately from the cellular layer. For the load coupling equation one can then just use the notation $\boldsymbol{l}^{\boldsymbol{W}}, \mathbf{d}^{\boldsymbol{W}}, \mathbf{p}^{\boldsymbol{W}}$ for the WiFi corresponding quantities of $\boldsymbol{l}, \mathbf{d}, \mathbf{p}$. Note that regardless of whether the supporting network is a

small cell or WiFi, the allocation of $\{d_{ij}, d_{\mu(ij)}{}'\}$ is coupled due to the total demand constraint $d_{ij} + d_{\mu(ij)}{}' \leq D_{ij}$.

4.4.6 Offloading Demands

We model demand serving gains in a network with offloading via two utility functions in and formulate an optimization problem maximizing the sum utility where the supporting network is either a WiFi or small-cell network. Following, we introduce an algorithm to bound the maximum optimal load by the unit.

We consider a feasible load such that $\mathbf{x}^* \geq 0$ for this section.

Our objective is to maximize the sum utility $U_{tot} = \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{J}_i} k_{ij} U(d_{ij}) + k_{\mu(i,j)} U(d_{\mathrm{i,j}})$, where $U(d)$ is the utility function for satisfying demand d. Positive weight factors k_{ij} and $k_{\mu(i,j)}$ model the priority of RD and the networks: The utility function can thus be used to measure the value of serving demand d to the cellular operator (or user), for example, in terms of revenue from (cost of) the access service, and the **fairness** of serving the demand of multiple users within each cell type. Note that the importance of serving in either cell type can be quantified via the weights k_{ij} and $k_{\mu(i,j)}$.

Two representative functions, a linear and a logarithmic are used for utility. Respectively $U_{lin}(d) = d$ and $U_{log}(d) = \log(d)$. These functions are monotonically increasing and hence 1-to-1. The former models a case where serving each additional demand unit would result in an additional unit for utility. While under logarithmic utility, serving an additional demand unit of a user with a low demand results in more utility resuling in a fairer demand distribution among users but could result in a smaller revenue to an operator as less demand is served in total. Notice also that a logarithmic utility function will not assign zero demand to any user, since it would lead to negative infinity utility.

We make the *same-demand assumption* that every RD j in the same regular cell i is served the same demand $d_{ij} = \tilde{d}_i$. Corresponding to the regular-cell user j in cell i, we denote the supporting-cell user as $a(i,j)$ in complementary cell $b(i,j)$, i.e., $(a(i,j), b(i,j)) = \mu(i,j)$. For the supporting network, we also make the same-demand assumption, i.e., $d_{\mu(i,j)}{}' = \tilde{d}_{a(i,j)}{}'$ for all i,j. We thus focus on varying the cell-level demand vectors $\tilde{\mathbf{d}} \triangleq [\tilde{d}_1, \cdots, \tilde{d}_n]^T$ and $\tilde{\mathbf{d}}' \triangleq [\tilde{d}_{1'}, \cdots, \tilde{d}'_{n'}]^T$. From the demand constraint, we get that:

$$d_{ij} + d_{\mu(i,j)} = \tilde{d}_{\mathrm{i}} + \tilde{d}_{\alpha(i,j)} \leq D_{ij} \ \forall j, \forall i$$

Since all cells are active with power vector $\mathbf{p} > 0$, we also have $\tilde{\mathbf{d}} > 0$ and $\tilde{\mathbf{d}}' > 0$.

For logarithmic utility, the same-demand assumption can be relaxed. Instead we assume more generally that each user $j \in \mathcal{J}_{\mathrm{i}}$ in cell i is allocated a demand of $d_{ij} = \alpha_{ij}\tilde{d}_i$, where $\sum_{j\in\mathcal{J}_{\mathrm{i}}} \alpha_{ij} = 1$. Here, α_{ij} is a fraction of the total demand $\tilde{d}_i$ served in cell i. Thus, the user's achieved utility is $U(\alpha_{ij}\tilde{d}_i) = \log(\tilde{d}_i) + \log(\alpha_{ij})$. Having α_{ij}'s fixed, we need only consider the first term $\log(\tilde{d}_{\mathrm{i}})$ for the sum utility U^{tot}. Hence the optimization problem is similar to the case under the same-demand assumption.

4.4.7 Optimizing offloading

Here we formulate the optimization problem with WiFi as the secondary network.

$$\max_{\tilde{\mathbf{d}}\tilde{\mathbf{d}}'} \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{J}_i} k_{ij} U(\tilde{d}_i) + k'_{\mu(i,j)} U(\tilde{d}'_{a(i,j)})$$

$$s.t. \quad \tilde{\boldsymbol{d}} \in \{\tilde{\boldsymbol{d}} > 0 : r(\boldsymbol{\Delta}(\tilde{\boldsymbol{d}})) < 1\}$$

$$\tilde{\boldsymbol{d}}' \in \{\tilde{\boldsymbol{d}}' > 0 : r(\boldsymbol{\Delta}'(\tilde{\boldsymbol{d}}')) < 1\}$$

$$\tilde{d}_i + \tilde{d}'_{a(i,j)} \leq D_{ij} \ \forall j, \forall i$$

where $\boldsymbol{\Delta}(\boldsymbol{d})$ is a n × n matrix with

$$\delta_{i,k} = \begin{cases} \sum_{j \in \mathcal{J}_i} g_{kj} d_{ij} / g_{ik} \,, i \neq k \\ 0 \qquad\qquad\qquad\quad , i = k \end{cases}$$

For the cellular network, and $\boldsymbol{\Delta}'(\boldsymbol{d})$ similarly defined for the WiFi case, and $r(\boldsymbol{A})$ the spectral radius (maximum eigenvalue) of matrix A.

We can transform $\tilde{d}_i$ to $y_i = U(\tilde{d}_i)$ for $i \in \mathcal{N}$ and let $\tilde{\mathbf{d}} = [U^{-1}(y_1), \cdots, U^{-1}(y_n)]^T \triangleq \mathbf{g}(\mathbf{y})$. Note that since $U(\cdot)$ is a monotonic function the inverse $U^{-1}(\cdot)$ does exist. Likewise for WiFi cells, we let $y'_i = U(\tilde{d}'_i)$ for $i \in \mathcal{N}$ and $\tilde{\mathbf{d}}' = [U^{-1}(y'_1), \cdots, U^{-1}(y'_n)]^T \triangleq \mathbf{g}(\mathbf{y}')$. Let $k_i \triangleq \sum_{j \in \mathcal{J}_i} k_{ij}$. We make similar definitions for $y_{i'}$ and $k_{i'}$ corresponding to the complementary cells. Our *transformed data offloading* problem written down as:

$$\begin{aligned} &\max\nolimits_{\boldsymbol{y},\boldsymbol{y}'} \sum_{i \in \mathcal{N}} k_i y_i + k'_i y'_i \\ &s.t. \quad y \in \{\boldsymbol{y} \in \mathcal{Y}^n : r(\boldsymbol{\Delta}(\boldsymbol{g}(\boldsymbol{y}))) < 1\} \\ &\qquad\; y' \in \{\boldsymbol{y}' \in \mathcal{Y}^{n'} : r(\boldsymbol{\Delta}'(\boldsymbol{g}(\boldsymbol{y}'))) < 1\} \\ &\qquad\; U^{-1}(y'_1) + U^{-1}\left(y'_{a(i,j)}\right) \leq D_{i,j}, \forall i,j \end{aligned}$$

where $\mathcal{Y}^n$ of dimension n denotes the set of positive vectors for the linear utility, and as the set of real vectors for the logarithmic one. Here $\widetilde{\boldsymbol{\mathcal{F}}} \equiv \{\boldsymbol{y} \in \mathcal{Y}^n : r(\boldsymbol{\Delta}(\boldsymbol{g}(\boldsymbol{y}))) < 1\}$ (and similarly $\widetilde{\boldsymbol{\mathcal{F}}}'$) are the transformed feasibility sets. Observe that the objective function is always linear in $\mathbf{y}$ and $\mathbf{y}'$. For any of the two utility functions, the set of $\tilde{\mathbf{y}}, \tilde{\mathbf{y}}'$ subject only the second constraint is convex. Therefore if $\widetilde{\boldsymbol{\mathcal{F}}}$ and $\widetilde{\boldsymbol{\mathcal{F}}}$ are then we have a convex optimization problem for which numerically efficient solvers exist.

4.4.8 Algorithm bounding the maximum load

In our analysis, we consider feasible load $\boldsymbol{l}^* \geq 0$, which holds in the exact formulation of our optimization problem. By the definition, the load cannot exceed one, even in practice, due to limited availability of network resources. To impose a constraint $0 \leq \boldsymbol{l}^* \leq 1$ explicitly in the optimization problems above can lead to solution instabilities. Here we propose an iterative algorithm that reduces the demand ensuring $0 \leq \boldsymbol{l}^* \leq 1$.

For the presentation of the algorithm consider the following generalized problem: Assume $0 \leq \rho \leq 1$, an arbitrary but fixed constant. Replace in the the first optimization problem of 4.4.7 the spectral-radius constraints by $r(\boldsymbol{\Delta}(\mathbf{g}(\mathbf{y}))) < \rho$ and $r(\boldsymbol{\Delta}'(\mathbf{g}(\mathbf{y}'))) < \rho$, respectively and denote the corresponding feasibility sets as $\boldsymbol{\mathcal{F}}(\rho)$ and $\boldsymbol{\mathcal{F}}'(\rho)$, respectively. Obviously, the newly defined optimization problem specializes to our original problem if $\rho = 1$. Now, the optimal demand vector and load vector are denoted respectively as $\mathbf{d}^*(\rho)$ and $\boldsymbol{l}^*(\rho)$ for the regular cellular network, and similarly $\boldsymbol{d}'^*(\rho)$ and $\boldsymbol{l}'^*(\rho)$ for the WiFi network. Finally, we denote the maximum optimal load as $l^*_{\max}(\rho) \triangleq \max\{l^*_i(\rho), l'^*_j(\rho), i \in \mathcal{N}, j \in \mathcal{N}'\}$. Thus, $\boldsymbol{l}^*(\rho) \leq 1$ if and only if $\boldsymbol{l}^*_{\max}(\rho) \leq 1$. Independent of the actual value of ρ, the numerical solution for the new problem can be obtained similarly as for the original. If $x^*_{\max}(1) \leq 1$, then $\mathbf{x}^*(1)$ is an optimal solution for the original problem satisfying the required constraint $0 \leq \mathbf{x} \leq 1$.

To ensure the load is limited by one, we propose to use the demand vector $\mathbf{d}^*(\rho)$ corresponding to the load vector solution $\boldsymbol{l}^*(\rho)$ in the problem we just formulated where ρ is determined by the solution of the following optimization problem:

$$\begin{aligned} &\max\nolimits_{0 \leq \rho < 1} \rho \\ &s.t. \quad l^*_{\max}(\rho) \leq 1 \end{aligned}$$

Although $l^*_{\max}(\rho)$ is not always a monotonic function of ρ, having reduced the optimization to only one variable, even an exhaustive search based on a finely-quantized interval over $0 \leq \rho < 1$ could be run to solve this. For each ρ the generalized problem at the start of this section is solved.

4.4.9 Setup and Results

Having end user fairness in mind, we present numerical results assuming the utility function is the logarithmic utility. The optimal demand vectors $\tilde{\mathbf{d}}^*, \tilde{\mathbf{d}}'^*$ can be solved efficiently by standard numerical solvers. Specifically, we use the active-set algorithm with the $\mathrm{fmincon}$ function in the Matlab software. Our work applies regardless of where base stations, gateways, and RDs are deployed.

We consider a regular cellular network consists of $n = 9$ cells, where each cell is of two unit radius. The cells are arranged uniformly in a 3x3 lattice. A base station is placed in the centre of each cell and we assume there are four WiFi enabled gateways covering the cell completely. For each WiFi cell, there are 5 RDs to serve. A GWis placed in the center of each WiFi cell.

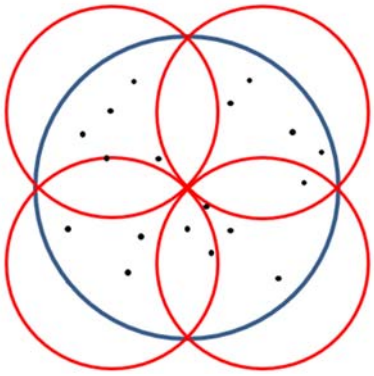


Figure 70: A simplified building block of a cell (blue) and four WiFi GWs (red) and 20 RDs (black)

Every user is served by one WiFi cell (subject to a fixed SINR criterion) and the base station cell that it resides in. Thus, every access point can support up to 5 users while every base station can support up to 20 users. The building block of our 3x3 topology is shown in figure? We make the same-demand assumption that all users in the same (regular or WiFi) cell are allocated the same demand.

We use the same weight $k_{ij} = 1$ for all base stations, and the weight $k_{ab}{}' = 0.25$ for all GWs; We set the (normalized) transmission power of every serving cell as 100, the transmission power of every GW cell to be 1. The noise has variance of 0.01. The channel gain from the ith regular cell to the jth user is fixed as $g_{ij} = x_{ij}^{-\kappa}$ where x_{ij} is the distance between transmitter i and receiver j, and $\eta = 4$ is the path loss exponent. The channel gains for the WiFi are obtained similarly.

Fixing the demand to 0.1 for all cells we solve the problem numerically and obtain the optimal load applications as below, for which we end with demand allocation 0.050 for all serving BSs and GWs. In Table 26 **Optimal load allocation sample** we see that all cells opearte below maximum load and the non-uniform load distribution is a result of the non-uniform RD distribution. The equal demand allocation follows from an equal share per RD from the BS and AP serving it and achieves the desired fairness.

Table 26 Optimal load allocation sample

	BS' load (sorted)	NE GW	NW GW	SE GW	SW GW	Average AP Load	Min AP Load	Max AP Load
Min	**0.135**	0.021	0.023	0.038	0.032	0.029	0.021	0.038
	0.147	0.032	0.040	0.029	0.035	0.034	0.029	0.040
Avg. Load	0.152	0.037	0.040	0.035	0.024	0.034	0.024	0.040
0.187	0.188	0.023	0.035	0.034	0.034	0.032	0.023	0.035
	0.192	0.030	0.031	0.039	0.037	0.034	0.030	0.039
	0.212	0.030	0.042	0.026	0.054	0.038	0.026	0.054
	0.212	0.045	0.025	0.037	0.038	0.036	0.025	0.045
	0.214	0.037	0.027	0.027	0.026	0.029	0.026	0.037
Max	**0.227**	0.044	0.041	0.038	0.056	0.045	0.038	0.056

4.4.10 Discussion

Having end user fairness in mind, we used as utility function the logarithmic utility and presented a sample result, indicating an equal allocation of loads. The optimal demand vectors $\tilde{\mathbf{d}}^*$, $\tilde{\mathbf{d}}'^*$ were derived efficiently by standard numerical solvers. Specifically, results were obrained using he fmincon function in Matlab, which uses the active-set algorithm. Our work applies regardless of where base stations, gateways, and RDs are deployed. For RERUM we have provided an accurate and efficient tool to identify the optimal allocation of loads to an RD network supported by small cell or wifi infrastucture.

5 Scalability

5.1 Scaling of the leak-resilient message authentication codes

5.1.1 Introduction

In order to quantify how well the secure communication supported by the low-powered devices scales with their increasing number and how it adapts to various mobility ratios within the underlying sensor network – mainly represented by the RDs and the RD-to-RERUM Gateway connections – simulations with active and inactive security mechanisms were performed and analyzed. More specifically, we assumed a simulation scenario replicating various expected number of RDs, mobility models and message rate of selected public use cases under both unsecured and security-enhanced communication routed by the RPL protocol, while considering appropriate network performance metrics (as described in Section 5.1.5).

Deliverable D5.3 has presented and evaluated two mechanisms, which required significant per message computational efforts during the message generation time: (i) the tag-generation of the leakage-resilient messages authentication protocol (LR-MAC), and (ii) the signature generation of the Malleable Signatures (MS). We do not consider the case of the DTLS mechanisms, for example, since the asymmetric DTLS handshake should be performed at a bootstrapping or idle stage of the network setup and only requires symmetric encryption when sending, in addition to a small overhead in the number of transferred bytes.

From the two mechanisms, the MS is taxing in both computational time and size overhead (D5.3 reported of a factor nine increase in message length, and a factor greater than three over classic ECC signatures), and are therefore to be only invoked for selected data and requests. We will therefore focus on this chapter on the scalability of the proposed LR-MAC protocol.

5.1.2 Relation to RERUM UCs

The data integrity of messages and the proof of their authenticity are two of the key requirements for security mechanisms in a communicating network, achieved through the usage of message authentication codes (MACs).

In the context of IoT devices forming such communicating networks, as those implied by the four use cases defined by RERUM – UC-O1 (Smart transportation), UC-O2 (Environment monitoring), UC-I1 (Home energy monitoring) and UC-I2 (Comfort quality monitoring) – their exposure significantly increases the risk of them being subjected to side channel attacks, allowing the extraction of secret data (referred to as *leakage*). Towards managing such attacks, the University of Bristol has proposed a leakage resilient MAC (the LR-MAC).

Security and privacy-enhancing techniques produce a measurable overhead when performed on resource-restricted nodes such as the RDs. The scale of the overhead depends on various factors, such as the mathematical concepts on which the techniques are based (e.g. elliptic curves, factorization challenges such as RSA, substitution-permutation networks such as AES, etc), the used key sizes (from 128 to 4096 bits), and the length of the data on which they are applied; and is measurable in the time needed to perform the required computations, in the resulting message size and in the resource consumption (RAM, code size).

For a network of devices, the security- and privacy-imposed overhead extends beyond the previously described device-specific metrics, towards network-specific measurable metrics - of which some are presented in Section 5.1.5: average delivery ratios, average network connectivity, or the average end-to-end delays. Through carefully constructed simulations and such metrics we can furthermore observe and evaluate the scalability of the network in aspects such as the increasing number of security-enabled nodes or their mobility is concerned.

This is of critical relevance to all of RERUM's use cases, which assumes a good scalability of the security mechanisms, having to adapt and perform under a dynamic network of RDs: (i) input is expected from infrastructure and user nodes, both of which can be either stationary or mobile (e.g. users voluntarily providing data in outdoor environments, in-transit busses etc.). (ii) And second, a variable number of nodes is plausible when considering a house or a housing complex hosting multiple smart metering and personal area network-service oriented nodes, as well as a busy junction containing tracked busses, observation stations and commuters with enabled data-sensing devices (e.g. mobile phones).

5.1.3 Simulation Scenario

The definition of the simulation scenario proved challenging, as it is supposed to provide a context in which a variable number of RD nodes, mobility and exchanges are justified. We also tried, while retaining relevance to the four use cases defined in D2.1, to provide a scenario that exemplifies the practicality of the security mechanisms outside of RERUM also.

We therefore consider the following simulation scenario: a public or open space of up to 250 by 250 m, with a single active sink, represented by the RERUM Gateway. Within this area, 20 to 40[3] RDs with partial mobility (15, 30 and 50 percent) periodically send readings to the RERUM Gateway.

5.1.4 Simulation Parameters

The simulation was performed using the Contiki Cooja simulator. Table 27 provides an overview of the overall simulation duration and on the number or repetitions: 600 seconds per simulation, in order to allow the mobility to reach a steady-state distribution, and 25 repetitions per scenario.

Table 27 Simulation description

Simulation Parameter	**Value**
Simulation time	600 seconds
Simulation repetitions	25
Simulator	Cooja

A description of the network settings is provided in Table 28: we performed simulations for a total of 21, 26, 31, 36 and 41 communicating RD devices – in each of the five cases, one of devices was a static data sink representing a RERUM Gateway. The traffic was simulated in CBR mode, with the frequency of a message per minute.

Table 28 Network simulation parameters

Simulation Parameter	**Value**
Number of nodes	21, 26, 31, 36, 41
Number of sources	20, 25, 30, 35, 40
Number of sinks (gateways)	1
Traffic pattern	CBR – every minute

[3] The range (20 to 40) represents an estimated number of communicating devices available per a RERUM Gateway at one point in an indoor scenario, under real-world conditions.

MAC	CSMA
Duty cycling	ContikiMAC
PHY	IEEE 802.15.4
Radio medium	Unit Disk Graph Medium (UDGM) – distance loss
Transmission range	TX: 50 *m*, Interference: 100 *m*

Finally, Table 29 provides an overview of the chosen mobility settings. Since we did not assume that the RDs are moving in groups, we opted for the Random Waypoint Model (Christian Bettstetter, 2004). Since the *naive* Random Waypoint Model might lead to unreliable routing metrics, we included the recommendation made in (Jungkeun Yoon, 03) concerning simulation time, speed and settling time.

Table 29 Mobility simulation parameters

Simulation Parameter	**Value**
Mobility	Partial
Number of mobile nodes	15%, 30% and 50%
Mobility pattern	Random waypoint model
Mobility speed	0.2 - 1 m/s
Pause time	5-60 seconds

We chose to simulate the overhead induced by the 128 bit LR-MAC. The timings results and overhead are presented in D5.3. In order to present a worst-case scenario, we chose not to consider the multitasking nature of Contiki, and therefore to block all radio communication for the entire duration of the computation for the LR-MAC (at 4.2 seconds). The main reason behind the decision is the fact that Contiki itself is not required by a RERUM-compliant node, and that multitasking support is still incomplete or absent from the libraries or the OS' used by many of the commercial nodes.

5.1.5 Simulation metrics and results

The simulations are focused on three relevant network metrics:

- The average delivery ratio,
- The average network connectivity,
- The average end-to-end delay.

We do not include the energy consumption of the nodes, since preliminary results were already presented in D5.3.

5.1.5.1 Average Delivery Ratio

The average delivery ratio denotes the percentage of messages which arrive at their destination. For our chosen scenario, the sink represents the only chosen destination throughout all simulations.

The results depicted in Figure 71 were obtained by averaging the delivery ratios of each simulation run (composed of a triple of number of nodes, a mobility percentage and a LR-MAC state). The per-simulation run delivery ratio was computing by dividing the number of messages reaching the sink to the number of messages sent by the reporting nodes. Since isolated nodes are not able to detect a route to the sink, these would not be sending any messages. The average delivery ratio contains therefore no information on the number of connected nodes; this is instead depicted by the network connectivity presented in Section 5.1.5.2.

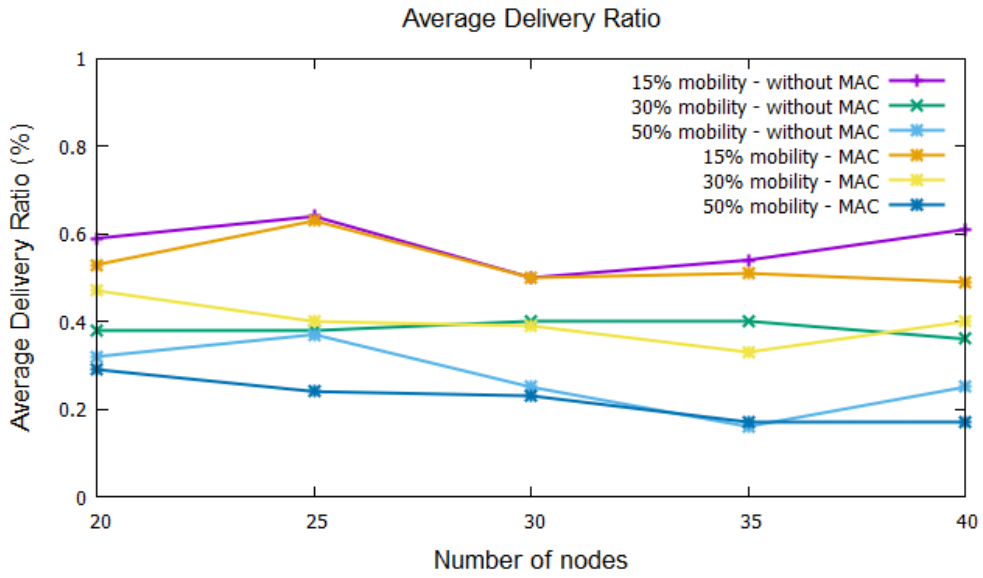


Figure 71 Average delivery for 25 to 40 nodes and a sink at 15, 30 and 50 percent mobility and two LR-MAC states (enabled and disabled)

5.1.5.2 Average Network Connectivity

The average network connectivity denotes the number of nodes which were able to reach the sink and deliver at least one of their readings.

As for the average delivery ratio, the results depicted in Figure 72 represent average values computed for each simulation triple run.

It is intuitively expected to reach a high connectivity over a long period of time when having partial mobile nodes – mainly due to the statistical chances of the mobile nodes bridging the originally isolated nodes with the sub-network of which the sink is part of. The main driver factors are the number of mobile nodes, their mobility pattern, their speed, and the allocated simulation time. Since the goal of this chapter is to evaluate the effect of the LR-MAC-mechanisms upon the network, we did not consider extending the simulation time.

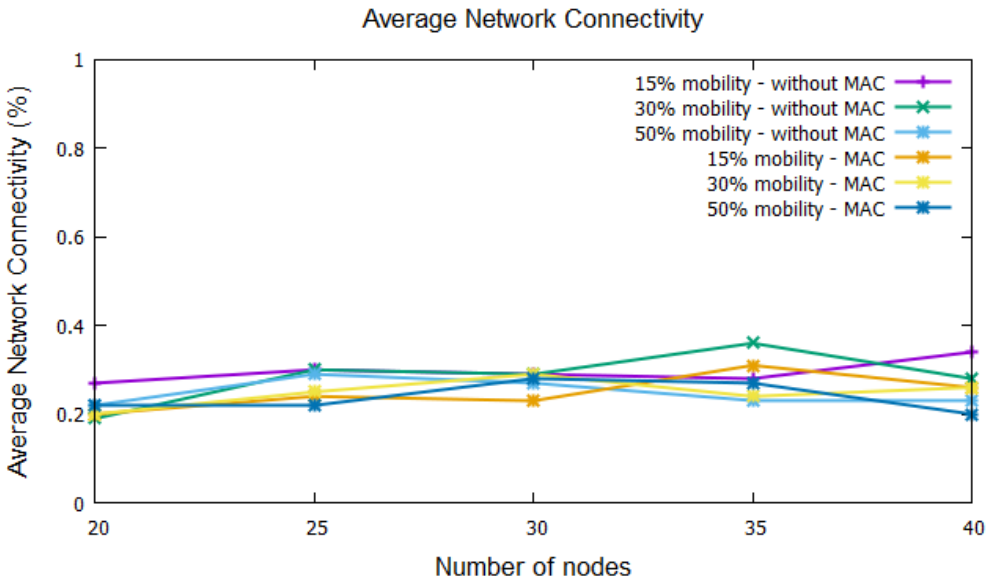


Figure 72 Average network connectivity for 25 to 40 nodes and a sink at 15, 30 and 50 percent mobility and two LR-MAC states (enabled and disabled)

5.1.5.3 Average End-to-End Delay

The average end-to-end delay denotes the average time (in milliseconds) necessary for a message to reach its destination.

Factors influencing the end-to-end delay include retransmissions due to package losses, the time required to process and forward the message (including queuing issues), the transmission delay itself (e.g. medium checks), as well as the RPL-specific route error and alternative handling mechanism.

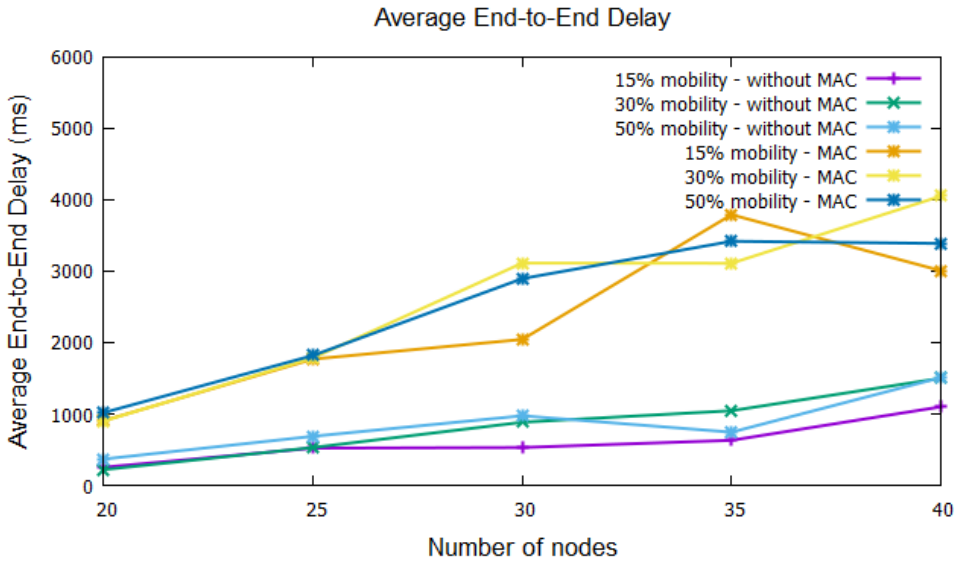


Figure 73 Average end-to-end delay for 25 to 40 nodes and a sink at 15, 30 and 50 percent mobility and two LR-MAC states (enabled and disabled)

5.1.6 Discussion

This section provides a brief discussion on the simulation results presented in Section 5.1.5, throughout which the focus will lie on how well the LR-MAC-enabled RDs scale compared to the nodes operating without (any) security mechanisms enabled.

As seen by the delivery ratio results in Section 5.1.5.1, the computational- and payload-overhead has minimum effects on the average delivery ratio in the simulated scenarios. The 15, 30 and 50 percent lines corresponding to the 20-to-40 RDs for both simulation scenarios (with and without LR-MAC) are very similar: they partially overlap and have a maximum offset of 7 percent.

The main reason behind this closeness in results is due to the random skew between the moments in which the messages are sent towards the RERUM Gateway, which leads to minor bottlenecks and package losses due to the temporary unavailability of some of the package forwarding RDs. A second reason is the way in which RPL is handling the forwarding of messages, through the use of alternative routes, of message buffering mechanisms and of several retries in sending a message up one hop towards its destination.

Similar to the delivery ratio results, the network connectivity throughout all simulations shows little-to now perceivable changes when toggling the LR-MAC mechanisms on or off. These results alone do not provide a high information value – but in conjunction with the average delivery ratio, illustrate and support the fact that the effects of the LR-MAC mechanism, even on such resource restricted devices, are limited: this is mainly due to the information contained by the network connectivity results. By itself, the graph does not provide information on how often a node was able to reach the sink – only that it succeeded so at least once. The delivery ratio, on the other hand, provides a more complete picture alongside the connectivity results, in the sense that the delivery ratios report on direct and indirect connectivity problems (represented mainly by message losses). Since the delivery ratios are similar, the two results support the preliminary conclusion that, in the simulated scenarios, the LR-MAC did not have a significant measurable impact on the network connectivity.

The average end-to-end delay simulations in Section 5.1.5.3, however, do show a measurable impact on the latencies incurred by the computation of the LR-MAC on the forwarding nodes residing between a sender and the RERUM Gateway sink. These computations trigger – as previously mentioned – forwarding RDs unable to reach their main next-hop node, due to it performing computation intensive operations in preparation for their own messages, to perform retransmission retries due to the lack of acknowledgements and, finally, to switch to alternative routes or to the discovery of new ones. This is not the case for networks consisting of RDs communicating without the LR-MAC enabled, where such actions are triggered only by connectivity issues (i.e. due to the partial network mobility). The rate at which the routing actions are triggered account therefore to an overall end-to-end delay increase of about one to two seconds.

If we consider nodes that only report sensed environmental data (e.g. temperature, humidity, average water or power consumption), such latencies are not problematic. In other types of scenarios, in which a user is trying to trigger a certain action, e.g. of turning the light on, such a perceived delay might not prove practical or acceptable. The relevance of the end-to-end delays are therefore very use case specific, as opposed to the delivery ratio and connectivity metrics, which did not deviate significantly even with an increasing number of nodes.

For the use cases assumed by RERUM (UC-O1, UC-O2, UC-I1, UC-I2), there is no direct interaction between a user and the middleware implied– instead, the use cases rely on monitoring and correcting (in the case of the indoor use-cases) parameters and devices (such as the air conditioning). The LR-MAC mechanisms therefore scales well for its intended scope and provides a viable security option.

5.2 Performance of large scale Cognitive RERUM Networks

5.2.1 Introduction

RERUM's cognitive wireless personal and local area network technologies are based on 802.15.4 [ST154] and 802.11 [ST11] standards respectively [RD2.5]. Here, we present the analytical models necessary to study the performance of these protocols for large scale networks and derive analytic expressions for their expected performance when used in a RERUM Cognitive Wireless LAN (CWLAN)

setting. Finally, we investigate issues related to the computational complexity of the stochastic, two-stage spectrum assignment scheme, proposed for RERUM power and QoS constrained CWLANs.

5.2.2 Relation to RERUM UCs

The analysis presented in this subsection relates to RERUM deployments that comprise of either IEEE 802.15.4 or IEEE 802.11 networks, which covers the three out of the four use cases of the project (all apart from the smart transportation use case). In such deployments, the RDs will be connected mainly with IEEE 802.15.4 technology, but not limited to this, since many devices could also be connected with IEEE 802.11 technology. It has been proved in the past that due to interference issues, coexistence of many wireless nodes can cause severe congestion problems, which degrades significantly the performance of the networks. To avoid these issues, we proposed in [RD4.1] a framework and a set of mechanisms to adapt the Dynamic Spectrum Assignment concept to the needs of lightweight devices, creating cognitive RERUM networks. These can be highly beneficial for all RERUM deployments for keeping the interference very low, decreasing the congestion of the network and increasing significantly the scalability of the deployments, allowing the coexistence of large number of devices. Considering that current projections estimate that in the near future billions of devices will be connected to the IoT, it is evident that such a cognitive framework and the analysis we present here will be quite important.

5.2.3 Performance of large scale Cognitive 802.15.4 WPANs

5.2.3.1 Introduction

In the literature various models for 802.15.4 WPANs have been studied [LCMS08], [CR09],[ZSDJ06], [RA12], [MMMT06]. The analytic model for the performance of 802.15.4 PANs presented here is based on the work of Pollin et al. [P++08] for the slotted CSMA/CA mechanism of 802.15.4. We extend the analysis presented in [P++08] by adding a fixed number of delay slots between consecutive transmissions in order to model unsaturated traffic conditions. Although the authors of [P++08] do utilize this technique they do not present the corresponding analytic model. Furthermore, we focus on the slotted CSMA/CA mechanism because it provides synchronization services using beaconing, and an optional Contention Free Period (CFP) using the Guaranteed Time Slot (GTS) mechanism. The former mechanism permits the control and synchronization of RERUM devices via a gateway while the latter allows guaranteed time-slots for low latency applications and applications requiring a specific data bandwidth. These features make the slotted CSMA/CA mechanism a better fit for RERUM Services compared to the unslotted CSMA/CA mechanism. However, in contrast to the unslotted version, the slotted CSMA/CA mechanism has particular characteristics different from other well-known CSMA/CA schemes (e.g. DCF in IEEE 802.11) due to its slotted nature, its distinctive backoff algorithm and its Clear Channel Assessment (CCA) procedure [LCMS08].

5.2.3.2 802.15.4 MAC Layer

In the slotted CSMA/CA mechanism of 802.15.4 a coordinator (which in our case is the RERUM Gateway) transmits a beacon periodically to form a superframe structure. A superframe consists of a beacon that signifies the beginning of a Contention Access Period (CAP). The CAP is followed by an optional Contention Free Period (CFP) and an optional Inactive Portion. During the Inactive Portion all nodes may enter a sleep mode to reduce power consumption. Figure 74 [ST154], presents an example of the superframe structure.

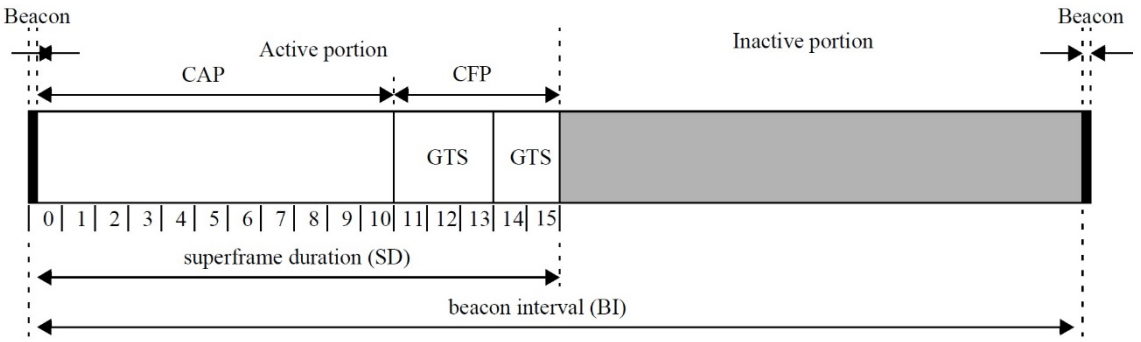


Figure 74: An example of the superframe structure.

The CAP and CFP together form the active portion of the superframe, during which all communication along the nodes should take place. In the CFP, the network coordinator alone controls entirely the contention-free channel access by assigning guaranteed time-slots to those nodes with their GTS requests granted. The assignment of the GTS to those nodes is determined by the scheduling scheme adopted by the network coordinator, which is open in the standard. Therefore, depending on the specific scheduling scheme used, the performance analysis of CFP is actually the same as that of well-studied centralized scheduling schemes in cellular systems.

In the following analysis we disregard the CFP and the Inactive Period and focus on the CAP period only. The impact of CFP and Inactive Period on the MAC performance can be incorporated in our analysis as a constant time cost, equal to the aggregate length of the CFP and the inactive portion of a superframe, measured in time-slots, for each CAP [LCMS08].

The standard defines a MAC layer for the CAP which operates as a non-persistent slotted CSMA/CA with binary exponential backoff. When using this MAC layer an RD that has a frame to transmit backs off for a random number of time-slots before performing a clear channel assessment (CCA). If this first CCA indicates an idle channel then a second CCA is scheduled for the next time slot. If the channel is sensed idle once more the station will proceed with the frame transmission. However, in case any of the CCAs would result in an occupied channel, a new backoff stage is initiated. There are three variables related to the transmission of each frame:

1. The first one, NB, counts the number of backoff stages. We have that $0 \leq NB \leq NB_{max}$ and if all the NB_{max} backoff stages end up with a busy channel indicated by the associated CCAs, a *Channel Access Failure* event will be reported to the upper layer and the node will proceed with the next frame, if one exists.
2. The second variable is the current backoff exponent (BE), where $BE_0 = macMinBE$, is the initial and minimum backoff exponent for each frame. The number of backoff slots in stage NB is drawn uniformly over the interval $[0, 2^{BE_{NB}} - 1]$. BE is incremented with each backoff stage increment according to the relationship $BE_{NB+1} = BE_{NB} + 1$ and is upper bounded by $aMaxBE$ which is the default value of the maximum backoff exponent.
3. Finally, the third variable is the contention window (CW) which is initialized to two (2) whenever a new backoff stage begins and is decremented by one whenever a CCA is performed by the station. The standard specifies the following default parameter values: $macMinBE = 3, aMaxBE = 5$, and $NB_{max} = 5$. Whenever the station succeeds in accessing the channel all three parameters are reset to their default values.

5.2.3.3 Formulation of the analytical model

Next we present the analytic model for the performance of 802.15.4. More specifically, Figure 75 presents a two-dimensional Discrete Time Markov Chain (DTMC) for the MAC layer of 802.15.4

described in the previous section. The evolution of the DTMC over time is described by a pair of discrete time stochastic processes $\{(s(t), c(t))\}$ where $s(t) \in \{-2, -1, 0, 1, \dots, m\}$ and $c(t) \in \{-2, -1, 0, \dots, \max_{i}(L-1, D-1, W_i - 1)\}$. Based on the value of $s(t)$ and $c(t)$, each state in the feasible set of states, as depicted in Figure 75, bears the following meaning:

1. When $s(t) = -2$ the RD is waiting for new data to arrive through its environment sensing equipment. The delay is assumed to be fixed [P++08] and equal to the expected time between data frame generation. This is indicated by the fact that $c(t)$ is always initialized with a zero value and incremented up to value $D - 1$(transitions between states occur with probability 1).
2. When $s(t) = -1$ the RD is transmitting the frame over the wireless link after successfully competing for medium access. Again the duration of the transmission is assumed fixed and equal to L time slots [P++08].
3. Finally, the values of $s(t) = 0, 1, \dots, m$ indicate the current backoff stage of the process, where $m = NB_{max}$ and the corresponding zero or positive values of $c(t)$ indicate the remaining number of time slots until the end of the backoff period. On the other hand, the negative values of $c(t)$, -1 and -2, indicate the first and second CCA respectively for the corresponding backoff stage, denoted henceforth with CCA^1 and CCA^2.

Let $\pi_{i,k}$ denote the steady state probability of state (i, k) of the DTMC of Figure 75, i.e.,

$$\pi_{i,k} = \lim_{t\to\infty} P\{(s(t), c(t)) = (i,k) | (s(t_0), c(t_0)) = (i', k')\}, \text{ for all states } (i', k')$$

where (i', k') is the initial state of DTMC at some initial point in time t_0 with $i, i' \in \{-2, -1, 0, 1, \dots, m\}$ and $k, k' \in \{-2, -1, 0, \dots, \max_{i}(L-1, D-1, W_i - 1)\}$.

We denote the long-term expected fraction of time, the station is in a $\pi_{i,0}$ state, $i \in [1, m]$, i.e., it has finished the backoff period of the i-th backoff stage and is ready to initiate CCA^1 in the next timeslot as,

$$\tau = \sum_{i=0}^{m} \pi_{i,0} \quad (6)$$

Assuming a star topology PAN of *'N'* RDs, with all nodes in the same collision domain, we can express the utilization S of the network as a function of the long term expected fraction of time, that a single RD is transmitting. To this end, a major assumption made by the authors of [P++08], is that the event of a station becoming ready to start listening to the channel, i.e., to enter a $\pi_{i,0}$, $i \in [1, m]$ state, occurs independently of the states of the rest $N - 1$ stations. Based on this assumption we can derive S by calculating the probability of only one, out of the N stations, to be to be ready to initiate CCA^1 in the next time slot, i.e., to be in any of the $\pi_{i,0}$, $i \in [1, m]$ states, times the probability that CCA^1 and CCA^2 will indicate an idle channel. The event described above combined with the assumption of a single collision domain guarantee that the station will access the medium successfully and that it will spend the next L time-slots transmitting. Thus we have the following relationship for the long term expected fraction of time, the medium is occupied by successful transmissions:

$$S = \binom{N}{1} L\tau(1-\tau)^{N-1}(1-\alpha)(1-\beta) = NL\tau(1-\tau)^{N-1}(1-a)(1-\beta)$$

In the next paragraphs we derive a system of three non-linear equations that relate α, β and τ.

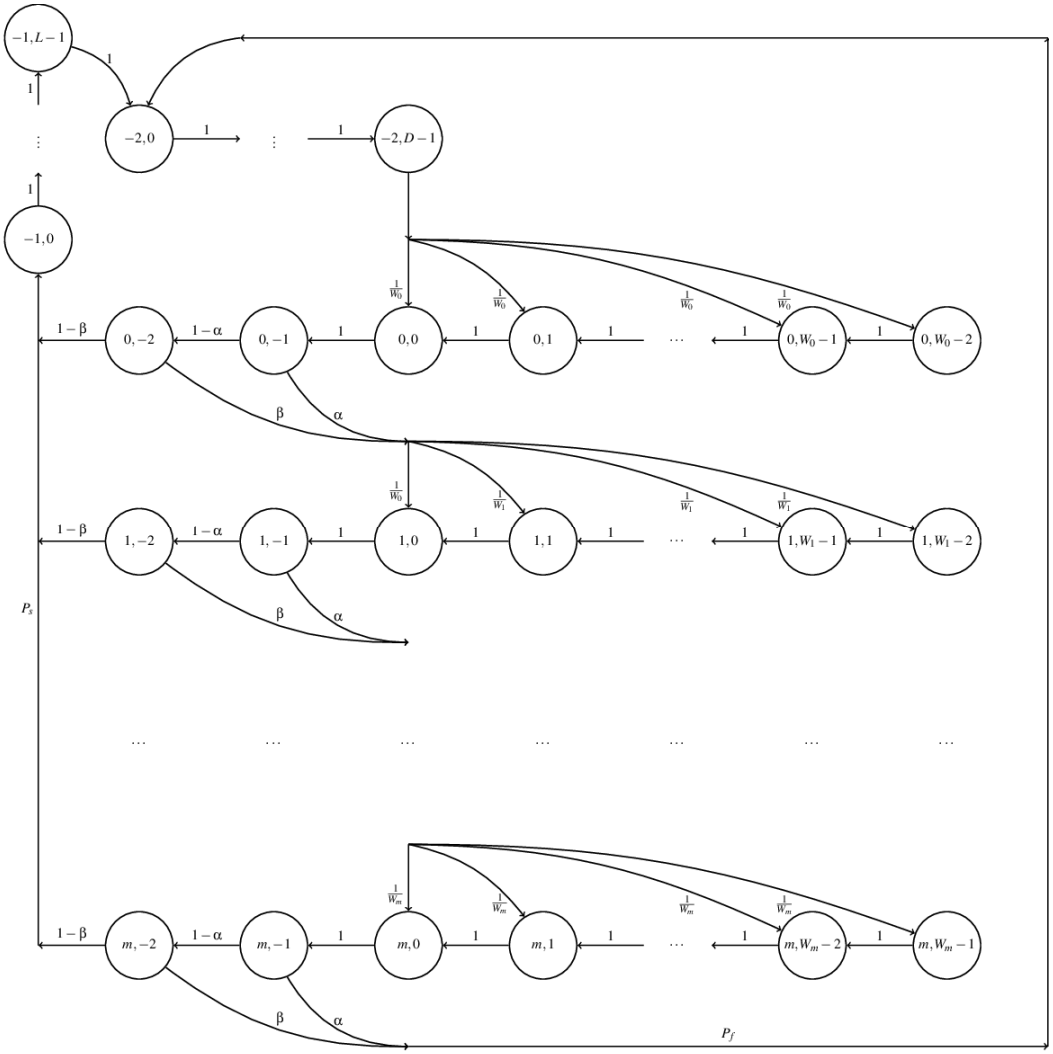


Figure 75 : Discrete Time Marckov Chain for the 802.15.4 MAC layer implemented by a single sensor

5.2.3.4 Performance analysis of Cognitive Wireless PANs.

In Cognitive PANs (CPANs), such as the ones considered in RERUM, the spectrum band b assigned to the PAN may change over time due to the appearance of its' Primary User (PU). Thus, the utilization of the assigned spectrum band will be less, compared to an idealized PAN that exhibits zero packet losses due to interference, because of the time spent in cognitive functions such as spectrum switching and spectrum sensing. Similar to [XJL08], [WA11], and [WA08] the expected aggregate throughput of the CPAN, assuming that all spectrum bands assigned to the CPAN over time are of equal bandwidth, will be given by,

$$C = h \cdot S^{max} \cdot r \tag{7}$$

where, S^{max} is the maximum theoretical utilization of the spectrum band by 802.15.4 given that no transmission fails due to interference, h is the long term expected percentage of time that spectrum will be available to the CPAN for transmission and r is the transmission rate used by the devices. The value of h depends on the frequency of spectrum switching, i.e., on the expected time each PU will remain idle once it stops transmitting, or busy once it starts transmitting, and on the frequency and duration of spectrum sensing, whereby all stations in the CPAN must remain silent and listen for possible PU transmissions. Furthermore, the value of h can be further reduced when multiple CPANs operate within interference

range of each other. In such a case, CPANs must be assigned orthogonal spectrum bands, which however may differ in terms of availability due to the different PUs' traffic characteristics. Finally, we emphasize that CPANs are expected to outperform typical PANs since the latter technology usually suffers from a high percentage of transmission failures due to interference in the congested ISM bands.

5.2.3.5 Performance Evaluation

In this section we present results for the performance of CPANs when the number of RDs in the network becomes large. Table 30 presents the default values for some of the parameters of the 802.15.4 model and related parameters defined by the 802.15.4 standard [ST154].

Table 30: Model parameters and their values along with related parameter values specified by the 802.15.4 standard.

Model Parameter	802.15.4 specification	Model Parameter Value
m	$NB_{max} = 5$	5
W	$macMinBE = 3$	3
timeslot	$aUnitBackoffPeriod = 20\ symbols$	1
L	$aMaxPHYPacketSize$ = 127 byte	15 timeslots

In the DTMC of Figure 75 each transition between states lasts for a single unit of time. We will refer to a unit of time as a timeslot. The 802.15.4 standard [ST154] defines the duration of a CCA, backoff, and superframe-slot to be 8, 20 and 60 symbols respectfully. In order to be consistent with the assumptions of the DTMC we define a timeslot to be a period of 20 symbols and we assume that CCA lasts for a period of 20 symbols instead of 8. We expect that this approximation will have a small impact on the performance evaluation of the PAN due to the small number of CCA states compared to backoff states.

To derive L we consider the maximum packet supported by 802.15.4 at the physical layer (PSDU) defined in Table 70 of [ST154] as $aMaxPHYPacketSize$ *= 127 octets or 1016 bit.* Now, each superframe is comprised of 16 superframe-slots (Section 5.1.1.1 in [ST154], *aNumSuperframeSlots=16*), and each superframe-slot corresponds to the time necessary for the transmission of 60 symbols. In case the transmission rate used is 250 Kbps each symbol carries 4 bits (Table 66 in [ST154]) so that $aMaxPHYPacketSize$ corresponds to 254 symbols and occupies $\frac{254}{60} = 4.23$ superframe-slots. Since superframe-slot number must be an integer we round the number of superframe-slots necessary for the transmission of 1016 bit to 5 superframe-slots or, equivalently, 15 timeslots. Furthermore, at the transmission rate of 250 Kbps we have 62.5 Ksymbols/sec (Table 66 in [ST154]) which means that a superframe-slot of 60 symbols lasts for 96 milliseconds while a timeslot lasts for 32 millieconds. We consider this duration in order to determine D, i.e., the number of timeslots between successive packets' generation by the sensor. The value of D plays a critical role in the performance of the RERUM PANs and their ability to scale. We expect that D will be large in a RERUM network where RDs take periodically measurements from the environment. However, in networks that support event driven applications certain events may trigger the generation of multiple packets on multiple sensors concurrently. In such a case the network becomes saturated and a large number of RDs will always have a packet to transmit. Thus, it is important to study the behaviour of the network for both large and small values of D.

Figure 76 presents the expected aggregate throughput of a PAN for parameter D values equal to 0, 10 and 104 timeslots, i.e., 0, 0.01 and 0.1 seconds respectively and h equal to 1. When D equals 0 timeslots the PAN achieves its maximum aggregate throughput when two stations operate in the network. This is due to the fact that these stations will transmit one packet after the other with no delay between consecutive packets generation and transmission, i.e., they always have a packet to transmit and very little contention for medium access. At this point we must note that the analytic model does not consider the case whereby a packet to be transmitted does not fit in the remaining timeslots of the superframe, e.g., if a sequence of packets takes 12 timeslots to be transmitted then if the next packet requires 5 timeslots for its transmission, it will have to wait for the next superframe to be transmitted since it doesn't fit in the remaining 4 timeslots of the current superframe. Obviously, this requirement will result in suboptimal utilization of the medium in the case of a small number of RDs. However, as the value of D increases the effect of this phenomenon of the PAN performance will become less significant. Finally, from Figure 76 we see that the aggregate throughput will diminish as the number of RDs in the network increases. This is due to collisions (see Figure 79) and the increased backoff delays that correspond to larger backoff stages suffered by stations that get a busy channel after either CCA^1 or CCA^2 (see Figure 77 and Figure 78). As D value increases a small number of RDs is not able to fully utilize the wireless channel any more since they do not always have a packet to transmit. Utilization of the wireless medium will increase as the number of RDs increases up to a certain point and then diminish as the number increases further due to collisions and increased backoff delays as can derived by Figure 77 and Figure 78.

Figure 80 presents aggregate throughput results for the same scenario as Figure 76, only for larger values of D. More specifically, D will take on values 1042, 10417, 20833 and 31250 that correspond to inter-packet generation delays of 1, 10, 20 and 30 seconds respectively. It can be seen from Figure 80 that as D increases so does the necessary number of stations to fully utilize the wireless medium and furthermore, that this number is in the order of hundreds of stations even for the smallest value of D that corresponds to 1 second. One must keep in mind though, as already mentioned, that rarely a CPAN will support a single application with a single sampling period. However, it is indicative of the fact that in a CPAN with a small number of nodes the channel will be idle most of the time, unless an event will trigger multiple concurrent transmissions by a large number of stations. Finally, Figure 81 and Figure 82 present the probability that CCA^1 and CCA^2 will indicate a busy channel while Figure 83 presents the collision probability for large values of D respectively. The results presented in these figures verify that the probability to find the channel busy and consequently initiate a new backoff interval as well as the probability to suffer a collision increase with the number of stations in the CPAN and eventually diminish the aggregate throughput of the CPAN.

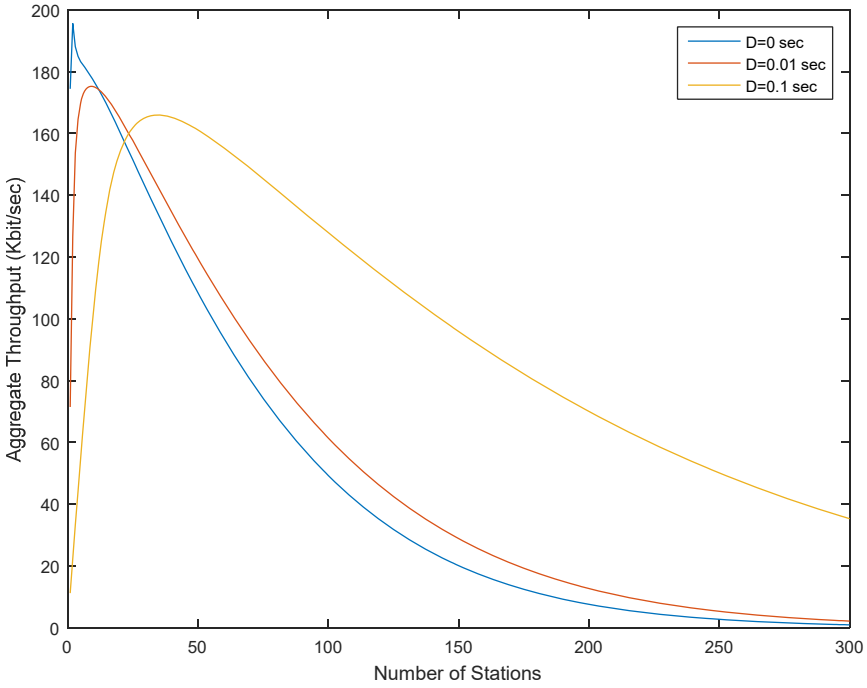


Figure 76: PAN aggregate throughput vs. the number of nodes in the PAN for small values of the packet generation delay $\boldsymbol{D}$ when $\boldsymbol{h = 1}$.

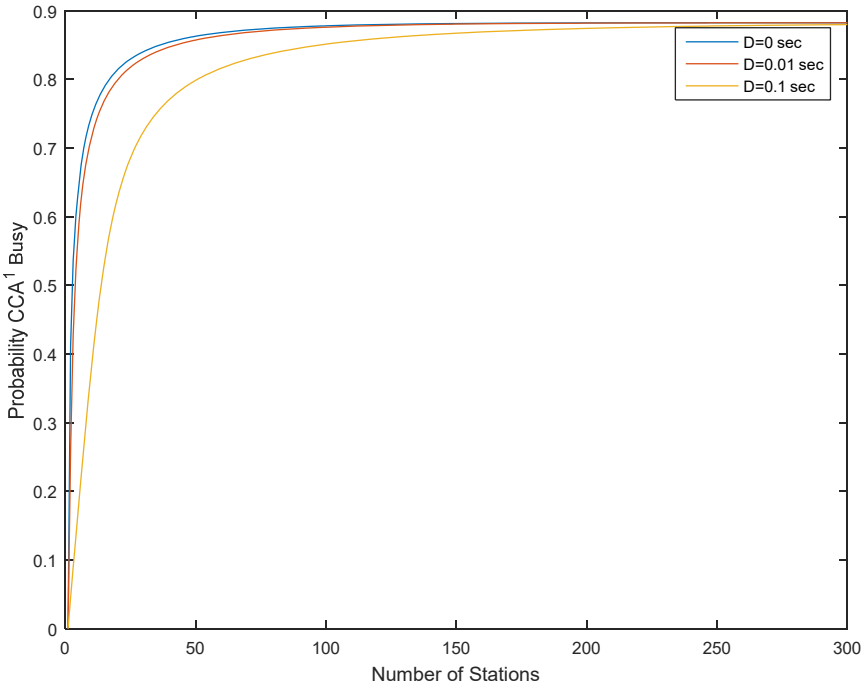


Figure 77: Probability for $\boldsymbol{CCA^1}$ to indicate a busy channel vs. the number of stations in the PAN for small values of $\boldsymbol{D}$.

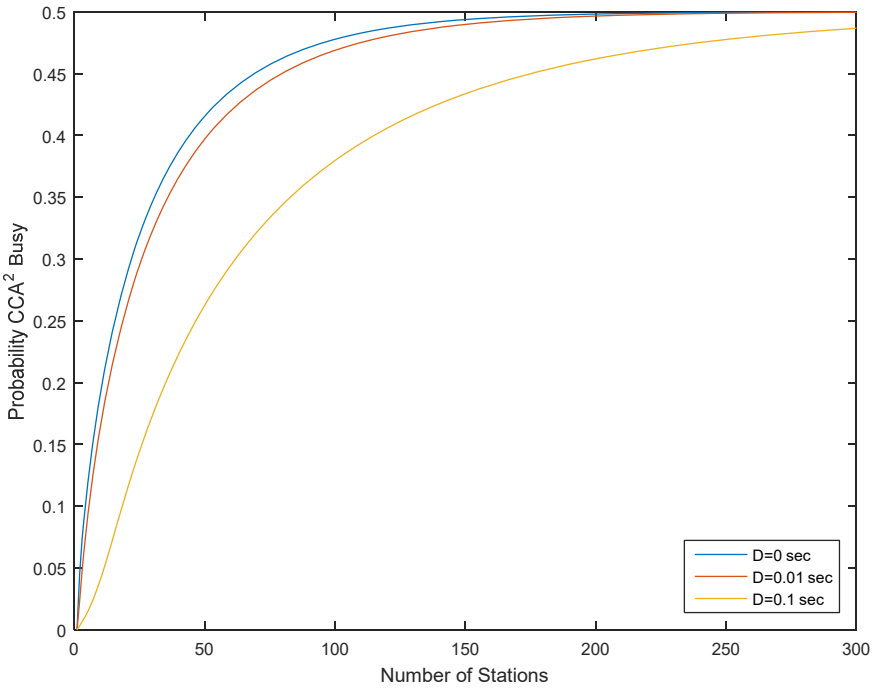


Figure 78: Probability for $\boldsymbol{CCA^2}$ to indicate a busy channel vs. the number of stations in the PAN for small values of $\boldsymbol{D}$.

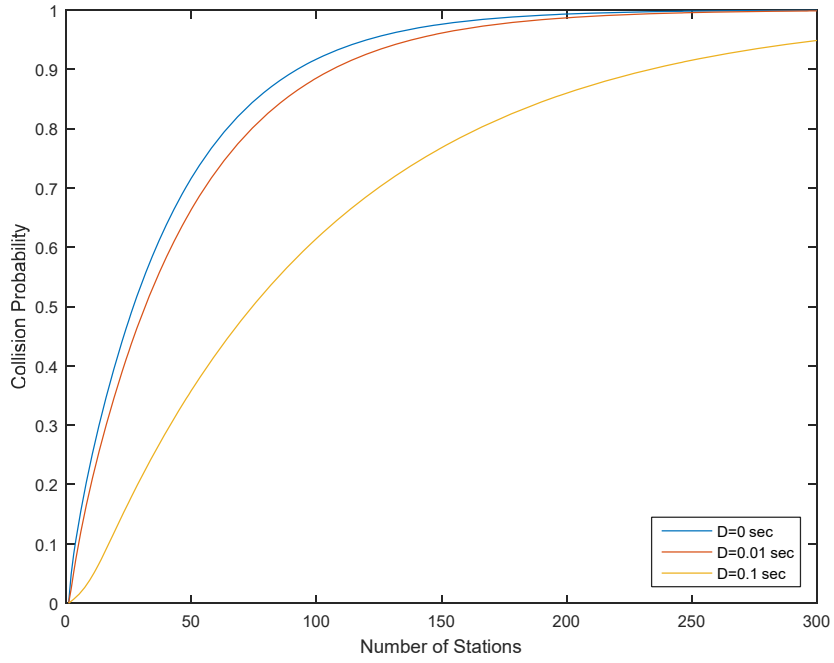


Figure 79: Collision probability among the stations of the PAN vs. the number of stations in the PAN for small values of $\boldsymbol{D}$.

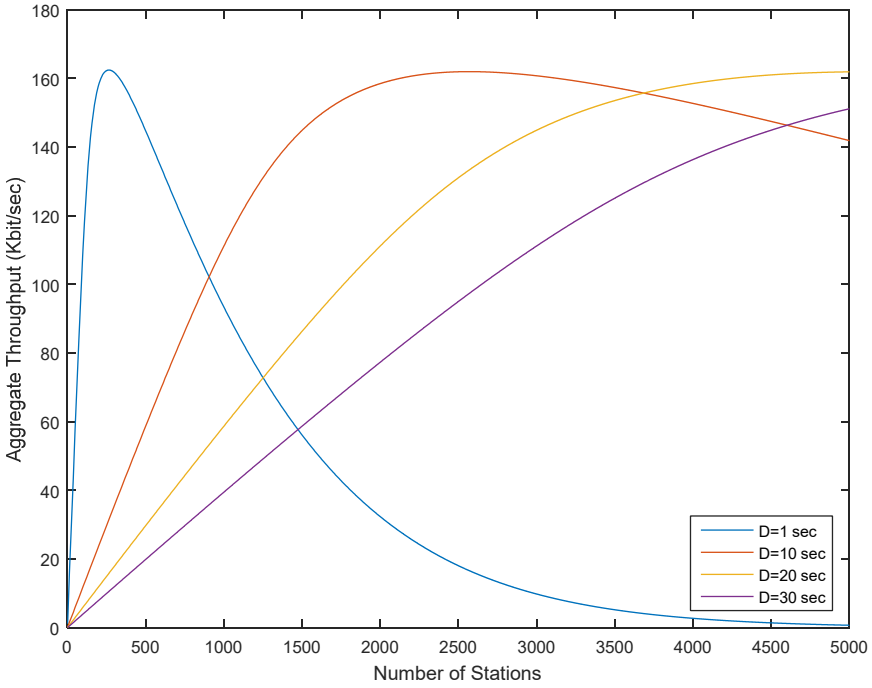


Figure 80: PAN aggregate throughput vs. the number of nodes in the PAN for large values of the packet generation delay $\boldsymbol{D}$.

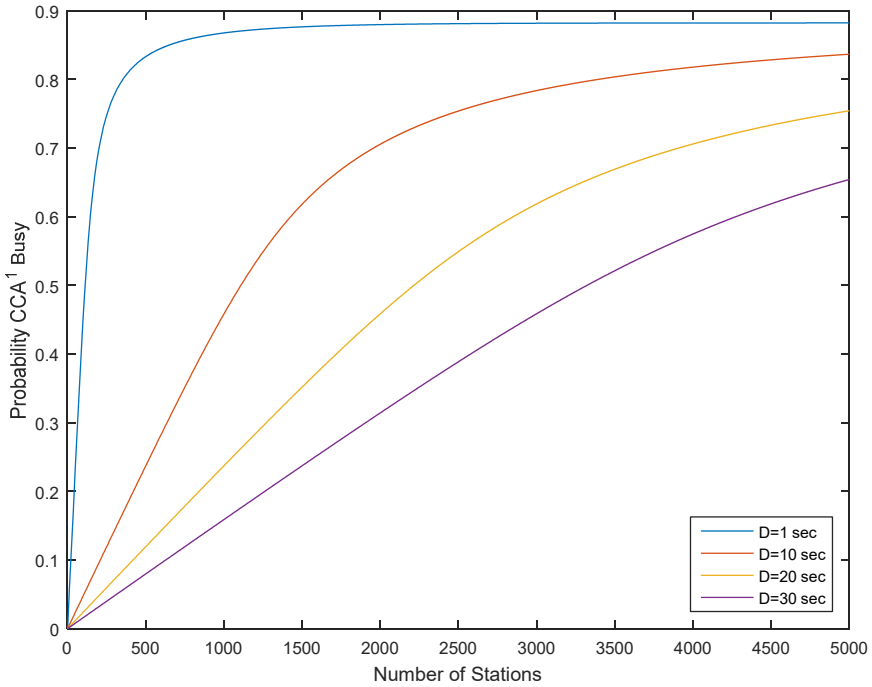


Figure 81: Probability that $\boldsymbol{CCA^1}$ will indicate a busy channel vs the number of stations in the PAN for large values of $\boldsymbol{D}$.

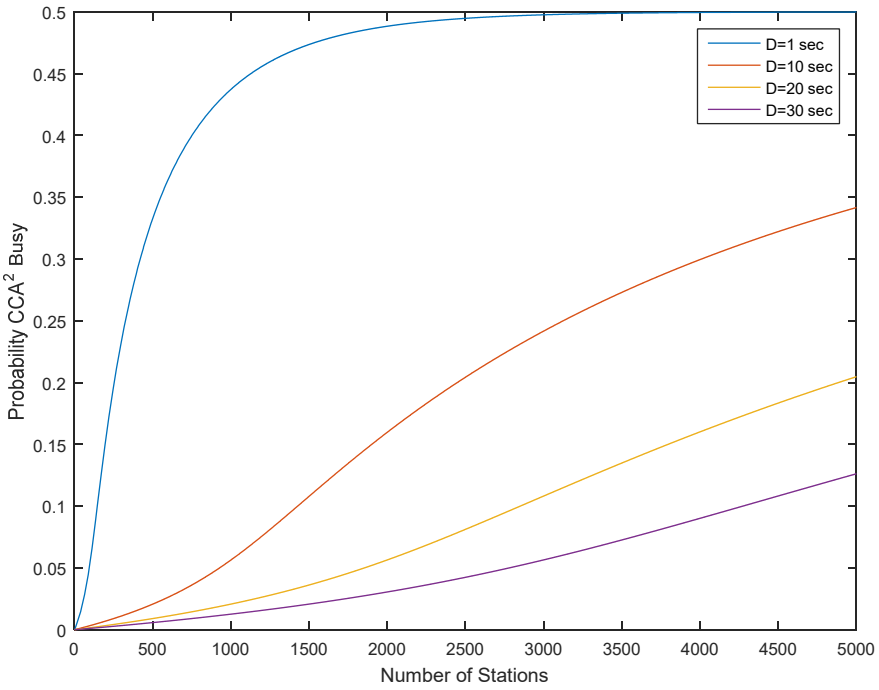


Figure 82: Probability that $\boldsymbol{CCA^2}$ will indicate a busy channel vs the number of stations in the PAN for large values of $\boldsymbol{D}$.

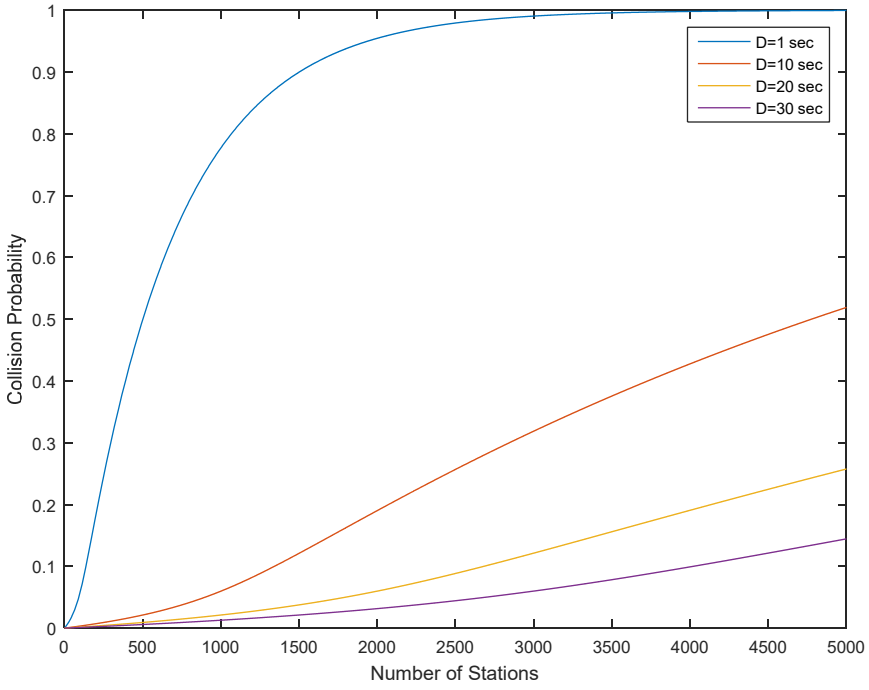


Figure 83: Collision probability among the stations of the PAN vs. the number of stations in the PAN for large values of $\boldsymbol{D}$.

5.2.3.6 Discussion

What is important to notice in the previous analysis is the much difference reaction of the network in different types of traffic. For almost saturated traffic flows, as seen in Figure 76 until Figure 79, the aggregated throughput drops significantly with the number of RDs in the network, which is expected due to the fact that the total capacity is reached very quickly. However, in standard IoT networks, this type of traffic is not usual and most traffic flows related with sensor measurements are more relaxed in terms of throughput. Thus, as seen in Figure 80 and on, in such a case, the network can tolerate and can support a large number of devices without significant performance issues or collisions. In this case, the total capacity of the network is not reached very quickly and even after 200-300 devices we can see that the possibility for collision can stay below 50%. Thus, the scalability of such networks can be quite high. It is evident though, that the scalability of the RERUM IEEE 802.15.4 deployment depends heavily on the type of traffic that the devices have. However, in the three use cases that have such deployments, the traffic of each device is only related with a few measurements per minute from each device, which can be mapped to the case with D=20 or 30seconds in the above graphs. Thus, the overall traffic in such a case will be very low and the network can scale quite well, with extremely large (>3000) number of RDs per gateway without a significant drop in the network performance.

5.2.4 Performance of large scale 802.11 networks

5.2.4.1 Introduction

RERUM services that require transmission rates greater than 250 Kbit/sec, the maximum transmission rate provided by the 802.15.4 standard, may utilize the 802.11 standard [ST11]. Although both standards are based on the CSMA/CA protocol, the Distributed Coordination Function (DCF) of 802.11 and the CSMA/CA procedure for the CAP in 802.15.4 are different and thus a different analytical model is required for 802.11. The performance analysis of 802.11 WLANs presented here is based on the work of Laufer and Kleinrock [LK13]. The model presented in [LK13] is especially attractive for sensor based WLANs since it supports both saturated and unsaturated multi-hop flows.

More specifically, the authors in [LK13] present a model for the CSMA/CA protocol that efficiently approximates 802.11 DCF. According to their model, each station, before transmitting a packet, verifies whether the medium is idle via carrier sensing. In case the medium is found busy, the station will defer its own transmission until the channel becomes idle. On the other hand, if the medium was found idle the transmitter would have independently sampled a random backoff interval from a given continuous probability distribution and it would have waited at least that long before transmitting. We note here that although the model deviates from the Binary Exponential Backoff model used in 802.11 DCF no assumptions are placed on the distribution of the backoff intervals. Furthermore, the authors in [LK13] assume that each station has a unique transmission queue for each flow. As a packet arrives at the station it is routed and placed in the corresponding queue. Packet scheduling across the different queues is realized with CSMA/CA. Each queue of the station operates as a different collocated transmitter and modeling complications, due to head of line blocking, are thus avoided. Based on the work by Liew *et al.* [LKLW10] whereby a saturated CSMA/CA network is proved to be a Markov field, and therefore detailed balance holds, the authors of [LK13] prove that the particular distributions of the backoff and transmission timew determine average throughput only through their expected values. Thus, each queue within a station has an individual backoff counter to store the remaining time until the scheduled transmission. If the medium becomes busy during the backoff interval, the transmitter freezes its counter and resumes the countdown only after the medium becomes idle again. When the counter is decremented to zero the packet is finally transmitted.

The duration of a packet transmission T_i depends on the length of the packet and the transmission rate used by the i-th station which is assumed to be fixed, i.e., the randomness of T_i comes only from the different packet sizes while dynamic transmission rates are not modeled. No assumptions are

made regarding the size of the packets and their generation process. Not even at relay nodes in the case of multi-hop traffic. The only assumption is that packet sizes are generated by the flow source according to a given discrete distribution, possibly different for each flow, and that packets retain their sizes as they traverse the network, i.e., packets may not be fragmented. Each flow source generates packets with a random interarrival time A_i, following a given probability distribution which can also be different for each flow. No assumptions are placed on these distributions and the sources are not required to be saturated.

Furthermore, the authors assume that carrier sensing is instantaneous and thus, the moment a transmission starts it is immediately detected by all neighboring stations. This assumption implies that collisions, due to stations finishing their backoff intervals at the same time, have zero probability to occur given that backoff intervals are continuous random variables. Finally, the authors do not consider the effects of hidden terminals on protocol's performance. We expect that the impact of both these phenomena will diminish as the saturation level of packet sources decreases. On the other hand, the model proposed by the authors captures high impact phenomena such as Flow In the middle (FIM), Information Asymmetry (IA) [GSK05] and Performance Anomaly (PA) [HRBD03] that will have a large impact on the performance of large scale WLANs [GSK05, HRBD03].

According to [LK13] the average throughput $x_i$of a transmitter i, i.e., of a single flow, in a WLAN is given by,

$$x_i = \frac{\gamma_i}{1 + \gamma_1 + \gamma_2 + \cdots + \gamma_n} r_i p_i \tag{8}$$

where, r_k and p_k are the transmission rate and the probability of successful transmission for the k-th flow respectively and γ_k is given by,

$$\gamma_k = \frac{\varrho_k \mathbb{E}[L_k]}{r_k \mathbb{E}[B_k]} \tag{9}$$

where $\mathbb{E}[L_k]$ is the expected packet length of the k-th flow, $\mathbb{E}[B_k]$is the expected backoff value of the k-th flow and $\varrho_k$is a parameter that characterizes the saturation level of the k-th flow by accounting for the fraction of time each transmitting flow reduces its backoff counter, i.e., it has a packet to transmit, given that the medium is idle. Parameter ϱ_k takes values in the interval $[0, 1]$, with value 0 denoting a transmitter that is always idle and value 1 denotes a saturated transmitter.

5.2.4.2 Performance analysis of Cognitive WLANs.

A RERUM Cognitive WLAN (CWLAN), compared to an idealized WLAN where $p_k = 1$ for all k, will exhibit lower utilization of its' assigned spectrum bands due to the time spent in cognitive functions such as spectrum switching and spectrum sensing. According to [XJL08], [WA11], and [WA08], the expected aggregate throughput of the CWLAN, assuming that all spectrum bands assigned to the CWLAN are of equal bandwidth, will be given by,

$$C = h \cdot \sum_k x_k^{max} \tag{10}$$

Where, x_k^{max} is the theoretical maximum aggregate throughput of the CSMA/CA protocol considered when $p_k = 1$ for all k and h is the long term expected percentage of time that spectrum will be available to the CWLAN for transmission. Finally, we emphasize that CWLANs are expected to outperform typical WLANs since the latter technology typically suffers from a high percentage of transmission failures due to interference in the congested ISM band.

5.2.4.3 Performance Evaluation

In this section we evaluate the performance of an 802.11 RERUM WLAN as the number of RDs increases. The default parameter values used for the evaluation of 802.11 WLAN performance are presented in Table 31. Furthermore, in the following sections we will use the terms flow and station interchangeably since, as explained in the previous section, we assume that each flow behaves as a collocated transmitter, i.e., as an independent station serving a single flow.

Table 31: Basic configuration of 802.11 WLAN parameters

Parameter	Value for all stations
$\mathbb{E}[L_k]$	1450 byte
ϱ_k	1
r_k	54 Mbit/sec
$\mathbb{E}[B_k]$	15 time-slots
p_i	1
timeslot	20 μsec

Figure 84 presents the aggregate throughput of all RDs in the WLAN, as a function of the number of RDs, for different values of saturation level. We assume that all stations in the network have the same saturation level. We note from Figure 84 that the larger the value of saturation level the fewer RDs are required to saturate the channel, i.e., to fully utilize the capacity of the network. We further note that, since collisions are not considered by the model assumed, the performance of the network does not degrade as the number of RDs increases. However, as seen in Figure 85, the throughput of each RD diminishes with the number of RDs in the WLAN and this occurs with a higher rate for larger values of saturation level. This result is clearly depicted in Figure 86 that presents the normalized throughput of each RD that is calculated by dividing the per flow throughput with the throughput achievable by a RD that operates alone on the channel and has the same saturation level. Thus, for example, given the default configuration of the WLAN, it can be seen from Figure 86 that when the number of RDs is 50 all flows achieve less than 10% of their maximum rate for saturation levels above 0.3. In case the applications generating the traffic are delay sensitive this means that they will not meet their requirements. Finally, Figure 86 indicates that scaling up the number of RDs in a WLAN by a dozen will lead to significant performance degradation.

Figure 87, Figure 88 and Figure 89 present aggregate throughput, per flow throughput and per flow normalized throughput results for various values of the expected backoff. All three figures indicate that the channel becomes saturated quickly with the addition of only a small number of RDs. Furthermore, the smaller the expected backoff value the larger the rate with which the channel becomes saturated. This becomes more evident in Figure 89 where we present normalized throughput results. The results presented in Figure 87, Figure 88 and Figure 89 are in accordance with those of Figure 84, 84, and Figure 86. This phenomenon is explained by the authors of [LK13] that prove the equivalence of saturated and unsaturated CSMA/CA networks when the latter networks exhibit a sufficiently larger expected backoff value compared to that of the saturated network.

Figure 90 and Figure 91 present aggregate throughput and per flow throughput results respectively, as a function of the number of stations, for varying percentages of low rate, i.e., stations transmitting with a rate of 11 Mbit/sec, among all stations' population. The purpose of these figures is to depict the severity of the Performance Anomaly [[HRBD03] problem on the flows' performance. As depicted in Figure 90 a single station within a WLAN of 10 RDs, i.e., 10% of the stations' population, is enough to diminish the aggregate throughput by 26% while the addition of another low rate station will result

in a 40% reduction of the aggregate throughput. The root of this problem is that the CSMA/CA protocol guarantees, in the long run, an equal number of transmission opportunities to all stations. However, low rate stations will keep the channel for a much longer period, whenever they gain access to the medium, in order to transmit their data at their low transmission rate, given that all stations have the same average packet length.

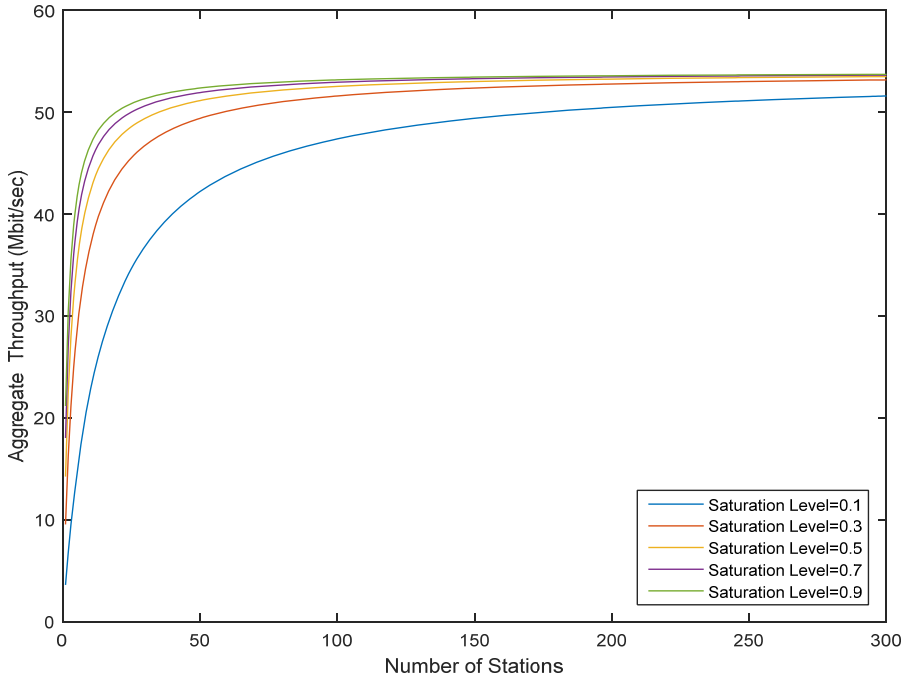


Figure 84: Aggregate throughput of a WLAN as a function of the number of stations for different source saturation levels.

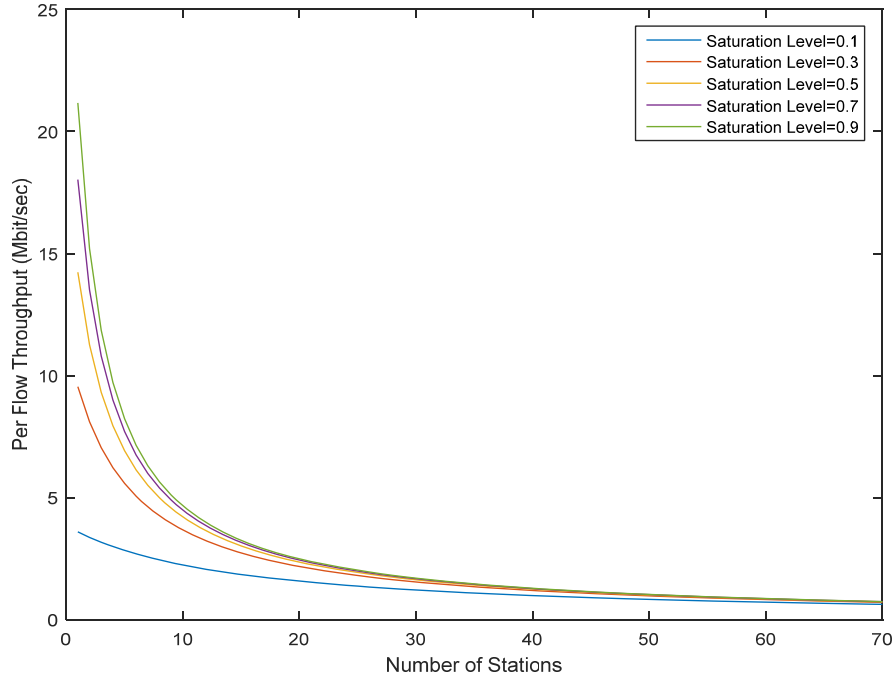


Figure 85: Per flow throughput as a function of the number of stations for different source saturation levels

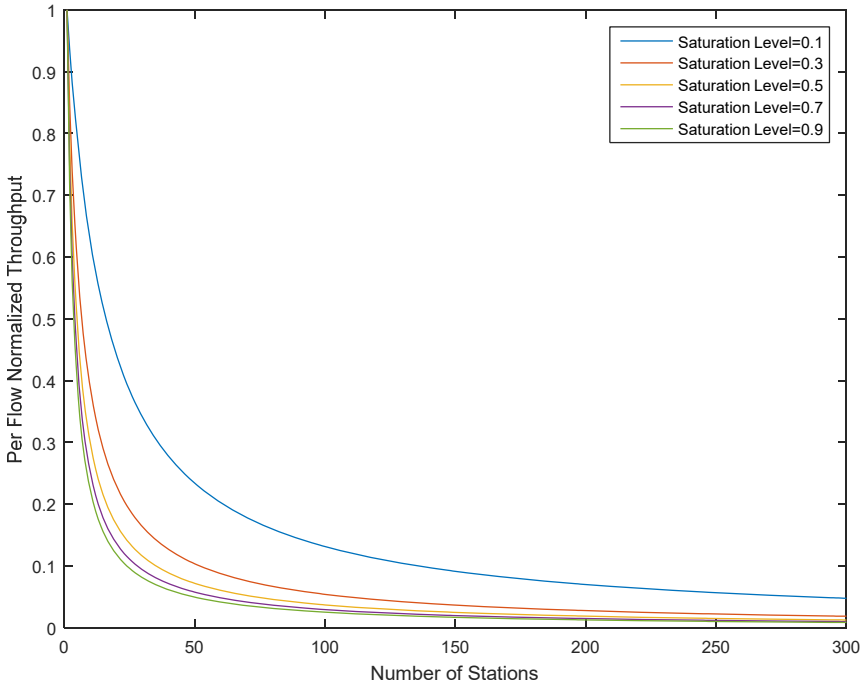


Figure 86: Per flow normalized throughput as a function of the number of station for different source saturation levels. Each flow's throughput is divided by the throughput achieved by a single flow in the network.

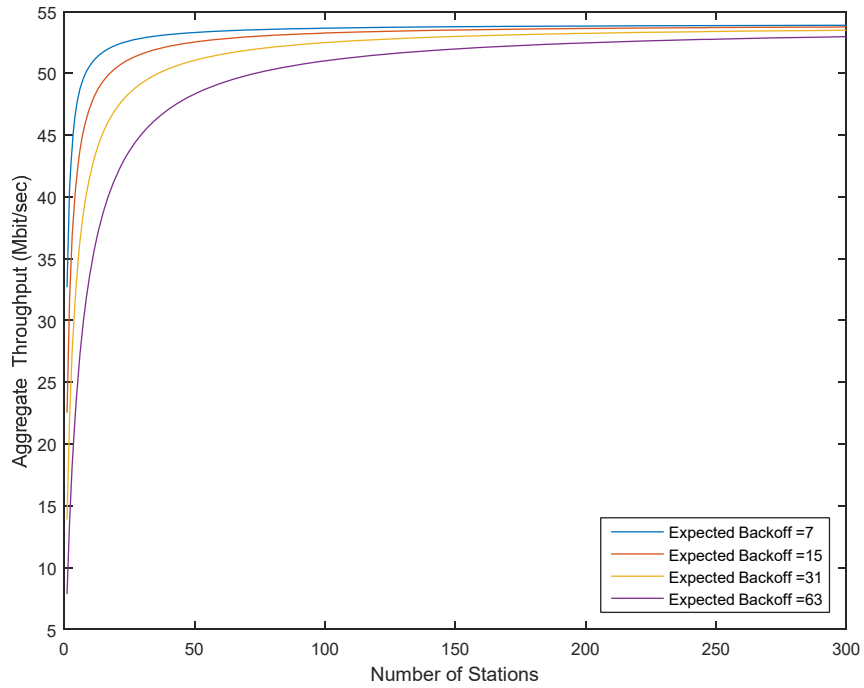


Figure 87: Aggregate throughput as a function of the number of stations for various values of expected backoff**.**

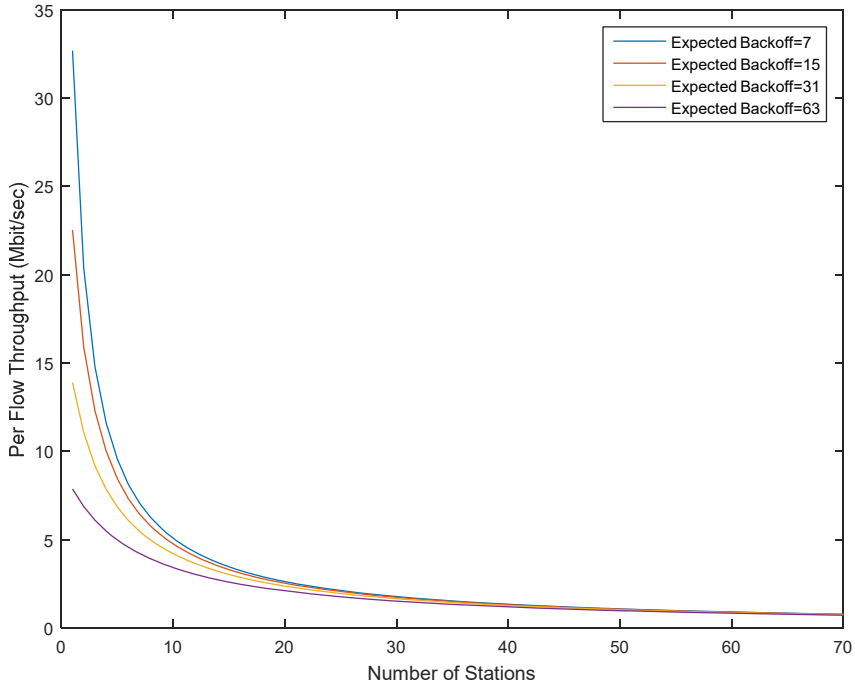


Figure 88: Per flow throughput as a function of number of stations for various values of expected backoff**.**

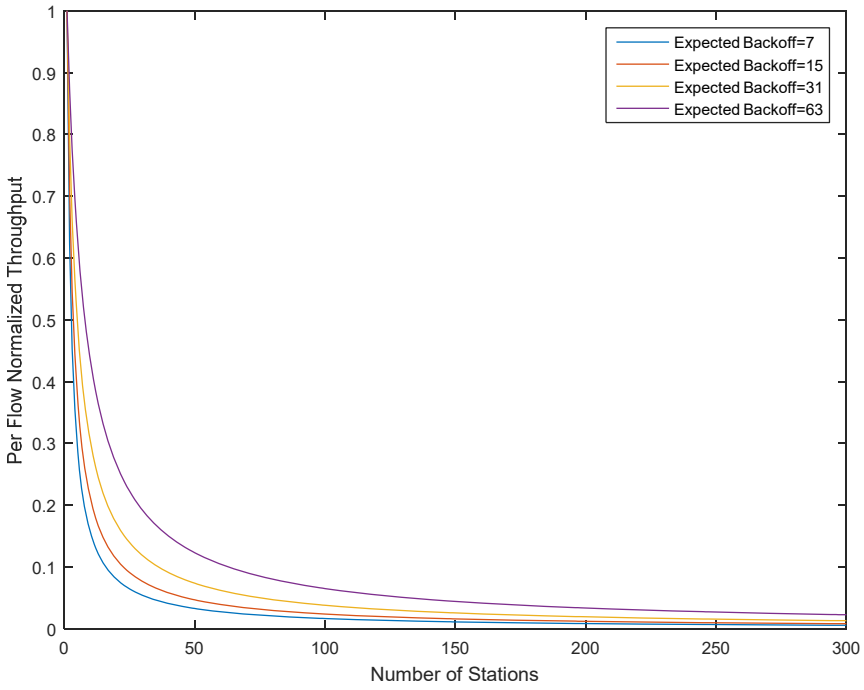


Figure 89: Per flow normalized throughput as a function of the number of stations for various expected backoff values

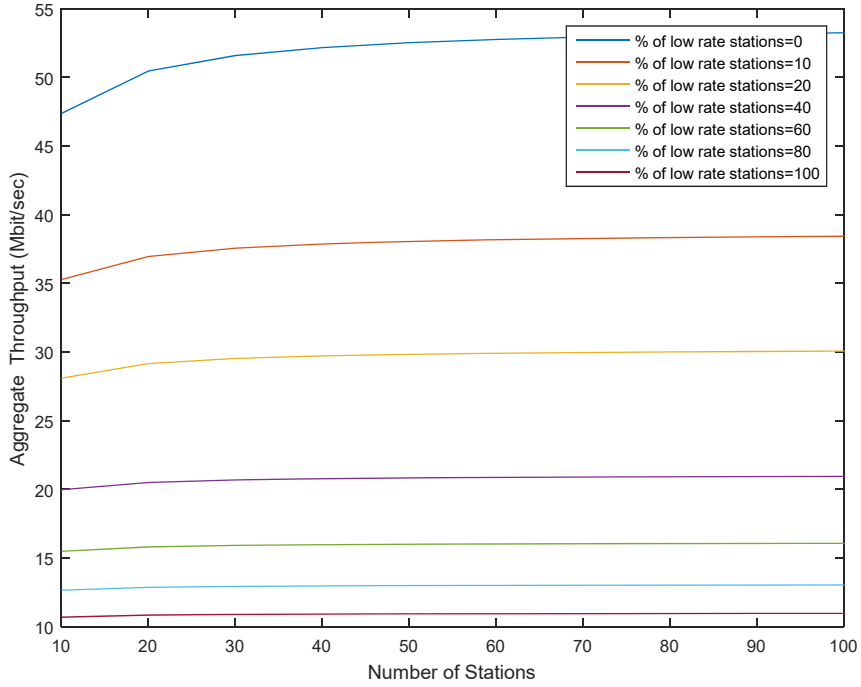


Figure 90: Aggregate throughput as a function of the number of stations for different mixtures of high and low rate stations. High rate stations use the default transmission rate while low rate stations use a transmission rate of 11 Mbit/sec.

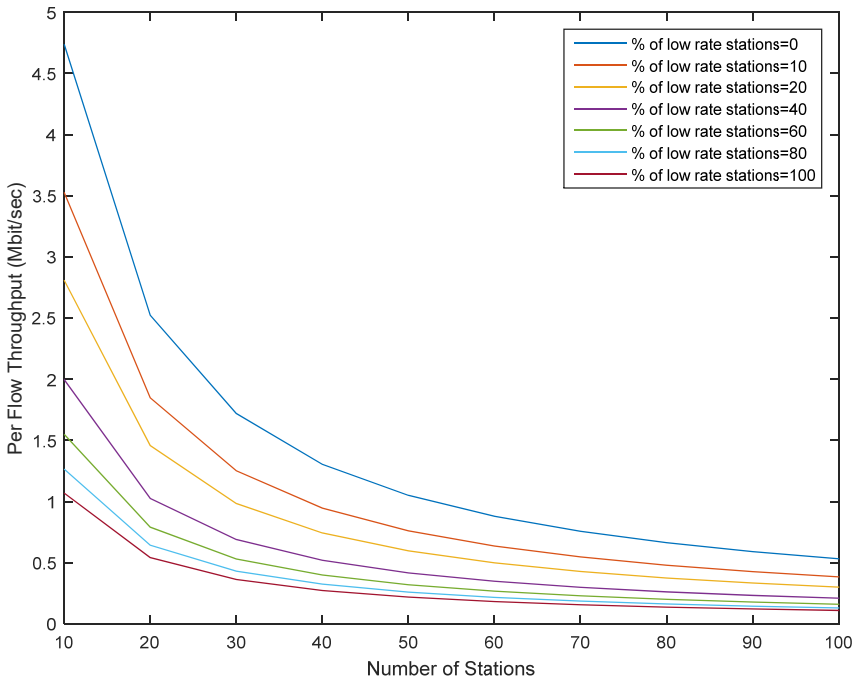


Figure 91: Per flow throughput as a function of the number of stations for different mixtures of high and low rate stations

5.2.5 Discussion

What is important to notice here is that for IEEE 802.11-based RERUM deployments, assuming that the traffic of such devices can be high, it is not really possible to support too many devices per gateway. If the traffic flows are demanding, then the network gets saturated quite quickly and cannot support more than 20-30 devices. However, this is not a very realistic scenario since demanding traffic flows here means that they will request ~20Mbps, which, even for high-definition video cameras is very high, but we wanted to show the network performance even for the worst case scenario. When the traffic flows are less demanding, i.e. for the saturation level=0.1, the network can support without problems more than 200 devices per gateway, and even in this case the throughput per device will be more than 1Mbps. It can be easily understood that for standard IoT applications that require only a few Kbps, the network would be able to support more than 500-1000 devices per gateway, without major drops in the network performance. This is very important when considering the large-scale requirement for IoT deployments in smart cities and it is proved that the proposed RERUM mechanisms can reasonably scale well even in high traffic scenarios.

5.3 QoS support in large-scale Cognitive RERUM WLANs

5.3.1 Introduction

Opportunistic spectrum access is a key technology for the development of novel applications with stringent QoS requirements on top of wireless sensor networks. However, the many challenges posed by opportunistic spectrum access itself are further exacerbated by the power constrained nature of wireless sensors and the QoS requirements of the applications to be supported. Within RERUM, as we described in RERUM Deliverable D4.1 [RD4.1], we considered the problem of dynamic spectrum assignment in the context of cognitive CSMA/CA based wireless sensor networks and proposed a two-stage dynamic spectrum assignment scheme that exhibits minimal power consumption and guarantees the QoS requirements of data flows in the network. The proposed scheme utilizes both opportunistic and free access bands, via cognitive radio technology, and is based on a two-stage stochastic integer program with recourse. The details of the proposed scheme and numerical results that verified its effectiveness were presented in [STTQ15] and [STTG15].

According to [STTQ15] and [STTG15] the optimal program that determines which channels will be used for each state of the Primary Users (PUs) that operate in the Opportunistic Access (OA) band is derived by the solution of a stochastic integer program. We tackled this stochastic integer program by transforming it into its deterministic equivalent Integer Program (IP) form. Let N be the total number of channels in the OA band. Furthermore, assuming that each RD cannot access all N channels of the OA band at once, let n be the number of accessible channels by the RDs and let h_d be the number of different priority classes for traffic flows in the network. Given N, n and h_d then the size of the deterministic IP is of $O(2^{n+h_d})$. We expect that n and h_d values will be small, given that wireless RDs are typically resource constrained devices and that the values of N and n are typically predefined by cognitive communication standards such as the 802.11af standard that sets N to be 40 [ETSI, Ofcom] and n to be at most 4. Thus, although IPs are known to be NP-hard, we expect that both the size and the solution space of the IP will deem it feasible by standard MILP solvers in typical scenarios. Furthermore, the computational complexity of the overall spectrum assignment problem increases linearly with the number of available channels since the IP is solved $(N+1)-n$ times in order to find the optimal spectrum assignment program. **Finally, we note that the complexity of the proposed scheme does not depend on the number of RDs in the CWLAN or CWPAN**.

5.3.2 Relation to RERUM UCs

This proposed technique can be useful for all RERUM use cases. The main objective is to identify in an energy efficient way the optimum spectrum to be assigned to the RDs, regardless of it's connectivity technology, assuming that the RD has cognitive radio technology capabilities. The RD will gather some spectrum occupancy information from a central point (either a gateway or a spectrum occupancy database) and at the second stage it will identify which will be the spectrum portion to occupy for the transmissions, considering the QoS requirements for the service(s) it supports and the requirement for minimum energy consumption. Furthermore, the RD will ensure that high-priority traffic flows will remain unharmed, aiming to achieve the highest network capacity. Thus, it is not use-case specific, but it is a very important contribution for achieving (as proved in [RD4.1]) a very high network performance.

5.3.3 Performance analysis of scalability of the network depending on the type of flows

The number of RDs as described above does not play a role on the performance of the two-stage spectrum assignment algorithm. On the other hand, the number and the type of flows in the network will have a profound effect on the number of flows with stringent QoS requirements that can be supported by the network. More specifically, assuming that the network supports two classes of flows with different priorities, termed the high and low priority class respectively [STTQ15], [STTG15], the transmission rate r_h required to guarantee an average throughput x_h to high priority flows will be given by,

$$r_h = \frac{N_h \cdot \phi_h}{\phi_h \cdot \frac{p_h}{x_h} - 1 - N_l \gamma_l} \tag{11}$$

where N_h and N_l are the number of high and low priority flows respectively and $\phi_h = \frac{\rho_h \mathbb{E}[L_h]}{\mathbb{E}[B_h]}$, where terms $\rho_h, \mathbb{E}[L_h], \mathbb{E}[B_h]$ and γ_l were defined in Equations (8) and (9). For transmission rate r_h to be a real number it must hold that,

$$0 \leq N_l \leq \frac{\phi_h p_h - x_h}{\gamma_l \cdot x_h} \tag{12}$$

We see in Figure 92 that the number of low priority flows in the network N_l rapidly diminishes as the average guaranteed throughput of high priority flows x_h increases. Furthermore, given x_h, N_l will decrease as the transmission rate used by low priority flows decreases. This is due to low priority flows occupying the channel for longer intervals when transmitting with a low rate and thus leaving less time for high priority flows to transmit and achieve their target average throughput.

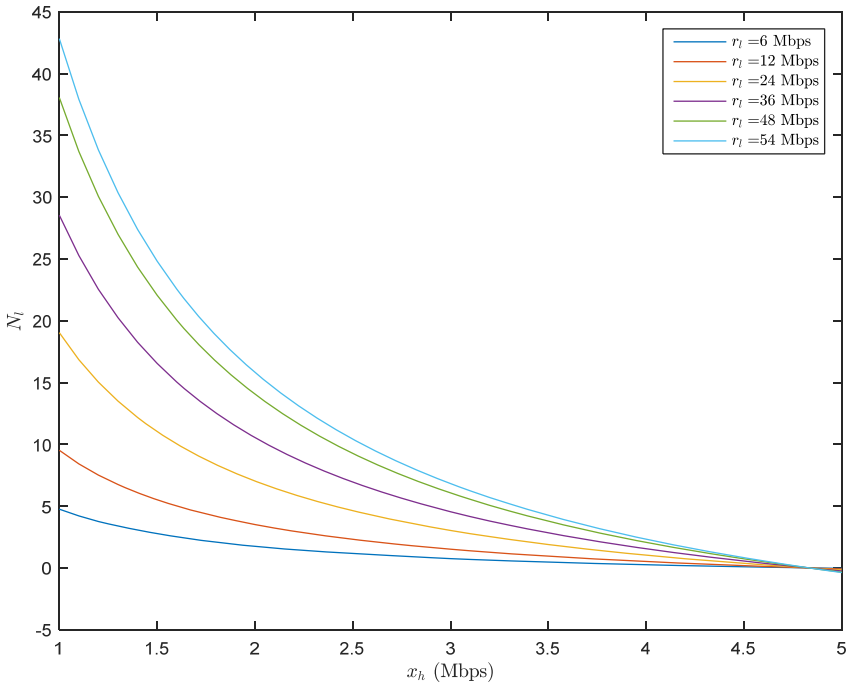


Figure 92: Maximum number of low priority flows as a function of the guaranteed average throughput of high priority flows for different values of low priority flows' transmission rate.

What is more, the constraint of Equation (12) is not the only one active in the network. Communication standards typically limit the available transmission rates to a specific set. Figure 93 presents the transmission rate, r_h, necessary for high priority flows to achieve a guaranteed average throughput of 2 Mbps when low priority flows use their maximum transmission rate (54 Mbps). Long before the number of low priority flows reaches its maximum, as determined by the constraint in Equation (12), r_h skyrockets to values that are not supported by any of 802.11 and 802.15.4 standards. Even CWLANs and CWPANs that are capable to support higher transmission rates by spectrum bundling will lack the necessary bandwidth to achieve the target transmission rate. Even service differentiation techniques, as the one presented in [STTG15] where high priority flows have a lower expected backoff value in order to access the medium with higher probability, will fail to provide the necessary transmission rate to high priority flows as the network scales up.

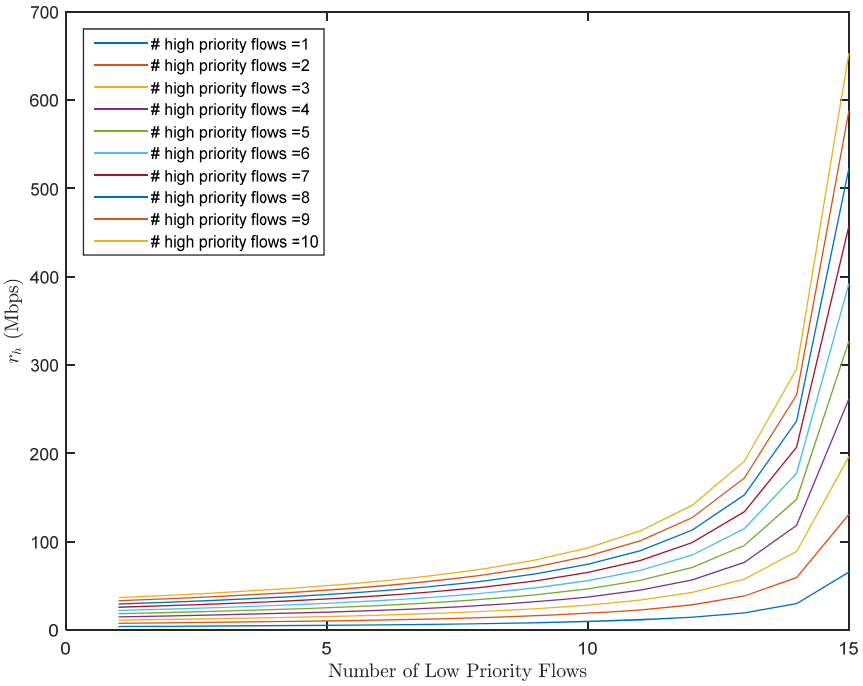


Figure 93: Transmission rate of high priority flows as a function of the number of low priority flows for different numbers of high priority flows supported by the network.

5.3.4 Discussion

It is important to notice here that since the two-stage spectrum assignment mechanism is running separately on each RD, the number of RDs in the network will not have an impact on each performance. The reason for this is that by nature this algorithm only considers the primary users when deciding for the spectrum assignment and not the other cognitive RDs. Thus, the number of the RDs indeed does not affect the performance of the mechanism itself. For the performance of the RERUM network when multiple RDs coexist, please refer to the section 5.2 above. What is also important to notice is that the type of the flows in the network is a significant factor that affects the performance of the RERUM network. This is proved by the fact that if there are many high-priority flows in the network, these have to be served immediately, increasing the delays or decreasing the throughput of the low-priority flows. This happens due to the fact that the high-priority flows have strict QoS requirements, thus the more they are the less is the remaining capacity of the network to serve the low priority flows. This means that if there are many high-priority flows in the network, the scalability of the system cannot be high. However, in the above-mentioned analysis, we have used a high required throughput for the high-priority flows, which is not very common in IoT scenarios. If the required throughput is much lower, even in such cases the scalability of the network can be quite high.

5.4 Scalability of HetNets

5.4.1 Introduction

State-of-the-art radio access solutions, such as the latest Long Term Evolution (LTE) releases, and their corresponding evolutionary paths (future LTE-Advanced releases) will not be able to fulfill these demands, as the recently started ITU vision work item (IMT 2020+) also suggests [ITU-R]. Traditional single-link PHY optimization techniques seem to have "hit the wall" in terms of achieved spectral efficiency levels. Multi-cell multi-link cross PHY- and MAC-oriented techniques (such as CoMP or Network MIMO), based on the key concept of transforming all the interference into useful signal via cooperation in order to enhance system capacity, reflect today's trends in mobile broadband, but target only cell-edge users and suffer from scalability issues. Dense multi-tier heterogeneous network deployments or "HetNets", empowered by Interference Coordination approaches for applying transmissions orthogonalization in various domains (frequency, time, space, power and code), are alternative measures to the capacity crunch.

It seems that the applicability of these concepts and technologies is inherently limited by the restraints of the existing cellular-based architecture. The cellular architecture conceived for stand-alone working units with limited processing and inter-communication signaling capabilities is intrinsically not suitable for large coordinated and cooperative systems. To be able to support cooperation, it requires heavy protocols and imposes stringent constraints. Hence the actual benefits of the proposed technologies prove negligible compared to their theoretically predicted potential. The existing cellular-based architecture is the actual limiting factor not only for applying existing technologies, but also for inventing new ones that would potentially further improve system capacity.

The limitations of the cellular networks, in terms of scalability, especially when considering scenarios like IoT, where a number of devices much larger than in conventional scenarios are expected to be deployed, are quite worrying. 5G networks, i.e., multi-RAT flexible networks are considered to be the solution to these limitations, where multiple technologies will cooperate (actually they will be coordinated at the core level) in order to provide the optimum access method to the devices, by taking in to account multiple factors, such as the type of devices, the QoS, the radio conditions over a wide range of bands, etc. RERUM aims at catching up with these technological advances and propose efficient techniques which will provide improved spectrum efficiency and QoS.

This section will investigate the scalability factor for such hybrid WiFi/Cellular hybrid network deployments. The main objective is to investigate how the access algorithm behaves in cases where the number of the things/devices within a typical cell is of the order of hundreds. The simulation tool is designed to take into account crucial parameters, such as the cell radius, the number of available WiFi access points, the number of devices, the transmission power of the eNB and the WiFi caccess points, the available bandwidth, the QoS of the devices, etc. The performance will be investigated mainly in terms of the achieved average signal-to-noise and interference.

Regarding the simulation setup, we consider a communication scenario with the following assumptions:

- The two cells utilize the same spectrum
- RDs in each cell are perfectly orthogonal
- WiFis in each cell are perfectly orthogonal
- RDs are uniformly distributed within a circular cell
- At the downlink, one of the cells interferers with another
- All RDs have the same received SNR

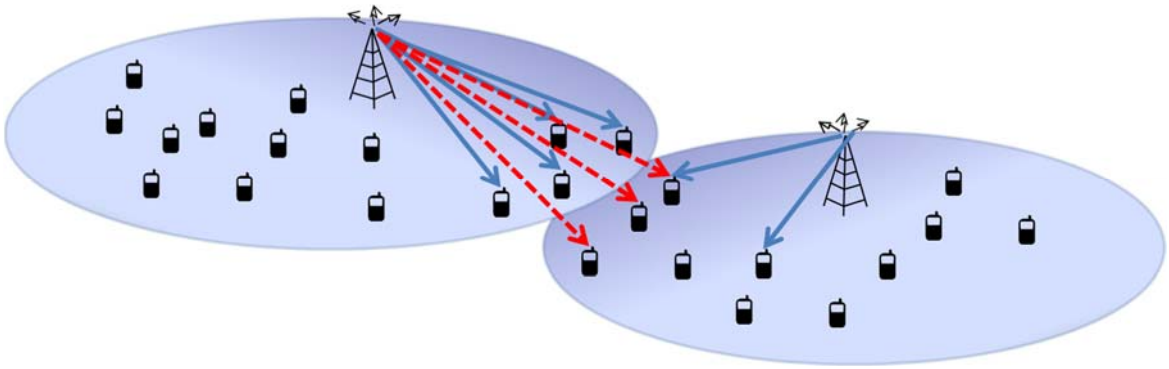


Figure 94: Simulation setup - conventional cellular

5.4.2 Relation to RERUM UCs

This network set-up directly applies to a typical RERUM deployment where the RDs are served (i.e., get connected to the MW via internet) through a RERUM GW, or directly through a public access network (see Figures below).

Although the proposed mechanisms could be directly applied to all RERUM deployments deployments that involve high mobility and hence increased possibility for frequent hand-overs should be avoided. In this sense, the proposed mechanism would ideally fit to the following RERUM use cases:

- Environmental monitoring (UC-O2)
- Home energy monitoring (UC-I1)
- Comfort quality monitoring (UC-I2)
-

In these deployments, assuming that the network provider owns both the WAN (e.g., LTE) and WLAN (e.g., WiFi) networks and that the RDs have both a WAN and WLAN radio interface, we can achieve considerably higher spectrum efficiency and fairness in the resource utilization by the RDs.

The simulation tool can support any scalability requirement in terms of the number of devices within the cells and any topology. Figure 94 and Figure 95 illustrate some deployment examples, demonstrating the scalability of the simulation tool. In those figures, two adjacent cells are considered where wifi access points coexist with cellular users that can also serve as access points by sharing their cellular broadband connections. The users are depicted with the "*" symbol, while the wifi access points with the "o" symbol. The access points of the one cell interfere with the access points of the other cell. The same assumption holds for the cellular users that serve as wifi access points.

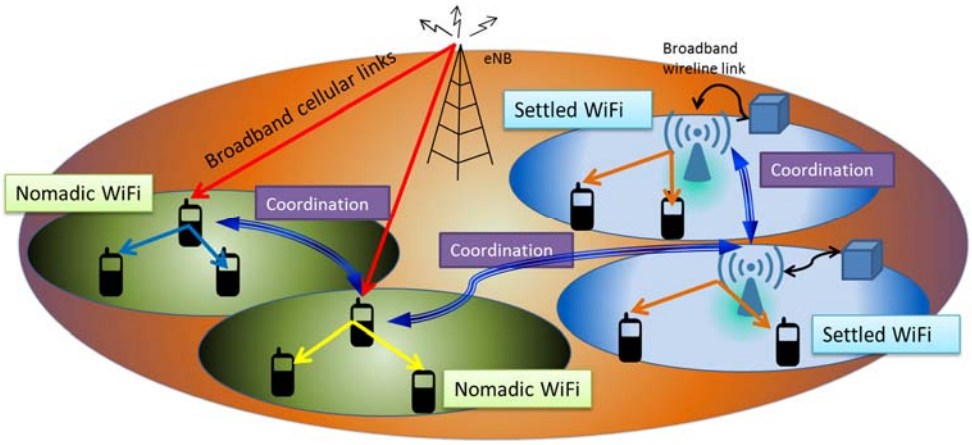


Figure 95: Simulation setup - hybrid cellular/WiFi

Assuming cellular network with cell radius equal to 500 meters and eNodeB distance equal to 1 kilometer, the total average signal to interference ratio (SIR) was plotted as a function of the number of WiFis within the cell, for various values of the WiFi range, as shown in Figure 100.

The graph shows that in all cases, utilization of the WiFi, when done using the proposed approach, increased the total SIR, irrespectively of the WiFi range. Also, with the increase in the WiFi range, the performance of the proposed algorithm also tended to improve.

When total SIR was plotted as a function of the cell radius as shown in Figure 96, it showed that in all cases, the utilization of the WiFi increased the total SIR, irrespective of the WiFi range. It is also noted that the performance decreased with an increase in the cell radius, while the best performance was obtained for higher values of the WiFi range.

Figure 96 assumes a cell radius equal to 500 meters, WiFi range equal to 20 meters and the number of WiFis equal to 20. The total average SIR was plotted as a function of the number of users for two communication scenarios, namely the proposed scheme (hybrid) and the conventional cellular scenario (no WiFis). It was shown that with the utilization of the WiFis, the total SIR increased linearly as the number of users increases.

5.4.3 Discussion

This subsection presented the scalability performance for the multi-RAT mechanisms presented in 4.1 and demonstrated how IoT deployments can benefit from them, and how network operators can exploit their WAN and WLAN networks towards a more efficient resource utilization. RERUM aspires to make those networking recommendations for IoT deployments, which will boost their performance in terms of spectrum utilization.

The presented results aim at pointing out important aspects regarding the network access technologies that are expected to play a crucial role in IoT deployments throughout the next years, i.e., 5G access networks. The first one is about the importance of hybrid network access for IoT, since it seems to be the most effective way to overcome the limitations of the existing networks and enable the deployment of a very large number of RDs. The additional degree of freedom that is added by such technologies (of course with an additional cost of complexity at the core level) provides numerous advantages in terms of spectrum utilization and improved QoS. On the other hand, it must become clear that that hybrid access without proper coordination algorithms may lead to the opposite results than those desired. Figure 97 and Figure 98 for example show that adding WLAN GWs may not always result to better performance.

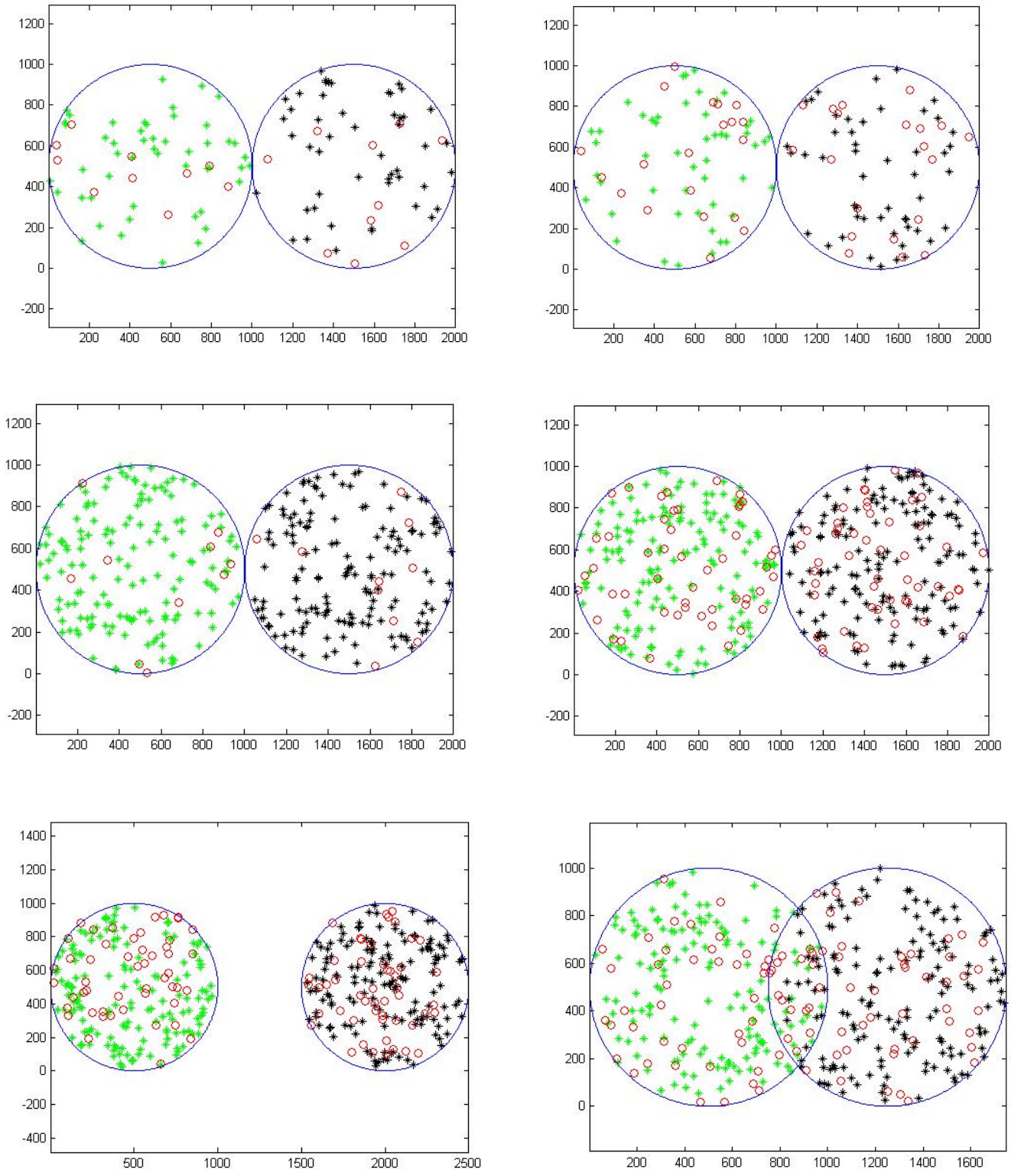


Figure 96: Different network deployments with different scalability factor, i.e., number of devices per km²

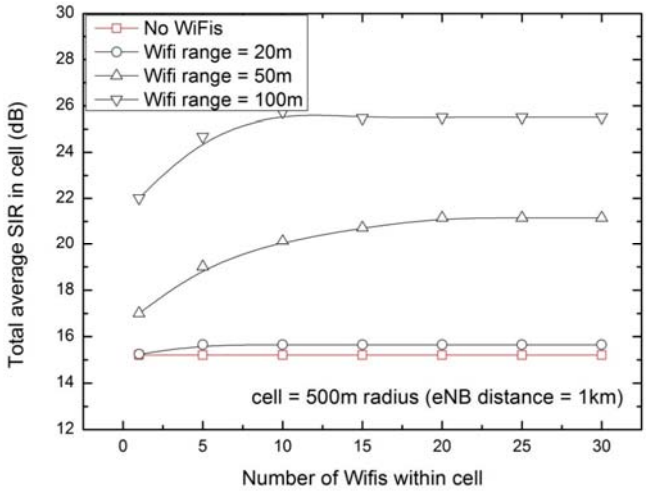


Figure 97: The scalability performance of the proposed mechanism

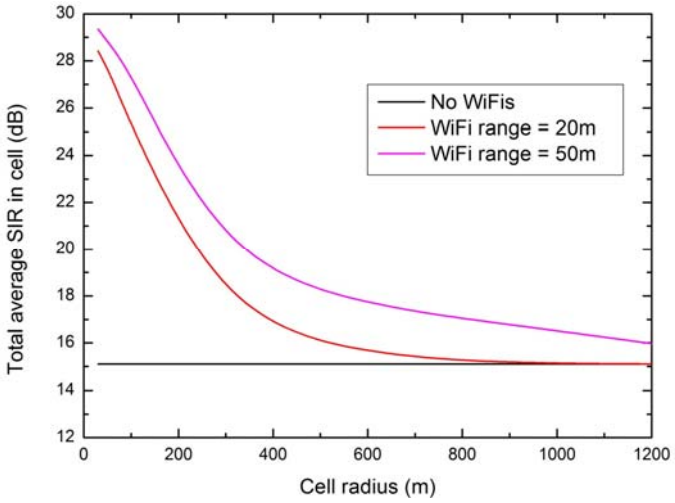


Figure 98: The scalability performance of the proposed mechanism in terms of cell size

5.5 Adaptive CS scalability analysis

5.5.1 Introduction

The Adaptation of the compression rate in CS acquisition and processing is essential due to the variable nature of sparsity in real-life signals that smart objects are sensing. In our proposed Adaptive CS (ACS) scheme, we use a Change Point Method (CPM) based on Kolmogorov-Smirnov (KS) statistic to detect changes in signal sparsity, by monitoring the residual of signal reconstruction algorithm, in our case Orthogonal Matching Pursuit (OMP). When a sparsity change is detected, additional CS measurements are requested from the smart object, until the required reconstruction error is achieved.

5.5.2 Relation to RERUM UCs

The adaptive CS mechanism belongs to the Data Encrypter/Decrypter component of the RERUM architecture, and can be used in all use cases apart from UC-O1. For example, environmental monitoring involves the collection of various diverse measurements (e.g. ambient temperature, ambient light, etc.) from multiple locations. As the sparsity of these data highly affects the performance of the CS encryption/compression, in terms of the reconstruction error, it is important to adapt the

compression rate so as to keep the error at a minimum level. The sparsity can be seen as a metric of the smoothness of the changes in the measurement signal. Thus, when for some reason, the signal starts to change for abruptly, the sparsity changes and the reconstruction error at the receiver might become very high. On the other hand, if the signal starts to become much smoother, the sparsity again changes and the reconstruction error becomes too low. In the first case, there is a need to adapt the compression rate so that more measurements are sent (less compression) in order to maintain a low reconstruction error. In the latter case, the transmitted measurements are too many and the sensor consumes more transmission energy, thus there is a need for higher compression. The adaptive CS scheme estimates the reconstruction error, and adapts the compression rate dynamically through a feedback mechanism based on the COAP protocol, between the RERUM Gateway and the RDs. Due to the diverse type of measurements gathered by the RDs in all use cases (apart from UC-O1), the ACS is very important to both have an accurate representation of the gathered measurements and maintain a low energy consumption at the RDs.

5.5.3 Performance evaluation

Obviously, usual requests for additional measurements due to sparsity changes can increase signaling overhead and, as a result, degrade network performance, especially for large network sizes. In the following, we propose ignoring part of the sparsity changes detected by the CPM, in order to decrease the required signaling for new sparsity estimation. We define p_i as the probability of ignoring a detected sparsity change, and we study the performance of the adaptive CS scheme in terms of reconstruction error e.

In particular, we generate blocks of $N = 64$ samples with sparsity levels varying in {5%-20%} and non-zero DCT coefficients independently drawn from a normal distribution $\mathcal{N}(0,1)$. Each sparsity level is chosen uniformly at random, and we investigate the behaviour of the proposed scheme for two different scenarios: (i) frequent sparsity changes (Scenario 1) where the interval (in number of blocks) between two successive sparsity changes is uniformly at random chosen in $[10, 15]$, and (ii) infrequent sparsity changes (Scenario 2), where this interval is uniformly at random chosen in $[20, 30]$. We change the sensitivity of the KS-CPM by varying the values of ARL (average number of observations between two false positives) in $\{100,\ 500, 1000\}$, and repeat each experiment 50 times. Finally, we vary the probability p_i in $[0, 0.7]$.

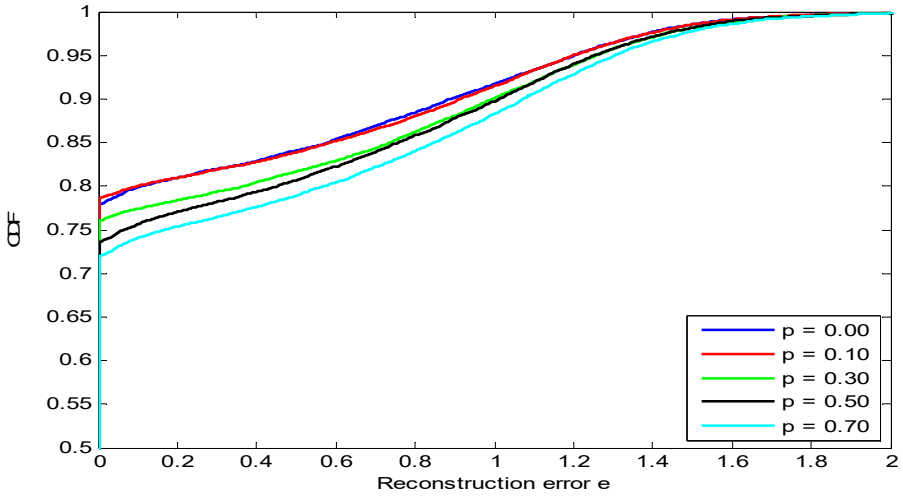


Figure 99: CDF of the reconstruction error for Scenario 1 and ARL = 100

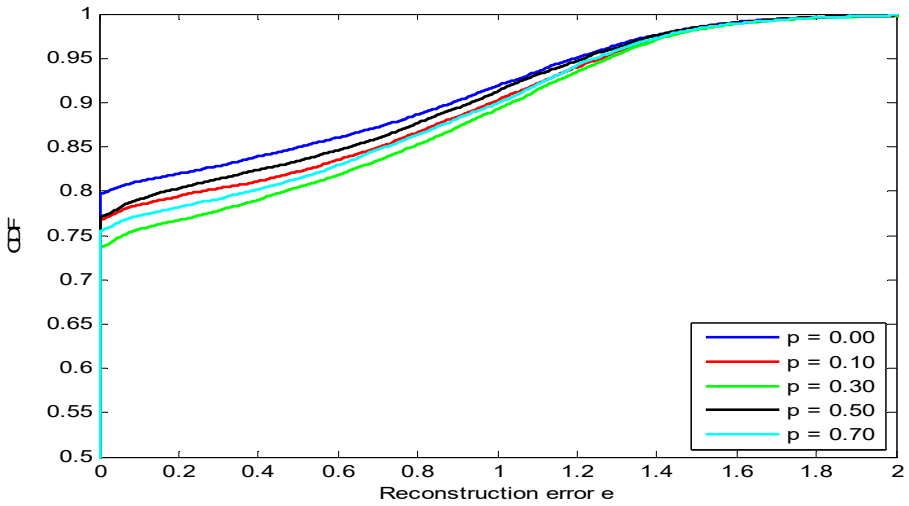


Figure 100: CDF of the reconstruction error for Scenario 1 and ARL = 500

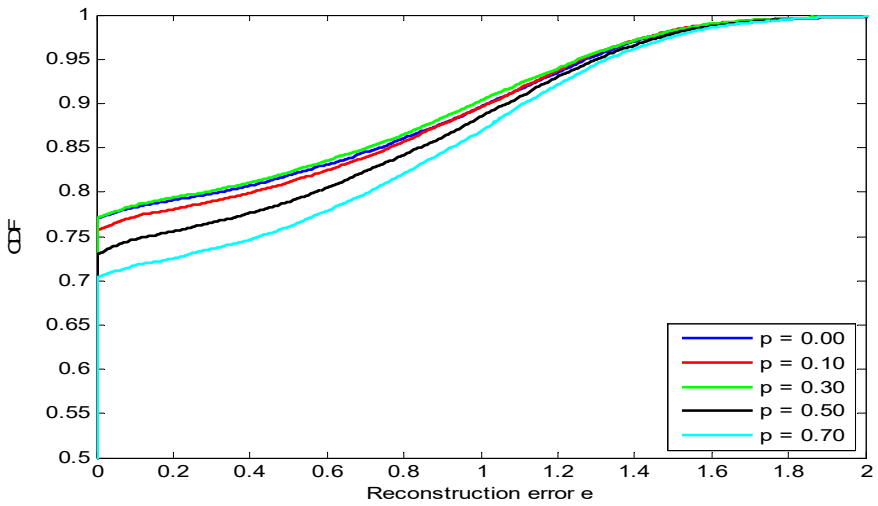


Figure 101: CDF of the reconstruction error for Scenario 1 and ARL = 1000

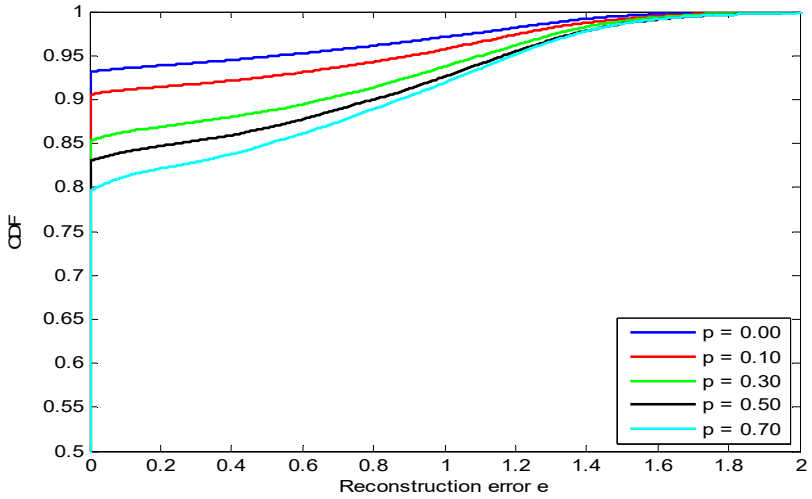


Figure 102: CDF of the reconstruction error for Scenario 2 and ARL = 100

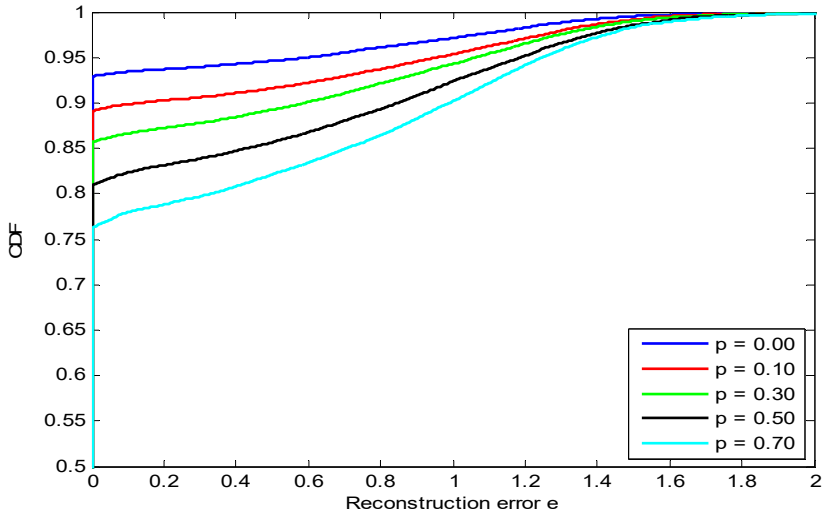


Figure 103: CDF of the reconstruction error for Scenario 2 and ARL = 500

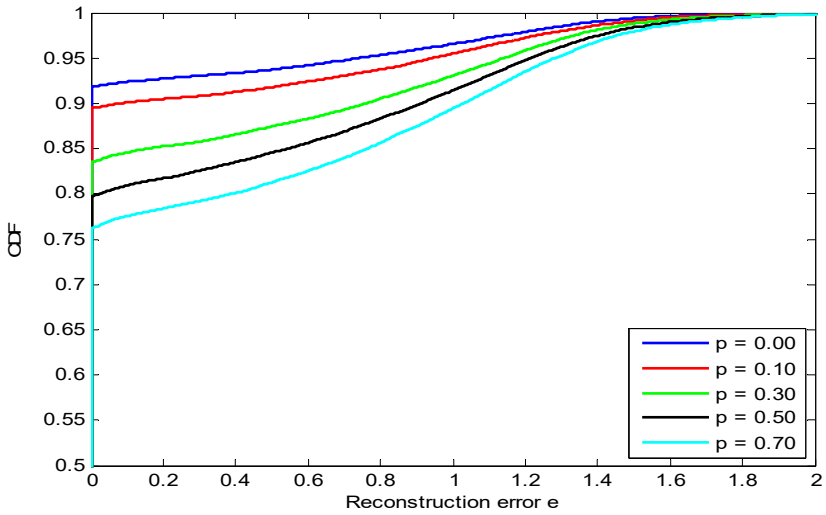


Figure 104 CDF of the reconstruction error for Scenario 2 and ARL = 1000

Figure 99, Figure 100, and Figure 101 display the CDFs of the reconstruction error for Scenario 1, while the error for Scenario 2 are shown in Figure 102, Figure 103, and Figure 104. As expected, the performance of ACS is better in case of Scenario 2 compared to that of Scenario 1. This happens because, for the latter case, the CPM statistic is unable to converge to a flat value, due to the low number of initialization samples, resulting in an increased number of false negatives.

The increase in probability p_i degrades performance of ACS, for all values of ARL. This is normal, since as p_i increases, ACS tends to behave like a static CS scheme that employs a non-changing compression rate. Thus, the compressed signal can be insufficiently sampled, compromising, that way, the reconstruction procedure. The difference is more profound for Scenario 2, where ACS performs the best. For example, in case of $ARL = 100$ (Figure 102), the probability of the reconstruction error lower than 0.01, drops from almost 93% ($p_i = 0$) to 79% ($p_i = 0.7$).

5.5.4 Discussion

As it is evident from the results presented above, whenever we identify that for a specific measurement signal there are frequent sparsity changes, we have to consider all sparsity changes in the ACS scheme, namely we have to maintain a very low probability of ignoring sparsity changes. This

will increase the signalling in the network, due to the requests for more measurements from the RDs to the GWs, but it is a necessary trade-off for maintaining a good security level and a good reconstruction error at the receiver. On the other hand, when some measurement signals are found to having sparsity changes very infrequently, we proved that even ignoring a small percentage of those changes (up to 30%), we can have a very low reconstruction error at the receiver. This contributes significantly to the decrease in the overall signalling and traffic in the network, improving the scalability of the system. Overall, using CS it is not always possible to maintain both a very low signalling and a good reconstruction error, but the proposed ACS scheme with the possibility to ignore some sparsity changes can improve the scalability of the system in most scenarios, even when there are many sparsity changes (which I s not very realistic, but here we tried to assess the scalability of the proposed scheme in the worst case).

5.6 Trust-based routing scalability analysis

5.6.1 Introduction

Trust management has been used as a simple and efficient approach to mitigate a wide number of attacks in routing of WSNs. Nodes create trust relationships based on the expectation that their neighbours will cooperate on particular tasks (e.g. packet forwarding) and decide on routing paths based on them. Our trust-based routing model has been built on the trust degree values that are computed based on the observed behaviour of an RD by its neighbour RDs. Next, we revise in brief the steps included in our trust-based routing scheme:

- Each node observes and records neighbour's forwarding behaviour each time there is an interaction between them. Malicious behaviour is captured by using (i) *packet drop rate* (PDR) and (ii) *packet modification rate* (PMR). Aggregate *misbehaviour rate* (MBR) of RD *j* as perceived by RD *i* is:

$$MBR_{i,j} = w \times PDR_{i,j} + (1-w) \times PMR_{i,j}$$

  where $w \in [0,1]$.

- Link quality between neighbours is quantified using the *expected transmission count* (ETX) metric. ETX, essentially, expresses the average number of transmissions needed for a packet to successfully reach its destination in cases when there are transmission failures due to degradation of link quality (e.g. interference, collisions, etc.)

- MBR views, reported for each node by its neighbours, are fused at the sink using *Dempster-Shaffer* theory, in order to handle uncertainty introduced by possibly unreliable values. Thus, a combined belief b_j^{CT} on trust degree of each node *j* is computed.

- Routing paths are computed/updated in a centralized fashion (at the sink) by using Dijkstra's shortest path algorithm, based on the following combined routing metric and calculated for each route *r*:

$$RM_r = \sum_{i,j \in r} ETX_{i,j} \left(1 - b_j^{CT}\right)$$

5.6.2 Relation to RERUM UCs

The trust-based routing algorithm belongs to the SPT Manager of the RERUM architecture, and can be used in all use cases apart from the smart transportation. As trust in IoT is of paramount importance, and the RERUM project focuses on security and trust issues, the proposed trust-based scheme can effectively detect malicious RDs that intentionally, or unintentionally, misbehave. The detected malicious RDs are isolated from the rest of the system, and their reports (measurements) are ignored. As described in RERUM Deliverable D2.5 [RD2.5], IoT deployments can be multi-hop deployments with many intermediate nodes between the RDs and the GWs. When sensitive information need to be transmitted from leaf devices to the RERUM MW, it is very important to try to find the most trusted routes from the RD to the GW. The notion of trust here incorporates both security and networking criteria. We try to find out the routes that are composed of devices that have not been hacked, do not alter measurements, do not drop packets or do not delay the transmissions significantly. For this, we have proposed a trusted routing mechanism which was evaluated for its performance in Deliverable D4.1 [RD4.1]. However, since the projection for the future is that IoT deployments will be quite large with many devices per gateway, here we will assess the scalability of the proposed algorithm with regards to the number of devices.

5.6.3 Performance analysis

Next, we study the behaviour of the proposed trust-based routing scheme for different network sizes, as regards number of nodes N, namely $N=\{50,100,200,300\}$ nodes. Nodes are placed uniformly at random in a square area of 100×100 meters and a sink node is placed at the centre of the area. Trust belief update interval is set to 20 rounds and we choose $w=0.5$. In order to account for some variance of wireless channel characteristics, in each round we vary randomly the ETX of all links in $[90\%,110\%]$ of their nominal values. We proceed by evaluating our scheme by means of packet delivery rate at the sink and assume two scenarios, namely *severe malicious behaviour* and *light malicious behaviour*. In the first case, malicious nodes exhibit PDR and PMR with values in $[0.8,1]$ while in the second case the values vary in $[0.3,0.5]$. For each scenario we execute 20 Monte Carlo runs each consisting of 2000 data aggregation rounds. Percentage of malicious nodes p_M is varied in $[0.1,0.4]$.

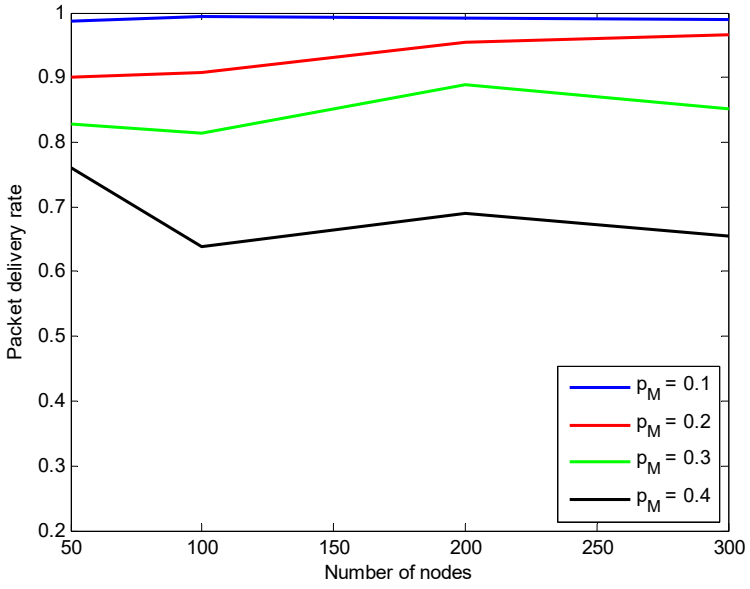


Figure 105: Packet delivery rate vs number of nodes for light malicious behaviour

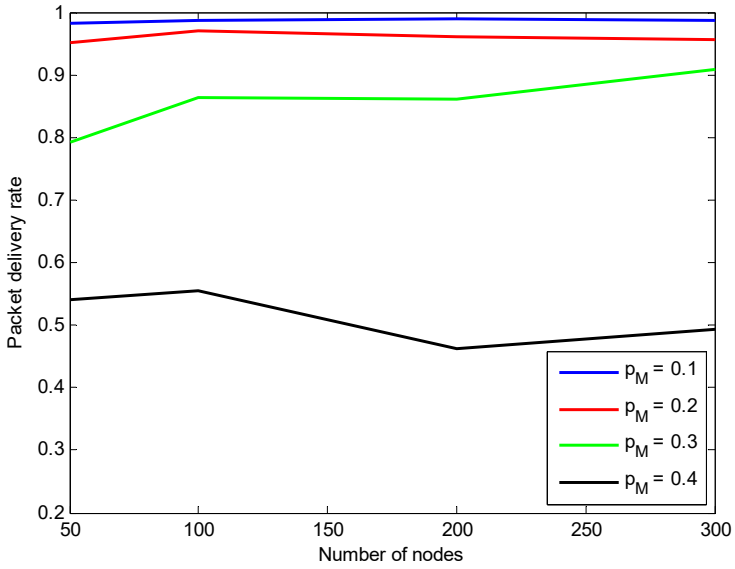


Figure 106: Packet delivery rate vs number of nodes for severe malicious behaviour

Figure 105 and Figure 106 depict the packet delivery rate versus the number of nodes in case of light and severe malicious behaviour, respectively. Obviously, packet delivery rate degrades as the percentage of malicious nodes increases, for both cases. This degradation is more profound in the case of severe malicious behaviour since high misbehaviour rate of malicious nodes leads to an increased number of dropped packets. For example, for $N = 200$ and $p_M = 0.4$, the packet delivery rate is almost 68% for light malicious behaviour but decreases to almost 45% for severe malicious behaviour. As regards network size, we observe no notable differences in packet delivery rate for the different values of N .

Furthermore, we study the behaviour of our scheme in case of packet loss in signal packets that fetch the observed MBR values to the sink. We vary MBR packet loss in $[0, 0.5]$ with a step-size of 0.1 and report packet delivery rate for the aforementioned scenarios, namely light malicious behaviour and severe malicious behaviour. Number of nodes is $N = 100$ and percentage of malicious nodes p_M is varied again in $[0.1, 0.4]$.

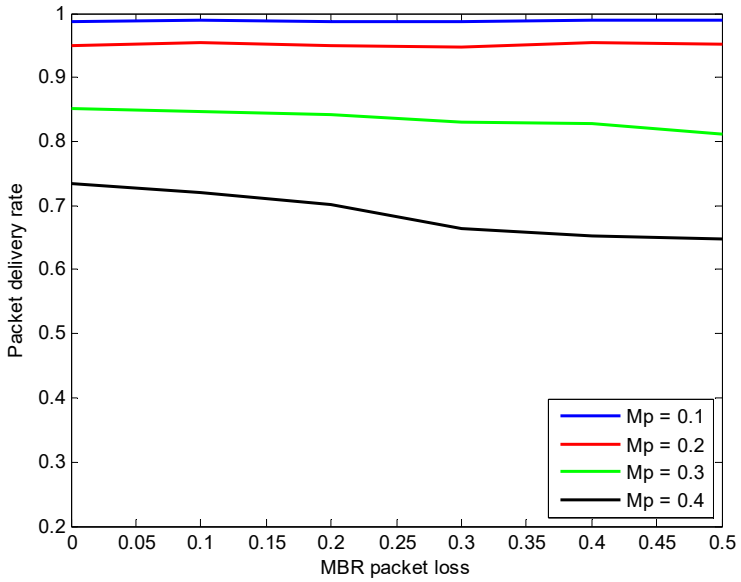


Figure 107: Packet delivery rate vs MBR packet loss for light malicious behaviour

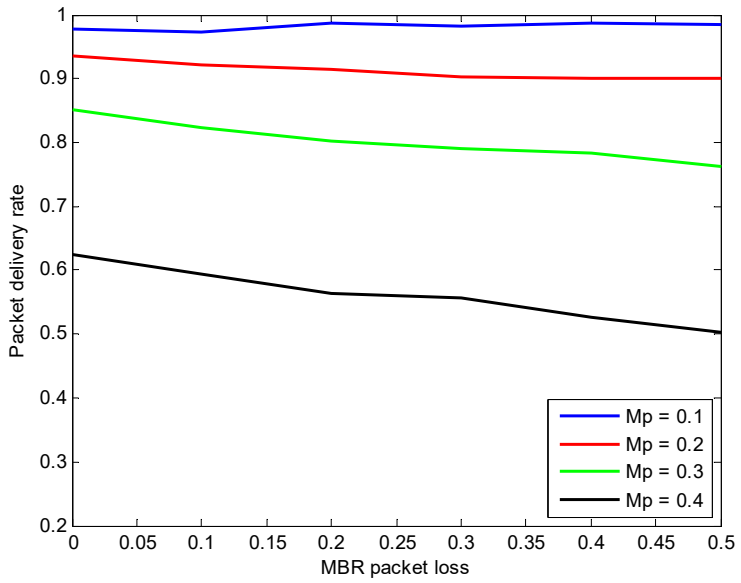


Figure 108: Packet delivery rate vs MBR packet loss for severe malicious behaviour

In Figure 107 and Figure 108 the packet delivery rate versus MBR packet loss is depicted. As a general trend, it can be seen that there is a decrease in packet delivery rate as MBR packet loss increases, with this behaviour being more profound in the case of severe malicious behaviour and high percentage of malicious nodes. This is normal since the more MBR values are lost due to packet loss, the less probable is that sink possesses sufficient information in order to isolate malicious nodes and omit them from routing paths.

5.6.4 Discussion

We presented the scalability analysis of our trust-based routing scheme for two attack intensities. The results show that the packet delivery rate degrades as the percentage of the malicious nodes increases, for both cases. Furthermore, when the network size increases, there are no notable differences

regarding the packet delivery rate. There is also a decrease in the packet delivery rate as the MBR packet loss increases, with this behaviour being more profound in the case of severe malicious behaviour, and with a large number of malicious nodes.

It is evident, considering the above presented results, that the proposed trust based routing scheme can work even in large scale deployments with many sensors per GW without any significant changes in its behaviour. It is proved that the successful percentage of packet delivery is kept almost constant with very small drops or even with increase in some cases when the number of devices per gateway are increased. This is quite reasonable, because the larger is the network of devices per gateway the larger is also the number of possible routes from the RD to the GW. Similarly, when there are packet losses to the signaling packets carrying the MBR from the nodes to the GW, the drop to the packet delivery rate is not significant. This means that even if there are nodes that are intentionally dropping signalling packets, the trust-based routing scheme can mitigate this behaviour and it can maintain a very high packet delivery rate, which means that proper trusted routes are identified even in this attack scenario. Thus, the trusted routing scheme scales very well not only with regards to the number of RDs per gateway, but also with regards to the packet loss percentage.

5.7 Scalability of Underlay network of RDs

Here we consider a cognitive network with one primary pair with bursty traffic and many randomly distributed RDs as secondary nodes having saturated queues as in the previous section. Receivers are assumed to have multipacket reception (MPR) capabilities so that the RD nodes operating as sceondary nodes of the CR network can transmit simultaneously with the primary under certain conditions. The MPR channel captures the effects of fading, attenuation, and interference at the physical layer in a more efficient way than the collision channel model.

We propose a delay-aware shared secondary network access scheme for the RDs with congestion control for the primary. A large-scale secondary network is considered in which the RD nodes are distributed according to a stochastic point process. We derive the average queue size and delay of the primary user as function of the secondary node access probability and transmit power. We introduce an optimization problem to maximize the secondary throughput subject to the delay constraints on the primary user. We analyze the impact of different network parameters on the behavior of secondary throughput and primary delay[4].

5.7.1 Relation to RERUM UCs

For the outdoor use cases of RERUM, it is expected that networks will need to support a potentially very high number of RDs and therefore scalability is critical. Especially for the case of the Smart Transportation (UC-O1), it can be the case that a large number of RDs (mobile phones) may need to be served by the same BS or Gateway. In such a case, we could consider that the networking architecture could allow for underlay networks to be deployed for which we provide the scalability analysis in this section. Therefore, this section can be viewed as an investigation towards an appropriate networking technology for the outdoor use cases of RERUM. Looking at the bigger picture of the IoT however, advocating for underlay (secondary) networks operating with predicted performance degradation cost to the primary as we do here is work highly relevant also in 5G, which has the IoT as one of the pillar use cases.

5.7.2 System Model

Topology

[4] This work has been accepted for presentation at the IoT SOS 2016 workshop (http://www.ics.forth.gr/tnl/IoT-SoS-2016/), co-organized by RERUM. An extended version, whose preprint is on arXiv [Z++16], has been also submitted for journal publication in April 2016.

We consider a network consisting of one primary source-destination pair at fixed location and randomly distributed secondary pairs, as shown in Figure 109 The primary device can be a regular subscribed user of a certain network operator, thus needs to be served with higher priority. The secondary devices are the RDs seeking to access the primary spectrum to communicate in a random manner with a certain access probability. The network region is a circular disk $\mathcal{C}$ with radius R. The primary receiver is centered at the origin of $\mathcal{C}$. The primary transmitter is located at fixed location with distance d_p to the primary receiver. We assume that the secondary RD transmitters are distributed according to a Poisson point process (PPP) $\Phi_s = \{x_i \in \mathbb{R}^2, \forall i \in \mathbb{N}^+\}$ with intensity λ_s. Their associated receivers are distributed at isotropic directions with fixed distance d_s from their transmitters. The time is slotted and each packet transmission occupies one time slot. We assume that all receivers have multipacket reception (MPR) capabilities and so RD nodes can transmit simultaneously with the primary node.

The primary source has an infinite capacity queue Q for storing arriving packets of fixed length. The arrival process at the primary transmitter is modeled as a Bernoulli process with average rate λ packets per slot. The secondary node queue is assumed to be saturated, i.e., it always has a packet waiting to be transmitted.

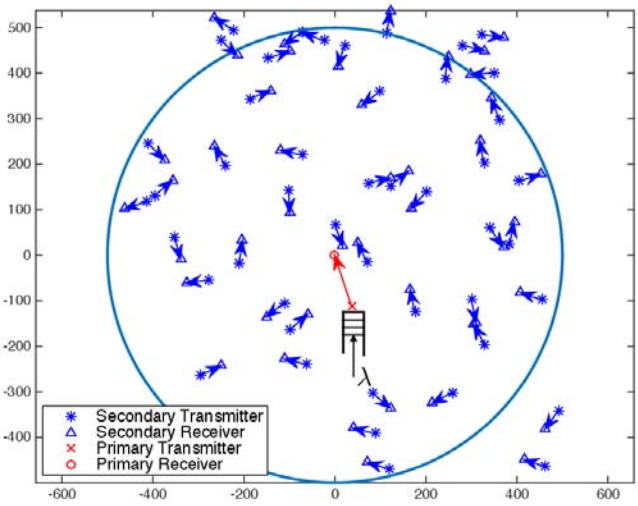


Figure 109: The cognitive network topology: one primary receiver centered at the origin with PPP distributed secondary transmitters under a given density, i.e., $\boldsymbol{\lambda_s = 5 \times 10^{-5}}$.

Priority –Based Protocol Model

We consider the following cognitive protocol. The primary node has bursty packet arrivals and transmits a packet whenerver backlogged. The secondary RD nodes access the channel with a probability that depends on the queue size of the primary node and such that the performance of the primary user is not severely degraded. Denote Q the queue size in the primary node, the activity of the secondary RDs in a time slot are as per the following cases:

• *Case 1*: When $Q = 0$, the primary transmitter does not have any packet to transmit. RD transmitters randomly access the channel with probability q_1.

• *Case 2*: When $1 \leq Q \leq M$, the primary transmitter transmits one packet. RD transmitters randomly access the channel with probability q_2.

• *Case 3*: When $Q > M$, the primary transmitter transmits one packet. RD transmitters remain silent.

The threshold M above plays the role of a congestion limit for the primary, meaning that when the queue reaches this size then, the secondary nodes defer from access to avoid collissions and increasing primary queue through failed transmissions. For brevity, we use PT and PR to denote the primary transmitter and receiver respectively, and ST and SR for the secondaries.

5.7.3 Physical Layer Model

With MPR capabilities enabled at the receivers, a packet can be decoded correctly by the receiver if the received signal-to-interference-plus-noise ratio (SINR) exceeds a prescribed threshold θ. Given a set $\mathcal{T}$ of nodes transmitting during the same time slot, the received SINR at the i-th receiving node is given by

$$SINR_i = \frac{P_i|h_{i,i}|^2 d_{i,i}^{-\alpha}}{\sum_{j\in\mathcal{T}\setminus\{i\}} P_j|h_{j,i}|^2 d_{j,i}^{-\alpha}+\sigma^2},$$

where P_i denotes the power of the transmitting node i; $|h_{j,i}|^2$ denotes the small-scale power fading from the transmitter j to the receiver i, which follows exponential distribution (Rayleigh fading) with with mean value equal to 1; $d_{j,i}$ denotes the distance between the transmitter j to the receiver i. Here we assume a standard distance-dependent power law pathloss attenuation $d^{-\alpha}$, where $\alpha > 2$ denotes the pathloss exponent; σ^2 denotes the background noise power.

Let P_1 and P_2 be the transmit powers of the PT and the STs, respectively. Denote x_0 the location of the PT and recall that the distribution of the STs is given by Φ_s, then we have $\mathcal{T} \subseteq \{x_0 \cup \Phi_s\}$. Note that in this work when we refer to the set of locations of the transmitting nodes, it means the set of transmitting nodes at these locations.

Following the description of our access protocol presented in Section 2.2, to derive the success probability of the primary and secondary nodes we need to consider three cases.

Case 1

When $Q = 0$, the PT is silent, STs attempt packet transmission with probability q_1. Denote Φ_a^1 the locations of active STs, as a result of independent thinning, Φ_a^1 is also a homogeneous PPP with intensity $q_1\lambda_s$. In this case the active transmitters are $\mathcal{T} = \Phi_a^1$. Without loss of generality, we consider an arbitrary (typical) ST x_i with the receiver at the origin. Denote $p_{2/2}$ the success probability of the typical secondary pair when only the STs from Φ_a^1 are active, we have

$$p_{\frac{2}{2}} = \mathbb{P}[SINR_i > \theta | \mathcal{T} = \Phi_a^1 = \mathbb{P}\left[\frac{P_2|h_{i,i}|^2 d_s^{-\alpha}}{\sigma^2 + \sum_{j\in\Phi_a^1\setminus\{x_i\}} P_2|h_{j,i}|^2 d_{j,i}^{-\alpha}} > \theta\right]$$

$$\overset{(a)}{=} \exp\left(-\frac{\pi q_1\lambda_s d_s^2 \theta^{\frac{2}{\alpha}}}{sinc(2/\alpha)}\right)\exp\left(-\frac{\theta\sigma^2 d_s^{\alpha}}{P_2}\right).$$

Here, (a) comes from $|h_{i,i}|^2 \sim \exp(1)$ and the probability generating functional (PGFL) of the PPP.

Case 2

When $1 \le Q \le M$, both the PT and part of the STs are active. Similarly, the locations of active STs follow another homogeneous PPP, denote Φ_a^2, with intensity $q_2\lambda_s$. Hence, the active transmitters are $\mathcal{T} = \{x_0 \cup \Phi_a^2\}$.

Denote $p_{1/1,2}$ and $p_{2/1,2}$ the success probabilities of the primary and secondary pairs in this case. Similarly, with the help of existing results on the interference distribution in PPP networks, we have

$$p_{1/1,2} = \mathbb{P}[SINR_0 > \theta | \mathcal{T} = \{x_0 \cup \Phi_a^2\}] = \mathbb{P}\left[\frac{P_1|h_{0,0}|^2 d_p^{-\alpha}}{\sigma^2 + \sum_{j\in\Phi_a^2} P_2|h_{j,0}|^2 d_{j,0}^{-\alpha}} > \theta\right]$$

$$= \exp\left[-\frac{\pi q_2 \lambda_s \left(\theta\frac{P_2}{P_1}\right)^{2/\alpha} d_p^2}{sinc(2/\alpha)}\right] \exp\left(-\frac{\theta\sigma^2 d_p^{\alpha}}{P_1}\right).$$

For the secondary RD network, we obtain the success probability in the following proposition.

Proposition 1 *The success probability of the typical secondary RD pair, when the active transmitters are $\mathcal{T} = \{x_0 \cup \Phi_a^2\}$, is given by*

$$p_{2/1,2} = \exp\left(-\frac{\pi q_2 \lambda_s d_s^2 \theta^{\frac{2}{\alpha}}}{sinc(2/\alpha)}\right) \frac{\exp\left(-\frac{\theta\sigma^2 d_s^{\alpha}}{P_2}\right)}{1+\frac{d_s^2}{\mathbb{E}[d_{0,i}]^2}\left(\theta\frac{P_1}{P_2}\right)^{\frac{2}{\alpha}}},$$

where $\mathbb{E}[d_{0,i}] = \int_0^{2\pi} \frac{1}{2\pi} \int_0^R \frac{2r}{R^2} \sqrt{r^2 + d_p^2 - 2rd_p \cos\varphi} \mathrm{d}r\mathrm{d}\varphi$.

Proof of the proposition can be found in [Z++16]

Case 3

When $Q > M$, only the PT is active. Denote $p_{1/1}$ the success probability of the primary pair, we have

$$p_{1/1} = \mathbb{P}[\mathrm{SINR}_0 > \theta | \mathcal{T} = x_0] = \mathbb{P}\left[\frac{P_1|h_{0,0}|^2 d_p^{-\alpha}}{\sigma^2} > \theta\right] = \exp\left(-\frac{\theta\sigma^2 d_p^{\alpha}}{P_1}\right)$$

Note that $p_{1/1} > p_{1/1,2}$ and $p_{2/2} > p_{2/1,2}$ always hold.

5.7.4 Network Performance Metrics

We define here several metrics for the performance evaluation of our shared access network.

Scalability of Secondary RD Throughput

The throughput of the secondary network, abbreviated as secondary throughput, is the number of packets per slot that can be transmitted by the active secondary nodes to their destinations. In order to be consistent with the PPP model where the secondary nodes are generated with a certain density λ_s, we define the secondary throughput as the the throughput of the secondary network per unit area, given by

$$T_s = \lambda_s \mathbb{P}[SINR_{i\in\Phi_s} > \theta].$$

Recall that the active STs are with density $q_1\lambda_s$ when the primary queue size is $Q = 0$, and with density $q_2\lambda_s$ when $1 \le Q \le M$. Hence, we have

$$\begin{aligned} T_s = \; & \mathbb{P}[Q = 0] \cdot q_1\lambda_s \mathbb{P}[SINR_i > \theta | Q = 0] + \mathbb{P}[1 \le Q \le M] \\ & \cdot q_2\lambda_s \mathbb{P}[SINR_i > \theta | 1 \le Q \le M] \\ = \; & \lambda_s\left\{\mathbb{P}[Q = 0] \cdot q_1 p_{2/2} + \mathbb{P}[1 \le Q \le M] \cdot q_2 p_{2/1,2}\right\}. \end{aligned}$$

Primary Service Rate

The service rate of the primary user given a certain SINR target can be defined as the percentage of successfully transmitted packets per time slot. When $1 \le Q \le M$, the primary service rate is

$$\mu_1 = p_{1/1,2}.$$

When $Q > M$, the service rate is

$$\mu_2 = p_{1/1}.$$

5.7.5 Primary Average Delay

The delay per packet at the primary node consists of the queueing delay and the transmission delay from the PT to the PR. From Little's law, we obtain the queueing delay which is related to the average queue size per packet arrival. The transmission delay is inversely proportional to the average service rate.

Denote $\overline{D}_p$ the primary average delay per packet, we have

$$\overline{D}_p = \frac{\bar{Q}}{\lambda} + \frac{1}{\bar{\mu}},$$

where $\bar{Q}$ and $\bar{\mu}$ are the average queue size and the average service rate of the primary.

5.7.6 Performance Analysis

We focus on the impact of the design parameters on the scalability of the secondary RD throughput with respect to the primary delay constraints. In order to do so, the primary queue and delay are firstly derived with closed-form expressions, then an optimization problem is formulated in terms of maximizing the total secondary throughput while keeping the primary delay below a certain threshold.

Primary Queue and Delay

We model the primary queue as a discrete time Markov Chain (DTMC), which describes the queue evolution and is presented in Fig. 2. Each state is denoted by an integer and represents the queue size. The packet arrival rate is always λ. The service rate is $\mu_1 = p_{1/1,2}$ when $1 \leq Q \leq M$, and is $\mu_2 = p_{1/1} > \mu_1$ when $Q > M$. All the metrics related to the rate are measured by the average number of packets per slot.

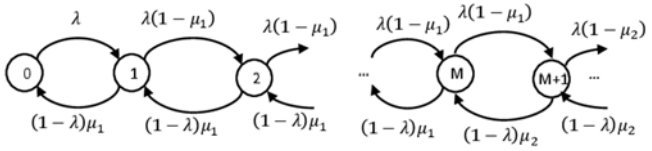


Figure 110: The Discrete Time Markov Chain which models the queue evolution at the primary node.

Denote π, the stationary distribution of the DTMC, where $\pi(i) = \mathbb{P}[Q = i]$ is the probability that the queue has i packets when it is in steady state. To simplify the equations, we define $\xi \triangleq \frac{\lambda(1-\mu_1)}{(1-\lambda)\mu_1}$. In the remainder we will assume that $\lambda \neq \mu_1$, however the general expressions of our results hold also for $\lambda = \mu_1$, but one should replace the $\pi(0)$ with the corresponding expression in this case. We have

$$\mathbb{P}[1 \leq Q \leq M] = \frac{\lambda(1-\xi^M)(\mu_2-\lambda)}{\mu_1\mu_2 - \lambda\mu_1 - \lambda\xi^M(\mu_2-\mu_1)},$$

$$\mathbb{P}[Q > M] = \frac{\lambda\xi^M(\mu_1-\lambda)}{\mu_1\mu_2 - \lambda\mu_1 - \lambda\xi^M(\mu_2-\mu_1)}.$$

The average queue size at the PT is given by

$$\bar{Q} = \frac{N_1+N_2}{\mu_1\mu_2-\lambda\mu_1-\lambda\xi^M(\mu_2-\mu_1)},$$

where

$$N_1 = \lambda(1-\lambda)\mu_1\frac{\mu_2-\lambda}{\mu_1-\lambda}[M\xi^{M+1} - \xi^M(M+1)+1],$$

and

$$N_2 = \xi^M \lambda(\mu_1 - \lambda)\left[M + \frac{(1-\lambda)\mu_2}{\mu_2 - \lambda}\right].$$

The average delay of the primary user is given by

$$\overline{D}_p = \frac{\bar{Q}}{\lambda} + \frac{\mu_2 - \lambda - \xi^M(\mu_2 - \mu_1)}{(1-\xi^M)(\mu_2-\lambda)\mu_1 + \xi^M(\mu_1-\lambda)\mu_2}.$$

Obviously, $\overline{D}_p$ is independent of q_1. When q_2 increases, $\mu_1 = p_{1/1,2}$ decreases, thus $\overline{D}_p$ increases because of the larger queue size and higher transmission delay. Similarly, we know that $\overline{D}_p$ is also an increasing function of P_2.

Secondary RD Throughput with Primary Delay Constraints

From the equations above, the secondary RD throughput can be derived as:

$$\begin{aligned} T_s &= \lambda_s\left(\mathbb{P}[Q=0]\cdot q_1 p_{2/2} + \mathbb{P}[1 \le Q \le M]\cdot q_2 p_{2/1,2}\right) \\ &= \lambda_s \frac{(p_{1/1}-\lambda)[q_1 p_{2/2}(p_{1/1,2}-\lambda) + q_2 p_{2/1,2}\lambda(1-\xi^M)]}{p_{1/1,2}p_{1/1} - \lambda p_{1/1,2} - \lambda\xi^M(p_{1/1}-p_{1/1,2})}. \end{aligned}$$

Considering T_s as a function of q_1, it is obvious that there exists an optimal value $q_1^* = \underset{q_1\in[0,1]}{argmax} T_s$, which is equivalent to $q_1^* = \underset{q_1\in[0,1]}{argmax} q_1 \cdot p_{2/2}$. We have then that the optimal access probability q_1 of the secondaries when the PT is silent, given by

$$q_1^* = \min\left\{\frac{sinc(\frac{2}{\alpha})}{\pi\lambda_s\theta^{\frac{2}{\alpha}}d_s^2}, 1\right\}.$$

Setting q_1^* in the T_s above when the PT transmit power P_1 and packet arrival rate λ are fixed, the secondary throughput depends only on the ST access probability q_2 and the transmit power P_2.

Since the primary user delay is an increasing function of q_2 and P_2. When the primary queue is stable, i.e., $\lambda < \mu_2$, taking into account the delay constraints of the primary user, we obtain the feasible region of the two variables $\{q_2, P_2\}$ defined as

$$\mathcal{R}_{\mathcal{F}} = \{(q_2, P_2): \overline{D}_p(q_2, P_2) < D_{\max}\},$$

where $D_{\max}$ is the threshold of the primary average delay.

In order to achieve the maximum secondary throughput with respect to the primary user delay constraints, we define an optimization problem as follows.

$$(q_2^*, P_2^*) = \underset{(q_2,P_2)\in\mathcal{R}_{\mathcal{F}}}{argmax}\, T_s,$$

subject to:

$$q_2 \in [0,1],$$
$$P_2 \in [0, P_{2,\max}],$$

where $P_{2,\max}$ is the maximum transmit power for the STs.

The optimization problem is hard to solve due to the involved analytical expressions related to the primary queue size. Hence, we resort to numerical evaluations.

5.7.7 Numerical Evaluations

Here we evaluate the secondary throughput as a function of the two variables (q_2, P_2) within their feasible region that satisfies the delay constraints of the primary user. The primary delay and the

feasible region boundary are also presented, showing the impact of the priority based protocol design on the network performance. The parameter values are given in Table 32: **Simulation Parameters**.Our results justify the scalability of the scheme we propose, given the obtained trhoughput for very tight constraints and large number of RDs.

Table 32: Simulation Parameters

Parameters	Values
ST density (λ_s)	2×10^{-4}
Secondary link distance (d_s)	40 m
Primary link distance (d_p)	300 m
Cell size (R)	500 m
Pathloss exponent (α)	4
PT transmit power (P_1)	100 mW
Noise power (σ^2)	-113.97 dBm
SINR target (θ)	0 dB
Delay threshold ($D_{\max}$)	3.5 time slots

In Figure 111 below the success probabilities $p_{1/1}$, $p_{1/1,2}$, $p_{2/2}$ and $p_{2/1,2}$ are depicted as a function of the ST access probability q_1 or q_2, when fixing the ST transmit power at $P_2 = 0.1$ mW. Recall that $p_{1/1}$ is a constant value, $p_{1/1,2}$ and $p_{2/1,2}$ only depend on q_2, $p_{2/2}$ only depends on q_1. As expected, when the secondary network is active, the success probabilities decline rapidly with q_1 and q_2 increasing, as a result of the increasing interference. In the simulations all the results are obtained with $\lambda < p_{1/1}$ so that the queue stability condition is satisfied.

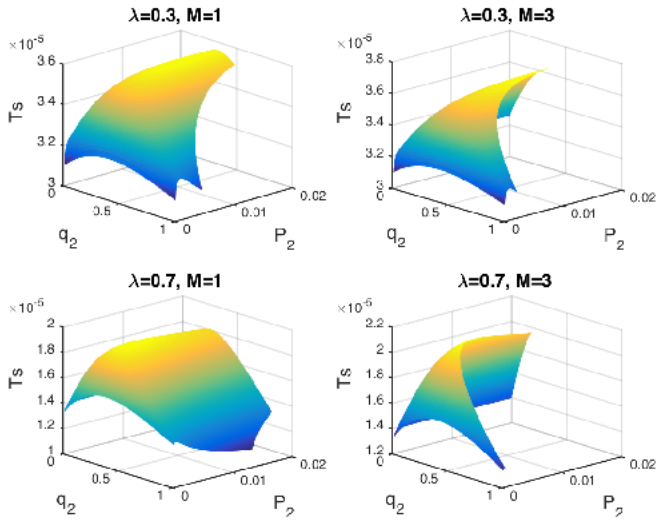


Figure 111: Success probabilities $\boldsymbol{p_{1/1}}$, $\boldsymbol{p_{1/1,2}}$, $\boldsymbol{p_{2/2}}$ and $\boldsymbol{p_{2/1,2}}$ vs the ST access probability $\boldsymbol{q_1}$ and $\boldsymbol{q_2}$, fixing the ST transmit power at $\boldsymbol{P_2 = 0.1}$

Below in Figure 112 we plot the secondary throughput under the primary delay constraints. The results are presented with the congestion threshold $M = \{1,3\}$ and the packet arrival rate $\lambda = \{0.3,0.7\}$. Our first remark is, the secondary throughput is not a monotonic function of q_2 and P_2. There exists an optimal point that gives the maximum T_s among the feasible choices of (q_2, P_2). We also observe that larger M provides higher potential improvement for the secondary throughput, as the secondary links are more likely to be active. In order to validate our conclusion, in Table. 2 we give the numerical values of the optimal solution (q_2^*, P_2^*) as well as the maximum SU throughput achieved with different λ and M. We confirm that for the same λ, larger M increases the maximum achievable secondary throughput.

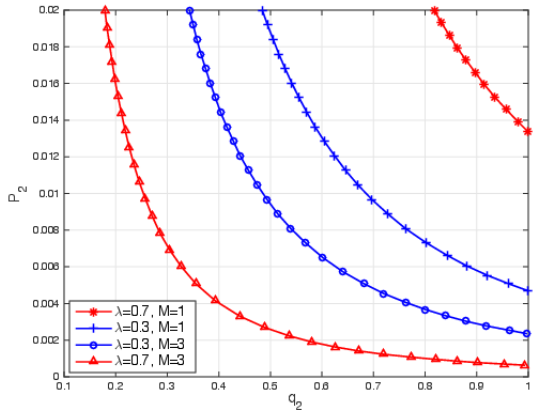


Figure 112: The boundary of the feasible region of $(\boldsymbol{q_2}, \boldsymbol{P_2})$ with $\boldsymbol{\lambda} = \{\mathbf{0.3}, \mathbf{0.7}\}$ and $\boldsymbol{M} = \{\mathbf{1}, \mathbf{3}\}$. Below each curve is the feasible region $\boldsymbol{\mathcal{R}_{\mathcal{F}}}$ with the the specific values of $\boldsymbol{\lambda}$ and $\boldsymbol{M}$.

5.7.8 Discussion

Concluding the numerical results, we have the following takeaway messages regarding the design parameters of our cognitive protocol for RERUM.

1. With larger congestion threshold M, the maximum secondary throughput for the RDs is expected to be higher. However, larger M put tighter constraints on the feasible values of (q_2, P_2). This means that for use cases without tight resource restriction on the network side (such as the UC-O1, where the network capacity is limited by the BS) one can scale to large expected throughputs with this type of underlay RDs. However although the average will be high the actual feasible values will be low.

2. With higher arrival rate λ at the PT, both the access probability and transmit power of the RD nodes should be lower in order to achieve higher service rate for the primary user. As a result, the primary queue size will decrease faster, which in turn gives more chance for the STs RDs to transmit during the next time slot.

6 Conclusions and overall discussion

6.1 Conclusions

Throughout the deliverable we have presented our discussions on the items investigated in each of the sections. Here we provide a distilled, concise summary.

With respect to the scalability and performance of the sensors available for the Trial implementations of RERUM UC-O2, I1 and I2, we have identified trade-offs between cost and measurement reliability. Specifically we identified temperature, light, humidity, and pressure sensors which are inexpensive and reliable enabling large deployments. Gas sensors investigated were found also reliable albeit more expensive, thus the scale of deployment can be good but not as much as of the previous group. Finally PM10 sensors are quite inexpensive but need additional processing for collaborative sensing among groups of them, since the single sensor measurement quality was found to be poor.

On other trade-offs, firstly between Security and energy consumption under DTLS we have exposed a fundamental trade-off between computation time, energy consumption and system security (key management). We concluded that to balance it one has to consider the actual requirements of the applications. To this end in Section 3.1 we argued for EEC using example of a typical RERUM UC-O2 – Environmental Monitoring), or dense scenarios, while our results indicated that in real-time applications Pre-Shared Key -based schemes are more suitable. For different attack models, we demonstrated the effectiveness of the detection and prevention, by two methods of validation. Trade-offs between network performance and its density (i.e., scalability) were uncovered with higher the network density leading to more successful attacks were observed in topologies, which implies a potential correlation between the attack model's effectiveness and the network density. Using Compressive sensing for security is meaningful in large scale instances, and the benefits significant in execution time and energy. With respect to energy, following on-off schemes addressed in D4.2, we investigated trade-offs between duty-cycling and 6LoWPAN performance and we concluded that we can by tackling the congestion on a low-power wireless network, decrease packet retransmissions, and thus reduce the amount of time an RD has to spend awake before it can exit its congested state. We concluded that CADC can be used to optimally tune the duty-cycles, as the actual optimal, being dependent on the network input generated by the running applications cannot be preset. Finally, we discussed how multiple access in 5G can lead performance (in terms of scalability throughput and service availability) to be tuned with fairness (guaranteeing a minimum QoS to all RDs).

For RD Networking, having considered already 5G as a potential, investigations took place over heterogeneous networks, in terms of not only access technologies, but also cell size. Clearly deploying hundreds of RDs in a specific geographical area without any provisioning on the access network, it will cause considerable interference, network congestion and bad QoS. Offloading has been investigated through considering load coupling of neighboring cells, for both small cells and WiFi deployments, assuming fairness ensuring utilities. An overhead reduction technique on D-MIMO was also presented applicable especially for large scale scenarios. Finally the timeliness of the received information is addressed again at MAC layer through a novel metric (AoI), in a specific set of time-critical messages.

In terms of the scalability quantified how well secure communication can be supported by the low-powered devices scales with increasing number and how it can adapt to mobility of RDs. Our results show that overall the LR-MAC did not have a significant measurable impact on the network connectivity, however there is a considerable delay toll. With respect to trust in the routing functions, the trusted routing scheme explored in RERUM, scales very well not only with regards to the number of RDs per gateway, but also with regards to the packet loss percentage. As for the scalability of different access technologies, traffic is the key, for WPANs extremely large (>3000) number of RDs per gateway can be supported without a significant drop in the network performance. On the other hand, classic IEEE 802.11 for RERUM would scale at an order of magnitude lower, as with requirements for only a few Kbps, the network would be able to support more than 500-1000 devices per gateway,

without major drops in the network performance. Introducing cognition, delivers scalability independence in the two-stage assignment, while if considered under the classic primary/secondary scheme, we have shown that RDs working in an underlay network can achieve large scales, even with stringent QoS constraints set by a primary user of the shared frequency.

6.2 Overall discussion

IoT deployments are expected to skyrocket in the next few years, since they have become an important trend especially towards enabling smart city applications. The municipalities among others aim to start investing on IoT infrastructures so that they can build upon them various smart applications for improving their bureaucracy and for improving the quality of life of their citizens. This has become evident through the excessive interest that is shown in everything that is IoT-related either on hardware or on software. It is obvious that this interest will only become larger as the IoT technologies become more mature and more market-ready.

However, there are various obstacles that have yet to be overrun before making the serious transition to large-scale IoT deployments. In RERUM we have split these obstacles to two different categories: (i) security, privacy and trust and (ii) networking. During the more than 30 months from the beginning of the project until the delivery of this document, the project partners have worked towards defining, designing and developing mechanisms and protocols for improving the security, privacy, trustworthiness and networking performance of IoT networks with a focus on the RERUM system requirements. The performance and the efficiency of the developed mechanisms have been presented in the previous project deliverables, and have been favourably reviewed in top-tier journals, magazines, conferences and workshops. Thus, it has been proved that the RERUM contributions have been quite important for the IoT domain. However, judging the performance of standalone mechanisms is not the only thing that is important in the IoT world. It has to be proved that the algorithms will perform equally well in large scale scenarios and this was the main goal of this document.

More specifically, the focus of this document was to provide answers to two key questions regarding the IoT deployments in general:

- What are the trade-offs between security, performance and energy consumption?
- How can the proposed mechanisms or deployment scenarios scale well when the number of devices becomes very large, i.e. in smart city deployments?

Although these are key questions for all IoT systems, we as RERUM have focused the analysis we have made solely on key mechanisms that we have developed within our project and we don't aim to provide answers for all IoT systems and deployment models.

Question 1:

With regards to the energy consumption, a detailed analysis of the RERUM mechanisms for improving the energy efficiency of the RERUM system has been provided in the RERUM Deliverable D4.2 [RD4.2]. In sake of completeness, we will remind here that the energy efficiency mechanisms were focused on

- Energy efficient and secure data gathering and transmission using Compressive Sensing,
- Congestion-aware duty cycling for 6LoWPANs
- Energy aware relay properties for RERUM gateways,
- Minimization of energy consumption of RDs with multicast forwarding
- Multi-radio selection for RDs aiming to minimize energy consumption

- Analysis of the energy consumption of RERUM authorization mechanisms
- Analysis of the energy consumption of JSON signatures
- RERUM IoT low-power hardware.

The above analyses have dealt with the energy efficiency of both the networking and security mechanisms of RERUM and we have proved that indeed we have designed and developed quite lightweight mechanisms that can run on constrained IoT devices. This is a major advancement of the project that had since its beginning the focus to embed lightweight mechanisms on the IoT devices and throughout its whole period until now promotes exactly this need, arguing that an IoT system can only be as secured as its devices are.

On the other hand, it is reasonable to assume that for making some mechanisms lightweight to be running on constrained devices, some compromises have to be made, which may result into a lower level of security or performance. In this deliverable we have tried in sections 2, 3 and 4 to discuss these trade-offs for various mechanisms and for various deployment cases examined in the RERUM system.

More specifically, with regards to trade-offs between security/trust/reliability and energy efficiency or performance, the key findings are:

- With regards to DTLS for real-time applications it is preferred to use the PSK mode which is less secure but it requires less computations and is much quicker. On the other hand, when security is at stake, ECC should be the preferred option.
- With regards to multicast forwarding, the proposed BMFA mechanism can achieve very low energy consumption compared to rivals like Trickle Multicast and can reduce end-to-end delays. However, when there is a need for high system reliability, the Trickle Multicast has to be preferred.
- With regards to the system duty cycling, the proposed CADC can significantly improve the energy efficiency of the RDs.
- With regards to the system trustworthiness based on the sensing reliability, a game theoretic approach can be used to identify attacks in the sensing in order to identify and mitigate outliers.
- With regards to sensing reliability of the hardware devices, it is proved that although most of the hardware is very low cost, it has been certified for its accuracy and reliability. Although it is known that the more expensive is the hardware its accuracy is much higher, for the majority of IoT applications, these proposed hardware sensors have an excellent ratio of value for money.
- With regards to encryption and compression, it is proved that there is an important trade-off between security and energy efficiency/memory/storage requirements. When there is a limitation in the storage and memory of the hardware devices, SRM matrices can be used which are very lightweight but their encryption strength has not been proved to be very high. This is though compared to Bernulli or Gaussian matrices who have higher encryption strength, but require much higher memory and storage.
- With regards to multiple access technologies in the RERUM network, we proved that the proposed selection mechanism can increase the number of served RDs and can improve the reliability of the network.

Question 2:

With regards to system scalability, the focus of the work in the deliverable was for assessing the scalability of the various states of the proposed RERUM network deployment and of some key mechanisms. The main findings can be summarized as below:

- When cognitive radio technology is used in the RERUM network, its scalability is improved significantly. We have shown that for cognitive RERUM WPANs, thousands of devices can be supported without significant drop in the network performance. This is of outmost importance for large scale deployments and it proves that any RERUM network can really work flawlessly (when outside interference are not considered though).
- Similarly, for cognitive RERUM WLANs, the system scales quite well when the requested throughput of the devices is kept in normal IoT levels (meaning not many Mbps).
- In terms of scalability of RERUM networks assuming that we provide QoS support, we proved that they also scale quite well when there are two priority classes and the throughput demands of the high priority class is not very large.
- With regards to heterogeneous RERUM network deployments, we provided an analysis of how these deployments can scale with the use of proper resource coordination mechanisms between them.
- With regards to using RDs as an underlay network, we proved that the RDs can scale to large expected throughput without affecting significantly the primary users.
- With regards to the adaptive CS scheme, it was evident that the algorithm scales adequately with regards to the number of sparsity changes and can indeed keep the reconstruction error at satisfactory levels even by neglecting some sparsity changes.
- With regards to the trusted routing mechanism, we showed that the mechanism scales very well with the number of RDs in the network and even with more than 300 RDs per gateway we can achieve a very high packet delivery rate.

From the above we can see that for the network deployment the proposed RERUM technologies contribute to very good system scalability, allowing the growing of the network to even high numbers of devices per gateway, without a severe degradation of the system performance.

Taking as an example one of the use cases, the smart transportation use case, we can also discuss briefly the scalability of the overall system. Regarding the Scalability aspects of the RERUM Middleware, in D5.3 we have already reported a host of lab trial measurements indicating the properties of components in the UC-O1 case.

Specifically in D5.3 Section 4 we have included the most critical aspect of scalability of the solution is that of the server side application. The aim of the experiments conducted was to evaluate the scalability of the server-side implementation of the smart transportation use case. The experiments evaluate jointly the middleware interface and the traffic application server, but not the middleware itself. The experiment was run in the Phase 1 of the trial of the UC-O1. However instead of considering the 30 busses deployed with the app, we performed an emulation of 50.000devices.
The main takeaway message of the experiment had been that CPU load increases clearly with the number of messages posing a potential bottleneck of the system using the current implementation if a vast amount of users would be included. However, the process that uses most of the CPU was identified to be the database, which has some clear potential in performance improvement by tuning polling intervals of the java application and the database structure. It could also be an alternative to use several threads to enable better use of the four CPUs that are available on the machine.
For details of the experiment conducted within WP5 we direct the reader to D.5.3 Section 4.

References

[ACFP2009]	Anastasi, G.; Conti, M.; Francesco, M.D.; Passarella, A. Energy conservation in wireless sensor networks: A survey. Ad Hoc Networks 2009, 7, 537 – 568.
[AKS2013]	Intrusion Detection Systems in Wireless Sensor Networks: A Review. International Journal of Distributed Sensor Networks (2013).
[AWBD2010]	Anwander, M.; Wagenknecht, G.; Braun, T.; Dolfus, K. BEAM: A Burst-aware Energy-efficient Adaptive MAC protocol for Wireless Sensor Networks, Seventh International Conference on Network Sensing Systems (INSS), 2010, pp. 195 –202.
[BLKS2014]	Bithas, P. S., Lioumpas, A. S., Karagiannidis, G. K., & Sharif, B. S. (2015). Hybrid Cellular/WLAN with Wireless Offloading: Enabling Next Generation Wireless Networks. *Submitted to IEEE Transactions Wireless Commun (available at http://arxiv.org/pdf/1311.2970v5.pdf)*
[BRS2011]	Game theory-based defense mechanisms against DdoS attacks on TCP/TCP-friendly flows. In: Computational Intelligence in Cyber Security (CICS), 2011 IEEE Symposium on 2011, pp. 129-136. IEEE
[BYAH2006]	Buettner, M.; Yee, G.; Anderson, E.; Han, R. X-MAC: a short preamble MAC protocol for duty-cycled wireless sensor networks. Proceedings of the 4th ACM International Conference on Embedded Networked Sensor Systems (SenSys); 2006; pp. 307–320.
[C2009]	Survey of network traffic models. Waschington University in St. Louis CSE 567 (2009).
[CB2010]	The Official Contiki OS Blog. Contiki 2.5 Release Candidate Available. Technical report, 2010.
[CH2010]	Clausen, T.; Herberg, U. Intelligent Sensors, Sixth International Conference on Sensor Networks and Information Processing (ISSNIP), 2010, pp. 7–12.
[CMPRS15]	V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti and G. Setti, ""On Known-Plaintext Attacks to a Compressed Sensing-Based Encryption: A Quantitative Analysis,"" in *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 10, pp. 2182-2195, Oct. 2015.
[CONET2011]	Cooperating Objects Network of Excellence. CONET newsletter The Contiki Operating System. Technical report, 2011.
[CR09]	"Performance Analysis of IEEE 802.15.4 Non Beacon-Enabled Mode", Chiara Buratti and Roberto Verdone, IEEE Trans. On Vehicular Technology, Vol. 58, No. 7, September 2009
[CR09]	"Performance Analysis of IEEE 802.15.4 Non Beacon-Enabled Mode", Chiara Buratti and Roberto Verdone, IEEE Trans. On Vehicular Technology, Vol. 58, No. 7, September 2009
[D2011]	Dunkels, A. The ContikiMAC Radio Duty Cycling Protocol. Technical Report T2011:13, Swedish Institute of Computer Science, 2011.
[DEFT11]	A. Dunkels, J. Eriksson, N. Finne, and N. Tsiftes, ""Powertrace: Network-level power profiling for low power wireless networks"" Tech. Rep.
[DOD2011]	Duquennoy, S.; Osterlind, F.; Dunkels, A. Lossy Links, Low Power, High Throughput. SenSys '11: Proceedings of the 9th international conference on Embedded networked sensor systems 2011.

[F++16]	A. Franco, E. Fitzgerald, B. Landfeldt, N. Pappas, V. Angelakis, LUPMAC: A cross-layer MAC technique to improve the age of information over dense WLANs 23rd International Conference on Telecommunications (ICT), 2016.
[FGK13]	"IEEE 802.11af: A Standard for TV White Space Spectrum Sharing", Adriana Flores, Ryan Guerra and Edward Knightly, IEEE Communications Magazine, October 2013
[FGK13]	"IEEE 802.11af: A Standard for TV White Space Spectrum Sharing", Adriana Flores, Ryan Guerra and Edward Knightly, IEEE Communications Magazine, October 2013
[GSK05]	"Modeling Media Access in Embedded Two-Flow Topologies of Multi-Hop Wireless Networks", Michele Garetto, Jingpu Shi and Edward W. Knightly, Mobicom, August 2005.
[GSK05]	"Modeling Media Access in Embedded Two-Flow Topologies of Multi-Hop Wireless Networks", Michele Garetto, Jingpu Shi and Edward W. Knightly, Mobicom, August 2005.
[H2005]	Generic modelingodelling of multimedia traffic sources. (2005).
[H2013]	An Agent-based Multi-model Tool for Simulating Multiple Concurrent Applications in WSNs. In: Journal of Advances in Computer Networks (JACN), 5th International Conference on Communication Software and Networks, Malaysia (June 2013) 2013
[HC2013]	Multi-Agent Support for Multiple Concurrent Applications and Dynamic Data-Gathering in Wireless Sensor Networks. In: Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2013 Seventh International Conference on 2013, pp. 320-325. IEEE.
[HC2014]	Sensomax: An agent-based middleware for decentralized dynamic data gathering in wireless sensor networks. In: Collaboration Technologies and Systems (CTS), 2013 International Conference on 2013, pp. 107-114. IEEE.
[HEKA15]	"A survey of MAC issues for TV White Space Access", You Han, Eylem Ekici, Haris Kremo, Onur Altintas, Ad Hoc Networks, 2015
[HEKA15]	"A survey of MAC issues for TV White Space Access", You Han, Eylem Ekici, Haris Kremo, Onur Altintas, Ad Hoc Networks, 2015
[HJB2004]	Hull, B.; Jamieson, K.; Balakrishnan, H. "Mitigating congestion in wireless sensor networks", in Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems, 2004.
[HRBD03]	""Performance anomaly of 802.11b,"" M. Heusse, F. Rousseau, G. Berger-Sabbatel and A. Duda, *INFOCOM 2003.* San Francisco
[HRBD03]	"Performance anomaly of 802.11b," M. Heusse, F. Rousseau, G. Berger-Sabbatel and A. Duda, *INFOCOM 2003.* San Francisco
[I2012]	Cryptographic Key Exchange in IPv6-based, Low Power, Lossy Networks. University of Bristol (2012).
[IHARMD2014]	Ishaq, I.; Hoebeke, J.; Van den Abeele, F.; Rossey, J.; Moerman, I.; Demeester, P. Flexible Unicast-Based Group Communication for CoAP-Enabled Devices. Sensors 2014, 14, 9833–9877.
[ITU-R]	ITU-R Working Party 5D Update on IMT VISION (2020+)," Technical Presentation, Aug. 2012.

[KGRK11]	S. Kaul, M. Gruteser, V. Rai, and J. Kenney. Minimizing age of information in vehicular networks. In 2011 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), pages 350–358, 2011.
[KYG12]	S. Kaul, R. Yates, and M. Gruteser. Real-time status: How often should one update? In 2012 Proceedings IEEE INFOCOM, pages 2731–2735, 2012.
[LCMS08]	"A Renewal Theory Based Analytical Model for the Contention Access Period of IEEE 802.15.4 MAC," Xinhua Ling, Yu Cheng, Jon W. Mark and Xuemin Shen, IEEE Transactions on Wireless Communications, Vol. 7, No. 6, June 2008
[LCMS08]	"A Renewal Theory Based Analytical Model for the Contention Access Period of IEEE 802.15.4 MAC," Xinhua Ling, Yu Cheng, Jon W. Mark and Xuemin Shen, IEEE Transactions on Wireless Communications, Vol. 7, No. 6, June 2008
[LK13]	"On the capacity of wireless CSMA/CA Multihop Networks", Rafael Laufer and Leonard Kleinrock, IEEE Infocom, 2013
[LK13]	"On the capacity of wireless CSMA/CA Multihop Networks", Rafael Laufer and Leonard Kleinrock, IEEE Infocom, 2013
[LKLW10]	"Back-of-the-Envelope Computation of Throughtput Distributions in CSMA Wireless Networks", Soung Chang Liew, Cai Hong Kai, Hang Ching Leung, Piu Wong, IEEE Trans. On mobile computing, September 2010.
[LKLW10]	"Back-of-the-Envelope Computation of Throughtput Distributions in CSMA Wireless Networks", Soung Chang Liew, Cai Hong Kai, Hang Ching Leung, Piu Wong, IEEE Trans. on mobile computing, September 2010.
[M2M]	"Machine-to-Machine (M2M) – The Rise of the Machines," Juniper Networks Whitepaper, 2011.
[MGOP2011]	Michopoulos, V.; Guan, L.; Oikonomou, G.; Phillips, I. A Comparative Study of Congestion Control Algorithms in Ipv6 Wireless Sensor Networks. Proc. 2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS), 2011.
[MGOP2012]	Michopoulos, V.; Guan, L.; Oikonomou, G.; Phillips, I. DCCC6: Duty Cycle-Aware Congestion Control for 6LoWPAN Networks. Proc. 8th IEEE International Workshop on Sensor Networks and Systems for Pervasive Computing (PerSeNS), 2012.
[MGP2010]	Michopoulos, V.; Guan, L.; Phillips, I. A New Congestion Control Mechanism for WSNs. Computer and Information Technology, International Conference on 2010, 0, 709–714.
[MKHC2007]	Transmission of Ipv6 packets over IEEE 802.15. 4 networks. Internet proposed standard RFC 4944 (2007).
[MLTK2008]	Musaloiu-E., R.; Liang, C.M.; Terzis, A. Koala: Ultra-Low Power Data Retrieval in Wireless Sensor Networks. Proceedings of the 7th IEEE International Conference on Information Processing in Sensor Networks (IPSN); 2008; pp. 421–432.
[MMMT06]	"Performance Analysis of IEEE 802.15.4 and ZigBee for Large-Scale Wireless Sensor Network Applications", Mikko Kohvakka, Mauri Kuorilehto, Marko Hannikainen, Timo Hamalainen, PE-WASUN, 2006
[MMMT06]	"Performance Analysis of IEEE 802.15.4 and ZigBee for Large-Scale Wireless Sensor Network Applications", Mikko Kohvakka, Mauri Kuorilehto, Marko Hannikainen, Timo Hamalainen, PE-WASUN, 2006

[MOGP2014]	Michopoulos, V.; Oikonomou, G.; Guan, L.; Phillips, I. CADC: A New Congestion Aware Duty Cycle Mechanism for 6LoWPAN Networks. 19th IEEE International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD), 2014.
[MR2004]	N. Modadugu and E. Rescorla. The design and implementation of datagram TLS. In NDSS, 2004.
[O06]	F. Österlind, "A Sensor Network Simulator for the Contiki OS", SICS technical report. Swedish Institute of Computer Science, 2006.
[O2006]	A Sensor Network Simulator for the Contiki OS. Swedish Institute of Computer Science Technical Report T2006:5, (2006)
[OOM2009]	Protecting Against Network Infections: A Game Theoretic Perspective. Paper presented at the IEEE INFOCOM (2009).
[OP2011]	G. Oikonomou, I. Phillips, "Experiences from Porting the Contiki Operating System to a Popular Hardware Platform" in Proc. International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS), 2011.
[OP2011]	G. Oikonomou, I. Phillips, "Experiences from Porting the Contiki Operating System to a Popular Hardware Platform" in Proc. International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS), 2011.
[OP2013]	Stateless Multicast Forwarding with RPL in 6LoWPAN Sensor Networks. In: Proc.2012 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops). Lugano, Switzerland (2012). 2012; 272-277
[OPT2013]	Ipv6 Multicast Forwarding in RPL-Based Wireless Sensor Networks. Wireless Personal Communications. 2013; 73(3): 1089-1116
[P++08]	"Performance Analysis of Slotted IEEE 802.15.4 Medium Accesss Layer," IEEE Transactions on Wireless Communications.
[PBGN]	G. Z. Papadopoulos, J. Beaudaux, A. Gallais, T. Noel and G. Schreiner, "Adding value to WSN simulation using the IoT-LAB experimental platform", in Proceedings of the 9th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob). 2013, pp. 485–490.
[PBGN2014]	G. Z. Papadopoulos, J. Beaudaux, A. Gallais and T. Noel, "T-AAD: Lightweight Traffic Auto-Adaptations for Low-power MAC Protocols" in Proceedings of the 13th IEEE IFIP Annual Mediterranean Ad Hoc Networking Workshop (MED-HOC-NET), 2014, pp. 79-86.
[PBPS2014]	Park, H.; Basaran, C.; Park, T.; Son, S.H. Energy-Efficient Privacy Protection for Smart Home Environments Using Behavioral Semantics. Sensors 2014, 14, 16235–16257.
[PGST2015]	G. Z. Papadopoulos, A. Gallais, G. Schreiner, and T. Noel, "Demo: Abstract: Live Adaptations of Low-power MAC Protocols", in the Proceedings of the 21st ACM Annual International Conference on Mobile Computing and Networking (MobiCom), 2015. Pp. 207-209.
[PH2012]	S. W. Peters, and R. W. Heath Jr. ""User partitioning for less overhead in MIMO interference channels."" *Wireless Communications, IEEE Transactions on* 11.2 (2012): 592-603.
[PKGCN2016]	G. Z. Papadopoulos, K. Kritsis, A. Gallais, P. Chatzimisios and Thomas Noel, "Performance Evaluation Methods in Ad Hoc and Wireless Sensor Networks: A Literature Study", in IEEE Communications Magazine, Vol. 54, 2016, pp. 122-128.

[PPGNA2015]	G. Z. Papadopoulos, N. Pappas, A. Gallais, T. Noel and V. Angelakis, "Distributed adaptive scheme for reliable data collection in fault tolerant WSNs", in the Proceedings of the 2nd World Forum on Internet of Things (WF-IoT), 2015, pp. 116-121.
[RA12]	"Performance Analysis of Beacon-Less IEEE 802.15.4 Multi-Hop Networks", Rachit Srivastava, Anurag Kumar, Communication Systems and Networks (COMSNETS), 2012
[RA12]	"Performance Analysis of Beacon-Less IEEE 802.15.4 Multi-Hop Networks", Rachit Srivastava, Anurag Kumar, Communication Systems and Networks (COMSNETS), 2012
[RD2.1]	T. Mouroutis, A. Lioumpas (Eds.), "Use-cases definition and threat analysis", Dec 2014.
[RD2.5]	E. Tragos (Ed), et. al, "Final System Architecture", Aug. 2015.
[RD3.1]	D. Ruiz Lopez (Ed.) et al., "Enhancing the autonomous smart objects and the overall system security of IoT based Smart Cities", RERUM Deliverable D3.1, March 2015.
[RD4.1]	E. Tragos (Ed.) et. al. "Introducing CR elements into smart objects towards enhanced interconnectivity for Smart City applications", Feb. 2015.
[RD4.2]	G. Oikonomou (Ed.) et al., "Advanced techniques to increase the lifetime of smart objects and ensure low power network operation", RERUM Deliverable D4.2, September 2015.
[RERUM-D4.1]	FP7 RERUM, "Deliverable 4.1: Introducing CR elements into smart objects towards enhanced interconnectivity for Smart City applications", June, 2014.
[RESDSQ2010]	A Survey of Game Theory as Applied to Network Security. In: System Sciences (HICSS), 2010 43rd Hawaii International Conference on, 5-8 Jan. 2010 2010, pp. 1-10
[Sensinode2008]	Sensinode. http://www.sensinode.com/EN/products.html, 2008.
[SKTO2013]	A game theoretic defence framework against DoS/DdoS cyber attacks. Computers & Security(0) (2013). Doi:http://dx.doi.org/10.1016/j.cose.2013.03.014
[SOTG2013]	Game Theoretic Approach for Cost-Benefit Analysis of Malware Proliferation Prevention. In: Janczewski, L., Wolfe, H., Shenoi, S. (eds.) Security and Privacy Protection in Information Processing Systems, vol. 405. IFIP Advances in Information and Communication Technology, pp. 28-41. Springer Berlin Heidelberg, (2013).
[ST11]	"Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications", IEEE Std 802.11-2012, http://standards.ieee.org/getieee802/download/802.11-2012.pdf
[ST11]	"Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications", IEEE Std 802.11-2012, http://standards.ieee.org/getieee802/download/802.11-2012.pdf
[ST154]	"Low-Rate Wireless Personal Area Networks (LR-WPANs)", IEEE Std 802.15.4-2011, http://standards.ieee.org/getieee802/download/802.15.4-2011.pdf
[ST154]	"Low-Rate Wireless Personal Area Networks (LR-WPANs)", IEEE Std 802.15.4-2011, http://standards.ieee.org/getieee802/download/802.15.4-2011.pdf

[STTG15]	"A two-stage spectrum assignment scheme for power and QoS constrained Cognitive CSMA/CA networks", George Stamatakis, Elias Z. Tragos and Apostolos Traganitis, IEEE Globecom, December 2015,.
[STTG15]	"A two-stage spectrum assignment scheme for power and QoS constrained Cognitive CSMA/CA networks", George Stamatakis, Elias Z. Tragos and Apostolos Traganitis, IEEE Globecom, December 2015,.
[STTQ15]	"A two-stage power and QoS aware dynamic spectrum assignment scheme for cognitive wireless sensor networks", George Stamatakis, Elias Z. Tragos and Apostolos Traganitis, QoMEX 2015, May.
[STTQ15]	"A two-stage power and QoS aware dynamic spectrum assignment scheme for cognitive wireless sensor networks", George Stamatakis, Elias Z. Tragos and Apostolos Traganitis, QoMEX 2015, May.
[TA2011]	Game Theory for Security: A Real-World Challenge Problem for Multiagent Systems and Beyond. Association for the Advancement of Artificial Intelligence (2011).
[TED2010]	Tsiftes, N.; Eriksson, J.; Dunkels, A. Low-power wireless Ipv6 routing with ContikiRPL" in Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks, 2010, pp. 406–407.
[VD2010]	Vasseur, J.P.; Dunkels, A. Interconnecting Smart Objects with IP; Morgan Kaufmann, 2010.
[W98]	"Integer Programming", Laurence A. Wolsey, Wiley 1998
[W98]	"Integer Programming", Laurence A. Wolsey, Wiley 1998
[WA08]	"Optimal Spectrum Sensing Framework for Cognitive Radio Networks", Won-Yeol Lee and Ian F. Akyldiz, IEEE Transactions on Wireless Communications, Vol 7., No 10, October 2008
[WA08]	"Optimal Spectrum Sensing Framework for Cognitive Radio Networks", Won-Yeol Lee and Ian F. Akyldiz, IEEE Transactions on Wireless Communications, Vol 7., No 10, October 2008
[WA11]	"A Spectrum Decision Framework for Cognitive Radio Networks", Won-Yeol Lee, Ian F. Akyldiz, IEEE Transactions on Mobile Computing, Vol 10, No 2, February 2011
[WA11]	"A Spectrum Decision Framework for Cognitive Radio Networks", Won-Yeol Lee, Ian F. Akyldiz, IEEE Transactions on Mobile Computing, Vol 10, No 2, February 2011
[WLSC2006]	Wireless Sensor Network Security: A Survey. In: Security in Distributed, Grid, and Pervasive Computing. P. 50. (2006).
[WSRED2010]	On modelinmgodelling and simulation of game theory-based defense mechanisms against DoS and DdoS attacks. Paper presented at the Proceedings of the 2010 Spring Simulation Multiconference, Orlando, Florida,
[W++2012]	RPL: Ipv6 Routing Protocol for Low-Power and Lossy Networks. (2012).
[XJL08]	"Optimal Bandwidth Selection in Multi-Channel Cognitive Radio Networks: How Much is Too Much?", Dan Xu, Eric Jung and Xin Liu, Dyspan 2008
[XJL08]	"Optimal Bandwidth Selection in Multi-Channel Cognitive Radio Networks: How Much is Too Much?", Dan Xu, Eric Jung and Xin Liu, Dyspan 2008
[XSW2014]	Xu, G.; Shen, W.; Wang, X. Applications of Wireless Sensor Networks in Marine Environment Monitoring: A Survey. Sensors 2014, 14, 16932–16954.

[YHE2002]	Ye, W.; Heidemann, J.; Estrin, D. An energy-efficient MAC protocol for wireless sensor networks. INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, 2002, Vol. 3, pp. 1567 – 1576.
[Z++16]	Z. Chen, N. Pappas, M. Kountouris, and V. Angelakis. On the performance of Delay Aware Shared Access with Priorities, arXiv:1603.08885v1 [cs.NI], 2016.
[ZSDJ06]	"Performance Analysis and a Proposed Improvement for the IEEE 802.15.4 Contention Access Period", Zhifeng Tao, Shivendra Panwar, Daqing Gu, Jinyun Zhang, WCNC 2006
[ZSDJ06]	"Performance Analysis and a Proposed Improvement for the IEEE 802.15.4 Contention Access Period", Zhifeng Tao, Shivendra Panwar, Daqing Gu, Jinyun Zhang, WCNC 2006

