CREW

Cognitive Radio Experimentation World

SEVENTH FRAMEWORK PROGRAMME

Project Deliverable D4.1

Definition of Test Scenarios and Benchmarks

Contractual date of delivery:	31-09-11
Actual date of delivery:	31-09-11
Beneficiaries:	IBBT – IMEC – TCD – TUB – TUD – TCF – EADS
Lead beneficiary:	TCD
Authors:	Danny Finn (TCD), Justin Tallon (TCD), Luiz DaSilva (TCD), Jono Vanhie - Van Gerwen (IBBT), Stefan Bouckaert (IBBT), Ingrid Moerman (IBBT), Christophe Heller (EADS), David Depierre (TCF), Sofie Pollin (imec), Peter VanWesemael (imec), Jan Hauer (TUB), Daniel Willkomm (TUB), Nicola Michailow (TUD)
Reviewers:	Stefan Bouckaert (IBBT), Daniel Willkomm (TUB)
Workpackage:	WP4 – Benchmarking the Federation
Estimated person months:	38
Nature:	R
Dissemination level:	PU
Version	4.0

Abstract: This document describes the various test scenarios that have been constructed for each test bed and gives the benchmarks that have been outlined for each test case within each usage scenario. One of the main goals of CREW is to establish a set of benchmarks for experiments that are carried out within the facilities of the federated test bed. These benchmarks will establish a way in which experiments can be reproduced accurately so that results from different but similar experiments can be compared accurately without having to consider how the experimental set up of a particular instance of the experiment has influenced the results. In doing this we ensure that any experiments carried out within the federation can be useful to others looking to reproduce and replicate the experiment and its results.

Keywords: benchmarking, usage scenario, experimentation

Executive Summary

This document describes the benchmarks that have been developed for internal usage scenarios i.e.:

- Radio environment sensing for spectrum sharing in ISM, TV and LTE bands. It aims at gathering information about the actual occupancy of a defined frequency range.
- Horizontal sharing between heterogeneous networks in the ISM bands. It refers to a home radio frequency environment with interferences resulting from multiple Wi-Fi stations.
- Cooperation in heterogeneous networks in licensed bands. It refers to the case of wireless microphones usage in DVB-T bands.
- Cognitive sensor networks. Coexistence of a cognitive body area network with other wireless sensor systems operating in the same frequency bands is considered.
- Cognitive networks in next generation cellular systems. Efficient detection of radio resources that are not in use by the primary system relies on traffic models representing typical primary user behaviour.

These benchmarks define sets of configuration parameters and performance metrics that will ensure the reproducibility of the use cases within each usage scenario and ensure the seamless comparison of results from different instances of an experiment within a test bed, as well as ones carried out on different hardware and software, at different times and in different locations.

In order to do so, for each usage scenario, are described the test conditions, the network conditions (frequency bands, radio technologies, description of the nodes in the experiment and of their physical topology), the type of applications, the interference sources. The performance metrics, allowing to evaluate the results of each wireless experiment are also defined.

List of Acronyms and Abbreviations

3GPP	3^{rd} Generation Partnership Project
AP	Access Point
API	Application Programming Interface
BAN	Body Area Network
BLER	Block Error Rate
CBAN	Cognitive Body Area Network
CR	Cognitive radio
CREW	Cognitive Radio Experimentation World
CQI	Channel Quality Indicator
CSMA	Carrier Sense Multiple Access
CTP	Collection Tree Protocol
DL	Downlink
DSA	Dynamic Spectrum Access
DVB-T	Digital Video Broadcast – Terrestrial
Dx.x	Deliverable x.x
EADS	European Aeronautic Defence and Space Company
eNB	e Node B
EUTRAN	Evolved UMTS Terrestrial Radio Access Network
FIRE	Future Internet Research and Experimentation
FM	Frequency Modulation
FP7	Framework Programme 7
HD	High Definition
IBBT	Interdisciplinary Institute for Broadband Technology
IC	Integrated Circuit
IEEE	Institute of Electrical and Electronics Engineers
imec	Interuniversity Microelectronics Center
I/Q	In-phase/Quadrature
ISM	Industrial Scientific Medical
JSI	Jozef Stefan Institute
LTE	Long Term Evolution
MAC	Medium Access Control
OFDM	Orthogonal Frequency-Division Multiplexing
OFDMA	
PFA	Probability of False Alarm

PMD	Probability of Missed Detection
PRB	Physical Resource Block
PSD	Power Spectral Density
PU	Primary User
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RF	Radio Frequency
ROC	Reciever Operating Characteristic
RSRP	Reference Signal Receive Power
RSSI	Received Signal Strength Indicator
SDR	Software Defined Radio
SINR	Signal to Interference and Noise Ratio
SU	Secondary User
SUT	System Under Test
TCD	Trinity College Dublin
TCF	Thales Communications France
TCP	Transmission Control Protocol
TUB	Technische Universität Berlin
TUD	Technische Universität Dresden
TWIST	TKN Wireless Indoor Sensor Network Testbed
UDP	User Datagram Protocol
UE	User Equipment
UHF	Ultra High Frequency
UL	Uplink
US	Usage Scenario
USRP	Universal Software Radio Peripheral
VHF	Very High Frequency
WPx	Work Package x
WSN	Wireless Sensor Network

Table of contents

1 Introduction

The CREW project establishes a federation of cognitive radio (CR) network testbeds. It enables repeatable experimentation on cognitive radio and dynamic spectrum access for a broad range of test cases using the wireless testbeds in the federation. This deliverable outlines test conditions for these experiments and discusses the establishment of benchmarking guidelines for experimental and prototype-based research on cognitive radio.

1.1 Deliverable objectives, background

So that the new CR concepts in WP6 can be evaluated, the test conditions need to be precisely defined. For each test case the following must be specified:

- The network conditions, where the deployment characteristics of the network, such as the radio access technologies, physical topologies, etc. are defined;
- A set of applications, where we define applications in a broad sense (e.g. streaming media, file transfer, web applications, monitoring application, etc.);
- A set of interference sources (primary user traffic sources, other secondary users competing for the band, random interferers displaying no pattern, interferers with some underlying stationary pattern, etc.).

During benchmarking experiments, sufficiently accurate models and settings are needed to emulate realistic traffic scenarios and interferers within the different network deployment scenarios. Applications and interference sources will generate specific traffic patterns, which can be modelled as a random process, parameterized by the packet arrival rate (uniform versus bursty), packet size, average transmission rate, etc.

1.2 Motivation: discussion of benchmarking

For each test case one or multiple benchmarks can be defined, depending on the criteria. A criterion is the target of the benchmark (e.g. spectrum efficiency, energy efficiency, throughput, scalability, etc.). A single benchmark may also combine multiple criteria (e.g. spectrum efficiency and throughput).

A benchmark is fully defined by:

- A configuration scenario, specifying the variable parameters from the models and network conditions. Possible variable parameters are e.g. parameters of the traffic patterns (such as packet size, inter packet gap, percentage of nodes which are sensing, etc.), parameters of the interference sources (number of sources, power, waveforms, specific or random transmission patterns, etc.), or network parameters (such as network size, node density, node failures, etc.);
- The performance metrics. Possible performance metrics may include spectrum utilization, end-to-end delay, delivery rate, application throughput, energy consumption, memory footprint, etc. The performance metrics will determine what data need be logged during a test run. In the case of vertical spectrum sharing, performance metrics for both the primary and secondary network should be considered, as in addition to increased capacity in secondary networks, it is also important to measure how the primary network is affected.

Additionally, non-technical metrics such as “financial cost” can be included in a benchmark.

This task will create a generic suite of configuration parameters and performance metrics for benchmarks applicable to each CREW wireless environment. These will ensure the reproducibility of the use cases within each usage scenario and ensure the seamless comparison of results from different instances of an experiment within a test bed, as well as ones carried out on different hardware and software, at different times and in different locations.

.

2 Radio environment sensing for spectrum sharing

2.1 Test conditions, applications, interference sources

2.1.1 Context awareness in the ISM bands

The goal of this test case is to investigate the gathering of spectrum occupancy information. Since there exist a wide variety of wireless communications techniques spread over a wide frequency range, different spectrum sensing solutions will lead to multiple interpretations of the same spectrum. For a number of popular frequency bands different sensing solutions will be evaluated. A solution can be different is terms of the hardware component(s) used or the algorithms used for detection, but also in the way the measurement results are processed and combined (e.g. local sensing versus distributed sensing versus a database based solution).

In this test case the focus is on the 2.4 GHz ISM band and more specifically on the IEEE 802.15.4 and 802.11 standards. In addition to specific 802.15.4 and 802.11 signals, we can consider random sources of interference with a specific time and frequency pattern. Different experiments are possible, and they vary in:

- **Signal generation**: The test signal used is a DVB-T signal. Although this signal is not present is this frequency band it has characteristics that resemble 802.11g signals as it is also OFDM modulated. We deliberately use a DVB-T signal with a large cyclic prefix to enable detection with more advanced algorithms that plain power detection.
- **Wireless channel**: When the goal is ensuring repeatability and reliable comparison of different hardware solutions, the channel should be replaced by a coax cable. When the goal of the experiment is studying the impact of the environment such as a meeting room, office, outdoor of home environment, wireless experiments can be done. Typically, the frequency used for the experiments is the unlicensed 2.4 GHz ISM band.
- **Sensing hardware and algorithm**: different sensing devices can be used, e.g. 802.11 devices, 802.15.4 devices, and software defined radio devices. The usage of the DVB-T signal for this test enables the usage of simple power detection algorithms but also commonly used feature detection algorithms e.g. cyclostationary detection.
- **Sensing application**: scenarios are possible to compare solutions for local sensing, for distributed sensing and for comparing sensing with an off-line or previously generated database.

A signal generator is used to generate test signals. The connection between the signal source and the sensing device is, depending on the type of experiment, either a cable or a real wireless channel. Multiple sensing devices are used: devices targeted for 802.11 or 802.15.4 operation and SDR solutions which target multiple wireless standards. When repeatability and the absence of interference are key factors to obtain a reliable result the connection between the generator and the sensing devices is made by means of a cable. To get a general understanding on the performance of the various sensing solutions in real life situations a wireless channel is used. Results for these experiments will be discussed in D6.1.

2.1.2 Context awareness in TV White Spaces license bands

TV signals are located in the VHF and the lower part of the UHF band. Due to the digital switchover significant amounts of bandwidth are no longer used. Propagation losses in these frequency bands are much lower than at the higher parts of the spectrum, where for instance the 2.4 GHz ISM band is located. Therefore there is a lot of interest in using these frequencies for other wireless communication when these bands are not occupied by the TV channels. The primary users, i.e. the TV channels, may not suffer any degradation from this opportunistic spectrum use and hence it is critical to accurately detect the presence or absence of the TV channels. The solution needs to minimize the probability of not detecting a present TV broadcast, also known as the probability of missed detection, to avoid causing interference on the primary user. For the secondary user it is of capital importance to avoid missing the detection of the available channels, also known as minimizing the probability of false alarms, in order to optimally use the available spectrum.

Different experiments are possible in TV White Spaces. For Europe, the focus is on sensing DVB-T signals:

- **Signal generation**: The DVB-T signals can be generated using off-the-shelf devices, using a signal generator or using an SDR platform such as the USRP or imec platform.
- **Wireless channel**: When the goal is ensuring repeatability and reliable comparison of different hardware solutions, the channel should be replaced by a coax cable. When the goal of the experiment is studying the impact of the environment such as a meeting room, office, outdoor of home environment, wireless experiments can be done. Experiments can be done in an unlicensed band, to focus only on the features of the digital signal, or in the appropriate UHF/VHF bands.
- **Sensing hardware**: the band used for the transmission/sensing of the signals determines the hardware that can be used. For UHF/VHF bands, SDR solutions such as the USRP or imec sensing solution can be used. For other bands, more flexible choice of hardware is possible.
- **Sensing application**: scenarios are possible to compare solutions for local sensing, for distributed sensing and for comparing sensing with an off-line or previously generated database.

The experiments focus on the reliable detection of primary users. A DVB-T signal is generated with a signal generator and sent to the imec sensing engine and a signal analyzer through a cable to obtain a controllable and reproducible environment. The first goal of this experiment is to evaluate the detection performance of a low-cost and low-power SDR solution (i.e. the imec sensing engine) to the performance achievable with a signal analyzer. The second goal of this experiment is to investigate the detection improvement by applying a feature detection based algorithm compared to a straightforward power detection algorithm. More details and results on these experiments can be found in D6.1

2.1.3 Reliable sensing of cellular systems

For cellular systems we focus on the 3GPP LTE signals. As LTE supports OFDMA operation it makes sense to develop sensing solutions to estimate the actual usage of the Physical Resource Blocks in the LTE signal. In addition, it is possible to design algorithms that rely on statistical properties of the LTE ODFMA signals, and possibly rely on the use of multiple antennas to improve the sensing performance. Different experiments can be designed for LTE sensing within the CREW consortium:

- **Signal generation**: The LTE signals can be generated by transmitting the LTE baseband signals using an emulated eNB and UE nodes, a signal generator or using an SDR platform such as the USRP or imec platform.
- **Wireless channel**: When the goal is ensuring repeatability and reliable comparison of different hardware solutions, the channel should be replaced by a coax cable. When the goal of the experiment is studying the impact of the environment such as a meeting room, office, outdoor of home environment, wireless experiments can be done in the EUTRAN band VII in Dresden. Experiments can also be done in an unlicensed band, to focus only on the features of the digital signal.
- **Sensing hardware**: the band used for the transmission/sensing of the signals determines the hardware that can be used. For LTE bands, SDR solutions such as the USRP or imec sensing solution can be used. For other bands, more flexible choice of hardware is possible. In the current CREW federation, only the Thales solution can be used to do sensing with multiple antennas.
- **Sensing application**: scenarios are possible to compare solutions for local sensing, for distributed sensing and for comparing sensing with an off-line or previously generated database. Sensing with multiple antennas is also possible, using the Thales testbed.

Experiments conducted at the end of year one focus on testing the solutions that have been designed specifically for sensing of LTE, i.e., the imec and Thales sensing solutions. The imec solution focuses on low-complexity sensing of the LTE resource block allocation. To limit the power dissipation of the sensing operation, severe constraints have been imposed on the complexity of the digital processing hardware. An algorithm was developed to detect the PRB allocation whilst complying with the hardware limitations. The goal of this experiment is to evaluate this solution in a real-life environment. To conduct this experiment the sensing hardware was placed in the vicinity of an LTE test bed. The sensing engine needs to successfully synchronize to the LTE signal and determine the PRB usage. The algorithm was verified using the TUD testbed where an LTE signal was transmitted wirelessly in the 2.6 GHz band, and successfully sensed by the imec hardware.

The sensing solution developed by Thales focuses on the use of the synchronization sequences of LTE and the use of multiple antennas. The solution has been verified by means of simulations using LTE baseband signal files generated by the TUD LTE test-bed reference signal generator. These files were used as inputs to a spatial multi-signal propagation channel simulator to simulate realistic LTE network interference situations. The simulator and the algorithms are described in details in D3.1; the simulation results, in D6.1

2.2 Performance metrics

2.2.1 Context awareness in the ISM bands

Sensing in the ISM bands is about determining the amount and source of interference. For that, it is important to be able to characterize for instance the amount of interference in both time and frequency. To obtain this, the captures made by all devices need to be processed to obtain a periodogram which plots the received power level for a given frequency at a certain time. This periodogram needs to be compared to the reference, e.g., the sequence that was used to control the signal generator. This comparison will provide a measure of the detection performance of the different solutions and enable us to compare the solutions amongst each other.

In addition to the time-frequency-power information about the interference, it is useful to determine the source of information, e.g., the number of 802.15.4 frames or the number of 802.11 frames that can cause harmful interference.

2.2.2 Context awareness in the TV White Spaces licensed bands

For sensing in the licensed TV bands, it is important to test the accuracy of the detection of the primary users. This is typically done by measuring the probability of false alarm versus the probability of missed detection. These metrics together make the Receiver Operating Curves (ROC) that characterizes a fundamental property of any sensing hardware/algorithm combination. These curves show the probability of false alarms versus missed detection for different power levels of the signal. By selecting the same operating point, i.e., setting an upper limit on the allowed probability of missed detection and the allowed probability of false alarms, the performance of all solutions can be compared in a uniform way. The ROC is a good way to compare the performance of a sensing solution for the licensed TV bands.

2.2.3 Reliable sensing of cellular systems

When sensing for 3GPP LTE, one is typically interested in the resource use of a given cell, or in determining the number of cells that cause interference. For the first, a good metric is the difference between estimated PRB usage versus actual PRB usage.

The use of antenna array and reference-based antenna processing algorithm allows to highly increase the detection and the identification of LTE base stations in a network. It is able to solve more accurately every intra-system interference situation by detecting and identifying interfering LTE base stations with E_c/I_0 as low as -20 dB, which allows to detect all interfering LTE base stations having a significant impact on the overall network performance.

The detection allows determining the LTE BTS characteristics including:

- Physical layer cell identity,
- Cyclic prefix length,
- Duplex mode,
- BTS level and E_c/I_0
- Time channel impulse response
- Frequency channel impulse response

3 Horizontal sharing between heterogeneous networks in ISM bands

The goal of this test case is to show how the CREW federation can be used to implement and optimize cognitive networking protocols supporting the coexistence of wireless devices operating in the ISM bands. With an ever increasing number of wireless end-user devices operating in the ISM bands, development of cognitive networking coexistence protocols is very relevant today, whether at home, in the office, or e.g. at a conference.

The test case described below is specifically targeted at analyzing the performance of cognitive sensor network protocol stacks using IEEE 802.15.4 radio communication, in an environment that is interfered by multiple Wi-Fi stations connecting to a Wi-Fi access point.

Note that several variations are possible on the described test case, depending on the need of the experimenter.

The described "home environment" set-up is intentionally basic for illustrative purposes. However, the methodology, developed tools, and benchmarking framework are universal and may be used to support larger-scale scenarios. In what follows, the full description of the benchmark is given. Next, the different implementation choices are motivated.

3.1 Test conditions, applications, interference sources

- High level goal: determine performance of a WSN protocol stack
- Duration of experiment: 10 minutes
- Output parameters:
    - Reliability of the sensor stack in terms of packet loss
    - Impact of the sensor network on the primary Wi-Fi network in terms of difference of the total throughput realized by the Wi-Fi network
    - PSD values of the frequency band of interest
    - Quality indication of experimentation environment
    - Quality indication of the System Under Test

3.1.1 Configuration scenario

a. Network conditions

- Frequency band of interest: 2.4 GHz ISM band
- Radio technologies
    - Of interest to the experiment: Wi-Fi (IEEE 802.11g- based), IEEE802.15.4
    - Possible (uncontrolled) interfering signals in the experimentation environment include: Wi-Fi devices outside the experiment, microwave oven, Bluetooth devices
- Physical topology of nodes in the experiment: fixed, see Figure 1
    - 1 x Wi-Fi access point (node id 54)
    - 3 x Wi-Fi station (node id 46,52,53)
    - 10 x wireless sensor node (node ids 38,40,41,42,43,44,45,48,49,51)
    - 1 x imec spectrum sensing agent
- Description of nodes in the experiment:
    - Wi-Fi access point: default w-iLab.t embedded PC (Alix node, cf. CREW portal) configured in AP mode
    - Wi-Fi stations: default w-iLab.t embedded PC (Alix node, cf. CREW portal) configured in client mode
    - Wireless sensor node: default w-iLab.t TelosB node
    - imec spectrum sensing agent (cf. CREW portal)

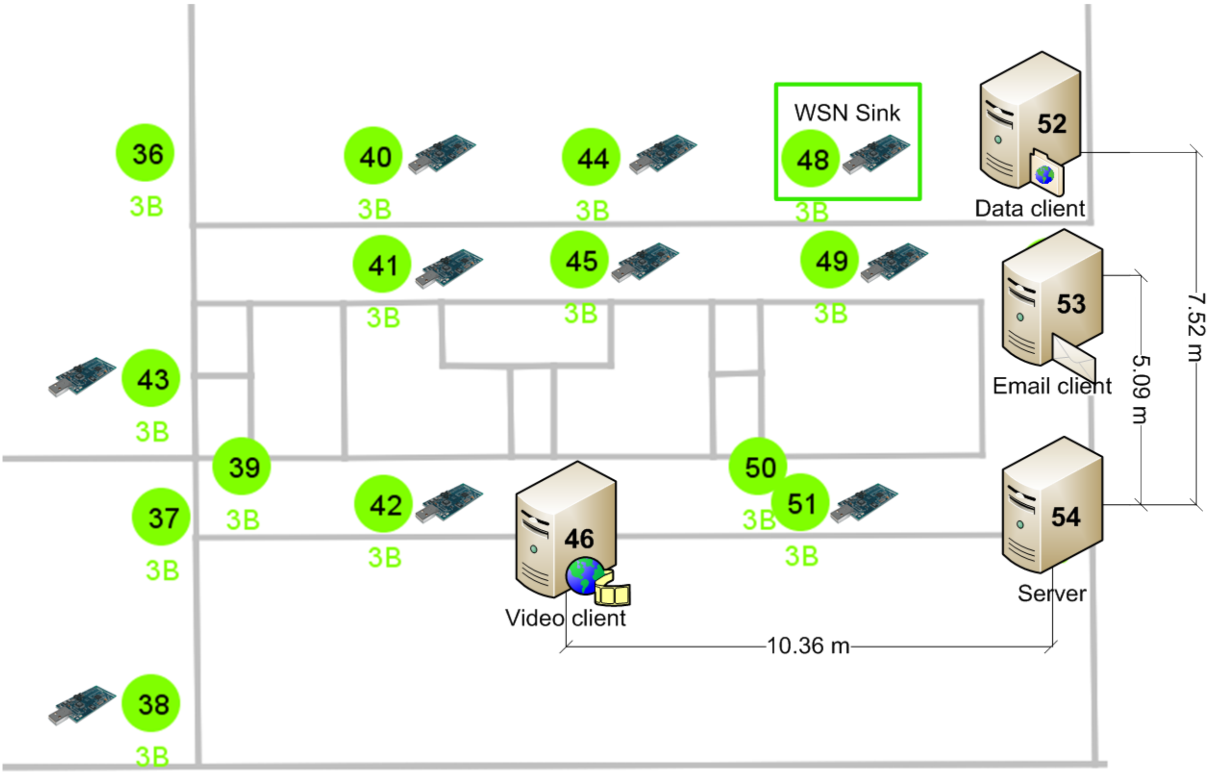


Figure 1: Topology of the IBBT test case

b. *Applications*

- Logical group 1: emulated home environment: Wi-Fi nodes
    - Node [a]: “checks emails”:
      TCP transfer of 1s at t= 0 + 1 * 60*x* s, *x*= 1..10
    - Node [b]: “browses the internet”
      TCP transfer of 90s at t = 180s and 180s at t = 420s
    - Node [c]: “watches HD video over the wireless network”
      UDP transfer of 300s at t = 300s and bitrate of 20Mbps
- Logical group 2: system under test: sensor network
    - The system under test is to be delivered by the experimenter.
    - To generate the performance metrics, each sensor node *must* send a packet of size 15 B to a gateway sensor node located at position 48, at times t= 0 + 1*15x s, x = 1..40 +- random delta to avoid collisions

c. *Interference sources*

The logical group 1 from the subsection above forms the interference source for the logical group 2. There are no other (known) interferers operational during the experiment

3.2 Performance metrics

The following performance metrics are considered during the benchmark:

- ***Aggregate throughput of the Wi-Fi network:***
  The access point is connected to a server instance of the *iperf* [1] tool; The emulated traffic generated by the Wi-Fi clients are implemented through *iperf* streams with different characteristics. On each node acting as a Wi-Fi client, *iperf* client connections are set up to the server. The total application layer throughput is measured at the server side using the Linux "*dstat*" tool.
- ***Packet loss of every wireless sensor link (source -> sink)***
  The number of packets that are sent is known from the application specification of the benchmark. The sink keeps track of which packets are received, and which are not.
- ***Reliability of the wireless sensor network in terms of aggregate packet loss***
  This is a simple calculation based on the packet loss measurement above
- ***Quality of the experimentation environment in terms of spectrum characteristics***
  The wireless spectrum is measured by the imec sensing agent just before, during and after the experiment. By comparing the sweeps with the average expected spectrum at the time and place of the experiment, outliers are detected. The location, duration and energy of these outliers will be used to determine a score from 0 to 10, with 10 being the average spectrum, and 0 fully interfered.

3.3 Detailed benchmark framework / channel selection experimentation

As indicated above, the tools and approach are universal and may be used for different experimentation set-ups as well. The following considerations explain the implementation choices for this benchmark and may assist experimenters in generating new, potentially similar benchmarking scenarios.

First, we address the choice for the specific "applications" that form the interference sources. While in this case the applications driving the Wi-Fi traffic were selected based on a short internal discussion on what activities could be ongoing in a home environment (based on personal experience), the realism of the traffic model can still be improved by basing the generated traffic on real traffic measurements, collected in real environments. Within the consortium, efforts are ongoing to record a "typical" 2.4/5GHz RF behavior at different locations under different circumstances, such as: use of Wi-Fi in a typical office environment, or, use of Wi-Fi at conference venues. Such real traces are currently being recorded, and will in the future be analyzed to come to a model of what the "typical" traffic patterns look like in these different environments. Based on these models, different random traffic patterns will be generated. This set of randomly generated instances of the model will then be saved. It will then be possible to test a single SUT against multiple Wi-Fi interference patterns that have similar characteristics, thus avoiding possible anomalies caused by accidental selection of specific traffic patterns.

Second, the choice was made to set up traffic streams with the iperf tool, emulating the "primary user" Wi-Fi traffic. This allows us to easily describe and trigger the scenarios and provides us with elegant opportunities to determine the impact of the secondary users (i.e. the sensor nodes) on the normal behavior of the primary user, simply by measuring the total throughput that was achieved during each run of the benchmark.

Other options could have been selected:

1. Recording actual traffic traces in a house environment and then replaying this trace packet by packet by what would be an “interference generator node”, or a set of such hypothetical devices. This approach would lead to complexities: a first option would be to make recordings in a wireless environment based on a single capturing node (possibly scanning multiple channels). However, replaying this exact information from a single interfering location cannot be compared to a real environment with multiple nodes. A second option would be to make recordings in a distributed way – however, the same complexity does exist: how to decide where these packets will be replayed from? Furthermore, replaying Wi-Fi traces in an exact way is complicated by carrier sensing algorithms that are implemented in the Wi-Fi hardware: without complex alterations of Wi-Fi hardware, there is no option to force packets to be put on the wireless medium “now”. Even if *exactly* replaying traces would be possible (at *microsecond* scale), then the packets that are sent in a forced way would not be representative loads: in real situations, as Wi-Fi devices will always perform carrier sensing.
2. Instead of a set of traffic streams starting and stopping at fixed moments in time, we could have opted to fix only the starting time of interfering Wi-Fi traffic and fix the size of the transfer instead of the duration. In this case, data transfers on the primary Wi-Fi network would go faster in case the nodes sense little interference, and slower in case of coexistence issues. This would also have been a feasible option, requiring a slightly different approach to characterize the performance of the primary Wi-Fi traffic. However, this modus operandi is slightly more difficult to implement, and may be harder to grasp for the experimenters not involved in the design process of the benchmark, since discrete on/off events are no longer used. We do not exclude benchmarks like these to be developed in the future, but more research would be needed to investigate the behavior of such benchmarking environment.

Third, regarding the reproducibility of the experimentation environment (i.e., the emulated home environment), to ensure that the traffic generated by the Wi-Fi devices is, to a certain degree, repeatable, the behavior of the Wi-Fi traffic scenario (or any ‘background scenario’ that would be part of a benchmark) is first to be tested multiple times in the absence of the actual system under test (which in this case is the sensor network). When comparing the output metrics that describe the behavior of the devices that are part of the benchmark setting (i.e., the logical group 1 of the benchmark above) between these different calibration experiments, there should only be minor differences. If there are only minor differences, this means that the reproducibility of the benchmark is good, which is a prerequisite before any performance analysis can happen using the benchmark. Furthermore, having a large set of these “calibration benchmarks” enables us to define a “normal behavior” of the benchmark. In this example, this means defining the “normal” aggregate throughput of the Wi-Fi network, and the “normal” RF spectrum occupation in the relevant ISM band. That such a degree of reproducibility can be achieved is illustrated through the outcome of the calibration experiments for the Wi-Fi throughput, which were performed in the topology used in this use case. These results are shown in Figure 2.

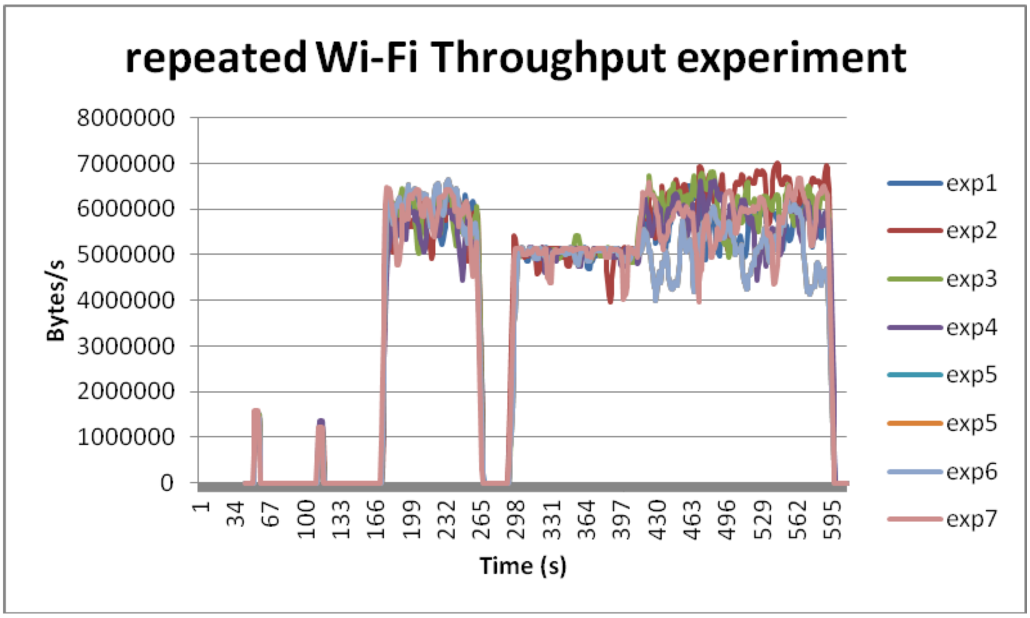


Figure 2 : Wi-Fi Throughput Vs Time

The results show that as long as a single Wi-Fi stream is established in the network the bandwidth results are very predictable at our experiment site, with few outliers. As expected, when combining a TCP and UDP stream, outliers increase due to the aggressive nature of the UDP stream. However, this is normal behavior in a wireless home environment, so larger outliers should be accepted in this timeframe of the experiment.

Fourth, it is important that the actual benchmark experiment (including the SUT) is repeated multiple times: although the "standard" behavior of the benchmark is known from the calibration experiments, no information is available on how the metrics will behave when the SUT is inserted in the benchmark. Multiple runs of the benchmark need to be compared in order to know whether the results are stable. Only in case the results are stable, the benchmark can be seen as valuable. If the outcome of different benchmarks is each time completely different, the only thing that may be concluded is that the SUT is too unstable to be tested in the considered benchmark environment.

Fifth, after running successful (thus stable) benchmarks, the results are stored. In the long run, this can lead to a complete library of performance results. The ultimate goal would be that researchers could easily compare the performance of their solution, simply by benchmarking their solution using "benchmark x", or a combination of benchmarks, on the CREW platform. Attracting experimenters to these reference benchmarks will evidently be easier if the benchmark infrastructure is user-friendly.

4 Cooperation in heterogeneous networks in licensed bands

The purpose of this test case is to explore how the CREW federation testbeds and benchmark framework can be used to investigate the possibility of networks of different topology, structure and function coexisting in the same band, i.e. in a primary user – secondary user relationship. The widespread prominence on digital TV and the demise of analogue TV have created what is referred to as "TV white spaces". These are large blocks of spectrum that were once occupied by analogue transmissions of TV but are now increasingly unoccupied. This has led to widespread debate and research into possible uses for this newfound available spectrum. There are many who argue that it is a perfect opportunity for dynamic spectrum access and cognitive radio principals to come into widespread use. However, before this can happen, there are a number of issues that must be explored and overcome.

This section outlines a benchmark for a generic set of experiments that investigate certain aspects of heterogeneous network coexistence. The experiments focus on the case of operation as a secondary user in a TV band.

4.1 Test conditions, applications, interference sources

- High level goal: successful coexistence between licensed and unlicensed users in the TV bands.

- Output parameters
    - A database of spectrum occupancy over a large geographical area.
    - Probability of false alarm vs. probability of missed detection statistics for wireless microphones.
    - QoS statistics for both primary and secondary users in a coexistence experiment.
    - Probability of false alarm vs. probability of missed detection statistics for primary user in a coexistence experiment.

4.1.1 Configuration scenario

a. Network conditions

- Frequency of interest: 470-860 MHz TV bands
- Radio technologies
    - Wireless microphones
    - DVBT transmitted signal
    - Cognitive/ software defined radio – algorithm (system under test)
    - Sensing platforms
- Physical topology of nodes in the various experiments
    - 1x wireless microphone
    - 1x signal generator
    - 1x imec spectrum sensing agent
    - 2x USRPs
    - Iris software defined radio

b. Applications

1. Wireless microphone detection experiment

a. Wireless microphone- FM modulated, 200KHz, varying dBm, operating in TV band.
b. Spectrum sensing agent (imec) sweeping TV band, windows of 20MHz

2 Geographical spectrum sensing experiment

a. Spectrum sensing agent (imec) sweeping TV band at location 1-N for 60 seconds and recording I&Q values.

3 Detection and transmission experiment

a. Signal generator – DVBT signal in TV bands, 8MHz, high dBm

b. Spectrum sensing agent (imec) sweeping TV band, windows of 20MHz

c. Iris software defined radio link – 2MHz, freq depending on information from spectrum agent.

c. Interference sources

- DVBT signal (3)

4.2 Performance metrics

The following performance metrics are considered within the benchmarks:

- ***Probabilities of false alarm/missed detection***
  Can be calculated in applications (1) and (3), as the transmitted signals are known and a good estimate can be gotten in (2) as DVB-T transmissions are relatively static.
- ***Level of structure in primary user signal readings***
  For the purposes of learning the presence of patterns and structure in recorded primary user signals are highly beneficial, if not completely necessary. Metrics such as the entropy of a set of readings can give a reasonable estimate of the level of structure it contains.
- ***Aggregate network throughput***
  This is the sum of the throughputs of both primary and secondary users within the network. Ideally through the addition of secondary users to the network this should increase.
- ***Degradation in performance of primary user***
  An increase in the bit error rate of the primary user signifies a degradation in performance. If the bit error rate in a received frame is too high the frame is dropped. The rate at which frames are dropped is monitored.
- ***Quality of Service (QoS)***
  For both PUs and SUs. (This can be measured in a similar way to degradation in performance.)

Table 1 states which of these performance metrics apply to each of the above applications for heterogeneous networks in licensed bands.

Application	1 Wireless microphone detection	2 Geographical TV White Space sensing	3 TV band detection and transmission
PFA/PMD	✓	✓	✓
Level of structure in PU signal	✓	✓	✓
Achieved network throughput			✓
Degradation in PU performance			✓
QoS			✓

Table 1: Performance metrics for heterogeneous networks in licensed bands and the applications cases to which they apply.

5 Cognitive sensor networks

This test case is an example of how to use the CREW federation to implement and benchmark solutions for robust cognitive networks. In this test case we focus on cognitive sensor networks, more specifically cognitive body area networks (CBANs).

The benchmark described below is specifically targeted at analyzing the performance of a CBAN in an environment with other (building automation) sensor networks. Note that several variations are possible on the described benchmark, depending on the need of the experimenter.

5.1 Test conditions, applications, interference sources

- High level goal: determine the robustness of a CBAN in the presence of interference from other communications in the ISM band of interest, specifically other WSNs
- Output parameters:
  - Robustness of the CBAN in terms of packet loss
  - PSD values of the frequency band of interest
  - Quality indication of experimentation environment
  - Quality indication of the System Under Test

5.1.1 Configuration scenario

a. Network conditions

- Frequency band of interest: 2.4 GHz ISM band
- Radio technologies
  - IEEE802.15.4
  - Possible (uncontrolled) interfering signals in the experimentation environment include: Wi-Fi devices outside the experiment, microwave oven, Bluetooth devices
  - Physical topology of nodes in the experiment: fixed, see Figure 3
  - Fixed TWIST testbed infrastructure
  - Mobile CBAN consisting of shimmer nodes (see Figure)
  - WiSPY spectrum sensors for monitoring
- Description of nodes in the experiment:
  - TWIST nodes: CC2420 radio with TinyOS CSMA MAC, CTP routing
  - Shimmer nodes: standard IEEE 802.15.4 MAC in beacon enabled mode; with reduced sending power

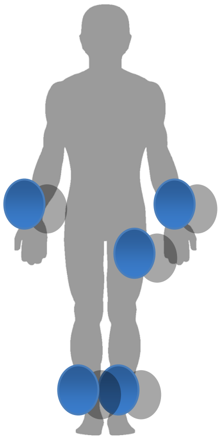


Figure 3: CBAN example

b. Applications

- TWIST testbed: typical building automation applications: periodic sending of sensor data to a sink (data collection)
- CBAN (system under test) : Client nodes sending periodic packets to the master

c. Interference sources

The TWIST nodes from the subsection above form the interference source for the CBAN. There are no other (known) interferers operational during the experiment.

5.2 Performance metrics

The following performance metrics are considered during the benchmark:

- ***Packet loss of the CBAN***
  The number of packets that are sent is known as it is specified in the benchmark. The sink keeps track of which packets are received, and which are not.
- ***Quality of the experimentation environment in terms of spectrum characteristics***
  The wireless spectrum is measured by the WiSPY spectrum sensors just before, during and after the experiment. By comparing the sweeps with the average expected spectrum at the time and place of the experiment, outliers are detected. The location, duration and energy of these outliers will be used to determine a score from 0 to 10, with 10 being the average spectrum, and 0 fully interfered.

6 Cognitive networks in next generation cellular systems

This test case revolves around the accuracy of detecting and tracking unused radio resources in an LTE-like network. The purpose of this benchmark is to provide traffic patterns that represent typical primary user behaviour.

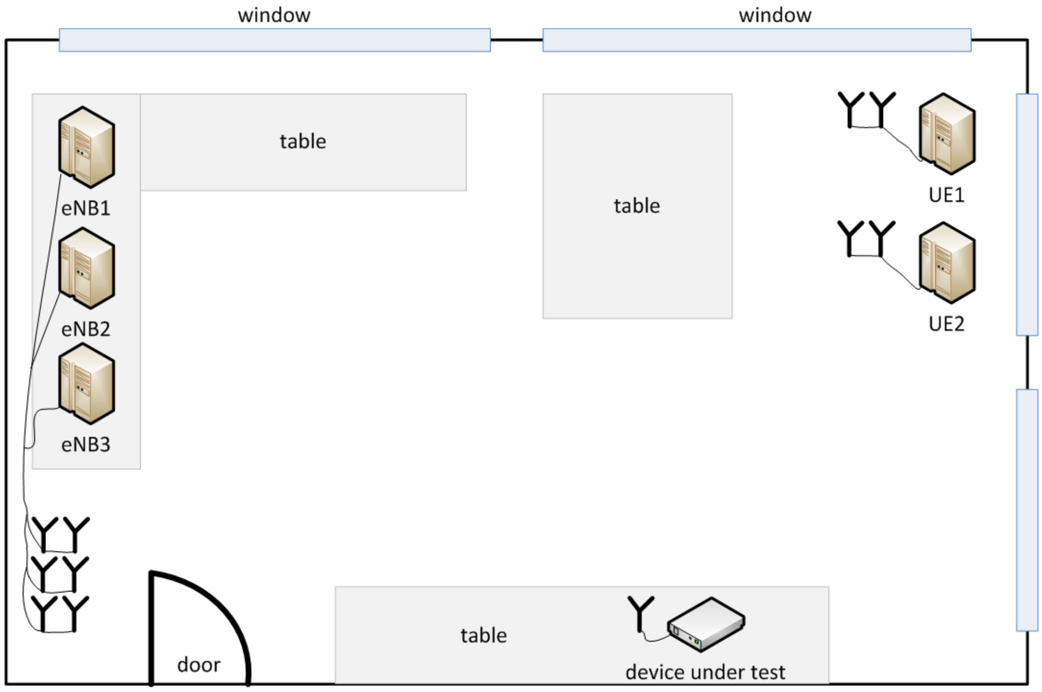
6.1 Test conditions, applications, interference sources

- High level goal: Accurate sensing of a primary cellular system with traffic patterns that represent typical primary user behaviour
- Duration of the experiment: 5 minutes
- Output: Performance of device under test (cf. Applications)

6.1.1 Configuration Scenario:

a. Network conditions

- Frequency band: EUTRAN band VII (DL: 2670-2690MHz, UL: 2550-2570MHz)
- Radio technologies: LTE Rel. 8 like UL/DL
- Topology: Cellular network
- Nodes involved: 1-3 eNB (Sorbas602 eNodeB Simulator), 1-2 UE (Sorbas 202 Test UE), spectrum sensing device under test (provided by the experimenter, e.g. Thales sensing device, imec sensing device)



- Interference sources: Testbed has exclusive access to above frequencies, interference can be created by operating two eNBs on the same resources

- Traffic models (for the primary cellular system): Three different ‘modes’ for traffic generation are considered. The exact numbers in what follows are deduced from [2] and [3].

➢ Mode 1: Voice

For voice traffic an ARM 12.2kbps Codec with encoding frame length of 20ms is assumed. The user behaviour is modelled according to the state chart in Figure 4. The transition probabilities are assumed as a = 0.01 and c = 0.01.

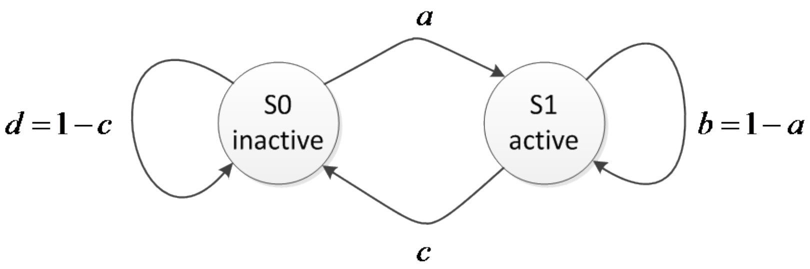


Figure 4: State diagram for voice traffic model.

While in active state, the payload consists of 244 bits and protocol overhead is 76 bits, which leads to a total of 320 bits every 20 ms.

While in silent sate, the silence indicator payload is 120 bits transmitted every 160 ms, as illustrated in Figure 5

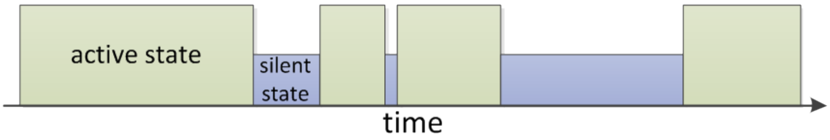


Figure 5: Active and silent states

➢ Mode 2: Web browsing

Browsing a website is modelled as requesting a main object of size S_M which has N_E embedded objects of size S_E. After the content is delivered, a reading time D is assumed before the next website request is made. The values are drawn from the following random distributions:

$S_M \square \mathit{Lognormal}(\mu = 8.37, \sigma = 1.37)$ truncated to $100 \text{ Bytes} \leq S_M \leq 2 \text{ MBytes}$

$N_E \square \mathit{Pareto}(\text{shape} = 1.2, \text{scale} = 2.5\text{ms})$ truncated to $0 \leq N_E \leq 53$

$S_E \square \mathit{Lognormal}(\mu = 6.17, \sigma = 2.36)$ truncated to $50 \text{ Bytes} \leq S_E \leq 2 \text{ MBytes}$

$\mathrm{D} \square \mathit{Exponential}(\lambda = 0.033)$

This model is illustrated in Figure 6.

Figure 6: Traffic model for web browsing

- Mode 3: Streaming

This traffic mode represents all kind of streaming content (e.g. a high definition video), where a continuous transmission with constant bit rate can be assumed, as shown in Figure 7.

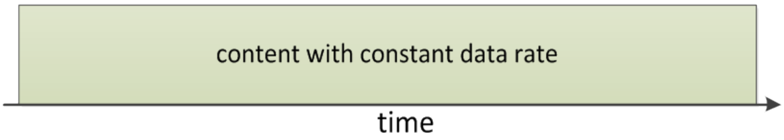


Figure 7: Constant data rate traffic

b. ***Applications:***

Primary system (consisting of at least one eNB and at least one UE) transmits random data with traffic patterns representing typical primary user behaviour (e.g. in downlink). Sensing device under test detects used and unused resources.

If the device under test is capable of spatial sensing, an interferer can be added (cf. Interference sources). Then the goal of the sensing node is to distinguish and possibly separate both sources.

c. ***Interference sources:***

Optionally, interference can be added to the setup by letting additional eNBs transmit on the same radio resources.

6.2 Performance metrics

- Sensing accuracy

  Validation of sensing results versus primary system's resource scheduling

- Primary system’s performance

  Real-time - directly from receiver

  - Received Signal Strength Indicator (RSSI)
  - Reference Signal Receive Power (RSRP)
  - Path loss
  - Channel Quality Indicator (CQI; derived from SINR)

  Non-real-time

  - eNB/UE signal dump
  - Via MATLAB post-processing
    - Block error rate (BLER)
    - QAM constellations

7 Conclusion

This deliverable presented test cases and benchmarks for cognitive radio experimentation on the CREW federated testbed. It has been pointed out that such kind of experiments require a precise definition of the applying test conditions regarding network deployment characteristics, network usage and also potential external interference sources. Besides the precise description of the experiments, benchmarks have been pointed out as a mandatory requirement for reproducible and measureable experimentation results. These benchmarks consist of the wireless scenario including all parameters of the system under test and of the available interference sources used in the experiment. The benchmarks furthermore comprise performance metrics to assess and evaluate the result of each wireless experiment in an impartial and comparable manner.

The purpose of the test case concerning Radio Environment Sensing is to outline a set of benchmarks for experiments that will give us a better idea of the wireless environement we intend to operate in, through both independent and collabourative spectrum sensing. To do this we have outlined experiments that reproduce ISM, TV band and cellular band environments and demonstrate how different sensing techniques and technologies within CREW can improve our performance and cooperation in these bands.

The test case and benchmark addressing the "Horizontal resource sharing between heterogeneous networks in the ISM bands" described in this deliverable refers to a typical home radio frequency environment, since interferences resulting from multiple Wi-Fi stations and one access point are present. By monitoring data throughput and packet losses, this test case aimed at delivering robustness in wireless sensor networks in the presence of such interference.

To address the "Cooperation in heterogeneous networks in licensed bands" usage scenario, a setup consisting of wireless microphones and a DVB-T signal source is considered. A cognitive radio setup based on the imec sensing agent and a USRP platform, with the Iris software radio, are used to monitor the spectrum and establish a wireless link. The applying performance metrics refer to the probabilities of false alarms and missed detections in the spectrum sensing, as well as to the network throughput of the secondary system and the performance degradation of the primary system.

The scenario presented in the "Cognitive sensor networks" usage scenario presents a methodology for comparing IEEE802.15.4 cognitive body area network, in terms of their robustness. This is done by monitoring the CBAN under test as it moves through the controlled TWIST network environment. The performance of the CBAN is benchmarked as a function both of i) its packet loss, and ii) the quality of the environment in terms of spectrum characteristics.

The foundations of any kind of "Cognitive networks in next generation cellular systems" rely on reliable detection of radio resources that are not in use by the primary system. In order to obtain accurate results and draw reliable conclusions, it is important to consider traffic models that represent typical primary user behaviour, thus the three activities 'voice', 'web browsing' and 'streaming' are differentiated. Indicators for the performance are the sensing accuracy of the device under test as well as the primary system's SINR and BLER.

8 References

[1] iperf open source tool, [online] iperf.sourceforge.net/

[2] 3GPP2 C.R1002-0: CDMA2000 Evaluation Methodology

[3] 3GPP TSG-RAN1, R1-070674: LTE physical layer framework for performance verification