Proceedings of
SDR'12-WInnComm- Europe
Wireless Innovation European Conference on Wireless Communications Technologies and Software Defined Radio
27-29 June 2012, Brussels, Belgium

Editors: Gerd Aschied , Marc Adrat, Lee Pucker, Kuan Collins, Stephanie Hamill

SDR'12-WInnComm-Europe Organization

Kuan Collins, *SAIC* (Program Chair)

Thank you to our Technical Program Committee:

SDR Track Chair: Linda Davis, *Institute for Telecommunications Research*
CR Track Chair: Paul Sutton, *CTVR*
Guest Editors, Springer Journal: Gerd Aschied, *RWTH Aachen University* and Marc Adrat, *Fraunhofer FKIE*

Anwer Al-Dulaimi, *Brunel University*
Masayuki Ariyoshi, *NEC Corporation*
Claudio Armani, *Selex ELSAG*
Kamran Arshad, *University of Surrey*
Gerd Ascheid, *RWTH Aachen University*
Fabio Casalino, *Selex ELSAG*
Alexander Chemeris, *Fairwaves*
Luca De Nardis, *Sapienza University of Rome*
Simon Delaere, IBBT-SMIT, *Vrije Universiteit Brussel*
Panagiotis Demestichas, *University of Piraeus*
Daniel Devasirvatham, *SAIC*
Maria-Gabriella Di Benedetto, *University of Rome La Sapienza*
Christoph Heller, *EADS*
Jens Elsner, *Karlsruhe Institute of Technology*
Jair Gonzalez, *TELECOM Paristech*
Jalaa Hoblos, *Kent State University*
Dr. Oliver Holland, *Center for Telecommunications Research*
Joseph Jacob, Objective Interface Systems
Ruediger Leschhorn, *Rohde and Schwarz*
Dake Liu, *Linkoping University*
Fa-Long Luo, *Element CXI*
John McAllister, *Queens University*
Natalia Milio, *Institute of Accelerating Systems and Applications (IASA)*
Klaus Moessner, *Mobile VCE*
Eric Nicollet, *Thales*
Dominique Noguet, *CEA-LETI (Minatec)*
Keith Nolan, *CTVR*
Jorge Pereira, *European Commission*
Venkatesh Ramakrishnan, *Intel Mobile Communications GmbH*
David Renaudeau, *Thales*
Michael Street, *NATO C3*
Paul Sutton, *CTVR*
Paul Tindall, *Cognovo Limited*
Tom Vander Aa, *IMEC*
Dariusz Wiecek, *National Institute of Telecommunication*
Daniel Willkomm, *Technische Universitaet Berlin*
Olga Zlydareva, *Eindhoven University of Technology*

Table of Contents

DEVELOPING GNU RADIO SIGNAL PROCESSING BLOCKS

André F. B. Selva (University of Campinas, Campinas, São Paulo, Brazil; andrefselva@gmail.com); André L. G. Reis (University of Campinas, Campinas, São Paulo, Brazil; andre.lgr@gmail.com); Karlo G. Lenzi (CPqD, Campinas, São Paulo, Brazil; lenzi@decom.fee.unicamp.br); Luis G. P. Meloni (University of Campinas, Campinas, São Paulo, Brazil; meloni@decom.fee.unicamp.br); Silvio E. Barbin (University of São Paulo, São Paulo, São Paulo, Brazil; barbin@usp.br)

ABSTRACT

GNU Radio is a popular toolkit for SDR development applications. Despite offering a large library of signal processing blocks, it is usually necessary to develop other blocks for new system requirements. Furthermore, complex applications, like Digital TV transceivers or LTE networks, for example, demand more complex digital signal processing blocks than the ones available in the GNU standard library. Thus, it is essential to know how to build custom blocks on GNU Radio if we desire to make full use of the GNU Radio features and Software Defined Radio platforms. This work aims to present an in-depth study on how to create new blocks in GNU Radio, clarifying dubious points generated from the lack of documentation available on the subject, which is based mainly on forums discussions and tutorials, and to serve as future reference for new developers. To better understanding this process and to cover all stages of development involved in the creation of blocks, we present a step-by-step procedure on how to build a byte interleaver, commonly used in wireless communication systems, such as WiMAX, LTE and DTV.

1. INTRODUCTION

GNU Radio is an open-source toolbox kit that provides a development environment and some processing blocks that can be used to create software defined radios. The software provides full integration with USRP boards beyond the device drivers specially designed for them [1].

GNU Radio applications are developed using the language Python, where the connection between signal processing blocks are made. A graphical user interface, GNU Radio Companion (GRC), is also provided. GRC could be used to simplify system design, just like Simulink does with projects developed with Matlab [2].

Despite GNU Radio Library provides more than a hundred signal processing blocks, the development of larger projects, like a Digital TV transmitter, or a LTE system, may probably require more complex operations.

In GNU Radio, the blocks are developed in C++. Despite there are already preliminary versions of Python-built blocks, this article will cover only the C++ approach.

Required knowledge to develop blocks would be some Object Oriented Programming concepts, signal processing techniques, familiarity with GNU/Linux operating system and good C++ programming skills (although some C familiarity could be enough to begin coding).

In the following sections we will present the block directory structure, the role of each file and library that we need to create with a block (and their structures), and finally develop a signal processing block fromscratch, as an example.

The examples shown here, as well as the shell scripts used on the compilation steps are tested on a Ubuntu 11.04, using GNU Radio 3.5.0.

2. CREATING A SIMPLE BLOCK

For the first step of this article, the basics of the block development will be discussed. The examples herein are based in [3]. For better understanding, we strongly recommend the download of the gr-my-basic file [4]. The file contains the example library we will study in depth from now on.

In this text and in the block's files, we follow the naming convention mainly adopted in GNU Radio. The name of a block is compounded by the library name, followed by its name, and a reference for its input and output data type. For example, a block from the library "gr_lib", named "add", that add two integer (i) inputs and converts the sum to float (f) will be named as gr_lib_add_if.

2.1. Library Example

Once the example has been downloaded and uncompressed, it should contain a series of directories and configuration

files. The creation process does not include the modification of all of them. Our work focuses on the following directories:

2.1.1 lib directory
This directory contains the source files for the block. Basically, it contains a header file (.h) and a C++ file (.cc) for each block in our library.

2.1.2 swig directory
SWIG (Simplified Wrapper and Interface Generator) is the tool that automatically generates a Python interface for the C++ blocks, so that they can be used in the Python flowgraphs. In the swig directory, there is a series of .i files, each one containing the data needed to create the Python interface. Additionally, the directory contains a .i file for the library.

2.1.3 grc directory
This directory contains the files necessary to create the GRC interface for the block. These files consist of .xml files describing the connections that the tool needs to do between GRC and the Python flowgraph.

2.1.3 Other directories
The other directories include a directory to examples and simple test applications (app), a directory containing a series of configuration files, which cannot be modified (config), and a directory for general python scripts (python).

2.2. Library Compilation

In this paper, we will cover the compilation method using autotools. Alternatively, cmake can be used, as indicated in [5]. Autotools are a group of programming tools GNU uses to make source code packages portable to different Unix-like systems and to reduce the amount of Makefiles that should be edited in the installation process. Due to the use of this tool, the library root folder contains some files, like AUTHORS, bootstrap, configure.ac, Makefile.am, Makefile.common, among others. A Makefile.am file is included on every folder of the library to provide Autotools with the configuration needed to compile the folder's content. Besides the block files themselves, this is the only extra code edition needed to make the library compilation by the addition of a new block.

```
1 $ sudo ./bootstrap
2 $ sudo ./configure
3 $ cd swig
4 $ sudo make generate-makefile-swig
5 $ cd ..
6 $ sudo make
7 $ sudo make install
8 $ sudo ldconfig
```

Fig. 1 – Library Compilation

The compilation process is quite simple. After opening a terminal window and changing the current directory to the library root folder, just execute the commands shown on Fig. 1.

The command executed on line 8 is only required at the first library compilation.

Now we will present in more details the three folders listed in this section and their corresponding files. Each folder contains some *Makefiles* that must be edited after the creation of each block. Only *Makefile.am* files need to be edited , and it is only necessary to add a reference for the new block's header and source file.

3. LIB DIRECTORY

As mentioned in the previous section, the lib directory is the one that contains all source files for the block which is a header file that contains the block's class declaration and definitions, and a C++ file that contains the definition of the block work.

In short, every block is created as a C++ class that inherits from a more comprehensive class, the gr_block class. Our job is to specify this new class, overwriting some of its methods and attributes, and, if needed, creating new ones.

Now, we will study the library downloaded as an example: gr_my. This library contains only one block, an amplifier that multiplies the amplitude of the input signal by a factor determined by the user. In the next subsections, we will take a closer look to each source file.

3.1. Header file

The header file (in the example gr_my_amplifier_ff.h) contains the class declaration. All GNU Radio blocks have their classes inheriting from a common class, the gr_block. The gr_block has three subclasses for simplification of three very usual kind of blocks: a) the gr_sync_block which is a block that consumes and produces an equal number of items per port, b) the gr_sync_interpolator where the number of input items is a fixed multiple of the number of output items (this constant factor between input and output must be determined in the source file) and c) the gr_sync_decimator, which is another fixed rate block, where the number of output items is a multiple of the number of input items predetermined by a constant.

In our example, the block is a synchronous block and so gr_sync_block will be used. After initial declarations, we declare the function gr_make_my_amplifier_ff (int k), whose type is an alias to the boost shared pointers, a type of

```
01  class gr_my_amplifier_ff : public gr_sync_block
02  {
03   private:
04   friend gr_my_amplifier_ff_sptr
          gr_make_my_amplifier_ff (int k);
05   int d_k;
06  gr_my_amplifier_ff (int k);  // private constructor

07   public:

08   int k () const { return d_k; }
09   void set_k (int k) { d_k = k; }
10  ~gr_my_amplifier_ff ();   // public destructor
11  int work (int noutput_items,
12            gr_vector_const_void_star &input_items,
13            gr_vector_void_star &output_items);

14   };
```

Fig. 2 – gr_my_amplifier_ff header file

pointers used to simplify storage issues. This function acts like the public interface for the block on other files (like swig files). More information about boost shared pointers can be obtained in [6].

So far, we have the class initialization. The private members of the class include a friend declaration of the make function, allowing gr_make_my_amplifier_ff to access the class private members, in special, the constructor. Besides, we have the declaration of a new attribute, d_k, which will store our gain information, and the declaration of the private constructor. A stretch of the header file is shown on Fig. 2.

The public members of the class are two methods that provide access to the private attribute, d_k, the destructor declaration, and the declaration of a function named work, where the entire job actually happens.

3.2. .cc file

The .cc file is where the signal processing routine is defined.

After the implementation of the make function, defined on the header file, there is the declaration of four constants: MIN_IN, MAX_IN, MIN_OUT and MAX_OUT, that indicate the minimum and maximum number of input and output streams the block will have.

Following this definition, the constructor's definition takes place. Here, we can see a gr_make_io_signature function. This function associates the block's input and output with the number of streams previously determined. Also, it indicates the data types of each stream. In this example, we do not need any code execution at the constructor. It is usual to use the constructor to allocate data structures we need to use on the signal processing. Further examples will cover this use.

Then, the destructor is overridden, but no code is required in this example.

Finally, we override the work function, where all the signal processing takes place. Let's take a deeper look at its code, shown on Fig. 3.

The parameters of the work function are the number of output items, a vector of pointers to input streams and a vector of pointers to output streams. In lines 4 and 5, we extract from these vectors a reference for the input and the output streams. After that, for every item produced (noutput_items), the output value is calculated: the input value multiplies by the gain parameter. For last, the function returns the number of output items, telling the runtime system how many items were produced.

4. SWIG DIRECTORY

SWIG , as already mentioned, is the tool that generates the interface between C++ blocks and Python. So, the library has some SWIG files (.i) associated, containing the information the tool will need to do its job.

The library contains a general file, which contains references to every block of the lib and it's corresponding header file and swig file. In the example, gr_my.i is the library swig file.

Besides the library information, each block has a .i file associated to its own. In the given example the file is gr_my_amplifier_ff.i. This file must contain shortened information about the block's class, reproducing the private and public class members we want to access from Python, and the declaration of the make constructor. Also, this file must have a call to GR_SWIG_BLOCK_MAGIC, which permits us to access the block as gr_my.amplifier_ff from Python.

5. GRC DIRECTORY

If we want to provide a graphical interface for the block, we

```
01  int gr_my_amplifier_ff::work (int noutput_items,
02  gr_vector_const_void_star &input_items,
03   gr_vector_void_star &output_items)
04   {
05  const float *in = (const float *) input_items[0];
06   float *out = (float *) output_items[0];

07   for (int i = 0; i < noutput_items; i++){
08     out[i] = in[i] * (float)d_k;
09   }

10  return noutput_items;
11}
```

Fig. 3 – Example of work function

need to create a .xml file specifying the correct interface for GNU Radio Companion.

The XML file contains a description of the visual appearance of the block on GRC, like input and output streams names, and some major structural information, like the parameters that should be passed to the code (like the gain, in the given example). It's also possible to propose a standard value for the parameter, or impose some restrictions on its value.

Further information about the XML interface is available on [7].

6. GR_BLOCK CLASS

After studying an introductory example, we are now able to take a deeper look on the gr_block class and its methods.

In the previous example, our block's class inherits from a subclass of gr_block: gr_synck block, already discussed in section 3.1. gr_block is a more comprehensive class, and the relation between the number of input and output items may not be constant or explicit. Instead of telling the run-time system the exact relation between input and output, we indicate, in the work function (general_work, actually, as we are working with gr_block class) the number of input items we consume from each input stream, using the *consume* function. The function needs as parameters the number of items consumed, and from which stream they come from. Alternatively, one may use consume_each function, which consumes the same amount of data from all input streams. In Fig. 4, we can see part of the implementation of the constant gain amplifier previously studied using gr_block class instead of gr_sync_block. One highlight is the use of general_work function, and the use of consume_each.

GNU Radio, during the execution of a Python flowgraph, bufferizes a certain amount of information and calls the general_work function. One very useful thing to know is how to control the quantity of data bufferized

```
01 int gr_my_amplifier_ff::general_work (int
02 noutput_items, gr_vector_int &ninput_items,
03 gr_vector_const_void_star &input_items,
04 gr_vector_void_star &output_items)
05 {
06  const float *in = (const float *) input_items[0];
07   float *out = (float *) output_items[0];
08
09   for (int i = 0; i < noutput_items; i++){
10     out[i] = in[i] * (float)d_k;
11 }
12   consume_each(noutput_items);
13   return noutput_items;
14 }
```

Fig. 4 - Example of the work function using gr_block

before the execution of general_work. To do so, we just have to override another gr_block method: forecast.

Forecast, by default, indicates a 1:1 relationship between input and output. In other words, the required number of input items will be the exact same number of output numbers to be produced. By overriding the forecast method we could implement a decimator or an interpolator by a way that not inheriting from gr_sync_decimator or gr_sync_interpolator, or we could require an exact amount of items to be buffered. This could be very useful when the input/output relation is not very simple, but could be determined with the processing of a fixed quantity of data. Fig. 5 illustrates the override of forecast.

Another important question is what to do if is required to preserve some information between the multiple bufferized executions of the signal processing routine. For example, let's suppose the block to be implemented is a randomizer, which uses a shifter register as a stage of the processing. Let's also suppose the data to be randomized is too large to be buffered, or it's even generated in real-time, making impossible to require a buffer large enough to contain all the information. If the multiple bufferized calls of the function are not considered by the programmer, every time one new call to the working function takes place, the shifter register will be cleared, leading to a periodic loss of information.

The simplest way to solve this problem is to remember that every block is a class, and we can define new attributes to it. So, in the randomizer example, one could declare a pointer to the shifter register's data type as an attribute of the randomizer's class. In the constructor, memory allocation could be placed, as correspondent memory deallocation could take place in the class' destructor. By this, our data will be preserved until the flowgraph stops its execution, eliminating the data loss problem.

7. AN ADVANCED EXAMPLE

In order to summarize most of the issues discussed in this paper, we will present a more complex signal processing

```
01   void gr_block:: forecast (int noutput_items,
02   gr_vector_int &ninput_items_required)
03   {
04    unsigned ninputs = ninput_items_required.size
();
05   for (unsigned i = 0; i < ninputs; i++)
06   // to create a 1:1 relation
07   // ninput_items_required[i] = noutput_items;
08   // to require 100 items on the input
09   ninput_items_required[i] = 100;
10)
```

Fig. 5 - Example of use of the forecast method

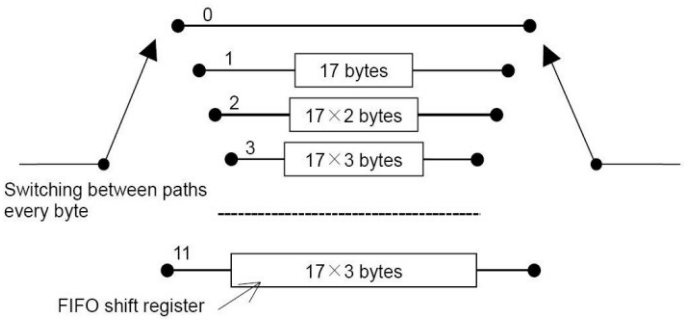


Fig. 6 - Byte interleaver

block example: a byte interleaver.

The byte interleaver here implemented is very common in digital communication systems, such as LTE and DTV standards [8]. The block's work is very simple: the input stream is redirected to different registers, from different sizes. Then, we recombine the shifters output, introducing some kind of shuffle on the bytes. This is very useful in the radio communication field, because it provides some data loss prevention in the transmission process. Fig. 6 illustrates the byte interleaving work.

The files herein referenced are available on [9].

7.1. lib/gr_my_byteinter_bb.h

First, let's take a look on the header file, in the lib folder. The block's class inherits from gr_block, so there is not a simple relationship between the input and output size. Besides similar structures we have already seen on section 3.1, the header files declares the public class members that can be seen on Fig. 7.

Each pointer to byte (v1 .. v11) indicates one path from the byte interleaver. The content of the shifters must be maintained between bufferized executions of the signal processing routine, so it is essential that this information is stored as attributes instead of as local variables. We also must preserve the information about the current shifter that must receive the next byte from the input stream. This information will be stored in the *state* variable.

```
1 public:
2
3 ~gr_my_byteinter_bb (); // public destructor
4  byte *v1, *v2, *v3, *v4, *v5, *v6, *v7, *v8, *v9,
5 *v10, *v11;
6
7  int state;
```

Fig. 7 - declaration of gr_my_byteinter_bb attributes

7.2. lib/gr_my_byteinter_bb.cc

After the declaration of the class, the .cc file is responsible to override the methods we need to use in the block.

First, we create a function called *shifter*, responsible to shift the data contained in each vector previously declared, returning the output number of the process, and receiving as parameters the size of the vector, the shift direction and the feedback input to the vector.

In the constructor, the shifter vectors are allocated, and properly initialized with zero. The *state* variable is also set to zero. In the destructor, all the vectors are freed from memory.

The general_work function makes use of the functions previously defined and implements the byte interleaving. As gr_byteinter_bb inherits from gr_block, we must indicate the number of items consumed from the input. Fig. 8 shows the general_work function code.

7.3. Other files

The example also requires the inclusion of a SWIG file and a XML file for GRC. It is also necessary to modify all Makefile.am files, and the SWIG file correspondent to the block's library.

```
01 int
02 gr_my_byteinter_bb:: general_work(int
03 noutput_items, gr_vector_int &ninput_items,
04 gr_vector_const_void_star &input_items,
05 gr_vector_void_star &output_items)
06 {
07  const byte *in = (const byte *) input_items[0];
08  const int ninput = (const int) ninput_items[0];
09  byte *out = (byte *) output_items[0];
10  for (int i = 0; i<noutput_items; i++){
11          out[i]=shifter(v1,state*M,'R',in[i]);
12          state++;
13          if (state == 12) state = 0;
14 }
15 this->consume(0,noutput_items);
16
17  // Tell runtime system how many output items we
18 produced.
19
20  return noutput_items;
21
22}
```

Fig. 8 – general_work function from gr_byteinter_bb.cc

8. FURTHER EXAMPLES

Readers willing to read further examples of block implementations are encouraged to visit GNU Radio's repository, where all natively included blocks' codes are available [10].

9. CONCLUSION

The development of a signal processing block with GNU Radio is very simple and allows the exploration of the full potential of this software. The examples illustrated here are very general and cover most part of user's common needs.

The final example, a byte interleaver, is a very common signal processing block which is part of many communication systems such as LTE, WiMAX and ISDB-T. Despite its real world application, the interleaver implementation is very simple.

10. REFERENCES

[1] Selva, A; Reis, A; Lenzi, K; Meloni, L; Barbin, S - A Software-Defined Radio Approach: GNU Radio. I2TS 2011 – Information and Telecommunication Technologies Conference Proceedings.

[2] Kaszuba, A. - MIMO Implementation with Alamouti Coding Using USRP2. Progress In Electromagnetics Research Symposium Proceedings 2011.

[3] How to program a new block for GNU Radio and GNU Radio Companion. http://gnuradio.org/redmine/attachments/download/271 – accessed on Apri 10, 2012.

[4] Library example used on this paper: http://gnuradio.org/redmine/attachments/download/276 – accessed on April 10, 2012.

[5] GNU Radio's official website Wiki page http://gnuradio.org/redmine/projects/gnuradio/wiki/CMakeWork – accessed on April, 10, 2012.

[6] Boost Smart pointers web page http://www.boost.org/libs/smart_ptr/smart_ptr.htm – accessed on April 10, 2012.

[7] GNU Radio's official website Wiki page – GRC section http://gnuradio.org/redmine/projects/gnuradio/wiki/GNURadioCompanion#Creating-the-XML-Block-Definition – accessed on April, 10, 2012

[8] ABNT NBR 15601:2007, first edition – Digital terrestrial television – Transmission system. Brazilian Association of Technical Standards.

[9] http://www.rt-dsp.fee.unicamp.br/

[10] GNU Radio Repository - http://gnuradio.org/redmine/projects/gnuradio/repository/ - accessed on April10, 2012.

SPECTRUM SENSING IN THE VEHICULAR ENVIRONMENT: AN OVERVIEW OF THE REQUIREMENTS

Haris Kremo (Toyota InfoTechnology Center, Tokyo, Japan; hkremo@jp.toyota-itc.com);
Rama Vuyyuru (Toyota InfoTechnology Center USA, Mountain View, CA, USA; rama@us.toyota-itc.com)
Onur Altintas (Toyota InfoTechnology Center, Tokyo, Japan; onur@jp.toyota-itc.com);

ABSTRACT

This paper overviews the challenges related to spectrum sensing in the vehicular environment, with emphasis on sensing in the TV licensed band. In the vehicular environment the cognitive radio can help to: 1) satisfy capacity demand for Intelligent Transportation Systems (ITS) applications; and 2) offload time insensitive applications from the ITS dedicated spectrum. However, neither sensing, nor geolocation database lookup alone can provide sufficient incumbent protection. Collaboration among the sensors to take advantage of spatial diversity is difficult due to the rapidly changing network topology. Nevertheless, mobility provides the opportunity to use time diversity at each sensor. We also discuss the influence of sensing subsystem design on the vehicular cognitive network medium access (MAC) sublayer. Whenever applicable, we compare sensing requirements for vehicular cognitive networks to the requirements provided in the IEEE 802.22 standard.

1. INTRODUCTION

Opportunistic utilization of scarcely used spectrum is seen as the way to provide bandwidth that can accommodate continuously increasing number of wireless applications. The TV broadcast band appears to be the "perfect" candidate for this purpose due to low utilization (in particular in rural areas) [1], combined with favorable propagation characteristics which provide relatively long range in comparison to higher frequency bands.

Recent efforts to exploit TV white space are progressing in two directions. On the standardization side IEEE 802.22 wireless regional area networks (WRAN) targeting rural areas [2] was published in July 2011. Other standards are in different phases of development [3] [4] [5] [6] [7] [8]. Recent examples of the implementation efforts are [9] and [10]. Typical applications are Internet access in remote rural areas and remote utility meter reading.

In this document we discuss some of the system engineering issues related to spectrum sensing in the vehicular environment, with emphasis on sensing in the TV licensed band. Although the vehicular environment, due to mobility, introduces significant implementation challenges, we believe that the cognitive radio in this environment can serve twofold purpose: 1) to satisfy capacity demand for Intelligent Transportation Systems (ITS) applications; and 2) as an aid to offload the time insensitive ITS applications from the dedicated spectrum.

The cognitive network nodes must be aware of the spectrum holes in order to utilize them. The two approaches being considered for spectrum awareness are incumbent user signal sensing and geolocation database lookup. We discuss advantages and drawbacks of both approaches and argue that none of them separately can provide sufficient primary user (PU) protection from interference created by vehicular cognitive networks. To corroborate this claim we provide examples in which either one or both of these methods fail.

In the research community collaboration among the sensing nodes, which exploits spatial diversity, is seen as the way to alleviate requirements on PU detection sensitivity [11]. Although mobility causes difficulties in implementing collaboration, it also introduces temporal diversity. Since reliable and timely fusion of sensing information is challenging because of rapidly changing network topology, we argue in favor of utilization of temporal rather than spatial diversity.

We also discuss the influence of sensing subsystem design on the medium access (MAC) sublayer of the vehicular cognitive network protocol stack. This includes: 1) scheduling of quiet periods for sensing; 2) exchange of sensing information among collaborating nodes; 3) coexistence with other cognitive networks of the same or different type; and 4) rules to establish and maintain connection, and to deal with disruptions in connectivity.

In the following, whenever applicable, we compare sensing requirements for vehicular cognitive networks to the requirements provided in the IEEE 802.22 standard.

In Section 2 we motivate our interest in cognitive radio applications for the vehicular environment. Section 3 addresses the issues with sensing and geolocation database lookup under high mobility. Section 4 treats the time diversity. Influence of sensing on the MAC design is addressed in Section 5. In Section 6 we point out some

issues related to differences in regulatory domains. Section 7 concludes the paper.

2. PURPOSE OF THE COGNITIVE RADIO IN THE VEHICULAR ENVIRONMENT

In this section we first argue that the current spectrum assigned for ITS applications is not sufficient. The first step to overcome the spectrum shortage is to use the TV white space to offload time insensitive applications like travel advisory from the ITS dedicated band, and consequently provide more bandwidth for delay intolerant applications.

Second, the cognitive radio can provide alternative to ITS traffic safety and information applications in the dedicated bands, provided that channel switching when a primary user is detected is performed sufficiently fast.

Third, from the perspective of the radio waves propagation, any infrastructure–to–vehicle (I2V) service would benefit from the extended range in the TV bands in comparison to the 5.9 GHz band. Many studies point out difficulties in maintaining WiFi connectivity in the vehicular environment because of small cell radius and wired infrastructure limitations [12] [13] [14].

2.1. ITS Spectrum Scarcity

The dedicated short range communication (DSRC) is introduced to improve safety, enhance travel experience, and even provide access to Internet. In the USA it utilizes seven 10 MHz channels around 5.9 GHz [15] and allows for bitrates ranging from 3 Mb/s to 27 Mb/s. It is based on the carrier sense multiple access with collision avoidance (CSMA/CA) 802.11–like MAC, labeled 802.11p. It is reasonable to assume that, due to the protocol overhead, only a half of the designated rates represent the actual goodput. In other words, the users benefit from a half of the spectral efficiency, which is between 0.3 bit/s/Hz and 2.7 bit/s/Hz. The available bandwidth is shared between a variable number of users. Assignment of one 10 MHz channel exclusively for vehicle–to–vehicle (V2V) communication is under consideration [15].

The protocol analysis provided in [16] considers one-dimensional array of stationary vehicles with variable density. In such a configuration, given a typical sedan length, the maximum density is 200 vehicles per kilometer, or one vehicle every 5 m. The protocol fails to deliver 90% of packets, even in the case of short 200 byte messages and arrival rate of only 2 packets per second per node. Of course, further performance degradation is expected on a multilane highway with mobile terminals because of increased number of nodes and time-varying radio channel.

The authors of [17] consider similar one–dimensional topology, and propose a dedicated transmission queue for time critical safety information. With this modification the allowed number of message retransmissions becomes the key parameter in providing reliability. A wrong setting can have detrimental effect on the packet delivery rate. Too small value results in a low delivery rate, and too large value causes network saturation.

In Japan, the Association of Radio Industries and Businesses (ARIB) standardized under code T75 a physical (PHY) layer in the 5.8 GHz band, which is different from DSRC PHY in the USA [18]. It is designed to provide range up to 30 m across seven pairs of 5 MHz uplink and downlink channels, each pair separated by 40 MHz. The channels support 1 and 4 Mb/s. The primary purpose of the system is electronic toll collection (ETC) and travel assistance. In addition to 70 MHz in the 5.8 GHz band, supplementary 10 MHz segment is assigned for ITS applications between 755 and 765 MHz [19].

2.2. Emerging Vehicular Applications

Since it is very difficult to predict which future vehicular applications will consume the spectrum, we can only provide a few examples of recent trends.

An application using speech recognition software can remotely, through voice commands, perform simple tasks like starting the car engine, opening the trunk, locking the doors, or activating the car alarm [20].

Car manufacturers consider open source hardware and software platforms as an opportunity for developers around the world to create novel vehicular applications [21]. Another example is wireless access to the controller area network (CAN) bus which connects on–board computer unit (OBU) to the vehicle sensors and actuators [22].

2.3. Propagation Advantages of the TV Band

In the free space the received power P_r as a function of transmit power P_t at distance d is

$$P_r = \frac{G_t G_r \lambda^2}{(4\pi d)^2} P_t, \quad (1)$$

where λ represents the wavelength, and G_t and G_r represent transmit and receive antenna gains, respectively. Linear antennas, like a half-wavelength dipole or a monopole, have their size (and thus the aperture) adjusted so that the gain is constant irrespective of the wavelength [23]. Using (1) it can be shown that in free space, for a fixed distance and transmit power, the received power at 5.9 GHz is almost two orders of magnitude smaller than the power at 700 MHz:

$$20\log_{10}\left(\frac{5900}{700}\right)=18.5\,\text{dB}. \quad (2)$$

It should be noted that this naive model neglects two important factors: 1) the intricacy of multipath propagation; and 2) diffraction, which favors lower frequencies over higher frequencies. The latter is usually referred to as the property of lower frequencies to easier "bend" around corners. Still, (2) points out general advantage of TV bands over the microwave DSRC bands: due to extended range the mobile terminal can maintain connectivity to the roadside unit for longer time. Consequently, less frequent handoff and simplified routing is needed as the cars traverse roadside units' coverage area. In addition, decreased density of roadside units results in lower deployment cost.

3. WHITE SPACE AWARENESS IN THE VEHICULAR ENVIRONMENT

3.1. Spectrum Awareness in IEEE 802.22

The IEEE 802.22 standard [2] is designed to provide wireless Internet access through a cellular–like centralized system in remote rural areas without wired infrastructure. The TV spectrum occupancy is obtained either by collaborative sensing coordinated by the base station (BS), and/or by access to a geolocation database with the channel allocation of primary users.

The maximum time interval allocated for sensing can be set between 1 and 160 ms. Channel occupancy is assessed for 30 s before establishing communication. In the USA, once the channel is declared free of incumbents and used for communication, it is reassessed at least once a minute. Whenever a primary user is detected the nodes are required to vacate the current channel after at most two seconds, and have less than 100 ms to transmit coordination messages. Sensing must be able to detect a PU in 2 s with detection probability P_d = 0.9, and with probability of false alarm P_f = 0.1. The BS performs fusion of sensing results by applying OR rule.

In the case of the geolocation lookup access to the database is required at least once in 24 hours.

3.2. Impact of Mobility on Sensing

Multipath fading occurs because many signal replicas are arriving at the sensor with different delays. In the most extreme cases severe fading can be observed when terminals move by a fraction of wavelength. In the ultra–high frequency (UHF) TV bands between 470 MHz and 890 MHz the wavelength is less than a meter. A car which is, for instance, traveling at 40 km/h (25 Mph) passes 1 m in 90 ms.

The time variations of a mobile radio channel are related to the mobile terminal speed. For a sinusoidal wave

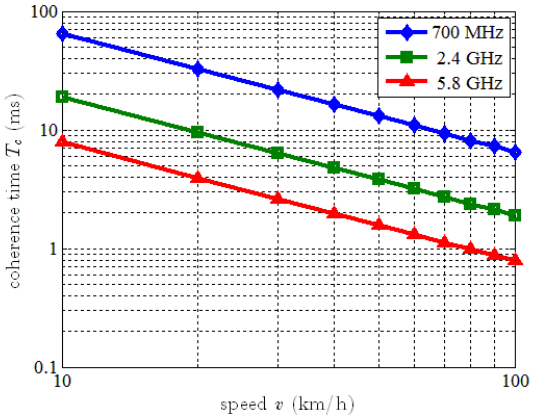


Fig. 1. Flat Rayleigh fading coherence time as a function of the mobile receiver velocity for different carrier frequencies.

of frequency f the maximum dispersion f_D in the Doppler frequency domain is directly proportional to speed v

$$f_{D\max}=\frac{v}{c}f, \quad (3)$$

with c being the speed of light in the vacuum. The time scale of random channel variations is inversely proportional to the maximum Doppler shift [24]. For flat Rayleigh fading this relationship, which is illustrated in Fig. 1, is approximated with

$$T_c=\frac{0.423}{f_{D\max}}=\frac{0.423\cdot c}{v\cdot f}. \quad (4)$$

In general, at lower frequencies the channel coherence time T_c is larger and the channel appears as approximately time invariant for longer time intervals. This means that the channel estimation and tracking, as well as synchronization, are less demanding tasks for the receiver.

Even after the multipath fading is averaged in the "local area" in which its mean does not change significantly, slowly varying fluctuations can be observed. These remaining variations are frequently modeled with a lognormal random variable perturbing the median path loss. Somewhat arbitrarily these variations are termed "shadow fading" because they decorrelate with changes in sensor position on the order of the size of objects in its vicinity. The most popular simple shadowing correlation model is the one presented in [25].

In Table 1 we present comparison of the system design properties of cognitive WRANs and vehicular cognitive networks.

Table 1: Comparison of system design properties of cognitive networks.

	Cognitive WRANs	Cognitive vehicular networks
Application	Internet access	ITS, possibly Internet access
Range	~ 30 km	At most a few kilometers
Mobility	Low: stationary and pedestrian	Can exceed 100 km/h
Topology	Centralized with base station	I2V: centralized V2V: ad–hoc
Target population density	~ 5 users per km^2	Could be larger by two orders of magnitude
Propagation environment	Line–of–sight (LOS) Large delay spread Large propagation delay Slow time variations	Both LOS and NLOS are possible Rate of time variations depends on vehicle speed

3.3 Link Budget

As an example, let us consider Advanced Television Systems Committee (ATSC) digital TV broadcast signal common in the USA. The 6 MHz ATSC channel noise floor is $N_0 = -106$ dBm. In practice, the required sensing threshold for ATSC signal is set below this floor. For instance, the empirical study presented in [26] recommends the sensing threshold between $-118.5 \le N_s \le -108.5$ dBm, depending on the secondary user transmit power limit. In the IEEE 802.22 standard [2] it is set to $N_s = -114$ dBm.

The required sensitivity threshold results in sensing being performed under a very low signal–to–noise ratio (SNR). In [27] it is determined as

$$\rho = N_s - N_0 + G_a - N_f, \tag{5}$$

where N_f represents the sensor front end noise figure (typically 5 to 10 dB), and G_a represents the antenna gain. In the vehicular environment it is reasonable to assume a quarter-wavelength antenna positioned on the car roof with approximately 5 dB gain. Assuming 10 dB noise figure, the required SNR is −13 dB.

To accommodate for fading fluctuations the fading margin must be incorporated into (5). This can be done in two ways.

The first approach would be to start with a system model which incorporates both white noise and fading, and devise P_d and P_f. However, this must be repeated for each sensing algorithm. An interested reader can for instance find in [28] that required SNR for the incoherent matched filter detecting a deterministic signal in flat Rayleigh fading is

$$\rho_0 = \frac{\ln(P_f)}{\ln(P_d)} - 1. \tag{6}$$

This requirement can be relaxed by sensing for multiple occurrences of a known signal feature like synchronization sequence, that is, by increasing the sensing interval duration.

In a rather simple but more generic approach which is independent of the selected sensing algorithm, we can assume certain *outage probability* and lower the sensing threshold by the corresponding *fading margin*. Let us consider flat Rayleigh fading, which represents the most undesirable situation in practice, with only indirect multipath signal replicas propagating toward the receiver/sensor. Under this assumption the SNR is exponentially distributed with some mean $\bar{\rho}$. The outage probability P_0 is the probability that the instantaneous SNR will fall below the threshold ρ_0

$$\begin{aligned} P_0 &= \Pr(\rho < \rho_0) \\ &= \int_0^{\rho_0} \frac{1}{\bar{\rho}} \exp\left(-\frac{\rho}{\bar{\rho}}\right) d\rho \\ &= 1 - \exp\left(-\frac{\rho_0}{\bar{\rho}}\right). \end{aligned} \tag{7}$$

If we assume that only P_0 fraction of time the SNR should be below the minimum sensing requirement ρ_0, the average SNR must be increased by the factor [29]

$$\frac{\bar{\rho}}{\rho_0} = \frac{1}{-\ln(1 - P_0)}, \tag{8}$$

which represents the fading margin. Therefore, for 10% outage the average SNR should be increased by approximately 10 dB, and for 1% outage rate it should be increased by approximately 20 dB. Alternatively, the sensing threshold should be decreased by the same quantity. For instance, an ATSC sensor operating in flat Rayleigh fading with –23 dB SNR should be able to detect incumbent users 90% of time with 90% success rate.

3.4. Database Lookup versus Sensing

In the USA, as well as in the United Kingdom, the database lookup is accepted as the primary method for spectrum awareness [30] [31]. A location aware secondary user is

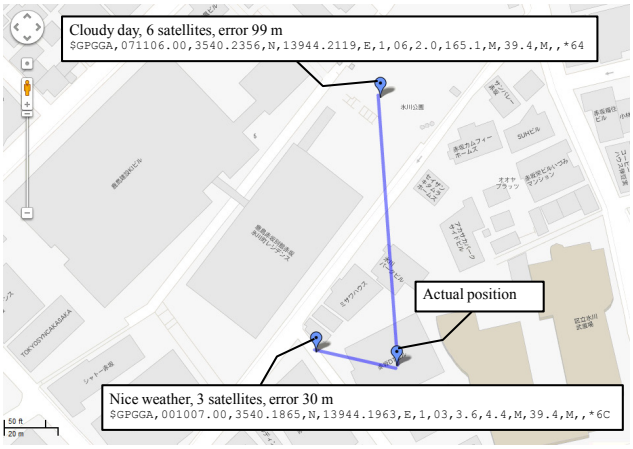


Fig. 2. Example of GPS localization accuracy.

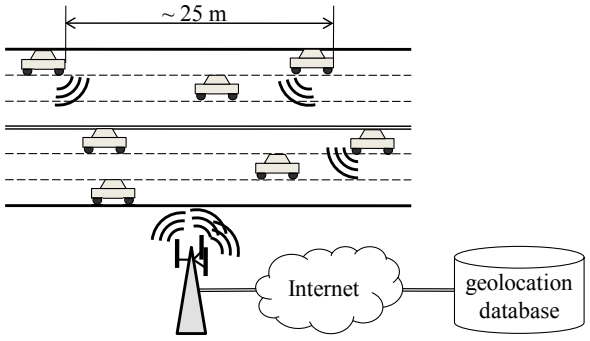


Fig. 3. Probing of the geolocation database in the vehicular environment.

required to periodically access the database of available white space. Although this solution seems very attractive given that it is much simpler than sensing (and in particular cooperative sensing), this approach is not without drawbacks.

First, the secondary user must be location aware. This typically assumes that it has attached a Global Positioning System (GPS) receiver. Alternatively, IEEE 802.22 allows for sophisticated ranging techniques to be used in order to determine user device locations.

Having a GPS receiver device does not guarantee that the secondary will have accurate position at all times. The accuracy depends on the signal strength which deviates with the cloudiness and amount of water vapor in the atmosphere as well as with shadowing due to trees, buildings, etc.

To illustrate the issues with accuracy we put an Ettus Research Universal Software Radio Platform (USRP) N210 [32] equipped with a GPS driven oscillator (GPSDO) [33] on the roof of a building, and measured coordinates on two occasions: during a cloudy day and when the sky was clear. The results in Fig. 2 illustrate that accurate location can be determined with only three satellites when the weather is favorable. When the weather is rainy even six satellites result in an error which exceeds the 50 m geolocation accuracy requirement of FCC [30] by almost 100%. The IEEE 802.22 requires that the location should be determined within 100 m with 67% reliability [2]. In both documents a mobile user is required to perform database query whenever it moves more than 50 m.

The need for Internet access in order to query the database can cause "chicken and egg" problem whenever the secondary users are actually looking for white space in order to access Internet. Centralized network topologies are less prone to this problem, because the base stations, acting as the gateway to Internet, can query databases and obtain spectrum occupancy charts. Since the base stations are fixed, their coordinates can be determined accurately during the commissioning phase. The clients can then simply listen to the BS beacons and associate on the advertised channel. An ad–hoc V2V network must have additional wireless connection to access the database and means to distribute spectrum occupancy between peer nodes.

Sending a database query over Internet involves a variable round–trip delay. To handle this delay Ofcom initially required that the database must reply to the secondary users in ten seconds. This requirement is currently under consideration and it might be abolished [31]. In the vehicular environment a similar restriction would be very important, since a mobile terminal can significantly change position before the response arrives. For a car traveling at 100 km/h this means that the occupancy information is obsolete by current FCC rules if it arrives later than 1.8 s because by that time the vehicle already moved more than 50 m.

To illustrate possible problems with database access congestion, let us assume a six–lane freeway as in Fig. 3. Let us also assume that a 10 km section of the freeway is covered by a single mobile communication system cell. At 100 km/h and heavy traffic involving one car in a lane passing every second, the average distance between the cars is approximately 25 m. Consequently, there are 2400 vehicles in the 10 km section.

If the rules given in [2] and [30] are adopted for vehicular networks, a car traveling at 100 km/h would create one database query every 1.8 s. This means that the considered base station must deal with more than 1300 queries per second in addition to already existing data and voice traffic.

3.5. Examples of Spectrum Awareness Failures

In some situations the sensing, the database lookup, or both can fail.

3.5.1. Sound Barriers

Sound barriers are used to reduce noise in settlements which are close to roads with heavy traffic. However, they can also conceal primary transmitters low on the horizon (Fig. 4) and

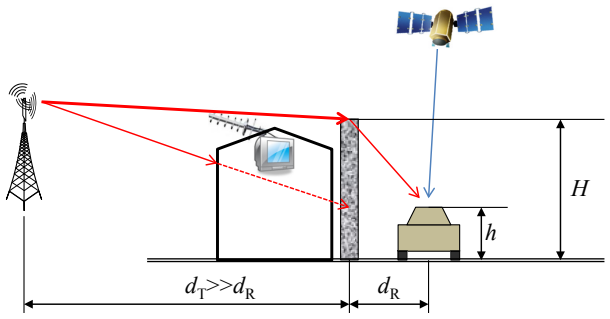


Fig. 4. Geometry of the primary user signal diffraction over the sound barrier.

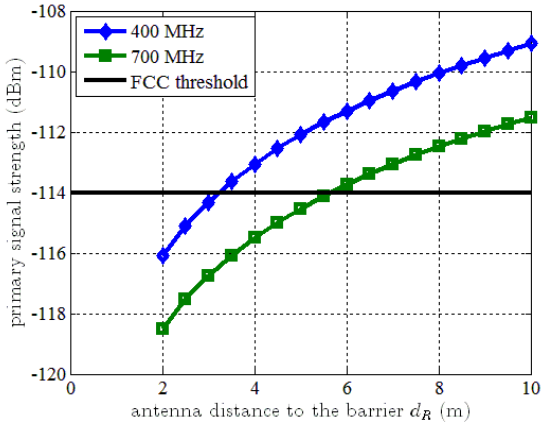


Fig. 5. Primary signal strength as a function of distance from the sound barrier for H = 5 m and h = 1.5 m.

create a classic hidden node problem. We assume that the PU signal cannot penetrate through the adjacent buildings and the barrier.

In such a case the signal strength at the sensor can be determined from a simple knife edge diffraction model [24]. Since $d_T >> d_R$, the Fresnel–Kirchhoff diffraction parameter κ is [34]

$$\kappa = (H-h)\sqrt{\frac{2(d_T+d_R)}{\lambda d_T d_R}} \approx (H-h)\sqrt{\frac{2}{\lambda d_R}}, \quad (9)$$

For $\kappa > 1$ the diffraction loss is approximately [34]

$$P_{\text{diff}} = 20\log_{10}\left(\frac{0.225}{\kappa}\right), \quad (10)$$

with less than 1 dB error.

Let us assume a TV set positioned in a house in front of the sound barrier with 7 dB noise figure and six–element

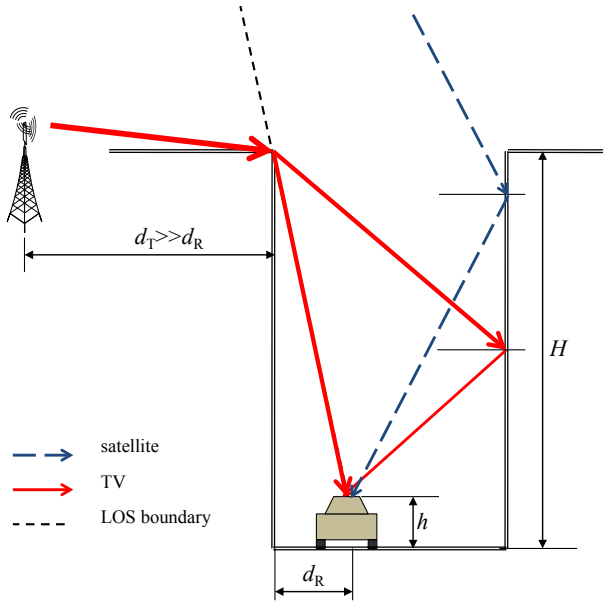


Fig. 6. Urban canyon propagation geometry.

Yagi–Uda antenna with 10 dB gain [23] on the roof. Since the set typically requires at least 15 dB SNR [35], the minimum signal strength at which it can operate $P_{\text{r min}}$ is –94 dBm:

$$P_{\text{r min}} = \rho_{\text{min}} + N_0 - G_a + N_f . \quad (11)$$

Therefore, on the other side of the barrier, depending on the distance between the barrier and the vehicle, the signal could be below the FCC imposed detection threshold (Fig. 5). At the same time, the sky above multilane freeways is usually free of obstructions and an on–board GPS device can calculate accurate position.

3.5.2. Urban Canyon

The satellite link budget is detailed in [36]. The budget takes into account the transmit power, free space path loss, atmospheric losses, user antenna gain, and polarization mismatch. The signal strength at the receiver is –130 dBm, way below the –114 dBm noise floor. In an urban canyon at 1.575 GHz the tight GPS link budged suffers from additional loss since the receiver is shadowed by buildings (Fig. 6).

In the first approximation the diffraction of the satellite signal can be neglected, and only reflection from the opposite wall of the canyon taken into account. Depending on the angle of incidence, the reflection loss can be several decibels, and can cause outage of the GPS signal.

On the other hand, the primary user signal diffracted from the canyon edge can be modeled as the superposition of the direct and the reflected ray [34] (Fig. 6). Similar to the reflected satellite signal, the reflected ray is scaled by

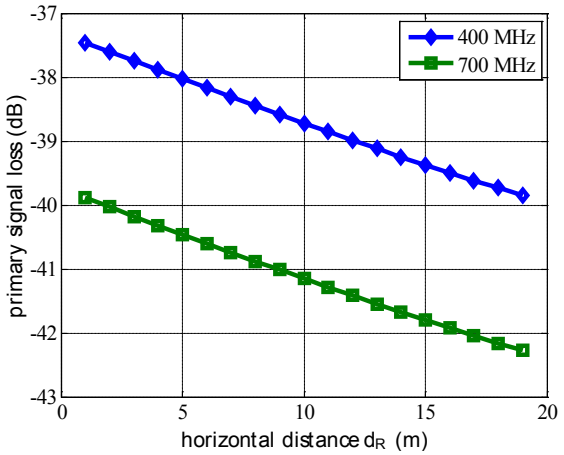


Fig. 7. Primary signal diffraction loss in the urban canyon assuming vertically polarized sensing antenna, H = 100 m, h = 1.5 m, and 20 m wide street.

the wall reflection coefficient. Although the primary signal diffraction loss (illustrated in Fig. 7) can exceed 40 dB, it can be accommodated for, depending on the TV signal strength at the edge of the canyon.

3.5.3. Tunnels

The example in which both sensing and database lookup fails is presented in Fig. 8. Usually, wireless connection is not available in tunnels. Even if the connection is available, the GPS localization cannot determine the vehicle position. One solution to this might be to use the gyro and yaw sensors within the vehicle to correct/estimate the position information when GPS signals are not available.

Spectrum sensing can correctly conclude that there is abundance of available spectrum inside the tunnel. However, once the vehicle leaves the tunnel, spectrum occupancy can almost instantaneously change dramatically.

4. UTILIZATION OF TIME DIVERSITY FOR SPECTRUM SENSING

Common engineering approach to overcome poor detection performance due to fading, described in Section 3, is to apply diversity technique. When a set of uncorrelated channel realizations is observed, it is less likely that the outage will occur on all the realizations. As more and more approximately independent realizations are taken into account, the performance of the overall system approaches (in the limit) the performance of a sensor exposed to only the thermal white noise [29]. Since the wireless channel is a function of frequency, time, and space, in the most general case all these dimensions can provide diversity:

1) Across time a single sensor repetitively searches for the primary user transmissions.
2) Collaboration is achieved across spatial dimension by exchanging information among the sensors.
3) Combined spatial and temporal diversity utilizes both of these approaches.

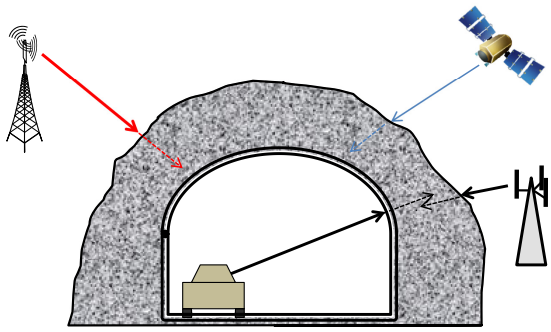


Fig. 8. In a tunnel both sensing and database lookup fail to provide information about available spectrum at the tunnel exit.

In order to satisfy the requirement for uncorrelated channel realizations the samples must be sufficiently separated in time and/or space. The channel realizations are practically uncorrelated when their separation is larger than the coherence time T_c or coherence distance D_c. These are, respectively, the time and the distance at which the magnitude of the correlation coefficient of the channel response reduces and remains below a suitable constant, typically 0.5. As the vehicle moves by an average speed v (regardless of direction in an isotropic environment) it takes

$$T_c = \frac{D_c}{v} \tag{12}$$

to traverse the coherence distance.

Provided that the data fusion can be readily performed, the advantage of spatial diversity is low latency. However, in order to aggregate sensing information, some form of communication must exist. This is difficult in the V2V case, due to the volatile nature of the network topology. This is particularly important for soft sensing algorithms, which generally require exchange of more information than hard fusion algorithms.

In [37] a slotted secondary communication with three phases is proposed: 1) sensing; 2) fusion; and 3) data communication among secondary nodes. The authors of [38] consider electing a node to take responsibility for sensing coordination. A method to select sufficiently uncorrelated spatial samples in a distributed manner is presented in [39]. Reference [40] analyzes the tradeoff between temporal and spatial diversity in a mobile environment.

Let us assume that a relevant regulatory authority requires that the sensing decision must be made every D meters or T seconds (Fig. 9). Whichever is the case, the parameters are coupled by the average sensor speed

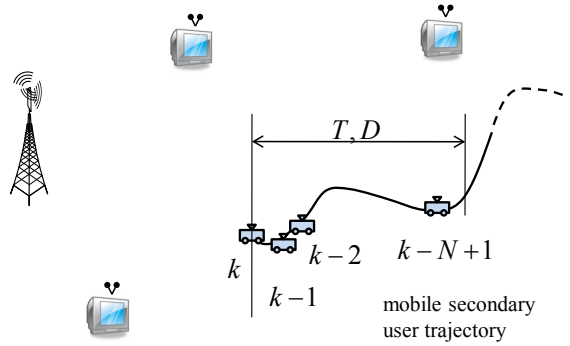


Fig. 9. Utilization of time diversity.

$$D = T \cdot v. \tag{13}$$

Inside this interval there are N uncorrelated sensing periods of duration $\Delta T < T_c$ (Figs. 9 and 10)

$$N = \left\lfloor \frac{T}{\alpha T_c} \right\rfloor = \left\lfloor \frac{D}{\alpha D_c} \right\rfloor, \tag{14}$$

with $\alpha > 1$ representing a suitably selected coherence margin.

Thus, sensing is performed for a time shorter than the coherence time, and the remainder until the next sensing period can be used for communication. The decisions made after N sensing intervals can be infrequently exchanged between the nodes to reinforce accuracy through collaboration.

The coherence time and/or distance can be coarsely estimated by means of crude geolocation information or environmental perception. For instance, the decorrelation distance D_c with typical range from 10 m to 100 m [29], can be tabulated into a handful of values, each suitable for a different terrain topography and urbanization level. The inaccuracy in estimation can be accommodated by α.

5. INFLUENCE OF SENSING ON THE MAC LAYER DESIGN

Design of the MAC sublayer and design of the sensing subsystem are tightly coupled. Sensing related activities controlled by MAC are:

- Scheduling of quiet periods for sensing.
- Selection of the sensing duration.
- Exchange of sensing related messages, including data fusion for cooperative sensing.
- Keeping track of unused available channels for backup.
- "Pushing" of spectrum availability information from the database to the end user terminals which are without sensing capability, like FCC Mode I devices [30].

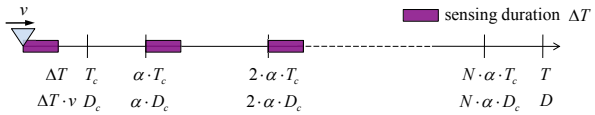


Fig. 10. Scheduling of sensing to utilize time diversity.

Sensing design is greatly influenced by the network architecture. In a vehicular environment the centralized architecture is usually associated with the I2V scenario. A V2V network connects a *swarm* of vehicles on the road which travel in the same direction with similar speeds. Such a network is more likely to be ad–hoc due to volatile routes. From the perspective of protocol design, and in particular sensing, centralized architectures offer many advantages:

- *Quiet period synchronization*: The BS is responsible for traffic management and quality of service (QoS). Thus, it can also order nodes to cease transmissions during sensing periods.
- *Sensing information fusion*: In collaborative sensing the information gathered by all sensors: 1) must be aggregated; 2) a decision about the presence of the primary user must be reached; and 3) the decision must be distributed to all secondary users. In a centralized network it is natural to assign these roles to the BS.
- *LOS propagation*: BSs (or RSUs) are usually positioned with careful planning to provide optimum coverage. BS antennas are mounted on towers which provide line–of–sight (LOS) propagation. This also makes sensing task easier.

In the vehicular environment these tasks must be distributed between the nodes. The nodes must be able to agree on quiet periods for sensing, as well as to perform data fusion in a temporally unstable topology.

6. STANDARDIZATION AND REGULATORY DOMAIN ISSUES

Crude geolocation information is required for a vehicular cognitive device irrespective whether it uses database lookup or active sensing to discover spectrum opportunities. This information is needed to determine regulatory domain (a country or a region) in which the device is located for the following reasons:

- In different regulatory domains different bands might be available for white space utilization.
- Even if the same band is used, the primary users can occupy channels of different width. For instance, the TV channels across regulatory domains could be 6, 7, or 8 MHz wide.
- Communication systems deployed in the same band and on the same channel in different regulatory domains can vary significantly with respect to the employed standard. Analog TV standards are NTSC, PAL and SECAM. Widespread digital standards include ATCS, DVB, and

ISDB–T. Unless the sensing algorithm falls in the class of so called blind algorithms, its design depends on the distinctive features of the signal. For instance, presence of the distinctive pilot tone is characteristic for ATSC signal, the DVB-T signal includes pilot symbols, etc.

- Even for the same technologies in the same band the sensing requirements could be different in different countries. For instance, in IEEE 802.22 sensing is optional in the USA, and not required in Canada [2].
- The design of vehicular MAC can be different across regulatory domains.

7. CONCLUSION

We presented some system design concerns related to the utilization of white space in the vehicular environment. Although the vehicular environment poses many implementation challenges, we conjecture that the ITS licensed spectrum scarcity makes the deployment of vehicular cognitive networks inevitable. Through a handful of realistic scenarios we show that neither spectrum sensing, nor geolocation database lookup alone can provide sufficient protection for incumbent users. While collaboration between the sensors is usually recognized as the method to improve sensing performance, we believe that due to mobility of the nodes, use of temporal diversity must have precedence over collaboration.

8. REFERENCES

[1] S. Chen, A. Wyglinski, S. Pagadarai, R. Vuyyuru and O. Altintas, "Feasibility analysis of vehicular dynamic spectrum access via queuing theory model," *IEEE Communications Magazine,* pp. 156-153, November 2011.

[2] *Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Policies and Procedures for Operation in the TV Bands,* Piscataway, NJ: IEEE, 2011.

[3] *P802.11af - Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification; Amendment: TV White Spaces Operation,* IEEE.

[4] *P802.15.4m - IEEE Standard for Local and Metropolitan Area Networks Part 15.4: Low Rate Wireless Personal Area Networks (LR-WPANs) Amendment: TV White Space Between 54 MHz and 862 MHz Physical Layer,* IEEE.

[5] *IEEE Standard for Architectural Building Blocks Enabling Network-Device Distributed Decision Making for Optimized Radio Resource Usage in Heterogeneous Wireless Access Networks,* Piscataway, NJ: IEEE, 2009.

[6] *P1900.7 - Radio Interface for White Space Dynamic Spectrum Access Radio Systems Supporting Fixed and Mobile Operation,* IEEE.

[7] *P802.19.1 - Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirements - Part 19: TV White Space Coexistence Methods,* IEEE.

[8] *IEEE DySPAN-SC ad hoc on Dynamic Spectrum Access in Vehicular Environments (DSA-VE),* IEEE.

[9] *A Wireless System Designed to Exploit the Full Potential of the TV White Space Spectrum,* Cambridge: NEUL, 2011.

[10] O. Altintas, M. Nishibori, T. Oshida, C. Yoshimura, Y. Fujii, K. Nishida, Y. Ihara, M. Saito, K. Tsukamoto, M. Tsuru, Y. Oie, R. Vuyyuru, A. Al Abbasi, M. Ohtake, M. Ohta, T. Fujii, S. Chen, S. Pagadarai and A. Wyglinski, "Demonstration of Vehicle to Vehicle Communications over TV White Space," in *Vehicular Technology Conference* , 2011.

[11] S. Mishra, A. Sahai and R. Brodersen, "Cooperative Sensing among Cognitive Radios," in *IEEE International Conference on Communications*, 2006.

[12] V. Bychkovsky, B. Hull, A. Miu, H. Balakrishnan and S. Madden, "A Measurement Study of Vehicular Internet Access Using In Situ WiFi," in *The Annual International Conference on Mobile Computing and Networking (MobiCom)*, 2006.

[13] A. Balasubramanian, R. Mahajan, A. Venkataramani, B. N. Levine and J. Zahorjan, "Interactive WiFi Connectivity For Moving Vehicles," in *Special Interest Group on Data Communication (SIGCOMM)*, 2008.

[14] H. Soroush, P. Gilbert, N. Banerjee, B. N. Levine, M. Corner and L. Cox, "Concurrent Wi-Fi for Mobile Users: Analysis and Measurements," in *Conference on Emerging Networking Experiments and Technologies (CoNEXT)*, 2011.

[15] J. B. Kenney, "Dedicated Short-Range Communications (DSRC) Standards in the United States," *Proceedings of the IEEE,* vol. 99, no. 7, pp. 1162-1182, July 2011.

[16] M. I. Hassan, H. L. Vu and T. Sakurai, "Performance Analysis of the IEEE 802.11 MAC Protocol for DSRC Safety Applications," *IEEE Transactions on Vehicular Technology,* vol. 60, no. 8, pp. 3882-3896, 2011.

[17] X. Ma, X. Chen and H. H. Refai, "Performance and Reliability of DSRC Vehicular Safety Communication: A Formal Analysis," *EURASIP Journal on Wireless Communications and Networking,* 2009.

[18] *Dedicated Short Range Communication System,* 1.0 ed., Tokyo: ARIB, 2001.

[19] "Ministry of Internal Affairs and Communications," 2011. [Online]. Available: http://www.soumu.go.jp/main_content/000134495.pdf.

[20] C. Velazco, "http://techcrunch.com/," Aol Tech, November 2011. [Online]. Available: http://techcrunch.com/2011/11/28/new-siri-hack-will-start-your-car-if-you-ask-nicely/.

[21] Ford Motor Company, "Ford Takes OpenXC Research Platform Global, Engaging Local Developers for Market-Specific Apps in India," Ford Motor Company, February 2012. [Online]. Available: http://media.ford.com/article_display.cfm?article_id=36005.

[22] "nikkei.com," Nihon Keizai Shimbun, February 2012. [Online]. Available:

http://www.nikkei.com/tech/personal/article/g=96958A9C93819499E2E1E2E3E08DE2E1E2E0E0E2E3E0E2E2E2E2E2E2;p=9694E0E7E2E6E0E2E3E2E2E0E2E2.

[23] S. J. Orfanidis, Electromagnetic Waves and Antennas, Draft ed., 2007.

[24] T. S. Rappaport, Wireless Communications: Principles and Practice, 2 ed., Prentice Hall, 2002.

[25] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *IET Electronic Letters,* vol. 27, no. 23, pp. 2145-2146, November 1991.

[26] M. McHenry, K. Steadman and M. Lofquist, "Determination of Detection Thresholds to Allow Safe Operation of Television Band "White Space" Devices," in *IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, 2008.

[27] R. Balamurthi, H. Joshi, C. Nguyen, A. Sadek, S. Shellhammer and C. Shen, "A TV white space spectrum sensing prototype," in *IEEE Symposium on New Frontiers in Dynamic Spectrum Access Network*, 2011.

[28] S. Kay, Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory, Prentice Hall, 1998.

[29] A. Goldsmith, Wireless Communications, Cambridge University Press, 2005.

[30] *Second Memorandum Opinion and Order In the Matter of Unlicensed Operation in the TV Broadcast Bands, Additional Spectrum for Unlicensed Devices Below 900 MHz and in the 3 GHz Band,* Federal Communication Commission, 2010.

[31] *Implementing Geolocation: Summary of Consultation Responses and Next Steps,* Ofcom, 2011.

[32] Ettus Research, *USRP N200/N210 Networked Series,* Mountain View, CA: Ettus Research, 2012.

[33] Ettus Research, *Installing the Ettus Research™ GPSDO Kit for USRP™ N200 Series & E100 Series,* Mountain View, CA: Ettus Research, 2011.

[34] F. Ikegami, S. Yoshida, T. Takeuchi and M. Umehira, "Propagation Factors Controlling Mean Field Strength on Urban Streets," *IEEE Transactions on Antennas and Propagation,* vol. 32, no. 8, pp. 822-829, 1984.

[35] S. R. Martin, "Tests of ATSC 8-VSB Reception Performance of Consumer Digital Television Receivers Available in 2005," Federal Communications Commission, 2005.

[36] S. C. Fisher and K. Ghassemi, "GPS IIF - The Next Generation," *Proceedings of the IEEE,* vol. 87, no. 1, pp. 24-47, 1999.

[37] H. Li and D. K. Irick, "Collaborative Spectrum Sensing in Cognitive Radio Vehicular Ad-Hoc Networks: Belief Propagation on Highway," in *IEEE Vehicular Technology Conference*, Spring 2010.

[38] W. X. Yu and H. Pin-Han, "A Novel Sensing Coordination Framework for CR-VANETs," *IEEE Transactions on Vehicular Technology,* vol. 59, no. 4, pp. 1936-1948, 2010.

[39] M. D. Feliche, K. R. Chowdhury and L. Bononi, "Cooperative Spectrum Management in Cognitive Vehicular Ad-Hoc Networks," in *IEEE Vehicular Networking Conference*, 2011.

[40] A. W. Min and K. G. Shin, "Impact of mobility on spectrum sensing in cognitive radio networks," in *ACM workshop on Cognitive radio networks (CoRoNet '09)*, 2009.

PERFORMANCE EVALUATION OF A SPECTRUM-SENSING TECHNIQUE FOR LDACS AND JTIDS COEXISTENCE IN L-BAND

Giulio Bartoli (Università degli Studi di Firenze, Florence, Italy; giulio.bartoli@unifi.it); Romano Fantacci (Università degli Studi di Firenze, Florence, Italy; romano.fantacci@unifi.it); Dania Marabissi (Università degli Studi di Firenze, Florence, Italy; dania.marabissi@unifi.it); Luigia Micciullo (Università degli Studi di Firenze, Florence, Italy; luigia.micciullo@unifi.it); Claudio Armani (SELEX Elsag S.p.A.; claudio.armani@selexelsag.com); Roberto Merlo (SelexElsag S.p.A.; roberto.merlo@selexelsag.com).

ABSTRACT

This paper deals with a cognitive approach able to guarantee the coexistence of new data link for air-ground aeronautical communications LDACS and military JTIDS systems. Future LDACS shall coexist with current systems operating in the same frequency band for this reason coexistence issues must be carefully investigated. In particular JTIDS transmissions can affect the LDACS performance acting as disruptive impulse noise. JTIDS exploits frequency hopping to protect information, hence its interference on LDACS system cannot be foreseen and avoided. In addition the bandwidth of the two signals results to be completely overlapped in case of collision. The disruptive effects of JTIDS interference on LDACS can be mitigated if the collisions can be detected and hence suitable processing techniques can be activated. This paper proposes a method to detect the presence of JTIDS interference exploiting an energy detection spectrum sensing technique based on sliding windows and packets retransmission. The performance of the proposed method is presented in terms of missed detection probability of the JTIDS interference and error rate of the LDACS system showing a significant capability to counteract JTIDS interference.

1. INTRODUCTION

The increasing demand for advanced communication services in civil aviation leads to the need for a new management and communication framework able to support the capacity and security requirements of the air transportation system [1]. The European project SESAR (Single European Sky ATM Research) aims to develop and validate a new communication system capable of satisfy the requirements specified in Communication Operating Concept and Requirements (COCR) document [2].

Since the COCR operational requirements cannot be fulfilled by a single technology, the Future Communication Infrastructure (FCI), constituting the communication part of the framework, will be implemented as a system of systems, integrating existing as well as new communication technologies. The FCI should support both digital voice and data communications. Particular emphasis is given to the data, since in case of failure, voice would not be able to maintain the operations with the same reliability.

Since the VHF frequencies are congested, the communication system in charge of supporting the air/ground data link will operate on the aeronautical L-Band (960-1213 MHz) and it will be named L-Band Digital Communication System (LDACS). This technology is currently under development and, at the present time, two options have been identified. LDACS1 is the former option. It is a Frequency Division Duplex (FDD) system exploiting the OFDM (Orthogonal Frequency Division Multiplex) technique, that is very effective against the inter symbols interference. The latter option, LDACS2, is a Time Division Duplex (TDD) system utilizing the CPFSK (Continuous-Phase Frequency-Shift Keying) modulation, that allows to reduce the out-of-band emissions. At the end of the SESAR program studies one of them will be selected as the key technology for air/ground communications. This work focuses on LDACS1.

LDACS will operate in the L-Band where several legacy systems are already present (e.g., DME, SSR, UAT, JTIDS/MIDS), hence the spectral compatibility is an important issue that needs to be addressed. In particular in this paper we consider the coexistence between LDACS1 and the Joint Tactical Information Distribution System (JTIDS), also known as Multi-functional Information Distribution System in the NATO implementation.

JTIDS is military system used for several purposes, like identification in surveillance and mission management. It exploits an impulsive signal and frequency hopping, in order to make the system interference-tolerant.

Unfortunately the frequency hopping is performed on a large range of frequencies spanning almost the whole L-Band, hence the probability to have collisions between the LDACS and the JTIDS signals is very high.
Since LDACS is an OFDM system and the decoding is performed in the frequency domain, the impulsive noise affects all the bits carried by the interfered OFDM symbol. This represents an advantage until the power of the interference does not exceed a certain threshold [3], but it becomes very disruptive since leads to high error probability on each OFDM subcarrier. The problem of impulsive interference in OFDM system is a known problem in Power Line Communication (PLC), where the man-made noise, as the turning on/off of an electrical switch, can deteriorate the performance. This topic has only recently been investigated in wireless networks context, where the co-channel interference can assume an impulsive nature. The solutions available in the literature deal with non-linear elaborations of the interfered signal through clipping or blanking schemes [4]-[6]: when the received signal amplitude exceeds a certain threshold the signal level is set to the threshold value or to zero, respectively. Another interesting solution relies on a close loop scheme based on data detection and successive noise estimation and reduction [7]. Unfortunately this scheme has a high computational cost. However, none of these methods takes under consideration any advanced technique to distinguish which samples are affected by interference.
Coexistence between systems through spectrum sensing is a typical topic of the Cognitive Radio networks, where an unlicensed (Secondary) system coexists with a licensed (Primary) system in a transparent manner: the Secondary user's radio identifies the free frequency channels for secondary usage and transmits in these frequencies in a non-interfering manner. The scenario considered in this paper is different: sensing operation is used to detect the presence of the JTIDS signal and then to perform suitable operations at the receiver side. In particular the proposed scheme is based on energy detection and packets retransmission. The energy detector is a well-known sensing algorithm that allows the detection of interfering signals with low complexity and works well when the interference features are not known. In our scheme sensing operation is combined with packets retransmission and is performed by exploiting two signal replicas in order to have a more efficient detection. Furthermore the replicas are combined to improve the signal detection after an interference blanking.
The paper is organized as follow: Section 2 presents the system model and Section 3 provides an analytical evaluation of the impact of JTIDS system on LDACS performance. Section 4 presents the proposed scheme while in Section 5 the numerical results are shown and in Section 6 the conclusions are drawn.

2. SYSTEM MODEL

In the scenario under consideration the LDACS and JTIDS systems operate in the same area and LDACS frequency band is part of one of the hopping bands used by JTIDS.
We considered a system where LDACS operates within one of the band used by JTIDS, so, due to the frequency hopping technique adopted by the latter, occasionally the two systems interfere each other. In particular, LDACS exploit OFDM modulation technique, where the data stream is divided in many orthogonal sub-streams, referred as sub-carriers, each one with reduced rate. This allows to reduce the negative effects of Inter Symbol Interference (ISI) without decreasing the total data rate. According to the latest specification in [8], the total LDACS available band is equal to B_{LDACS}=498.05 kHz, that is oversampled with sampling rate f_s=625 kHz. This spectrum is divided in N=64 subcarriers, of which only 50 are active. In addition, a cyclic prefix (CP) with duration equal to 11 samples is inserted in order to avoid ISI. In the considered system each active subcarrier is modulated following a QPSK scheme. The LDACS main parameters are outlined in Table 1, where t_s is the sampling period, Δf is the frequency spacing between two contiguous subcarriers and T_u, T_g and T_{TOT} are the useful symbol time, the guard interval duration and the total OFDM symbol duration, respectively.

B_{LDACS}	498.05 kHz	**N**	64	**T_u**	102.4 μs
f_s	625.00 kHz	**CP**	11	**T_g**	17.6 μs
Δf	9.76 kHz	**t_s**	1.6 μs	**T_{TOT}**	120.0 μs

Table 1: LDACS system parameters

JTIDS exploits the TDMA (Time Division Multiple Access) technique [9]. The transmission in each time slot is composed by a minimum of 258 pulses of duration T_p=6.4μs and spaced by T_i=6.6μs. The active part of each pulse is multiplied for a spreading sequence containing 32 chips modulated through a CPFSK scheme and representing a combination of 5 informative bits: in particular, JTIDS specifies only one spreading sequence and the particular combination of informative bits introduces a circular offset in that. The chip duration is equal to T_{chip}=200 ns and hence the signal bandwidth is equal to 5MHz. However, since the JTIDS system operates in a large range of frequencies, [960 - 1215] MHz, divided in bands of 3 MHz, the signal is filtered to fit the bands. Each pulse is transmitted on a different frequency according to a certain hopping sequence. In particular there are N_{ch}=51 possible carriers. Since the hopping sequence is a classified information, in our model we assume every carrier has the same probability to be selected, hence we consider a hopping pattern randomly generated with uniform probability on all frequencies. Considering that the LDACS spectrum is

significantly smaller than the JTIDS bands, for the sake of simplicity we can assume it is completely contained in one of the these bands. This means that only one hopping frequency affects the LDACS signal but on the whole spectrum. Finally, since the JTIDS sampling rate is much higher of that of LDACS, the JTIDS signal is down-sampled. This makes the interference even harder to be detected, since the signal power is spread on a longer period.

We assume both the signals (LDACS and JTIDS) are transmitted on fading channels and AWGN noise is added at the receiver. Fading coefficients are modeled as random variables with Ricean distribution, with Ricean factor equal to K=4 dB and normalized power.

3. ANALYTICAL INTERFERENC EVALUATION

The development of a new system operating in a frequency band densely populated by legacy systems introduces the need of a careful evaluation of coexistence issue. The joint representation of the frequency domain signals of different systems operating on the same band gives a qualitative evaluation of the mutual interference permits to put in evidence potentially critical scenarios. For this reason we give a joint representation of JTIDS and LDACS spectrum in nominal band and in Out of Band and Spurious Domain. The signals have been generated in Time Domain (TD) taking into account the standards mandatory features [8],[9] that directly affect the spectrum shape. For LDACS1, an OFDM signal with 64 subcarriers has been generated in FD and transformed in TD by means of a IDFT (Inverse Discrete Fourier Transform). Then a raised cosine window that aims to reduce out-of-band radiations has been applied. Later, the signal has been transformed in FD and filtered with the spectral mask. For JTIDS an impulse of 6.4μs duration has been produced and modulated with a spreading sequence of 32 chips each of 200ns duration and CPFSK modulated. Once transformed the signal to FD, the spectrum mask has been applied. Figure 1 shows LDACS and JTIDS power spectra when an offset (Δf) of 5 MHz is assumed between the central frequencies of the two systems. It is evident that the JTIDS spectrum heavy interferers with the LDACS due to its high transmission power. It can be noted that the results are similar even with higher frequency offsets (i.e. translating the LDACS spectrum).

To quantify the effect of the interference a numerical analysis can be performed. The aim is to identify the conditions that permit to LDACS to operate in presence of JTIDS interference. The analysis consists of the computation of the interference power level at the victim receiver and on its comparison with the maximum tolerable interference power level obtained from a protection criteria typically the minimum Carrier to Interference ratio (C/I).

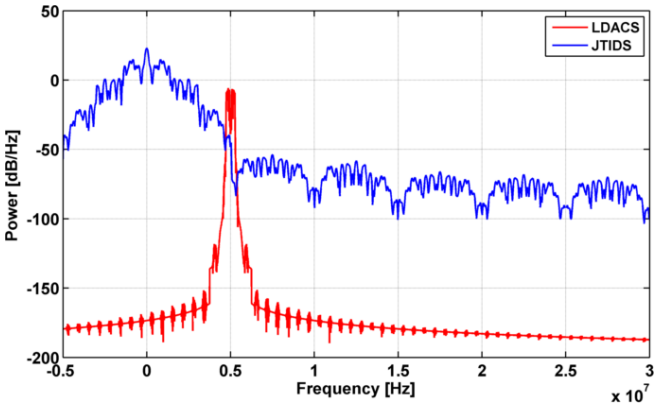


Figure 1 JTIDS and LDACS spectra with a frequency offset equal to 5 MHz

The analysis follows the procedure defined in the CEPT MCL (Minimum Coupling Loss) method [10]. The interference level is obtained by means of a link budget (1) that depends on different parameters, as frequency separation and distance:

$$I(d,\Delta f) = EIRP - PL(d) + Grx - Lrx + OCR(\Delta f) + DC \quad (1)$$

where:

- *EIRP* is the JTIDS Equivalent Isotropically Radiated Power;
- *PL(d)* is the free space path loss;
- *Grx* e *Lrx* are the LDACS receiver antenna gain and cable loss, respectively;
- *OCR(Δf)* (Off Channel Rejection) is a term that takes into account the ability of the victim receiver to reject the interferer signal. It depends on the power spectral density of the interferer and on the frequency separation between interferer and victim [11].
- DC is a term that takes into account the interferer duty cycle. For JTIDS, it depends on the Time Slot Duty Factor (TSDF) that represents the maxim number of slots assigned to an user in a frame.

From this analysis it is possible to determine the minimum distance between the interferer and victim equipment at which the interference results to be tolerable. The coexistence between the two systems is guaranteed if this distance is lower than the minimum operational distance: in a rough evaluation we can consider the minimum operational distance equal to the minimum vertical separation of the aircrafts: 300mt (on the ground this distance is lower).

The results of the analytical analysis are shown in Table 2 in a worst case assuming the following working hypothesis:

- Parameters of the link budget as antenna gains and cable losses are varied according to the scenario

taking into account typical values for ground or aircraft installations;

- The ratio C/I is fixed to 10 dB, in according to B-AMC (Broadband Aeronautical Multi-carrier Communication) specifications, from which LDACS has been derived, since this value is not available in the current LDACS specification.
- The value of maximum interference power acceptable is obtained from C/I ratio assuming the carrier power equal to receiver sensitivity.
- The JTIDS transmission power is fixed to 1000 W.
- TSDF is set to 50% and 5% that represent the maximum and minimum values.

	Scenario	Non Interfering Distance	
		TSDF=50%	**TSDF=5%**
1	Ground Station to Airborne Aircraft	d > 500km	d > 157km
2	Airborne Aircraft to Airborne Aircraft to	d > 500km	d > 500km
3	Airborne Aircraft to Ground Station	d > 500km	d > 260km
4	Ground Station to Ground Station	d > 500km	d > 500km
5	Aircraft on the Ground to Aircraft on the Ground	d > 26.6km	d > 18.5km
6	Ground Station to Aircraft on the Ground	d > 500km	d > 162km
7	Aircraft on the Ground to Ground Station	d > 46.5km	d > 25.5km

Table 2: Analytical Evaluation Results

These results show that minimum distance that allow the spectral compatibility is considerably higher than the minimum operational distance in all scenarios, therefore the problem of the interference of JTIDS transmission on LDACS1 requires the application of some countermeasures to ensure the coexistence between the two systems.

4. INTERFERENCE SENSING AND MITIGATION

In this Section the proposed sensing and mitigation scheme is explained. The basic idea is the retransmission of the Packet Data Unit (PDU) when the presence of JTIDS system is detected (it can be done through a first sensing phase).

The first copy of the packet is stored and combined with its retransmission: since JTIDS and LDACS transmissions are independent processes, even if both the copies of the PDU are affected by JTIDS interference with high probability different portions of the PDU are corrupted. Packet combining is used either to improve interference detection and signal decoding.

Interference detection aims to detect with significant reliability which samples of the PDU are affected by interference. This is done through the energy detector, that is a particular spectrum sensing algorithm which computes the energy of the received samples during a time interval called sensing period. This detector is well-known in the Cognitive Radio networks, where a secondary unlicensed user (SU) looks for spectrum holes that are not used by the primary licensed user (PU). Free spectrum portions can be used by the SU for transmission. Energy detector provides the test statistic (i.e., the energy of the signal coming from the sampler) used to decide between two binary hypothesis: the channel is free and only thermal noise is present or primary user signal plus noise is present. Accuracy of energy detector is proportional to the duration of the sensing period: increasing the number of collected samples is possible to improve the performance. The interference detection proposed here is slightly different: in our system the goal is to detect which samples are corrupted and not only if the interference is present. It means that the sensing period must be limited: differently from traditional approach increasing the sensing period does not leads to a performance improvement, since the energy of the pulse is spread on a longer interval making more difficult to discriminate which samples are affected.

For this reason we introduce a modified energy detector that exploits a sliding window which collects the energy of a part of the received signal.

In addition the interference detection is not performed on the received signal but on a sample by sample difference between the two copies of the signal that have been previously equalized to remove the channel effect.

This permits to reduce the false alarm probability due to by the a high peak to average ratio (PAPR) that characterizes the LDACS (i.e. OFDM) signal: the difference between the two replicas depends only on the JTIDS interference and noise power; the LDACS signal fluctuation does not affect the detection procedure.

Assuming a perfect channel equalization, the received signals difference on the *i*-th sample is:

$$\Delta r[i] = r_1[i] - r_2[i] \tag{2}$$

where $r_k[i]$ is the *i*-th received sample of the *k*-th PDU transmission (k=1,2).

The *n*-th output of the sliding window energy detector is the test statistic T_n:

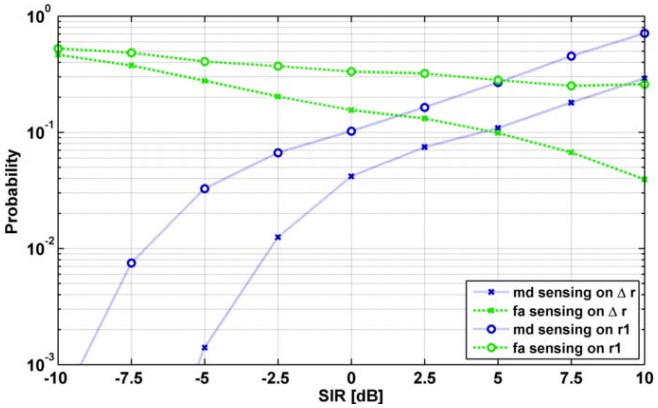


Figure 2 False alarm and miss detection probability for the proposed method and traditional sensing

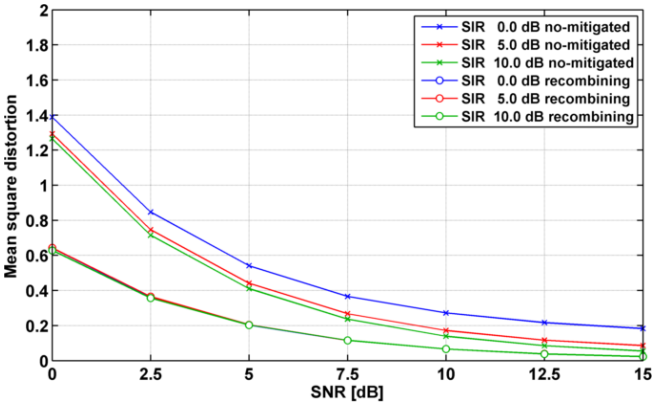


Figure 3 Mean square distortion of the received signal and the signal after the interference mitigation

$$T_n = \sum_{i=n-\frac{W}{2}}^{n+\frac{W}{2}} a_i \left\| \Delta r[i] \right\|^2 \qquad (3)$$

where W is the window width and a_i are the window weights (weights are selected in order to give more importance to central samples).

If T_n exceeds a certain threshold interference is supposed to be present in the sample n-th.

The window length depends on the duration of the interfering signal that cannot be exactly known because the pulse is filtered by the receiver, however a rough estimation of the JTIDS signal duration can be calculated as $L = T_p/f_s = 4$ samples, where f_s is the sampling frequency and T_p the pulse duration. We adopt a windowing size, W, equal to $L + 1$ samples. A further improvement to reduce the false alarm probability is obtained by observing M consecutive test statistics: if at least $M=L-1$ consecutive samples are over the threshold we assume the interference is present.

Observing the retransmissions difference is possible to know which samples are affected by interference but it is not known if the interference is introduced by the first, $r_1[n]$, or the second $r_2[n]$ copy of the received signal.

Indicating as n_i, with $i = 0,\ldots, I-1$ the samples affected by the interference, for each n_i the values of $r_1[n_i]$ and $r_2[n_i]$ are compared: the maximum is blanked while the minimum is doubled. The resulting signals are summed together, and the final signal can be expressed as:

$$r'[n] = \begin{cases} r_1[n] + r_2[n] & if \quad n \neq n_i \quad \forall i \\ 2r_1[n] & if \quad r_1[n] < r_2[n] \quad and \quad n = n_i \\ 2r_2[n] & if \quad r_2[n] < r_1[n] \quad and \quad n = n_i \end{cases} \qquad (4)$$

Demodulation and decision are taken on r' signal.

5. NUMERICAL RESULTS

This section shows the numerical results obtained to validate the proposed scheme by resorting to computer simulations.

We start our analysis by evaluating the performance of our interference detection method. The main drawback of this method is the higher noise power, since by considering the difference between two signal copies the noise power is doubled. On the other hand, it does not suffer the presence of the LDACS signal as traditional method. Figure 2 shows the comparison of the false alarm (*fa*) and miss detection probabilities (*md*) between our detection technique based on difference between two signal replicas (Δr) and the same sensing technique performed on one signal (r_1) when the SIR (Signal to Inference Ratio) varies. Looking at this figure we can see that by exploiting the two signal replicas it is possible to get a significant improvement for both false alarm and miss detection probabilities.

In order to evaluate the behavior of the mitigation scheme, we define the performance index D as the mean square distortion of the signal.

$$D = \frac{1}{P} \sum_{p=0}^{P-1} \left| r'[p] - s[p] \right|^2 \qquad (5)$$

Figure 3 presents the distortion of the received signal and the reconstructed one, when the SNR varies and for different SIR values. From this figure we can see that the proposed scheme is able to significantly decrease the distortion of the signal for the considered SIR values. In particular, the residual distortion is due to the AWGN noise.

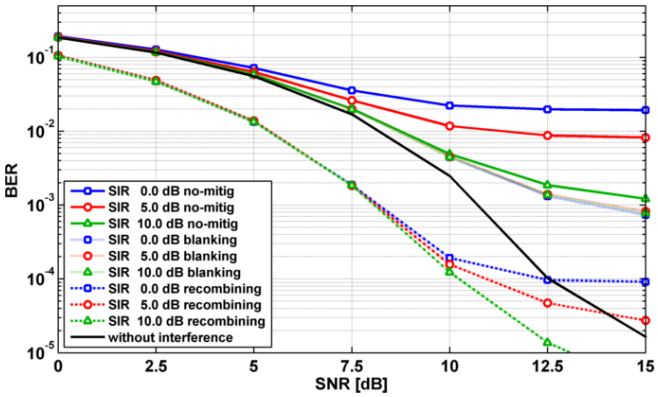


Figure 4 BER performance of the system without channel coding when interference is not mitigated and when blanking and the proposed technique are applied

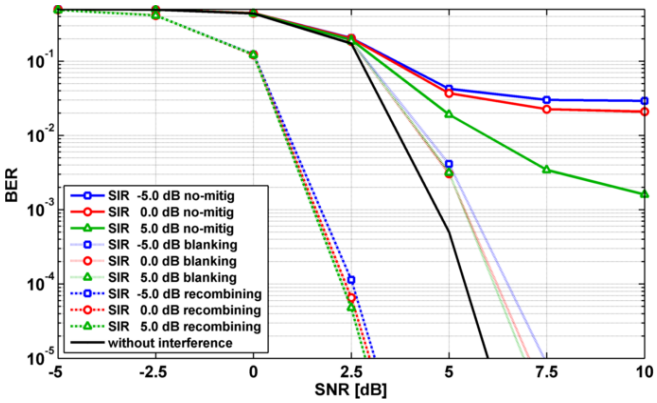


Figure 5 BER performance of the system with channel coding when interference is not mitigated and when blanking and the proposed technique are applied

This means that the presented mitigation scheme is able to reject the interference made by JTIDS and to bring an additional gain of 3 dB due to the soft combining of the packet's replicas.

The good behavior of the proposed scheme is even more evident in terms of BER (Bit Error Rate). When the interference is high the BER gain permits to counteract the reduction of throughput introduced by the retransmission assuring more reliable communications. Figure 4 represents the BER when the QPSK modulation is used for different SIR values and when the SNR varies. In this figure the proposed scheme is compared with traditional blanking method and the case without any elaboration. We can see that interference leads to a floor of the BER: this means that the performance does not improve even when the SNR gets higher. Even traditional blanking does not allow to reach satisfying performance. On the other hand, when the proposed scheme is applied we have excellent results: the performance shows the interference is almost completely removed and we have an additional gain of about 3 dB due to the soft combining of the packets. For high signal to noise ratio even the performance of the proposed method reaches a floor, due to those samples affected by interference in both the signal replicas. However this drawback is overcome when the channel coding is considered: Figure 5 shows the same performance of the previous case but considering the channel coding envisaged in LDACS specifications. In particular we considered an inner convolutional coder with coding rate equal to 1/2, an interleaver and an outer Reed-Solomon coder. We can observe that the floor concerns only the performance of the system when mitigation is not applied. Here the performance of the proposed scheme clearly overcome the traditional blanking solution. We can note from the previous results that performance does not depends on the considered SIR: when the interference power is very high it is easy to be detected and removed, while when the JTIDS signal strength is low it becomes more difficult to detect but it has also a lower impact on performance.

6. CONCLUSIONS

In this paper we proposed a novel scheme for impulsive interference detection and mitigation in the new LDACS system. The proposed method is based on interference detection and packet retransmission: by exploiting the difference between the signal replicas affected by two independent interference realizations it is possible to improve the interference detection, since it does not depend on the useful signal. Furthermore, the two signal copies can be used to reconstruct the useful signal, decreasing the distortion and the system bit error rate.

7. REFERENCES

[1] "Future Aeronautical Communications", ISBN 978-953-307-625-6, book edited by: Dr. Simon Plass, German Aerospace Center (DLR), Germany, September 2011. http://www.intechopen.com/books/future-aeronautical-communications

[2] Communications Operating Concept and Requirements for the Future Radio System, ICAO Std., Rev. Version 2, May 2007.

[3] M. Budsabathon and S. Hara, "Robustness of OFDM signal against temporally localized impulsive noise," in Proc. VTC 2001 Fall Vehicular Technology Conf. IEEE VTS 54th, vol. 3, 2001, pp. 1672–1676.

[4] S. V. Zhidkov , "Performance analysis and optimization of OFDM receiver with blanking nonlinearity in impulsive noise environment," vol. 55, no. 1, pp. 234–242, 2006.

[5] S. V. Zhidkov, "Analysis and comparison of several simple impulsive noise mitigation schemes for OFDM receivers," vol. 56, no. 1, pp. 5–9, 2008.

[6] Y.-H. Kim, K.-H. Kim, H.-M. Oh, K.-H. Kim, and S.-C. Kim, "Mitigation of effect of impulsive noise for OFDM systems over power line channels," in Proc. IEEE Int. Symp. Power Line Communications and Its Applications ISPLC 2008, 2008, pp. 386–390.

[7] J. Armstrong and H. A. Suraweera, "Impulse noise mitigation for OFDM using decision directed noise estimation," in in IEEE Int. Symp. Spread Spectrum Techniques and Applications, 2004, pp. 174–178.

[8] Updated LDACS1 System Specification, SESAR 15.2.4 ET - Task EWA04-1 T2 Std. 00.01.00, Aug. 2011

[9] JTIDS System Segment Specification (DCB79S4000C), Std.

[10] E. R. C. E. within the European Conference of Postal and T. A. (CEPT), "A comparison of the minimum coupling loss method, enhanced minimum coupling loss method, and the Monte-Carlo simulation," Tech. Rep.,1999

[11] Recommendation SM.329-10 Unwanted emission in spurious domain, ITU-R Std.

Software Radio Spectrum Analyzer

Jérôme PARISOT, Emilien LE SUR, Christophe MOY,
Daniel LE GUENNEC, Pierre LERAY

SUPELEC/IETR
Avenue de la Boulais, CS 47601, 35576, Cesson-Sévigné Cedex, France
christophe.moy@supelec.fr

Abstract — **This paper presents the implementation of a software radio spectrum analyzer using USRP platforms form Ettus Research as a frequency translator and Simulink environment for baseband processing. The algorithms that are used and the performance that are obtained are detailed. This work can be used as a pedagogic tool. It also illustrates that software radio is not only restricted to communication devices. Moreover, it gives insight on the performance bottlenecks of a software defined radio approach mixing hardware and software considerations, especially on the software environment aspects (operating system and software environment).**

Index terms— **USRP, GNU radio, Digital spectrum analyzer, Simulink, Welch method.**

I. INTRODUCTION

The aim of this paper is to assess the feasibility of a Spectrum Analyzer using an SDR (Software Defined Radio [1][2]) approach with COTS (Component Of The Shelves) low-cost platforms. A classical signal processing approachfor spectrum analysis is used in this paper. The originality consists in using an environment usually dedicated to functional simulation, as a real signal processing environment, in a high-level design philosophy [3].

The common equipment used for spectrum analysis of radio signals, proposed by Rohde & Schwarz, Tektronix or Agilent, are devices that rely on complex analogic circuitry and digital signal processing algorithms. Those devices are very effective, but bulky, heavy and very expensive. Using software defined radio, we know we cannot create a digital spectrum analyzer that exactly mimics professional spectrum analyzers. However we want to evaluate what can be done using USRP™ platforms (Universal Software Radio Peripheral) from Ettus Research [4] and Mathworks™' Simulink™ as a development environment. We first present in section II an overview of the system. Section III describes the spectral analysis algorithm used in the project. The implementation of each sub-function is detailed in section IV and finally the implementation results of the system in terms of performance are presented and discussed in section V.

II. SYSTEM OVERVIEW

A spectrum analyzer is nothing but a specific radio receiver that does not process radio signals to demodulate them, but to make an analysis of signals in the time domain so as to convert them in the frequency domain. We propose then to use USRP platforms as a front-end for a spectrum analyzer to convert signals form a chosen RF frequency to baseband and then to process baseband signals in a classic computer.

Radio signal sampling is handled by the USRP platform. We can split the sampling process into several functions: shifting the signal to an Intermediate Frequency (IF), sampling with an Analog-to-Digital Converter (ADC), Digital Down-Conversion (DDC), buffering into frames and transferring on the Ethernet link. Those functions are already implemented in the USRP platform and can be parameterized by the user on a host PC.

The GNU Radio project [5] provides a development interface on a Linux computer which could be used here. However our aim was to evaluate the possibility that has appeared recently to use Simulink™ on Windows operating system. The goal is to benefit from the user friendly Matlab™ interface to make digital signal processing. This is particularly suited to a educational context for instance, but we can imagine in the future that this context could be generalized further, as a possible technique for commercial products.

Signal processing for spectrum analysis is done in Simulink™ environment, running on a host computer. First operation consists in buffering frames and managing USRP platform configuration (carrier and down-sampling factor for instance). Then a processing chain computes the power spectrum density using Welch's method [6]. The result is then displayed using three techniques. Direct display plots the Power Spectral Density (PSD) as a curve refreshed at each buffer arrival. Spectrogram display represents the spectrum through time: the vertical axis corresponding to time,

horizontal to frequency and color to amplitude of the PSD. This kind of display is used to observe quick changes in the spectrum, often difficult to notice to the naked eye such as short periodic pulses. Persistence display uses color to indicate the continuing presence of a part of the PSD. It can be used to detect signal below the average level of noise. All these different sub-functions are described as a block diagram in Figure 1.

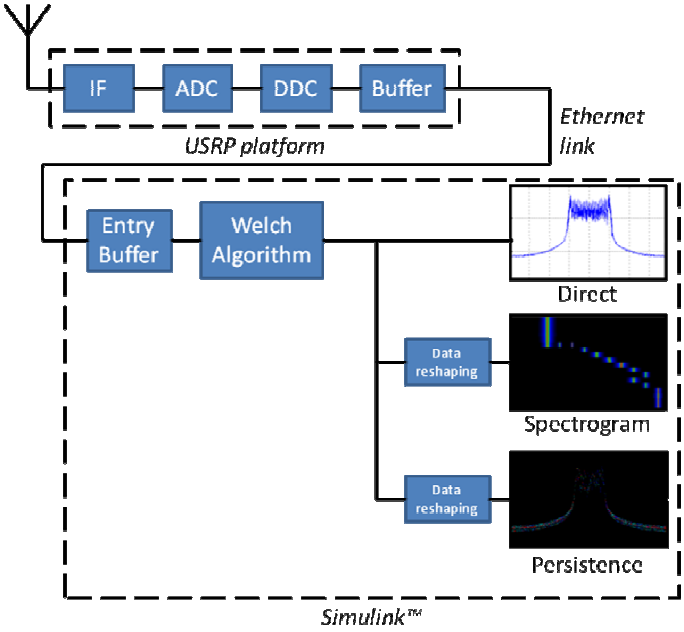


Figure 1: Block diagram of the entire System and functions distribution between USRP and Simulink™.

III. Welch's Method

A. Principle

The Welch method [6] is used to estimate the Power Spectral Density of a signal using time samples. A common approach is to estimate the PSD of the signal in one step using the FFT applied on every available samples. The originality of the Welch method is to divide the samples of the signal into overlapping sub-parts and then calculate the PSD for each segment. To each sub-parts is applied a window in order to reduce the effect of truncation. The sum of all the PSDs is then computed and normalized to obtain the final estimator of the PSD. The method is summarized in Figure 2.

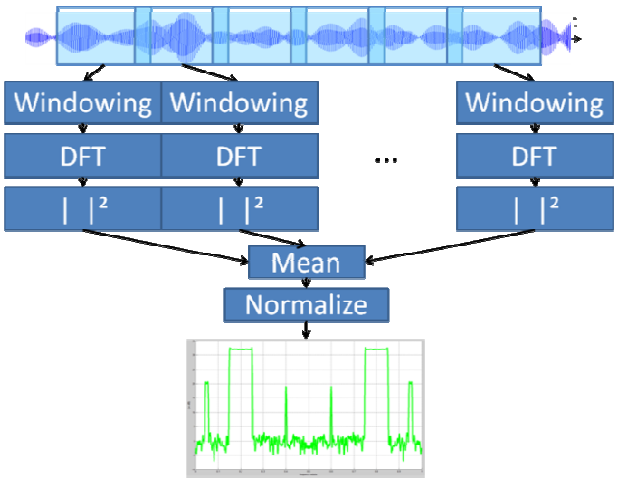


Figure 2: Diagram description of the Welch Method

B. Benefits

The main benefit of using this method is to minimize the noise observed on the spectrum. If we consider the K PSD obtained for each segment as independent random variables, the mean of all these variables has a variance K times less than the variance of the PSD computed using a periodogramm on the entire signal. Due to the overlapping parts, the variance will be slightly higher than the one obtained by equation (1) but gives an approximation on the behavior of the algorithm.

$$\mathrm{var}(\mathrm{PSD}_w) \approx \frac{\mathrm{var}(\mathrm{PSD}_p)}{K} \qquad (1)$$

Nonetheless, calculating the Discrete Fourier Transform (DFT) on the M samples from each segments instead of using the N samples from the whole signal means that the DFT obtained will have fewer points, thus the resolution R will be worse in the Welch method than with the classic periodogramm approach as shown in equation (2):

$$R = \frac{1}{M} \alpha \frac{K}{N} \qquad (2)$$

C. Results

In order to validate the performance of our algorithm, we feed our processing chain with a signal in the time domain with a known spectral density. The prototype spectrum contains two narrow-band signals and a Dirac impulse. Noise is then added: the resulting signal has a SNR of 10 dB. The algorithm performs well in placing the different elements on the frequency axis as can be

seen in Figure 3. The results on this figure show our algorithm approaching the theoretical spectrum. Moreover, results are strictly identical to the one of the *pwelch* Matlab™ function.

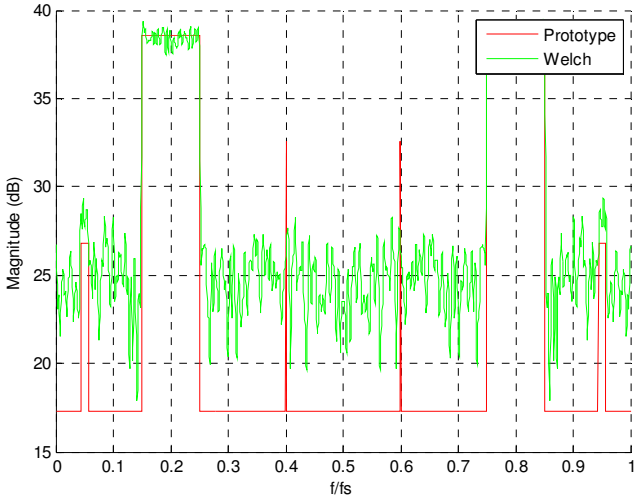


Figure 3: Our algorithm estimation of the prototype spectrum.

IV. Implementation

A. Hardware

The experiments have been performed with the USRP™ N210 from Ettus Research. This platform can be connected thought Ethernet to the computer. The ADC of the USRP platform runs at the constant sampling rate of 100MHz, the samples are then put into fixed-size frames [4].
A signal generator (SMY 01 from Rohde & Schwartz) has been used to generate known signals so as to test the device and the reception algorithms. Simulink™ runs on a computer with the following characteristics: 2.53 GHz Intel Core i5 with a 4 GB RAM.

During the experiments the limits of the n210 USRP platforms has been revealed: signs of aliasing appeared in some situations.

B. Matlab / Simulink

The software part was developed thanks to the Simulink support package for USRP (in Matlab 2011b). The typical reception architecture consists in the reception bloc (from the USRP™ library) associated with a subsystem presented in Figure 4. This subsystem contains a hand-made user-defined block that computes the PSD using the Welch's method described in section III. Our implementation of the Welch's method is purely sequential: the input signal is split, the DFTs for each part are computed and added to the output one at a time and finally the result is normalized. At each iteration the whole content of the entry buffer is treated.

In direct display mode, we use a basic vector display to plot the spectrum as a curve at each iteration.

In Spectrogram mode, starting with the spectrum computed by the Welch block, a hand-made block converts the vector into a color gradient [7]. At each iteration the colored line is displayed below the preceding one: the result is a spectrogram, showing the evolution of the frequency band according to time. The described steps are represented in Figure 5.

In persistence display, a fixed number of spectrum curves are superposed. The intersections of these curves bring change in color (green then red) of the intersection point. Thus, persisting features of the spectrum intersect themselves for a long time and change color whereas a short signal does not intersect itself and appears with a different color (blue).

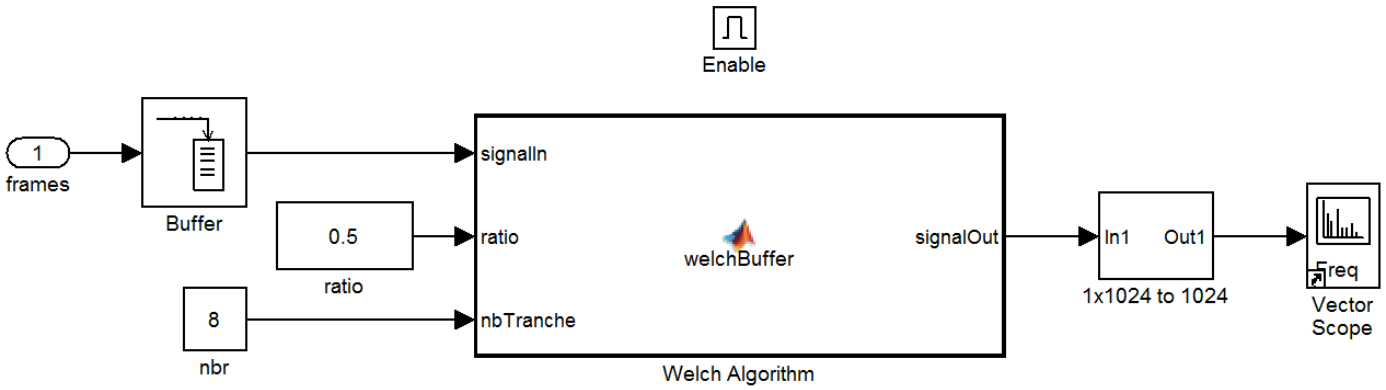


Figure 4: Simulink structure for the Direct Plot display

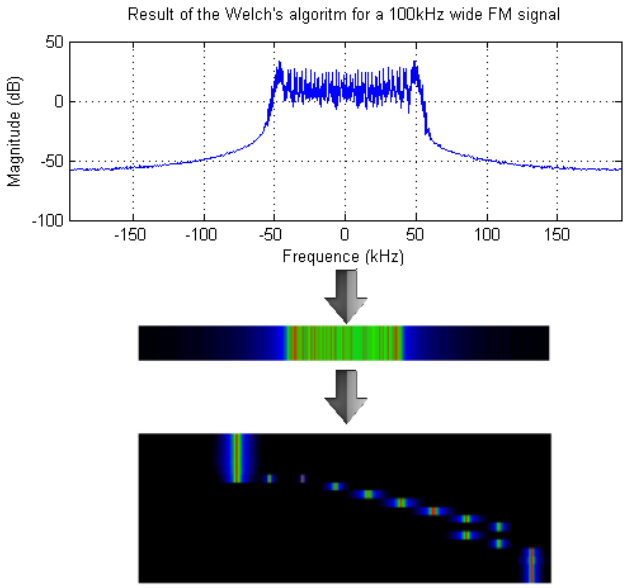


Figure 5: Displaying a spectrum through time.

C. Performance

For the system to run in real time, all computations must be finished before the next portion of signal is acquired - otherwise the system would accumulate unfinished calculations and will not be in real time. Therefore the duration of signal contained in the entry buffer has to be the display period. The display rate can then be expressed with the following equation:

$$f_{disp} = \frac{f_{ADC}}{DCR} \cdot \frac{1}{n_f \cdot n_b} \qquad (3)$$

Where f_{ADC} is the ADC sampling rate (100 MHz), n_f the number of samples per frame, n_b the number of frames in the entry buffer, and DCR the Down-Conversion Ratio. The time to compute the PSD depends on the Welch's algorithm's parameters and on the output size of the FFT. An output size of 1024 gives sufficient precision for this project.

On one hand, DCR of the USRP platform determines at which rate the sample packets reach the host computer, that is to say the computation rate. The higher the computation rate, the harder maintaining the real-time constrain is. On the other hand, the DCR determines the bandwidth that can be observed with the analyzer. Considering Shannon's theorem, we can deduce the maximum observed frequency f_{max}:

$$f_{max} = \frac{f_{ADC}}{2 \cdot DCR} \qquad (4)$$

V. Results

The following paragraphs present the results for each reception configuration (Direct plot, Spectrogram and Persistence). Moreover, it shows the limits of our software spectrum analyzer in terms of bandwidth, display rate and latency.

A. The Welch's estimator

Figure 6, shows a 100 kHz wide FM signal at a carrier frequency of 600 MHz displayed using our analyzer. This signal was observed on the computer and, in order to validate it, on an analogic spectrum analyzer.

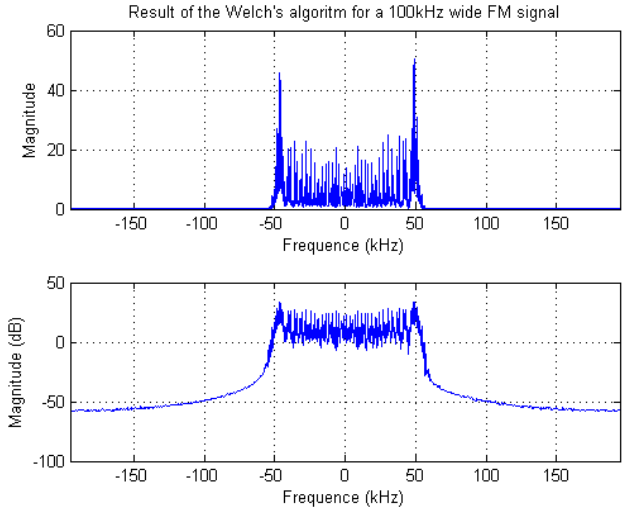


Figure 6: Display of the Welch's estimator for a FM signal of 100 kHz of bandwidth centered on $f_{carrier}$ = 600 MHz

The trade-off concerning the DCR with our configuration has been determined to be a DCR of 256 which yields an observed bandwidth of 390 kHz, which is comprised between $f_{carrier}$ - f_{max} and $f_{carrier}$ + f_{max}. The physical limit is a ratio of 64 which corresponds to an error-less transmission of data through an Ethernet link. Then limitation is coming from the execution speed of Simulink application on the host PC.

Maximum display frequencies for different configurations of the Welch algorithm are presented in Table 1. For this experiment, the buffer size has been

settled to 362 (default value) and the DCR to 256. The choice of the DCR depends on the required bandwidth for the application. 256 was a good trade-off between bandwidth (almost 400 kHz) and processing power (we achieved to reach 30 output frames per second).

Table 1 : Maximal frequency for the Direct Display

Welch's Ratio / Number of Slices	0.5 / 8	0.8 / 16
Size of the entry buffer	30	112
Maximal display frequency	36 Hz	9.8 Hz

Figure 7 presents the Relative Delay as a function of overlapping ratio and number of slices. Relative Delay is the computation time divided by the duration of signal in the entry buffer (i.e. number of samples x sampling period). A Relative Delay of 1 means that the system is running in real time.

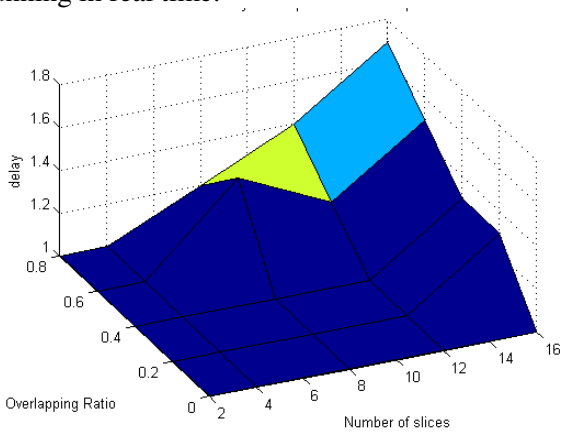


Figure 7: Influence of Welch's parameters on the system's performance

The number of slices is the number of successive FFTs that the system computes. The size of each FFT diminishes with the number of slices for a fixed total number of samples; it increases with the overlapping ratio for a fixed number of slices. These are the reasons why, as Figure 7 shows, increasing the number of slices and the overlapping ratio decreases the performances: the system has to compute a lot of long FFTs in a sequential manner.

We then study the impact of these same parameters on the PSD estimate's variance. For the measurements a noised empty band of 340 kHz has been observed at 801 Mhz. Each measurement of the variance is normalized according to the signal energy, and is compared to the variance measured for one slice as shown in Figure 8.

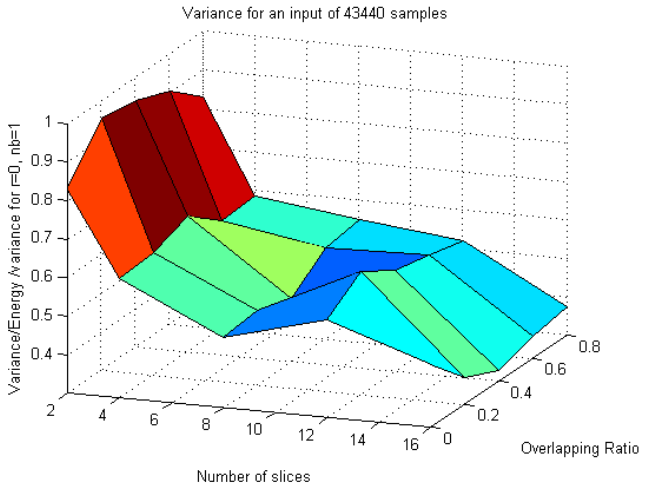


Figure 8: Influence of Welch's parameters on the variance

This graph shows that in order to minimize the variance, the number of slices has to be increased. Unfortunately, Figure 7 showes that there is a trade-off between quality and execution duration performances.

B. Spectrogram

The spectrum estimate is done with the same blocks than in previous sub-section A. Figure 9 shows the output of the spectrogram display when submitted to the following two-step test:

- Variation in frequency of a pure sine wave from 800 kHz to 860 kHz. (1)
- 100 kHz-wide FM modulation of the previous sine wave. (2)

Table 2 shows the operating limits of such a system in terms of disply refreshment.

Table 2 : Maximal display rate for the Spectrogram

Welch's Ratio / Number of Slices	0.5 / 8	0.8 / 16
Maximum display rate	5.7 Hz	2.6 Hz

Compared to the result obtained in the sub-part A, the display rate decreases due to additional calculations required for the display.

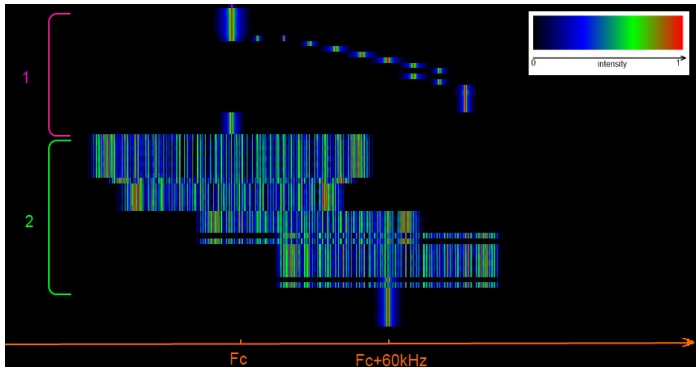


Figure 9: Spectrogram

C. Persistence

The persistence display is composed by 20 accumulated results from our frequency estimator. Figure 10 illustrates the obtained result for a wide-band signal.

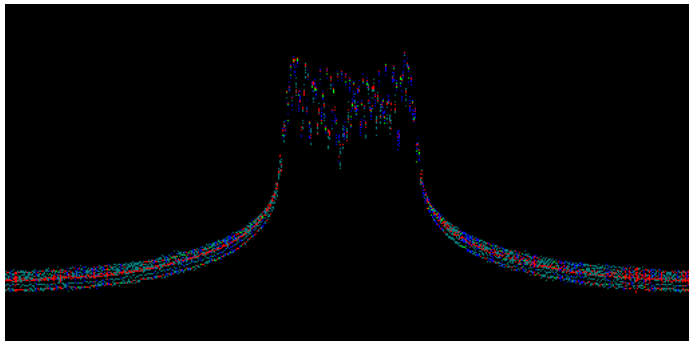


Figure 10: Remanence display

As a way to compare the results, the operating frequencies have been mesured and are given in Table 3.

Table 3 : Maximal display rate for Persistence

Welch's Ratio / Number of Slices	0.5 / 8	0.8 / 16
Maximum display rate	8.3 Hz	3.9 Hz

This algorithm is faster than the previous one, essentially due to the reduction of the memory size.

VI. Conclusion and Perspectives

This paper shows that a digital spectrum analyzer based on USRP™ platform interfaced with Simulink™ can be implemented. Several limitations and trade-offs have been pointed out and should be taken into consideration for future work. We demonstrated that estimating the PSD of the signal using successive calculations of DFTs does not allow the user to minimize the estimates variance without impacting the performances.

A parallel implementation of the Welch method on FPGA would be a way to make this trade-off disappear and at the same time tremendously increase performances without sacrificing the flexibility of our approach.

Simulink™ running on Windows™ showed to be a powerful tool in terms of flexibility during the design process; nonetheless such an environment is not sufficiently optimized to be used for a commercial application in the manner we propose in this paper.

This spectrum analysis could be in an extended manner used as a sensor for spectrum analysis in a cognitive radio perspective. We could indeed imagine in the long term that this technique could be integrated in future radio systems in order to detect frequency holes nad manage dynamically access to free spectrum resources [8].

This lab and others [9] can be found at:
http://www.rennes.supelec.fr/ren/perso/cmoy/SCEE-SERI/

VII. References

[1] J. Mitola, "The Software Radio Architecture," IEEE Comms. Mag., vol. 33, no. 5, pp. 26-38, May 1995

[2] J. Palicot, "Radio Enineering: From Software radio to Cognitive Radio", Wiley 2011; ISBN: 978-1-84821-296-1

[3] L. Godard, C. Moy, J. Palicot, "An Executable Meta-Model of a Hierarchical and Distributed Architecture Management for the Design of Cognitive Radio Equipments", Annals of Telecommunications special issue on Cognitive Radio, vol. 64, pp.463-482, n°7-8, Aug. 2009

[4] Ettus Research "Products" - accessed 02/04/2012 http://www.ettus.com/products

[5] GNU Radio project - accessed 31/03/2012 http://gnuradio.org/redmine/wiki/gnuradio

[6] P. D. Welch, "The use of Fast Fourier Transform for the estimation of power spectra: a method based on time averaging over short modified periodograms", IEEE Transactions on Audio Electroacoustics, AU-15, 70–73, 1967

[7] Tektronix, "Fundamentals of Real-Time Spectrum Analysis", White paper 2011.

[8] W. Jouini, D. Ernst, C. Moy, J. Palicot, "Upper confidence bound based decision making strategies and dynamic spectrum access", International Conference on Communications, ICC'10, Cape Town, South Africa, 26-29 May 2010

[9] A. Le Naour, O. Goubet, C. Moy, P. Leray, "Spread Spectrum Channel Sounder Implementation with USRP Platforms", SDR'11 Wireless Innovation Conference, Washington DC, USA, 29 November-2 December 2011

A SOFTWARE DEFINED RADIO APPROACH FOR DIGITAL TELEVISION ISDB-T TRANSMITTERS

André L. G. Reis (University of Campinas, Campinas, São Paulo, Brazil; andre.lgr@gmail.com); André F. B. Selva (University of Campinas, Campinas, São Paulo, Brazil; andrefselva@gmail.com); Karlo G. Lenzi (CPqD, Campinas, São Paulo, Brazil; lenzi@decom.fee.unicamp.br); Luis G. P. Meloni (University of Campinas, Campinas, São Paulo, Brazil; meloni@decom.fee.unicamp.br); Sílvio E. Barbin (University of São Paulo, São Paulo, São Paulo, Brazil; barbin@usp.br)

ABSTRACT

Although most countries already have Digital Television (DTV) available, there are several others migrating their infra-structure from analog to digital. One of these countries is Brazil, which intends to accomplish a full switch up to 2016. This initiative is seen as of extreme importance by local authorities since Brazil will host the FIFA World Cup in 2014 and the Olympics Games in 2016, events that will be seen by millions of viewers around the globe. The Brazilian DTV is based on the Japanese ISDB-T standard for the modulation scheme, which is an OFDM system capable of supporting services from mobile to full-definition. Since the relevance of the subject, this paper presents a software defined radio approach to implement a full ISDB-T transmitter using GNU Radio and USRP, covering a complete study from modeling to implementation. This papers aims not only to discuss and clarify several misinterpretations of the ISDB-T standard, but also to present an efficient implementation on a SDR architecture, serving as reference to new developers and as a roadmap for manufactures.

1. INTRODUCTION

The Integrated Services Digital Broadcasting – Terrestrial (ISDB-T) [1] was developed to support different services from mobile to full-definition. Occupying a channel of 6 MHz, it is a versatile and robust system, which integrates several protections against errors that can occur in the transmission.

The most common implementation of the ISDB-T transmitter is based on specialized hardware to perform the steps required to modulate the signal. Our goal here, however, is to present a Software Defined Radio [5] approach to implement such transmitter. SDR platforms are already widely used in scientific field, allowing several researches to work on areas like cognitive radios [10], OFDM modulation [11], among others. A research group in Italy has already successfully implemented the DVB (Digital Video Broadcasting) transmitter using SDR [13].

The main reason to implement a ISDB-T transmitter using a SDR platform is the flexibility offered by the latter. Characteristics of the system can be easily changed and upgrades can be performed in short time, without hardware modification.

This paper describes the ISDB-T Standard, trying to clarify some of its aspects and particularities. Also, it briefly discusses the implementation of the transmitter using Matlab and GNU Radio.

2. ISDB-T STANDARD OVERVIEW

The ISDB-T is an OFDM (Orthogonal Frequency Division Multiplexing) system with 13 segments per frame organized in up to 3 hierarchical layers. It supports 3 operation modes, which defines how many data carriers will be used in each OFDM segment.

The input of the ISDB-T modulator is a BTS (Broadcast Transport Stream) file which contains the TSPs (Transport Stream Packet) used to transport the audio and video data, among other informations. Fig. 1 below shows the processing chain of the modulator.

First we have a shortened Reed-Solomon (204,188) encoder, which includes parity bytes in TSPs, followed by the hierarchical divider, which is responsible to distribute TSPs across each hierarchical layer processing chain. Each hierarchical layer is composed by a sequence of signal processing chain responsible to adapt the TSPs to transmission format. This includes an energy dispersal, bit and byte interleavers, convolutional encoder and mapper.

The final steps include hierarchical layers combiner, OFDM frame adaptation (allocation of data carriers, pilots and Transmission and Multiplexing Configuration Control), IFFT and guard-interval addition. When all these tasks are completed, the signal is ready to be transmitted. It is worth noticing that a multiplex frame (which comes from BTS) is

aligned with a OFDM frame. All these tasks will be detailed in following sections.

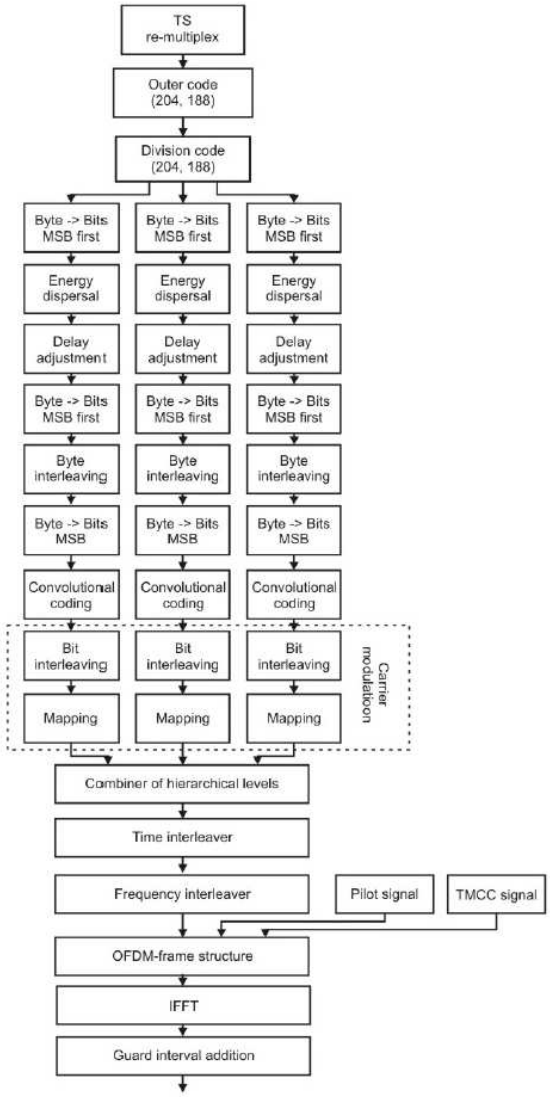


Fig. 1. Block diagram of channel coding [1].

The ISDB-T Standard defines several parameters that are crucial for good understanding of the system. Below we clarify some of those definitions presented in [1].

- 1 OFDM Frame = 204 OFDM Symbols
- 1 OFDM Symbol = 13 OFDM Segments
- 1 OFDM Segment = 1 Data Segment + Pilots + TMCC + AC
- 1 Data Segment = *n* Data Carriers (*n* depends on the operation mode, namely 2k, 4k and 8k)

To illustrate those parameters, in Fig. 2 we see in Frequency Axis a OFDM symbol composed by sub-carriers (Data Carriers + Pilots + TMCC + AC) and in Time Axis we see the propagation of OFDM symbols (204 of these will constitute a OFDM frame).

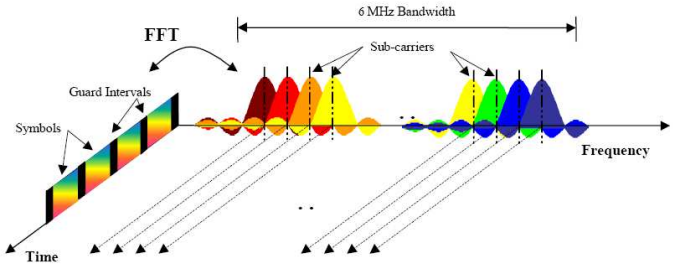


Fig. 2. OFDM Symbol in frequency and time domain [3].

2.1. Reed-Solomon

The Reed-Solomon algorithm is an error-corrector code developed by Irving S. Reed and Gustave Solomon in 1960 [2]. It is capable to correct random errors and burst noise. Due to its code properties, which makes it very robust to errors,, it is largely used in storage medias like DVD and CD, mobile communications, etc.

In the ISDB-T, a shortened version of RS encoder is used, namely RS(204,188). This is obtained by adding 51 null bytes in the beginning of the TSP and applying it to RS(255,239). After encoding, the inserted zeros are removed from the data stream. In the end of this process, the TSP will have 188 data bytes and 16 parity bytes, composing the 204 bytes BTS packet.

2.2. Hierarchical Divider

The BTS file contains all packets of hierarchical layers A, B and C. The Hierarchical Divider is responsible to redirect each TSP to its processing chain, to discard null packets generally used to maintaining a system's constant transmit rate, to store the IIP (ISDB-T Information Packet), which will configure the parameters used to modulate the signal, and to shift the sync byte (0x47) from beginning to end of the packets. More details of its operation are available in [1] and [12].

2.3. Energy Dispersal

This block randomizes the data within the TSP, in order to distribute its energy equally along the packet, reducing the inter symbol interference caused by transmitting repetitive information. Fig. 3 shows the circuit responsible to perform this task. It is a simple shift register with a XOR operation. This polynomial is also used in other OFDM transmission standards such as WiMAX.

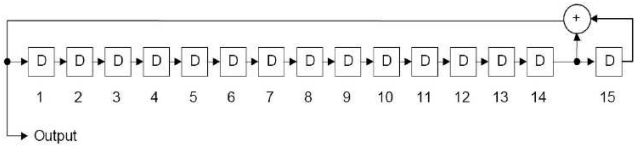


Fig. 3. PRSB-Generating Polynomial and Circuit [1].

It is worth emphasizing that the sync byte (0x47) should not be randomized along with the data bytes.

2.4. Byte Interleaving

This block is responsible of interleaving the data, so that if frequency selective errors occur in the channel, they do not corrupt the data block since it's scattered . The interleaver is composed by 12 paths with variable delays. Fig. 4 shows its scheme.

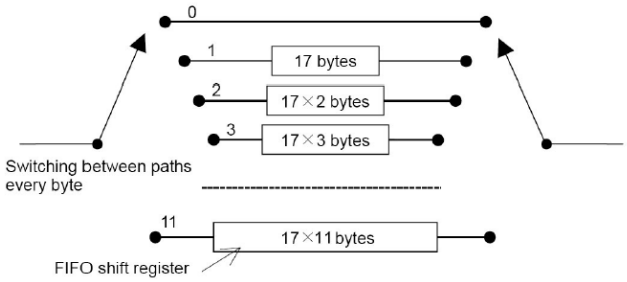


Fig. 4. Byte Interleaving circuit [1].

A delay adjustment is required here, in order to compensate the delay introduced by the circuit above. Details on the amount of delay adjustment required are available in [1].

2.5. Inner Code

This is a very common convolutional encoder with constraint length k = 7. This inner code allows multiple coding rates through bit puncturing patterns. Fig. 5 shows its circuit.

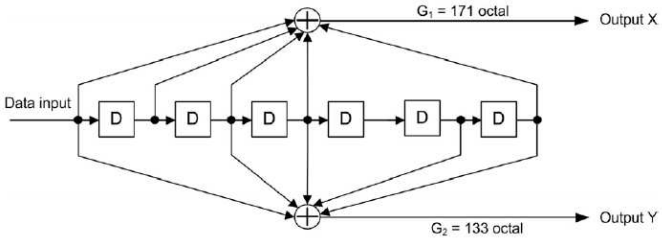


Fig. 5. Coding circuit of a convolutional code with constrain length *k* of 7 and a coding rate of ½ [1].

The puncturing pattern depends on the coding rate used. Table I shows puncturing pattern that should be performed to achieve the coding rates.

TABLE I. INNER-CODE CODING RATES AND TRANSMISSION-SIGNAL SEQUENCE [1]

Coding rate	Puncturing pattern	Transmission-signal sequence
1/2	X : 1 Y : 1	X1, Y1
2/3	X : 1 0 Y : 1 1	X1, Y1, Y2
3/4	X : 1 0 1 Y : 1 1 0	X1, Y1, Y2, X3
5/6	X : 1 0 1 0 1 Y : 1 1 0 1 0	X1, Y1, Y2, X3, Y4, X5
7/8	X : 1 0 0 0 1 0 1 Y : 1 1 1 1 0 1 0	X1, Y1, Y2, Y3, Y4, X5, Y6, X7

2.6. Bit Interleaving

The same concept used in Byte Interleaving is used here. The main difference is that now we should parallelize the data depending on how many bits each modulation scheme uses. For each modulation scheme we will have a different interleaver in terms of shift registers and paths. Fig. 6 shows the bit interleaving for 64QAM, which uses 6 bits per carrier.

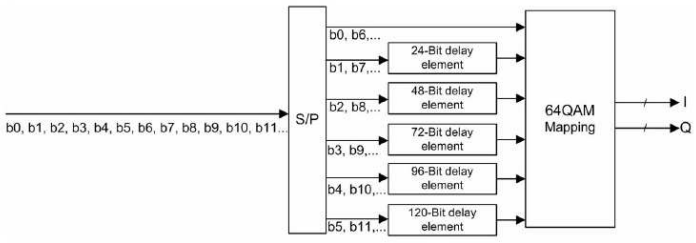


Fig. 6. 64QAM Modulation system diagram [1].

It is worth noticing that it is required a delay adjustment [1], as in Byte Interleaving.

2.7. Carrier Modulation

The ISDB-T supports DQPSK, QPSK, 16QAM and 64QAM. After the bit interleaving, the data is mapped in a point of the constellation. As an example, Fig. 7 shows the constellation of 16QAM modulation scheme.

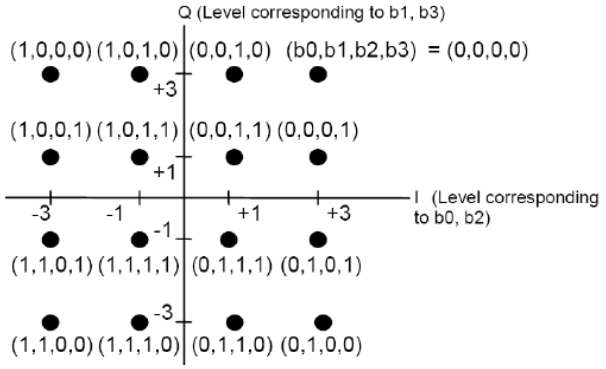


Fig. 7. Constellation of 16QAM [1].

After mapping the data in constellation, a modulation level normalization is required. Table II shows the normalization factor.

TABLE II. MODULATION LEVEL NORMALIZATION [1]

Carrier modulation scheme	Normalization factor
DQPSK and QPSK	$Z/\sqrt{2}$
16QAM	$Z/\sqrt{10}$
64QAM	$Z/\sqrt{42}$

2.8. Data segment configuration

This step just rearranges data into data segments to make easier the allocation of data carriers in OFDM symbols. Fig. 8 below shows the organization of data carriers.

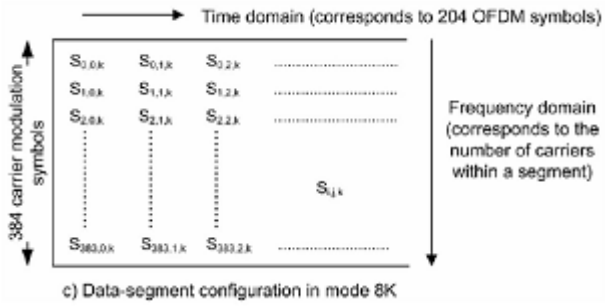


Fig. 8. Data-segment configuration [1].

Note that 96, 192 and 384 are the number of data carriers per OFDM segment used in modes 1, 2 and 3, respectively.

2.9. Hierarchical combiner

The Hierarchical combiner puts all 13 segments together in order to gather enough data to build a OFDM symbol. The sequence of segments is shown by Fig. 9.

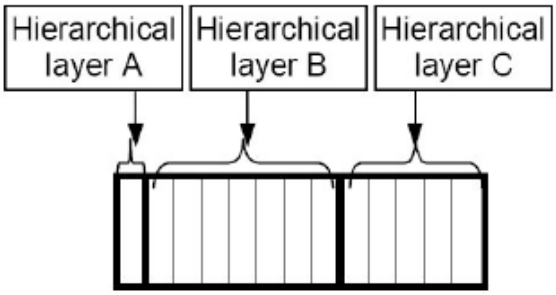


Fig. 9. Segments order [1].

2.10. Time Interleaving

The Time Interleaver operates separately on each of 13 data segments. Fig. 10 and Fig. 11 show the schematic of this block.

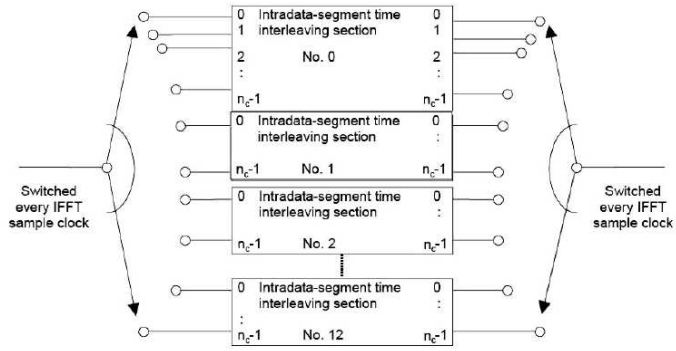


Fig. 10. Time Interleaver.

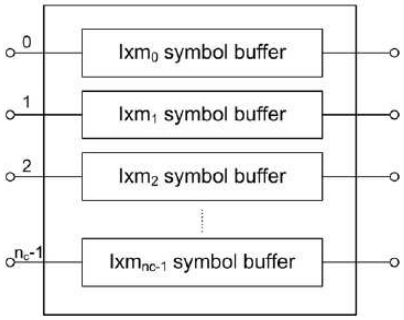


Fig. 11. Configuration of intradata-segment [1].

The parameter I corresponds to the interleave length and m_i is given by

$$m_i = (5i) \bmod 96, \qquad (1)$$

where i is the carrier number, that is, each carrier will have paths with different delays. Also, this step requires a delay adjustment [1], as Bit and Byte Interleaving.

2.11. Frequency Interleaving

This module is divided in three steps: inter-segment interleaving, intra-segment carrier rotation and intra-segment carrier randomizing.

2.11.1. Inter-segment interleaving

This step is performed only on hierarchical layers B and C, since layer A has only one segment. The ISDB-T Standard shows a confusing figure to describe this interleaving, but it is simply a matrix interleaver [9]. Suppose we have *k* segments and *n* data carriers per segment. The interleaving process consists in writing the segments column-wise in a *k* x *n* matrix and reading line-wise, so each line of the matrix is the interleaved segment. Fig. 12. illustrates this process.

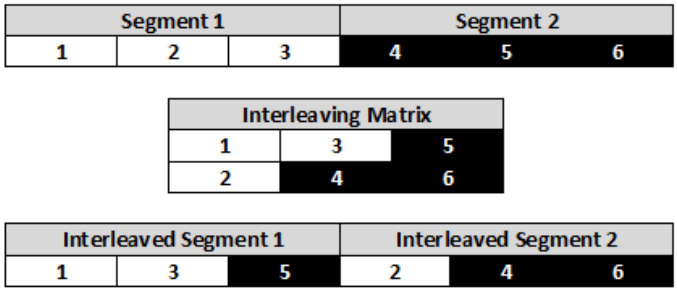
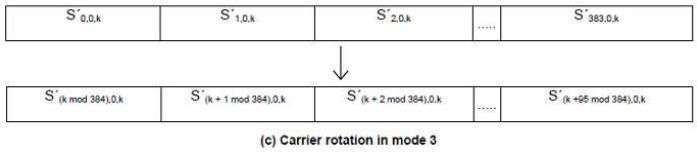


Fig. 12. Matrix Interleaver fictional example for 2 segments with 3 data carriers each.

2.11.2. Intra-segment carrier rotation

The carrier rotation, as its name says, rotates the carriers inside the segment according to the segment number. Fig. 13 shows the process for modes 1, 2 and 3.



NOTE: The symbol $S'_{I,j,k}$ represents the carrier symbol of the k_{th} segment following inter-segment interleaving.

Fig. 13. Carrier rotation [1].

2.11.3. Intra-segment carrier randomizing

This step just randomizes the carriers' position according to pre-defined tables available in ISDB-T Standard [1].

2.12. OFDM Frame Structure

The OFDM frame structure is shown by Fig. 13.

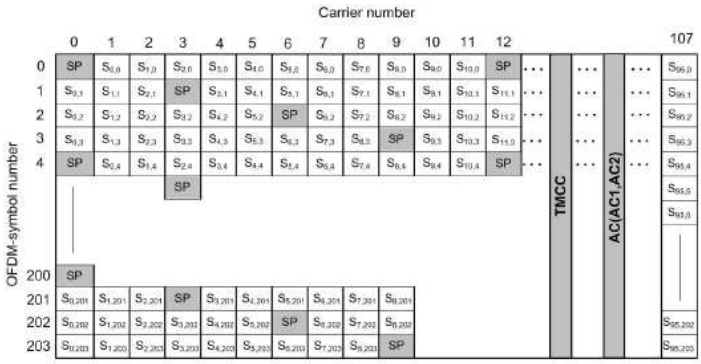


Fig. 13. Configuration of the OFDM frame for synchronous modulation at mode 1 [1].

It is important to notice that Fig. 13. above extends to the right 12 more times, so we have 13 segments per symbol.

It is also necessary to organize the segment as showed in Fig. 14.

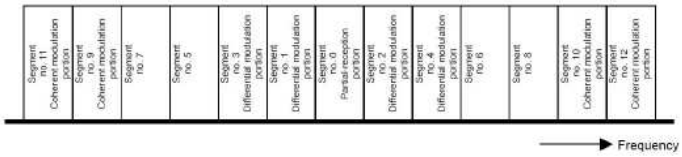


Fig. 14. OFDM segment numbers on the transmission spectrum [1].

To make up the entire transmission spectrum, a continuous carrier with its phase defined by W_i is allocated in right-hand end of the band. This means, each OFDM symbol will have an extra continual pilot in its structure.

2.13. Pilots, TMCC and AC

The SPs (Scattered Pilot), TMCC (Transmission and Multiplexing Configuration Control) and AC (Auxiliary Channel) are dependent on the output W_i of a PRBS-generating circuit, because it defines the modulation of these signals. Fig. 15. shows the circuit that generates de W_i and Table III shows the point on constellation according to W_i.

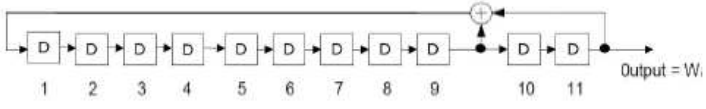


Fig. 15. PRSB-generating circuit [1].

TABLE III. W_i AND MODULATING SIGNAL [1]

W_i value	Modulating-signal amplitude (I,Q)
1	(- 4/3, 0)
0	(+ 4/3, 0)

In short, the SPs are used by the receiver to synchronize the signal, the TMCC contains all parameter necessary to configure the receiver and AC shall be used to transmit additional information on modulating signal-transmission control.

A SP is inserted every 12 carriers in OFDM segments direction. Then its index is shifted 3 positions for each new OFDM symbol. TMCC and AC carriers are fixed and defined in tables available in [1]. More details can be found in the ISDB-T Standard [1].

2.14. IFFT and Guard Interval

Once all processing tasks are completed, it is time to apply the IFFT (Inverse Fast Fourier Transform) to the signal. The configuration of the IFFT is showed in Table IV.

TABLE IV. IFFT PARAMETERS

Sample Frequency	8126984 Hz
Number of points for mode 1	2048
Number of points for mode 2	4096
Number of points for mode 3	8192

It is worth emphasizing that it is necessary to use the zero padding technique [8] in the symbol before applying the IFFT. The zeros should be inserted in the middle of the symbol, so that the total data is the same length of the number of points of the IFFT.

After the IFFT, it is necessary to insert the guard interval to the signal. This step consists in adding to the beginning of the effective symbol a part of its end. Fig. 16 illustrates this process. At this time, the signal is ready to be transmitted.

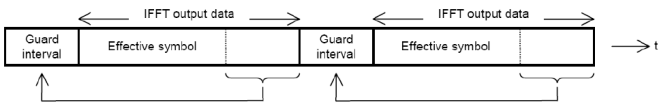


Fig. 16. Guard interval insertion [1].

3. SOFTWARE DEFINED RADIO

The main idea of a SDR (Software Defined Radio) is to transfer the tasks performed by hardware to software. By doing so, the implementation of a communication system becomes more flexible, allowing a radio to chance its operation characteristics during run time, such as modulation, bandwidth, output frequency, etc. Besides flexibility, SDRs have many other advantages over traditional radio implementations. They are more immune to temperature changes and component aging, since they transfer the processing task to digital domain, no longer depending on the precision of analog components. They also offer all advantages that a software environment provides to the development process, like simulation tools and error correction. All of these advantages can reduce time-to-market, since the same hardware platform can be used to implement different radios, just by replacing the software that controls the radio.

By transferring most part of processing task to software, the complexity of modern radios hardware is reduced, limiting it just to the RF front-end implementation. This also implies in high integration, given that many passive and active elements of the radio, once responsible to signal processing and modulation, are eliminated through the use of a processor. All these factors have an effect on the product final cost, which can be reduced.

There are several solutions in the market to implement a SDR system. In this paper, however, we will present a solution propose by Matt Ettus, the USRP (Universal Software Radio Peripheral) [4].

Composed by a motherboard, responsible for the base-band processing and by many RF daughterboards cards, which can be chosen according to the application's needs, the USRP abstracts the system RF front-end. A Ethernet interface establishes the communication between the PC and the USRP radio, allowing data exchange in a full-duplex manner [5].

GNU Radio is an open-source toolbox kit that provides a development environment and some processing blocks that can be used to create software defined radios. The software provides full integration with USRP boards [6]. It provides around a hundred signal-processing blocks and allows developers to design new ones using C++ as programming language. It also contains specific drives to communicate with the USRP.

To implement a ISDB-T modulator, though, it is necessary to create new blocks, since it has some specific processing steps that are not implemented in this package. The creation of blocks is not the scope of this paper, however.

GNU Radio applications are developed using the Python language, where the connection between blocks are made. The blocks are built in C++, for performance issues. A

graphical user interface, called GNU Radio Companion (GRC), is also provided. GRC could be used to simplify the system design, just like Simulink does with projects developed on Matlab.

4. MODELING AND IMPLEMENTATION

A model of the ISDB-T modulator was implemented in Matlab [7] as a reference for the SDR GNU Radio development. Each block of the system was implemented first as a Matlab Function and then, as a GNU Radio signal processing block. Comparisons between GNU Radio and the simulation outputs were made to validate each processing block.

In the simulation, the parallel processing of the three hierarchical layers were saved as states. Each component, when necessary, has a workspace to store its last state, so it is possible to reuse the same components in all 3 hierarchical layers without having data overwrite problem.

The simulator works with one OFDM frame for each iteration, once the multiplex frames that come from BTS are aligned with the OFDM frames. All configuration parameters of the modulator are automatically acquired from IIP (ISDB-T Information Packet) available in the BTS data stream. Each multiplex frame has one of these packets.

Concerning the implementation in GNU Radio, it was necessary to make an in-depth study of how to write blocks in this platform, since the open-source library does not offer all of the processing blocks needed to implement a ISDB-T modulator. Although this is not the scope of this paper, we will present part of the code of the Energy Dispersal in Matlab and GNU Radio (that is, in C++) in order to establish some comparisons between them.

```
% MATLAB CODE
function [tsp_ed] =
isdbt_energy_dispersal(tsp_rs, state ,
reset)

length_tsp_rs = length(tsp_rs) - 1; % The
sync. byte should not be randomized
tsp_syncByte = tsp_rs(end); % Sync Byte
is the last one of the input vector
% Convert data in a stream of bits
...
size  = length(data_in);

if reset == 1, % This value should be
initialized every OFDM frame.
    prbs_state = [1 0 0 1 0 1 0 1 0 0 0 0
0 0 0]; % Initial value of PRBS-gen
else
    load(state);
end
```

```
% randomize
for k=1:size
    fdB   =
bitxor(prbs_state(14),prbs_state(15)); %
calculate the feedback bit
    prbs_state = [fdB prbs_state(1:end-
1)]; % shift right and insert the fd bit
    data_out(k) = bitxor(data_in(k),fdB);
end

% rearrange bit stream in bytes
...
save(state,'prbs_state');
```

```
// C++ CODE
...
int
gr_isdbt_energy_dispersal_ff::work (int
noutput_items,
            gr_vector_const_void_star
&input_items,
            gr_vector_void_star
&output_items)
{
  const int *in = (const int *)
input_items[0];
  int *out = (int *) output_items[0];
  int x = 1000;
  for (int i = 0; i < noutput_items;
i++){
    cont_sinc++;
    if
((cont_sinc>INIBYTESINC)&&(cont_sinc<FIM
BYTESINC)) {
        out[i]=in[i];
        continue;
    }
    else if (cont_sinc==FIMBYTESINC) {
        out[i]=in[i];
        cont_sinc=0;
        continue;
    }
    else {
        out[i] = in[i] ^ prbs(v);
        shifter(v,1,'R');
    }
  }
  return noutput_items;
}
```

Analyzing both codes we see that they do not differ much. Except by syntax differences (Matlab Script and C++), the C++ code seems to be a direct transcription of the Matlab code. Though, in GNU Radio we need to care about input and output size precisely, data rate, etc, what makes the

porting process more complicated than expected. Despite those aspects, the simulation is still a good reference for implementation.

5. DISCUSSIONS

We have seen that the performance of the implementation in GNU Radio is far better than simulations. Maybe some improvements will have to be performed in order to achieve the real-time constraints of ISDB-T system.

Our next steps are toward concluding the implementation in GNU Radio. We are at the final stages of porting the code and then we have to integrate the blocks and perform tests in order to validate the implementation against the model.

6. FUTURE WORK

Currently, we are optimizing our ISDB-T simulator, using techniques such as parallel computing, GPU programming and hybrid implementation (C and Matlab), since the processing of one BTS for a full-seg transmission involves a great deal of data, which in turns make the simulation very slow. At the moment, the simulation time needed to generate 1 frame is around 7 minutes. It began taking 35 minutes. We believe that is possible to reduce the simulation time to 1.

We are also working on the conversion of the model, implemented through scripts to a Simulink model, since Mathworks recently offered support to communicate with USRPs via a specific driver, making possible to use the original model to generate data to the USRP, offering an alternative to the GNU Radio approach. For this reason, it is important to improve the simulation performance so it can emulate a real-time system for the ISDB-T transmission.

In terms of the SDR, it would be interesting to study some way to embed the ISDB-T transmitter within the USRP, so that the PC would be no longer necessary to the system, creating then a stand-alone application that still integrates a flexible hardware. There are indeed in the market some SDR platforms that incorporate a microprocessor, so this idea could become more feasible with such hardware.

7. CONCLUSION

This paper aimed not only to clarify some aspects of the ISDB-T Standard that can be confusing and lead to misinterpretations, but also to briefly discuss the implementation of a ISDB-T transmitter in GNU Radio and Matlab.

Offering great flexibility, SDR platforms show themselves as interesting solution to implement communication systems such as DVB [13] and ISDB-T.

8. REFERENCES

[1] "Digital terrestrial television – transmission system", Associação Brasileira de Normas Técnicas (ABNT), ABNT NBR 15601:2007, Dec. 2007.

[2] I. S. Reed and G. Solomon, "Polynomial code over certain fields", J. Soc. Ind. Appl. Math., 8:300-304, June 1960.

[3] HSCTechnicalWiki, "Orthogonal Frequency Division Multiplexing". [Online]. Available: http://wiki.hsc.com/wiki/Main/OFDM

[4] M. Ettus, "Universal Software Radio Peripheral". [Online]. Available: http://www.ettus.com.

[5] A. F. B. Selva, A. L. G. Reis, R. U. Labsch, K. G. Lenzi, L. G. P. Meloni and S. E. Barbin, "A software-defined radio approach: GNU Radio", 10th International Information and Telecommunication Technologies Conference, 2011, São José, SC. Proceedings of the 10th International Information and Telecommunication Technologies Conference, 2011. v. 1. p. 48-52.

[6] GNU Radio. [Online]. Available: http://gnuradio.org.

[7] Mathworks Matlab. [Online]. Available: http://www.mathworks.com/products/matlab/

[8] R. V. Van Nee and R. Prasad, "OFDM for Wireless Multimedia Communications Norwood", MA: Artech House, Jan. 2000.

[9] M. C. Paiva, "Uma implementação em software do subsistema de transmissão do padrão ISDB-TB". Master Thesis, INATEL, 2010.

[10] Z. Yan, Z. Ma, H. Cao, G. Li, and W. Wang, "Spectrum sensing, access and coexistence testbed for cognitive radio using USRP", 4th IEEE International Conference on Circuits and Systems for Communications, 2008. ICCSC 2008, 26-28 May 2008.

[11] A. Marwanto, M. A. Sarijari, N. Fisal, S. K. S. Yusof and R. A. Rashid, "Experimental study of OFDM implementation utilizing GNU Radio and USRP - SDR," Communications (MICC), 2009 IEEE 9th Malaysia International Conference, 15-17 Dec. 2009.

[12] "Digital terrestrial television – Operational guideline part 1: transmission system – guideline for ABNT NBR 15601:2007 implementation", Associação Brasileira de Normas Técnicas (ABNT), ABNT NBR 15608-1:2008, Aug. 2008.

[13] V. Pellegrini, G. Bacci, and M. Luise, "Soft-DVB, a Fully Software, GNURadio Based ETSI DVB-T Modulator," in 5th Karlsruhe Workshop on Software Radios, 2008.

Vehicle Power Line Communication (VPLC) implementation with USRP2 Platforms

Philippe Tanguy, Fabienne Nouvel, Clovis Lunel

prenom.nom@insa-rennes.fr

IETR-INSA –Rennes – France

Abstract **- This paper deals with the implementation of an embedded power line communication system for vehicle (VPLC) using USRP platforms and GNU Radio environment. This platform allows a modular design of VPLC. Many configurations of the PHY layer of can be tested in a real PLC environment, without developing them on a specific board. In conjunction with USRPs platform and daughter cards, the signal processing performed in software is transmitted over a real channel. Several frequency bands can be tested by using different daughter cards without changing the signal processing.**

Index terms – USRP, OFDM, PLC

I. INTRODUCTION

Software-defined radio (SDR) [1] has been defined to help the wireless operators to maximise their investments in multiple mobile standards and base stations. With such approach, many of the drawbacks involved with a classical approach (specify FPGA, DSP designs, RF designs) fall down as one can re-use sub-parts of an existing radio system, and just adapt its operation through code reprogramming.
Similar approach can be adopted for wired communication, and more specifically for power line communication (PLC) communication, the radio link is equivalent to the wired PLC.
Today the car manufacturers have to face with an increase of electronic nodes (ECU) connected each other. The ECUs already use networks like CAN and Flexray, but the number of specific wired always increase. One solution to reduce the amount of wires would be to use the PLC technology that is currently being developed for domestic networks to transmit information or at least some of it, over the 12V power distribution system found in cars. However, the PHY layers may be optimised according to the channel states and embedded environment.
Many projects have investigated SDR and cognitive radio using flexible platforms [2]. SDR platforms provide flexibility, reducing the amount of time it takes to develop and update communication systems. With software update, many communication systems can be achieved using the same hardware. SDR technology has first focus on wireless communications. Implementing the new ideas of cognitive radio on inexpensive and known hardware will allow for practical tests. Using software generated standard waveforms "on the fly" reconfiguration will allow radios to fit into rapid prototyping platform.
Today, wired systems are concerned by the issue of spectrum management and aim to provide a single device to network over any supported wire types and more specifically in home. Thanks to the advance made in access technologies, the triple play services become widespread [3]. For example, the recent VDSL2 technology provides high data rate and takes advantages of the higher frequency bands up to 30 MHz. At the same time, the indoor PLC standards take advantage of the power grid in the [2-30] MHz band. VDSL2 and PLC technologies are complementary to provide high data rates by using the telephone copper pair and the power cables.
Nevertheless, the band plans used by both technologies overlap: [0.138 - 30] MHz for VDSL2 and [1.8 - 30] MHz for PLC. Methods to improve the co-existence of different wired standards become necessary.
In addition, the ITU [4][5] has begun to address this "SDR wired" and has specified a PHY layer and the architecture of G.hn, the common name for a home network technology family. A single G.hn device will be able to network over any supported interfaces in home. The advantages are lower costs, lower equipment and lower time development.

In our paper, SDR can be considered to be a "wired" communication system, where some of its functional components, such as modulations, coding, synchronisation, etc., are generated in software and implemented on open hardware. This makes it possible to configure the signal according to the requirements of the application and the characteristics of the communication channel. The software generated signal is then applied to USRP [6] platforms that are connected to a host computer through USB or Ethernet connections. To help the designer, the GNU Radio Companion is a friendly environment for the development of baseband processing thanks to a set of pre-defined block sets, which enables to build a complete radio chain without developing specific code (most of the time).
Ettus Research (part of National Instrument) provides the USRP2 hardware platforms supporting any carrier frequency between DC and 5.9 GHz depending on the daughter board. Then RF to baseband conversion sub-system is ready for prototyping (except specific requirements in terms of transmitting power, depending on each application).
In addition, the GNU Radio Companion [7] is a friendly environment for the development of baseband processing thanks to a set of pre-defined block sets, which enables to build a complete radio chain without developing specific code (most of the time). More recently, MATLAB and Simulink connect to the USRP to provide a radio-in-the-loop design and modelling environment. With the support package, Communications System Toolbox and a USRP radio, it is possible to design and verify practical SDR systems rapidly. Furthermore, comparison between simulation and real environment can be performed using the same developed

algorithms and codes. In this case, the GNU Radio open-source software is not required to use this support package.
In [8] SDR-GNU Radio is used to implement a power line data transmission link between a motor and an inverter, but at a low data rate. In our study, SDR-GNU Radio is used to implement power line communication system in a vehicle (VPLC) for high data rates and non-safety applications. The orthogonal frequency division multiplex (OFDM) is used to efficiency fight against the fadings. The system operates at the center frequency of 12,5 MHz (VPLC) thanks to the daughter boards.

The reminder of the paper is organized as follows. In Section II, the communication system and channel are studied. The OFDM has been chosen for the VPLC communication as for domestic PLC based on different previous works. Using the measurements, we can now optimize some parameters like the CP length, the FFT size, etc. The main idea is to optimize the PHY parameters without developing all the TX/RX boards. Another objective is to be able to test the VPLC modem parameters on other channel (wireless has been tested), by using another RF interface while keeping the same PHY layers. Section III concentrates on GNU Radio and implementation of the data transmission link. The VPLC testbed is based on two USRP2 cards with the daughter board's LFTX and LFRX. Using these daughter boards, the possible operation frequency range is very modular (from DC to 5.9 GHz). Section IV details the results we obtained, both on indoor PLC and VPLC. Section V concludes the paper and perspectives are given.

II. COMMUNICATION SYSTEM AND PHY LAYER

A.Transmission scheme

The orthogonal frequency division modulation or Discrete Multi-Tone (OFDM or DMT) has been chosen for the VPLC communication as for domestic PLC. Several standards with different kinds of multi-carrier modulations like HPAV and HD-PLC [9] are already proposed and investigated. In the previous paper [10][11], different measurements have been performed and have shown the OFDM modulation combined with bit loading and equalization is convenient for VPLC. Those standards have been applied in car to measure data throughput on DC channels [10].
However, the indoor and car channels are different. We need to optimize some parameters like the CP length, the FFT size, and other PHY parameters. The main idea in this study is to optimize the PHY parameters without developing all the hardware of the modem. Another objective is to be able to test the VPLC modem parameters on other channels, by using another RF frontend.

The transmitted OFDM waveform can be expressed as:

$$s(t) = \frac{1}{\sqrt{N_p}} \sum_{m=0}^{N_p-1} R_{eal}\left\{c_m \prod(t) e^{2.j\pi F_m t}\right\} \quad (1)$$

With Np= M/2, M the FFT size, cm the complex symbol on sub-carrier Fm. The sub-carrier spacing Δ_{OFDM} is defined as $1/T_{OFDM}$ with T_{OFDM} the OFDM symbol duration. In order to generate a real OFDM signal, the hermitian symmetry is first applied before the IFFT of size M as depicted in Figure 2. As the indoor and vehicle channels are not completely similar, the CP and FFT size will be optimized in the reminder of the paper.

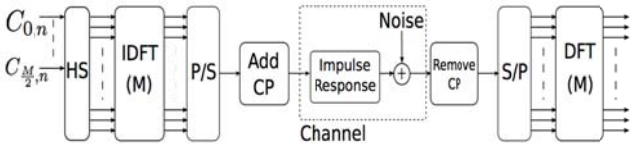


Figure 2. OFDM PLC transmitter and receiver

B.PHY layer specifications

All the PHY parameters have been defined according to the previous channel measurements carried out on different vehicles and described in details in [11]. For the embedded tests, we have used a Peugeot 407SW gasoline.
If we focus on PLC, results in [11] have shown it is possible to reduce the FFT size compared to the HPAV standard, as the coherence bandwidth is about 500 KHz, two times less than in indoor channel. Concerning the CP length, as the mean delay spread is about 135 ns, it can be shorter than the HPAV CP duration of 5.56 us. After software simulations using channel measurements, it is interesting to test different combinations in other VPLC environments like boat or aircraft. By using the VPLC demonstrator, it will be possible to manage rapidly different configurations. The parameters of the VPLC system are summarized in TAB I.

TABLE I. VPLC PARAMETERS

Parameters	**Values**
Transmission bandwidth/sampling rate	DC-25 MHz/25 MHz
M FFT size	1024/512/512
Subcarrier spacing Δ_{OFDM}(KHz)	24.414/48.428
Modulation per carrier(according to bit loading)	BPSK/ QPSK/16QAM/256 QAM
Number of used carriers	412/207/207
CP length(samples)	139/60/20

Theses parameters have been first simulated thanks to MATLAB and with a PLC channel model. In the next step, we will convert the simulations in reality, using the same PHY parameters and using a real channel.

The OFDM symbols are generated and organized in frames. However, it is necessary to generate a synchronization symbol in order to detect the beginning of the useful signal.

C.Time and frequency synchronization

The frame format illustrated in Figure 3 is based on (DMTN+1) OFDM/DMT symbols with a preamble for the time synchronization. Before the frame begins, null symbols are used to separate the different frames. The time synchronization symbol includes a pseudo noise sequence with good autocorrelation property of length equal to (L=M/4) and the CP. The synchronization symbol includes consequently four sequences. Then all the DMT symbols of the file are transmitted. The value of DMTN is configured according to the performances we want to test, with a default tvalue of 20. The synchronization symbol is organized as:

$$[\ P\ P\ -P\ -P\ ] \quad (2)$$

With P the IFFT of the pseudo random sequence of size L.

At the receiver, the synchronization is performed on the first symbol using the Minn & Bhargava algorithm [12] in the time domain. The four sequences of the synchronization symbol are correlated and generate a peak when the sliding window of size M includes these four sequences and zero everywhere else.

The correlation function is given by:

$$R(n) = \sum_{k=0}^{1} \sum_{m=0}^{N_p - 1} r(n + 2Lk + m) r(n + 2Lk + m + L) \quad (3)$$

Figure 4 illustrates the peak when the synchronization is detected. On the green curve, we observe the different peaks given by (3).

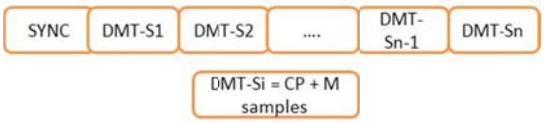


Figure 3. Frame - PHY Layer

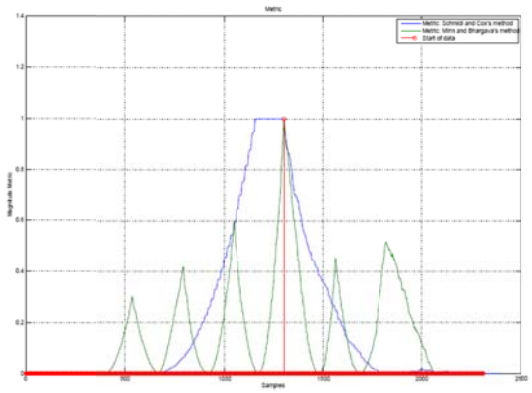


Figure 4 Synchronization detection

The received data begins with the red peak. The blue curve illustrates the algorithm proposed by Schmidl and Cox, which is weaker. This approach is simple and robust enough for our experiments.

This frame structure will be used both for the channel sounding and the data transmission. The length of the frame could be different.

III. USRP2 PLATFORM FOR VPLC

The VPLC platform is based on USRP2 boards combined with daughter boards [6]. Using these daughter boards, the possible operation frequency range is very modular (from DC up to 5.9 GHz). The mother board includes ADC, DAC, a FPGA (Xilinx Spartan XC3S2000) and a Gigabit-Ethernet interface used to transmit data or configuration to the USRP2 board. The two slots of the board are for the front-end daughter boards. As we can see on Figure 5, the two 14-bit ADC and the two 16-bit DAC are independent and can work up to 100 Msps.

For the PLC interface, we use the LFTX and LFRX cards because they allow transmitting and receiving signal from DC up to 30 MHz. These daughter boards include differential amplifiers and low pass filters for antialiasing.

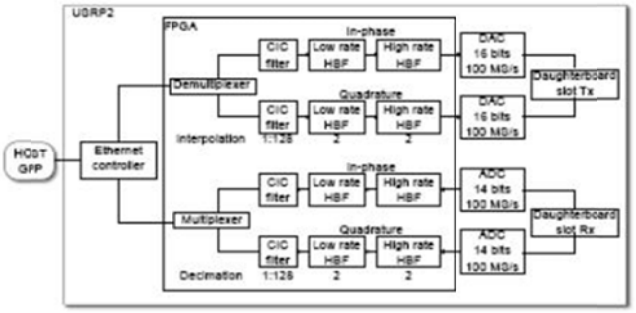


Figure 5. USRP2 board (from [2])

On the transmission side, the real baseband signal is generated thanks to MATLAB and saved in a file. It is then transferred through the Ethernet link to the USRP2 board, converted into analog by one of the two DAC converters and then transmitted over the PLC or RF channel. A passive coupler provides isolation between the output of the card and the PLC link. Both I and Q signal components are transferred to the USRP2 board and can be considered as independent. It is then possible to consider several configurations: real baseband signal (Q =0), complex baseband signal (I and Q), real IF-frequency signal.

At the receiver, the analog signal is translated to baseband and then digital converted. Samples transmitted by the Ethernet link are stored on the workstation. They will then be processed "off line". As above, the ADC conversion paths are independent but must be used consistent with the transmission part.

The maximum bandwidth of the OFDM signal depends on the interpolation and decimation factors N of the ADC and DAC. The sampling rate Fs of the ADC/DAC (Fs at 100MHz) is divided by N (4 to 512) and results in the output sample rate of the data sent from/to the host.

TAB II illustrates some PLC configurations of the USRP2 boards we can use for the transmission of the software generated signal.

TABLE II. USRP2 PLC CONFIGURATIONS

Parameters	Configurations
N =4, I channel	real signal, bandwidth = 12, 5 MHZ
N=4, I& Q complex channels	complex signal , bandwidth = 25 MHZ, external I&Q mixer
N=4, I& Q , two independent real channels	real signal , bandwidth = 12,5 MHZ, two independent transmissions , multiple outputs
N=512, I channel	real signal , bandwidth = 87,5 KHZ, for narrow bandwidth
N=512, I& Q complex channels	complex signal , bandwidth = 195 KHZ , two independent narrow bands transmission

IV. EXPERIMENTS AND RESULTS

The VPLC platform has been tested first for PLC transmissions, both in indoor PLC and VPLC environments. Two USRP2 board are used, one for the transmitter and one for the receiver, allowing a point to point transmission. Thanks to the GNU-Radio interface, we can change easily the configurations as presented in TAB II and adjust the transmitted power. Another important benefit is the option to probe and analyze the signal in the processing chain.
We have first tested our system overs the 220v electrical power. The Tx and Rx USRP2 boards were linked to the 220V power network through couplers plugged on the same multi-outlets. The aim of this test was to prove the concept of SDR over wires. We will describe the VPLC experiments below.

A. VPLC experiments

The two USRP2 boards are arranged in the vehicle according to Figure 5 in points G, F, H and D. We have considered the GH, GF and HD links as in [10]. We will focus on link GF.

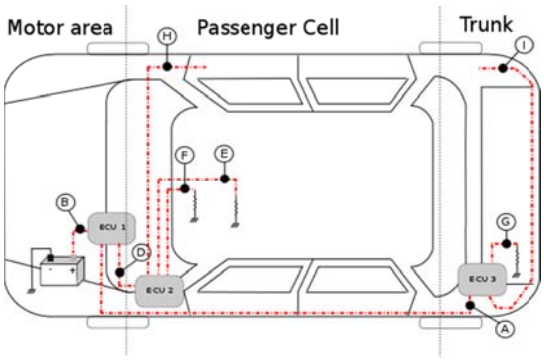


Figure 5. TX/RX VPLC platform

We perform point to point transmission, like from G point to H point. Two scenarios are considered: motor OFF and motor ON. They have been studied in the previous paper [10] with HPAV modified modems.

B. VPLC results

In Figure 6, we can observe the spectrum of the LFTX daughterboard output with a DSP of -80dBm/Hz. The bandwidth [2-12.5] MHz is used. We can observe there are notches compared to HPAV standard.

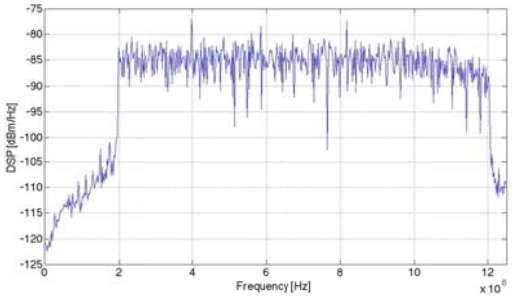


Figure 6. LFTX output spectrum

Before using the VPLC platform for BER measurements, it can be used first as a channel sounder. A sounder enables to point out all the propagation problems and will allow us to optimize the PHY parameters. We first observe the signal to noise ratio for each sub-carrier in the [2-12.5]MHz bandwidth. As the channel is considered as time invariant, the bit loading algorithm will allocate the bits on each sub-carriers at the TX side. Taking into account a referred bit error rate of 10^{-3}, a transmitted power of -80dBm/Hz, the channel capacity definition given in [13], the parameters given in TABLE I and the frame defined in Figure 3, we can achieve a data rate of about 30 Mbps/s.
Then, we performed data transmission according to the frame defined in Figure 3. In this frame, the first two DMT-S1 and DMT-S2 are the estimation symbols we use to equalize the received signal. A ZF equalizer at the receiver compensates the phase rotation due to timing error of the FFT window. However, as the PLC channel is noisy, especially when the motor is ON, the DMT symbols will be corrupted by the impulsive noises and results in false frame detection and bit errors. The SNR is also lower. In this case, the data rate falls down to 7 Mbps. Figure 7 illustrates the received signal during OFDM transmission when motor is ON.

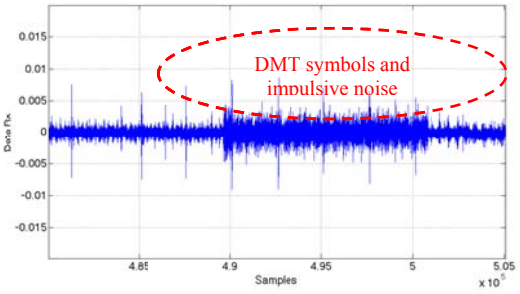


Figure 7. DMT transmission with impulsive noise

With these first results, it is possible to analyze the different combinations proposed in TABLE I.

C. SDR implementation aspects

The VPLC paltform can be used to experiment different PHY parameters using the same hardware.

We have first focus on the CP. During the simulation, we have analyse the CP length according to the link and the FFT size. Results in figure 8 show that the there is an optimum value for the CP according to the FFT size and according to the SNR. It is not necessary to extend the CP with the FFT size. Furthermore, we can adjust the CP according to the point to point communication. Three FFT sizes (256,512,1024) have been experimentend on the VPLC board. We can then generate DMT symbols and frames optimized to the link.

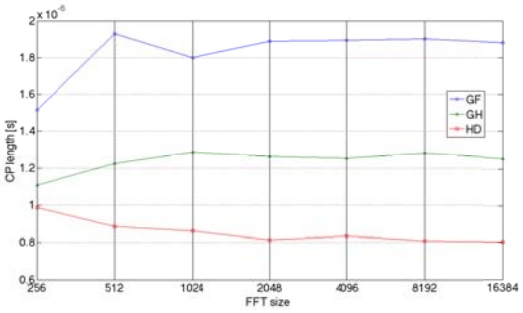


Figure 8. CP length according to FFT length

Additionnaly, the number of used sub-carriers can be modified by introducing notches in sofware. Notches can be introduced when the noise level is very important resulting in no bits allocation. The can be introduced to reduce the transmission bandwidth (upper or lower part) without modifying the analog part. In Section II, we have shown the PHY frame. In this one, it is possible to extend the number of DMT data symbols if the channel is not noisy. If the channel does not change in time, we can estimate it less often and it leads to an increased data rate. Once again, the parameters are software modified. We see that this with SDR approach multiple configurations can be tested in a real system rapidly using the same board.

If we focus on the analog part, it is possible to extend the bandwidth transmission up to 25 MHz by introducing external mixer linked to the outputs (I and Q) of the daughter boards. This work is currently in progress. Figure 9 illustrates the VPLC platform with the software and generic hardware parts.

However, the VPLC is not real time as the received signal is processed “off-line” with MATLAB. It could be possible tu use the GNU-radio interface both to generate the waveform and to configure the board.

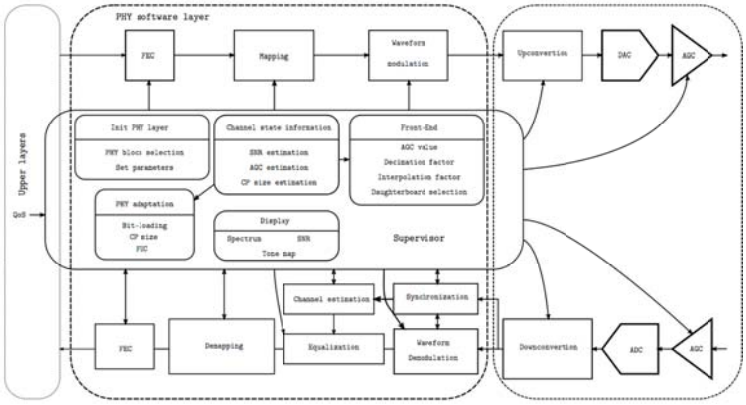


Figure 9. VPLC platform

V. CONCLUSION

We have studied a VPLC platform with USRP2 boards for PLC vehicle channels. Different experimental measurements show the entire possibilities offered by this kind of rapid prototyping. We performed PLC transmissions on real in-vehicle channels. In future work, the bandwidth could be extended and the software waveform generation will be enhanced in order to perform test in real time.

VI. REFERENCES

[1] J. Mitola, “The Software Radio Architecture,” *IEEE Communications Magazine,* vol. 33, n°5, May 95

[2] Michael Joseph Leferman. Rapid prototoyping interface for software defined radio experimentation. Master’s thesis, Worcester polytechnic institute, 2010.

[3] Praho, B.; Tlich, M.; Moulin, F.; Zeddam, A.; Nouvel, F.; , "PLC coupling effect on VDSL2," Power Line Communications and Its Applications (ISPLC), 2011 IEEE International Symposium on , vol., no., pp.317-322, 3-6 April 2011

[4] ITU http://www.itu.int/newsroom/press_releases/2009/46.html

[5] Home Grid Forum [Online]. Available : http://www.homegridforum.org/

[6] “Ettus research, llc.” [Online]. Available: http://www.ettus.com/

[7] “Gnuradio project.” [Online]. Available: http://gnuradio.org/

[8] A. Pinomaa, H. Baumgartner, J. Ahola, and A. Kosonen, “Utilization of software-defined radio in power line communication between motor and frequency converter,” in Power Line Communications and Its Applications (ISPLC),

2010 IEEE International Symposium on, march 2010, pp. 172 – 177.

[9] “Hd-plc standard”. [Online]. Available: http://www.hd-plc.org/

[10] P. Tanguy, F. Nouvel, and P. Maziéro, “Power line communication standards for in-vehicule networks,” in Intelligent Transport Systems Telecommunications,(ITST), 2009 9th International Conference on, 20-22 2009, pp. 533 –537.

[11] P. Tanguy, F. Nouvel , "In-Vehicle PLC Simulator based on Channel Measurements", in Intelligent Transport Systems Telecommunications ITST2010, 6 pages, 9-11 novembre 2010, Kyoto, Japon.

[12] Hlaing Minn, V.K Bhargava, "A robust timing and frequency synchronization for OFDM systems", IEEE Wireless Transactions on, vol. 2, no. 4, pp. 822 –839, july 2003.

[13] Hao Lin and P. Siohan. Capacity analysis for indoor plc using different multi-carrier modulation schemes. Power Delivery, IEEE Transactions on, 25(1) :113 –124, jan. 2010

IEEE 802.15.4 transceiver for the 868/915 MHz band using Software Defined Radio

Rafik Zitouni‡ †, Stefan Ataman‡, Marie Mathian‡ and Laurent George‡ †
‡ ECE Paris-LACSC Laboratory
37 Quai de Grenelle, 75015, Paris, France
† LISSI / UPEC
120, rue Paul Armangot
94400 Vitry S/Seine, France
Email: {zitouni, ataman, mathian, lgeorge}@ece.fr

***Abstract*—This paper reports an implementation of the PHY specifications of the IEEE 802.15.4 standard for the frequency band 868/915 MHz on a Software Defined Radio (SDR) platform. This standard is defined for low power, low data rate and low cost wireless networks. These specifications are used by the Zigbee technology for various applications such as home automation, industry monitoring or medical surveillance. Several hardware PHY 868/915 MHz band IEEE 802.15.4 transceiver implementations have been already reported on ASIC and FPGA [1] [2]. SDR offers one possibility to realize a transceiver with high flexibility and reconfigurability [3]. The whole –transmitter and receiver– chain has been defined in software using the GNU Radio software project [4] and the USRP (Universal Software Radio Peripheral) platform from Ettus Research [5]. Two new blocks have been added to the GNU Radio project, one for the Direct Sequence Spread Spectrum and the second for the reconstruction of the packets. The experimentations have been performed in a noisy environment and the PER, BER and SNR have been computed. The obtained results are coherent with what can be expected from the theory.**

***Index Terms*—Wireless communications, Software Defined Radio, IEEE 802.15.4, GNU Radio.**

I. Introduction

Most of the standards and protocols of lower layers of wireless transmissions (AM, FM, IEEE 802.11, IEEE 802.15.1, IEEE 802.15.4. etc.) are mainly implemented in hardware (HW). This lack of reconfigurability makes the adaptation to varying radio resources difficult, especially when multiple standards need to be often switched in order to take advantage of the scarce radio resources available. The purpose of Software Defined Radio (SDR) is to avoid these drawbacks of traditional wireless communications and replace the hardware equipment by software. The huge advantage of SDR platform lies in its flexibility, its multi-functionality and its low development cost. The reconfigurability of the platform ensures the reusability of the hardware [3], thus minimizing the design complexity of new RF terminals.

The ideal SDR allows the analog-to-digital (ADC) and digital-to-analog (DAC) conversion to be as close as possible to the antenna [6], eliminating the need of high-frequency radio subsystems. Subsequently, the CPU executes the software (SW) subsystem of the SDR, all signal processing operations are accomplished by SW. Unfortunately, today's technology is neither cost-effective for direct ADC conversion from the antenna nor enough power full to compute GSPS (Giga Samples-per-Second) in real-time. Therefore, the typical SDR platform available today uses HW high-frequency radio front-end, the SDR part being implemented in the baseband only. The HW supporting the SDR platform is typically based on FPGAs or DSPs (Digital Signal Processors) [7].

GNU Radio [8] and OSSIE [9] (Open-Source Software Communication Architecture Implementation Embedded) are the two open source software subsystems for the USRP (Universal Software Radio Peripheral) SDR from Ettus Research [5]. The USRP HW is available in different versions. In our implementation we used the USRP1 HW, featuring a sampling rate of 128 MSPS (Mega Samples-per-Second) for the transmitter and 64 MSPS for the receiver. By addition of different daughter-boards, the baseband signal can be transposed in frequency bands up to 6000 MHz. The USRP1 HW platform proves to be also cost-effective, compared to its competitors (Microsoft's SORA and Datasoft's Typhoon).

The IEEE 802.15.4 [10] standard defines the physical and link layers for low-rate Wireless Personnel Area Networks (LR-WPAN), used in wireless sensor networks applications with strong energy consumption constraints. The physical layer comprises three principal frequency bands allowing 49 channels: 16 channels in the 2450 MHz for the ISM (Industrial Scientific Medical) band, 30 for North America and 3 channels in the 868 MHz band for Europe [10]. The band of 2450 MHz operates at law data rates of 250 kb/s while the bands of 915 MHz and 868 MHz operate at 40 kb/s and 20 kb/s respectively.

A number of hardware implementations of the IEEE 802.15.4 have been reported on ASICs or FPGAs [1], [2], but they do not allow us to control the flexibility and the ability of all software stack layers. The first software implementation of the IEEE 802.15.4 using the GNURadio environment for the 2450 MHz band was reported in [11]. In wireless sensor networks, the transceiver in the 868/915 MHz band is more suitable when low data rate transmission are used between sensor nodes. Furthermore it presents a longer range than that of the 2450 MHz band for a given link budget. The objective of our work is to implement the specifications of the IEEE 802.15.4 standard for the 868/915 MHz band, which is not

yet reported in the literature.

Our software transceiver was developed by closely following the IEEE 802.15.4 specifications for the 868/915 MHz bands. The implementation is similar to the one of 2450 MHz band presented in [12], [11]. To evaluate the transmitter/receiver performances, the BER (Bit Error Rate) and SNR (Signal-to-Noise Ratio) have been computed by changing the input power signal at the transmitter.

The rest of the paper is organized as follows. Section II presents a description of the SDR platform used. In Section III, we present the description of the developed transmitter/receiver chain. Section IV discusses the experimentations and the obtained results. Finally, in Section V we formulate some concluding remarks.

II. USRP and GNU Radio

In the following two subsections we describe briefly the USRP1 HW [5], used in our implementation as well as the GNU Radio [4] toolkit.

A. Universal Software Radio Peripheral

The USRP1 HW consists of a motherboard and optional add-on RF daughterboards. It is connected to a host computer via USB 2.0. The USRP's motherboard supports up to four daughterboards: two for transmission (TX) and two for reception (RX). The motherboard has four 12-bits ADCs (with a maximum sampling rate of 64 MSPS), four 14-bit DACs (with a maximum conversion rate of 128 MSPS), and an Altera FPGA for simple but high-speed operations such as up-conversion, down-conversion, interpolation, and decimation [5]. The ADCs and DACs allow us to receive baseband signals up to 32 MHz and are able to generate baseband signals up to 50 MHz. Unfortunately, the USB tunnel limits these performances to 8 MHz. The USRP1 provides buffer in both the USB controller and the FPGA at 2 KB and 4 KB respectively. Fig. 1 depicts the USRP1 blocks from the motherboard.

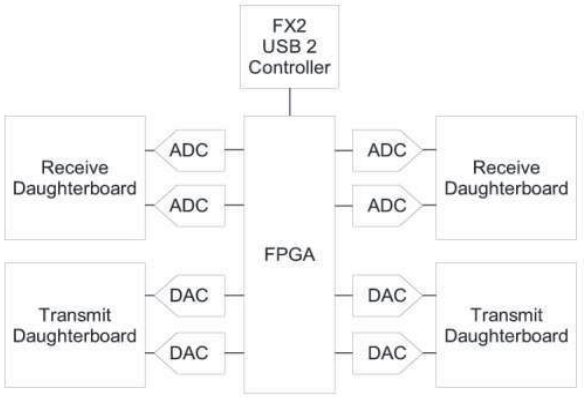


Fig. 1. USRP1 block diagram [5]

B. GNU Radio

GNU Radio is an open source project toolkit for building software radios that run on host computers [4]. It provides signal processing blocks for modulation, demodulation, filtering and various Input/Output operations. New blocks can be easily added to the toolkit. The software radio platform is created by connecting these blocks to form a *flowgraph*. The blocks are written in C++ and they are connected through a Python script. The Verilog HDL layer is dedicated to configure the FPGA.

The advantage of Python in connecting these processing blocks is that it allows the data flow to be at maximum rate, without being interpreted. The integration of the C++ blocks into the scripting language is provided by the SWIG (Simplified Wrapper and Interface Generator), which is an interface compiler. Many signal processing blocks are available to the GNU Radio community to facilitate the development. To create a flow graph we can proceed by the graphical interface called gnuradio-companion or directly through the python code. The C++ blocks are described by the XML code to facilitate the use and the visibility of the block chains, the XML is interpreted to the python code by the cheetah tools[1] . In Fig. 2 we depict the programming language layers of the GNU Radio.

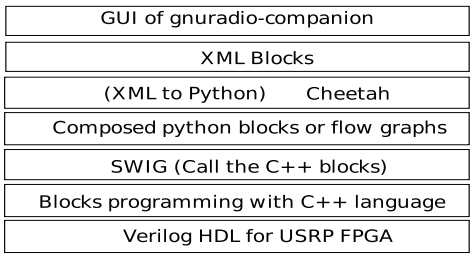


Fig. 2. Software layers of the GNU Radio

III. Transceiver Description

The IEEE 802.15.4 [10] standard is the definition of wireless physical (PHY) and medium access control (MAC) protocols for low-data rate and low power applications. It specifies two families of bands: the first one is centered at 868 and 915 MHz with 20 and 40 kbps, the second one at 2450 MHz with 250 kbps.

The specifications from [10] define the use of different modulation techniques and data rate for the specified channels. The D-BPSK (Differential Binary Phase Shift Keying) is one of the modulation techniques used in the 915/868 MHz. The symbol spreading is the Direct Sequence Spread Spectrum (DSSS), in which each symbol is represented by a Pseudo Noise sequence of 15 chips. The chips are modulated/demodulated by the D-BPSK encoding/decoding at rates of 300 kchips/s and 600 kchips/s for the 868 MHz and 915 MHz bands respectively.

[1] http://www.cheetahtemplate.org/

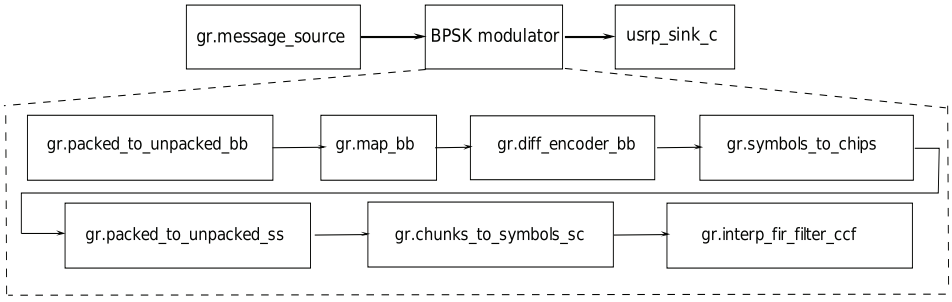


Fig. 3. Transmitter flow graph

A. Transmitter

Our transmitter comprises eight processing blocks, as depicted in Fig. 3. The definition of the packet messages is based on that of the IEEE 802.15.4 standard. The packet format is detailed in Fig. 4. At the output of the transmitter, the maximum packet size is 133 bytes. Due to the USB 2 tunnel, the packet size should be a multiple of 128 samples, therefore, zero padding with the x/00 (representing the "*NUL*" character) is performed. The number of padded bytes is conditioned by the parameter called $Byte\_Modulus$ which depends on the sampling rate and on the number of bits per symbol. The $Byte\_Modulus$ is given by:

$$Byte\_Modulus = \text{LCM}\left(\frac{128\ \text{MSPS}}{8\ \text{MSPS}}, sps\right) \cdot \left(\frac{bps}{sps}\right) \quad (1)$$

where

- 128 MSPS – DAC sampling rate of the USRP1
- 8 MSPS – Sampling rate of the USB tunnel
- sps – Number of samples per symbol
- bps – Number of bits per symbol
- LCM – Lowest Common Multiple of 16 MSPS and sps

To avoid padding and to get the same fields as in the IEEE 802.15.4 specifications, the packet size is set equal to 130 bytes. This size is obtained by reducing the address information field $AddressInf$. Moreover, a 16-bit CRC (Cyclic Redundancy Check) is attached to the packet payload, allowing the receiver to calculate the PER (Packet Error Rate).

Input bits	**Chip values** (c0 c1 . . . c14)
0	1 1 1 1 0 1 0 1 1 0 0 1 0 0 0
1	0 0 0 0 1 0 1 0 0 1 1 0 1 1 1

TABLE I
SYMBOL TO CHIP MAPPING

The packets are divided into chunks of symbols by the `gr.packed_to_unpacked` block, each symbol representing 1 bit. Since the C++ programming language does not allow us to have a data type of 1 bit, the bits in the bytes of an input stream are grouped into chunks of 1 byte. The MSB (Most Significant Bit) of 8 output bits represents the one bit at the input of `gr.map_bb`. After that, the differential encoder `gr.diff_encoder_bb` encodes a current symbol modulo-2 of the previous one. Then, the symbols are mapped by `gr.symbols_to_chips` into 15 Pseudo Number Sequence chip as specified in Table I. The output of mapping is short-type (16 bits carrying the 15 chips). With the same technique the stream is unpacked to a chunks of 16 bits representing the chips stream. Each chip is represented by a complex constellation point in 1 dimension for the BPSK modulator by `gr.chunks_to_symbols_sc`. The stream is then fed through a Root Raised Cosine `gr.interp_fir_filter_ccf` filter which up-samples the signal, after which it is sent from the host computer via USB to the transmitting USRP.

B. Receiver

The receiver begins with an USRP source connected to a squelch filter `gr.pwd_squelch` which admits only signals with a certain dB strength. The squelch filter in GNU Radio outputs 0 when the incoming signal is too weak. The stream result of the squelch is passed to the Automatic Gain Control `gr.agc_cc` (AGC) of a D-BPSK demodulator, it regulates the gain in a way that does not have a large or small amplitude and to avoid distortions. After that, the result enters to two filters in `gr.interp_fir_filter_ccf`, FIR (Finite Impulse Response) and RRC (Root Raised Cosine) allowing the receiver to process the change of the transmitted pulse and minimize symbol interference. The RRC filter makes the correlation between the received signal and the expected one. It calculates a FIR filter coefficient or a tap weight. The demodulator synchronizer is composed by two blocks, a Costas Loop `gr.costas_loop_cc` (Phase Locked Loop) and the Mueller and Müller `gr.clock_recovery_mm_cc` [13]. The Costas Loop recovers the carrier and improves the Bit Error Rate of BPSK demodulator. The Mueller-Müller Timing recovery block recovers the symbol timing phase of the input signal. After the demodulator, the stream is converted from complex to real in order to send it to our developed block `ieee.ieee802_15_4_packet_sink` which slices real stream from chips to bits. With the knowledge of the packet length field, the packets are decoded. The first information decoded is the preamble with four 0x00 bytes, it is followed by the rest of the fields. If the preamble is not detected, the preamble search is re-launched. The receiver performs the error detection without correction. After the packet construction,

Preamble	SFD	PHR	FCF	SeqN	AddressInf	Data payload	CRC-16
4 bytes	1 byte	1 byte	2 bytes	1 byte	[0 20] bytes	<= 104 bytes	2 bytes

Fig. 4. IEEE 802.15.4 packet format for the USRP

a CRC-16 value is processed and compared to that carried by the CRC field of the received frame. If they are not equal, the received packet is incorrect.

The packet queue is observed by an external python thread. When a message arrives to the queue, a thread starts to call a function that process the packet, e.g: like printing the packet content.

IV. Experimental conditions and results

The experimentations are performed in an indoor environment. We use two USRP1 platforms coupled with RFX 900 daughterboards covering a frequency range from 750 MHz to 1050 MHz. The GNU Radio software stack is executed on a host computer having one Core 2 Duo CPU running at 2.4 GHz and 2 GB of RAM. The distance between the two USRP1 boxes was greater than 2 meters.

The principal USRP1 parameters are the transmitter Interpolation I and receiver Decimation D, they are calculated according to a symbol rate r, DAC_s and ADC_s sampling, and a number of samples per symbol sps, such as:

$$I = \frac{DAC\_s}{r \cdot sps}, \qquad D = \frac{ADC\_s}{r \cdot sps} \tag{2}$$

where :

- $DAC\_s = 128$ MSPS
- $I \in [16,\ 20,\ 24, \ldots 508,\ 512]$
- $ADC\_s = 64$ MSPS
- $D \in [8,\ 10,\ 12, \ldots 254,\ 256]$

For 20 kbps, the transmitter and receiver parameters are respectively $I = 400$ and $D = 200$ with a $sps = 16$. Otherwise, when the data bit rate is equal to 40 kbps, the I and D take the same values but with $sps = 8$. The amplifier amplitude is defined by a dimensionless scalar with values ranging from 0 to 32767.

The results shown in Fig. 6 depict the power spectrum of the transmitted signal from the GNURadio transmitter. They correspond to the output of the FFT spectrum-analyser tool that is included in the GNURadio framework. A peak is visible with our software transceiver when we choosing the channel at 916 MHz, with a number of 35 samples per symbol which allow us to have an intermediate frequency of 1.5 MHz. This value is in concordance with the values taken by the transmitted power spectral density of the IEEE 802.15.4 standard (see Fig. 6). Furthermore, frequencies at the edge of the main band are visible but strongly attenuated. These imperfections may be due to the roll-off characteristics of the interpolation filter in the up-conversion processing of the FPGA.

We use a D-BPSK modulation and the receiver constellation is depicted in Fig. 7.

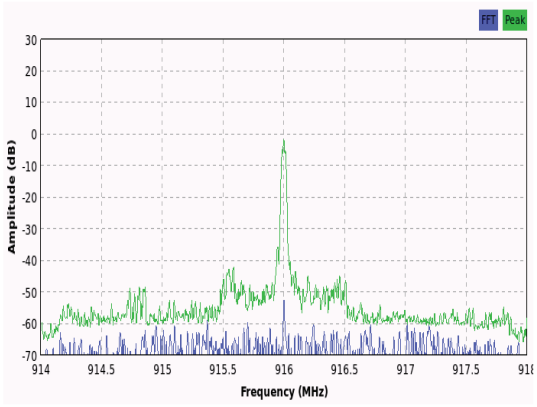


Fig. 6. Power spectrum of our software transceiver recorded with the USRP and drawn by FFT gnuradio plot

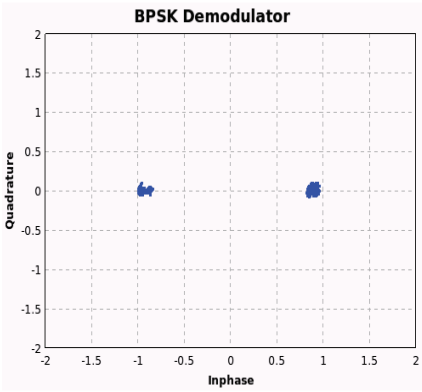


Fig. 7. Receiver symbol constellation

The performance of D-BPSK modulation is evaluated without packet generation. The flexibility of the GNU radio permits the reconfigurability of the transmitter/receiver chain by adding or replacing blocks. In a first experiment, we use the modulator and demodulator chains to measure the BER and SNR parameters. Fig. 8 illustrates the average BER versus the input SNR (dB) for the frequency 868.3 MHz and for the MFB Matched Filter Bound of D-BPSK modulation. The results have been computed by changing the amplifier amplitude values from 1000 to 12000 with the step of 100 for a time period of 400 seconds. Although noisy, the results are in concordance to the theory, proving that the implementation is working.

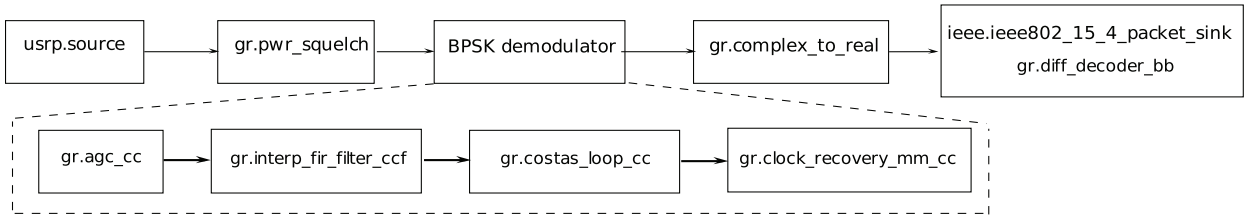


Fig. 5. Receiver flow graph

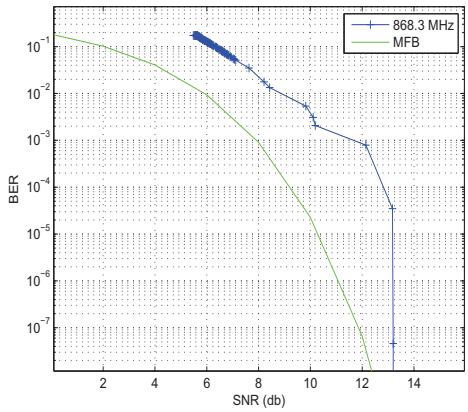


Fig. 8. The BER versus received SNR for central frequency 868.3 MHz and for the MFB

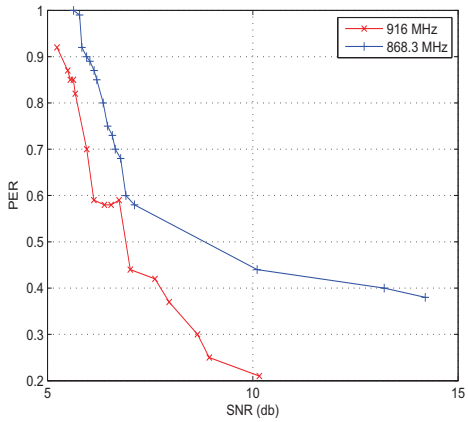


Fig. 9. The PER over received SNR using two central frequencies 916 MHz and 868 MHz

In the second experiment, the packet generator and packet sink are connected to the transmission chain and we measure a PER parameter as a function of SNR (dB). The measures are obtained by sending for each amplifier amplitude 100 packets apart from 0.2 s between two successive packets. The PER decreases when the amplifier amplitude increases. The shape of the curve is compliant to that of the BER (show Fig.9). The PER depends on the synchronization between the transmitter and the receiver. We noticed that the synchronization does not occur at every execution. This issue may arise when the USRP does not clear its buffer memory.

V. Conclusion

In this paper, we report the implementation of the IEEE 802.15.4 standard on a SDR transceiver for the 915/868 MHz band. The SW stack is based on the GNURadio open-source project and the HW is based on an USRP1 platform from Ettus Research. The BER and PER of the 802.15.4 have been calculated independently in an indoor environment by changing the signal amplitude. The results are coherent with the lower theoretical bound that is expected. The obtained performances of the PER are degraded compared to the BER because the successful receiving packets depend on the synchronization and the BER.

References

[1] J. Sabater, J. Gomez, and M. Lopez, "Towards an ieee 802.15.4 sdr transceiver," in *Icecs*. Ieee, 2010, pp. 323–326.

[2] N.-J. Oh, S.-G. Lee, and J. Ko, "A cmos 868/915 mhz direct conversion. zigbee single-chip radio," *IEEE Communications Magazine*, vol. 43, no. 12, pp. 100–109, 2006.

[3] T. Ulversoy, "Software defined radio: Challenges and opportunities," *IEEE Communications Surveys and Tutorials*, vol. 12, no. 4, pp. 531–550, 2010.

[4] E. Blossom, "Gnu radio: tools for exploring the radio frequency spectrum," *Linux J.*, vol. 2004, pp. 4–, Jun. 2004.

[5] Ettus, "About ettus research," Feb. 2011. [Online]. Available: https://www.ettus.com/product

[6] J. Mitola, "Software radios: Survey, critical evaluation and future directions," *IEEE Aerospace and Electronic Systems Magazine*, vol. 8, no. 4, pp. 25–36, Apr. 1993.

[7] M. N. O. Sadiku and C. M. Akujuobi, "Software-defined radio: a brief overview," *Ieee Potentials*, vol. 23, no. 4, pp. 14–15, 2004.

[8] gnuradio.org, "Gnu radio," Feb. 2011. [Online]. Available: http://gnuradio.org/redmine/projects/gnuradio/wiki

[9] ossie.wireless.vt.edu, "Sca-based open soruce software defined radio," Feb. 2011. [Online]. Available: http://ossie.wireless.vt.edu/

[10] "Ieee standard for local and metropolitan area networks–part 15.4: Low-rate wireless personal area networks (lr-wpans)," pp. 1–314, 2011, iEEE Std 802.15.4-2011 (Revision of IEEE Std 802.15.4-2006).

[11] T. Schmid, "Gnu radio 802.15.4 en- and decoding," Tech. Rep., 2006.

[12] T. Schmid, T. Dreier, and M. B. Srivastava, "Software radio implementation of short-range wireless standards for sensor networking," in *SenSys*, 2006, pp. 381–382.

[13] G. R. Danesfahani and T. G. Jeans, "Optimisation of modified Mueller and Muller algorithm," *Electronics Letters*, vol. 31, no. 13, pp. 1032–1033, Jun. 1995.

Fixed-Point Aspects of MIMO OFDM Detection on SDR Platforms

D. Guenther, T. Kempf, G. Ascheid
Institute for Communication Technologies and Embedded Systems, RWTH Aachen University, Germany
guenther@ice.rwth-aachen.de

Abstract - In IEEE 802.11n compliant receivers, MIMO detection causes a major part of the computational complexity. Various publications exist on the issue of ASIC design for MIMO detection. However, the increasing variety of mobile communication standards calls for more flexible platforms, implementing the different standards in software, hence called Software Defined Radios (SDRs). This work focuses on achieving quality of service close to floating point performance while using the limitted fixed-point precision typically available on SDR platforms. A suitable algorithm for QR decomposition of the channel matrices for real time processing is derived, and the consecutive spatial equalizing as well as the SINR calculation are presented. A software implementation on the maturing P2012 platform by ST Microelectronics is benchmarked with respect to timing- and error correction performance.

I. Introduction

With the increasing popularity of mobile communication, the amount of mobile communication standards is increasing equally. Furthermore, modern portable communication devices are supposed to support cellular networks like LTE [1] and wireless LANs, e.g. IEEE 802.11n [2]. Instead of creating an application-specific integrated circuit (ASIC) for each communication standard, the Software Defined Radio (SDR) approach implements each standard in software on a flexible, programmable platform instead. Simultaneously, the demand for higher data rates is answered by multiple-input and multiple-output (MIMO) systems with more than one antenna at the transmitter- and receiver side to benefit from spacial diversity. In LTE as well as in IEEE 802.11n, MIMO comes along with orthogonal frequency-division multiplexing (OFDM) as modulation technique, which minimizes the impact of frequency selective fading.

As discussed in [3], MIMO processing requires vector analysis. Hence single instruction multiple data (SIMD) capabilities are of central importance for an SDR platform. However, programmable platforms come along with fixed bitwidths for integer datatypes (typically 8, 16 and 32 bit), while in an ASIC solution, the internal bitwidth can be chosen freely. In [4], it has been shown how the QR decomposition of the MIMO channel matrices can be numerically stabilized for a limitted bitwidth. An approach which will be adopted and extended here.

In this work, the maturing P2012 platform [5] by ST Microelectronics is used as a reference. The P2012 is composed of several clusters of RISC cores. Each cluster comprises a maximum of 16 cores which can be tailored by application specific extensions like the VECx extension, offering single-instruction multiple-data (SIMD) instructions for vector analysis. Figure 1 gives an overview of the P2012 platform. The proposed detector along with all other parts of the inner modem is implemented on this platform.

This work has been supported by the UMIC Research Centre, RWTH Aachen University and by the EC under grant 2PARMA FP7-248716.

The remainder of this paper is organized as follows: Section II introduces the system model and the structure of an adequate MIMO OFDM receiver. Section III presents efficient implementations of the QR decomposition of the channel matrices, focussing on how to achieve good fixed point stability with the limited available bitwidth. Moreover, the calculation of the signal to interference and noise ratio (SINR), which operates on the same data, is explained. Section IV discusses the execution time of different algorithms, while Section V elaborates on the resulting error correction performance. Section VI concludes the paper.

II. System Model

OFDM divides the available bandwidth into a number of equally spaced subchannels (subcarriers). Hence, one may treat each subchannel separately as a frequency flat channel. For a system with N_t transmitter antennas and N_r receiver antennas, each subchannel is modeled by the transmission equation:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1}$$

Whereat $\mathbf{x}$ is the transmitted symbol vector of dimension N_t, $\mathbf{y}$ is the received symbol vector of dimension N_r, $\mathbf{H}$ is the channel matrix for the current subchannel of dimension $N_r \times N_t$ and $\mathbf{n}$ is the noise vector of dimension N_r.

The basic structure of a MIMO OFDM receiver is shown in Figure 2. The receiver is divided into an inner modem and an outer modem. The inner modem works on the complex baseband representation of the received data, while the outer modem gets a bitwise representation as input. Within the inner modem, OFDM processing comprises several steps: Firstly removing the cyclic prefix, which is inserted upfront every OFDM symbol to protect against inter-symbol-interference (ISI) and inter-carrier-interference (ICI). Secondly and mainly,

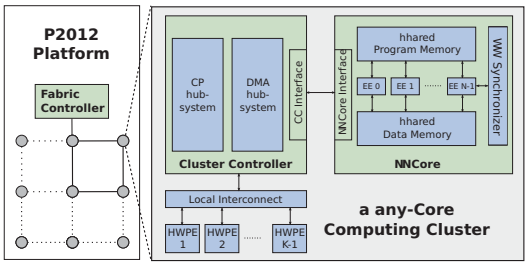


Fig. 1. P2012 platform [5]

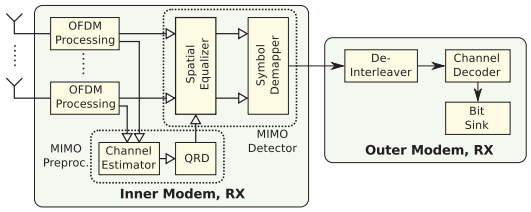


Fig. 2. MIMO OFDM receiver overview

it performs an OFDM demodulation, which is an FFT, to separate the OFDM subcarriers. Thirdly, it conducts a subcarrier demapping, since certain guard carriers at the channel borders are left empty to mitigate inter-channel interference. The channel estimator derives an estimate of the channel matrix $\mathbf{H}$, to which a QR decomposition (QRD) is applied. The combination of channel estimation and QRD is also referred to as MIMO preprocessing. The QRD of $\mathbf{H}$ is used by the spatial equalizer to mitigate the impact of the channel on the payload data. Soft symbol demapping finally converts the complex baseband representation of the received data back to a bitwise representation. Equalizing and soft demapping together are also referred to as MIMO detection, whose output is then fed to the deinterleaver and the channel decoder.

MIMO equalizing, which is the main focus of this work, derives an estimate $\hat{\mathbf{x}}$ for the originally transmitted symbol vector $\mathbf{x}$ by using $\mathbf{y}$ and the estimated channel matrix $\hat{\mathbf{H}}$. Two common approaches, which will be discussed in the following are **linear minimum mean square error** (LMMSE) equalizing and **successive interference cancellation** (SIC).

A. LMMSE Equalizing

LMMSE minimizes the expected value of the square error of the estimated and transmitted symbol vector by multiplying $\mathbf{y}$ by an equalizer matrix $\mathbf{G}$ of dimension $N_t \times N_r$.

$$\begin{aligned} \hat{\mathbf{x}} &= \mathbf{G}\mathbf{y} \\ \arg\min_{\mathbf{G}} \mathrm{E}\left\{ \left| \mathbf{x} - \hat{\mathbf{x}} \right|^2 \right\} &= \arg\min_{\mathbf{G}} \mathrm{E}\left\{ \left| \mathbf{x} - \mathbf{G}\mathbf{y} \right|^2 \right\} \end{aligned} \quad (2)$$

Assuming uncorrelated additive white Gaussian noise (AWGN) results in:

$$\mathrm{E}\left\{ \mathbf{n}^H \mathbf{n} \right\} = N_0 \cdot \mathbf{I} \quad (3)$$

With spectral noise density N_0 and $\mathbf{I}$ representing an identity matrix. Using this relationship in (2) delivers the LMMSE equalizer matrix:

$$\mathbf{G} = \left(\hat{\mathbf{H}}^H \hat{\mathbf{H}} + N_0 \mathbf{I} \right)^{-1} \hat{\mathbf{H}}^H \quad (4)$$

Using the regularized channel matrix $\bar{\mathbf{H}}$ of dimension $(N_r + N_t) \times N_t$,

$$\bar{\mathbf{H}} = \begin{pmatrix} \hat{\mathbf{H}} \\ \sqrt{N_0}\mathbf{I} \end{pmatrix} \quad (5)$$

(3) can be rewritten as:

$$\mathbf{G} = \left(\bar{\mathbf{H}}^H \bar{\mathbf{H}} \right)^{-1} \hat{\mathbf{H}}^H \quad (6)$$

However, (6) still contains a matrix inversion, which is not suitable for implementation on platforms with limited fixed point precision. Therefore, $\bar{\mathbf{H}}$ is decomposed into the product of $\mathbf{Q}$ and $\mathbf{R}$, so that:

$$\mathbf{Q}\mathbf{R} = \begin{pmatrix} \mathbf{Q_a} \\ \mathbf{Q_b} \end{pmatrix} \mathbf{R} = \bar{\mathbf{H}} \quad (7)$$

where the matrix $\mathbf{Q_a}$ is of dimension $N_r \times N_t$ and $\mathbf{Q_b}$ is of dimension $N_t \times N_t$.

$$\mathbf{Q}^H \mathbf{Q} = \mathbf{I} \quad \wedge \quad \mathbf{Q_b} = \sqrt{N_0}\mathbf{R}^{-1} \quad (8)$$

Now, (6) can be rewritten as an inversion-free equation:

$$\mathbf{G} = \mathbf{R}^{-1} {\mathbf{Q_a}}^H = \frac{\mathbf{Q_b}}{\sqrt{N_0}} {\mathbf{Q_a}}^H \quad (9)$$

B. MMSE-SIC Equalizing

In contrast to linear equalizing, as presented above, which derives all elements of the symbol vector estimate at once, MMSE-SIC equalizing calculates the estimate componentwise. For that purpose, (2) is reformulated:

$${\mathbf{Q_a}}^H \mathbf{y} = \mathbf{R}\hat{\mathbf{x}} \quad (10)$$

Due to the upper triangular structure of $\mathbf{R}$, the equation can be successively solved for the elements of $\hat{\mathbf{x}}$, starting from the element with the highest index. Before re-using an already determined element to derive another, it is quantized (sliced) to the closest constellation symbol. This technique improves detection quality, given the fact that elements are sliced to the correct constellation symbol. Otherwise, an error propagation is caused.

$$\hat{x}_i = \frac{\tilde{y}_i - \sum_{j=i+1}^{N_t} r_{ij}\, \mathrm{Q}\left[\hat{x}_j\right]}{r_{ii}} \qquad i = N_r \,..\, 1 \quad (11)$$

To circumvent error propagation, an ordered SIC may be used. Ordering is performed in a way that the symbol with the highest SINR is detected first. In [6], it is shown that this corresponds to reordering the QRD by a multiplication by a permutation matrix $\mathbf{P}$ so that $r_{ii} < r_{jj}$ for $i < j$.

$$\bar{\mathbf{H}} = \mathbf{Q}\mathbf{R}\mathbf{P}^T \quad (12)$$

This means instead of (10), one has to successively solve:

$${\mathbf{Q_a}}^H \mathbf{y} = \mathbf{R}\mathbf{P}^T \hat{\mathbf{x}} \quad (13)$$

Based on the above given equations, LMMSE and MMSE-SIC equalizing are divided into a **preprocessing** and an **actual equalizing** phase. Preprocessing comprises the QRD and for LMMSE also the calculation of $\mathbf{G}$. Note that preprocessing is performed for every subcarrier and is independent of the received data payload. Actual equalizing then uses the results from preprocessing to calculate $\hat{\mathbf{x}}$. For LMMSE, this phase solely consists of a number of matrix vector multiplications, while for SIC-MMSE, a pre-multiplication of the received symbol vectors by ${\mathbf{Q_a}}^H$ has to be performed. Then, finally, the estimated symbol vectors are demapped to a soft, bitwise representation.

III. QRD Investigation

Different approaches exist for performing a QR decomposition. Two methods suitable for a hardware friendly implementation are **Modified Gram-Schmidt** (MGS) and **Givens Rotation** (GR).

A. Modified Gram-Schmidt

The MGS algorithm operates column-wise on the regularized channel matrix $\bar{\mathbf{H}}$. It starts from the leftmost column which is first normalized and then projected on the right-hand columns. The result of the projection is subtracted from these columns. Then, the process starts again from the next, righthand column vector. Thereby, linear dependencies are removed. However, the repeated subtraction of projections may cause the righter column vectors to become too big or small for an accurate fixed-point implementation. For that reason, a technique called dynamic scaling (DS) is introduced. During the QR decomposition, DS performs bitwise left- and right shifts on the column vectors to keep them within fixed-point range. Similar to [4], this work uses a MGS with DS, which is described by the pseudo-code given in Algorithm 1.

DS is performed in lines 3 to 9, while the remainder of the algorithm is the ordinary MGS algorithm. Note that the algorithm does not deliver $\mathbf{R}$. Instead, one gets $\mathbf{Q_b} = \sqrt{N_0}\mathbf{R}^{-1}$ which can be used in (9) to derive $\mathbf{G}$ directly.

For high SNR regions, like the SNR-target-region for 64-QAM modulation, this algorithm may become critical, though. Due to limited fixed-point precisions, the lower scaled identity matrix of $\bar{\mathbf{H}}$ becomes too small for an accurate fixed-point representation. Thus, $\mathbf{Q_b}$ and the equalizer matrix $\mathbf{G}$ will degrade equally, compromising the entire detection. For this reason, this work uses the unity-regularized channel matrix (URCM) $\bar{\mathbf{H}}_u$ instead of $\bar{\mathbf{H}}$.

$$\bar{\mathbf{H}}_u = \begin{pmatrix} \mathbf{H} \\ \mathbf{I} \end{pmatrix} \tag{14}$$

Naturally, this affects the normalization and the projection in lines 10 and 12. However, the square norms of the initial $\bar{\mathbf{H}}$ matrix can be calculated from $\mathbf{H}$ and N_0. Subsequently these norms can be updated as presented in [7]. For the projection, though, the last N_t entries of the column vectors $\mathbf{v_j}$ have to be scaled down by a factor of $\sqrt{N_0}$. These considerations lead to Algorithm 2.

Algorithm 1 MMSE MGS-QRD with DS

1: $\mathbf{V} \leftarrow \bar{\mathbf{H}}$
2: **for** $i = 1$ to N_t **do**
3: **for** $j = i$ to N_t **do**
4: **if** $\max\{|\Re\{v_{j,1}\}|, |\Im\{v_{j,1}\}|, ..\} < B_l$ **then**
5: $\mathbf{v_j} \leftarrow 2\mathbf{v_j}$
6: **else if** $\max\{|\Re\{v_{j,1}\}|, |\Im\{v_{j,1}\}|, ..\} > B_h$ **then**
7: $\mathbf{v_j} \leftarrow \mathbf{v_j}/2$
8: **end if**
9: **end for**
10: $\mathbf{v_i} \leftarrow \mathbf{v_i}/\|\mathbf{v_i}\|$
11: **for** $j = i+1$ to N_t **do**
12: $\mathbf{v_j} \leftarrow \mathbf{v_j} - \left(\mathbf{v_i}^H \mathbf{v_j}\right)\mathbf{v_i}$
13: **end for**
14: **end for**
15: $\mathbf{Q} \leftarrow [\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_{N_t}}]$

Algorithm 2 MMSE MGS-QRD with DS and URCM

1: $\mathbf{V} \leftarrow \bar{\mathbf{H}}_\mathbf{u}$
2: **for** $i = 1$ to N_t **do**
3: $\xi_i = \left(\mathbf{H}^H\mathbf{H}\right)_{i,i} + N_0$
4: **end for**
5: **for** $i = 1$ to N_t **do**
6: **for** $j = i$ to N_t **do**
7: **if** $\max\{|\Re\{v_{j,1}\}|, |\Im\{v_{j,1}\}|, ..\} < B_l$ **then**
8: $\mathbf{v_j} \leftarrow 2\mathbf{v_j}$
9: $\xi_j \leftarrow 4 \cdot \xi_j$
10: **else if** $\max\{|\Re\{v_{j,1}\}|, |\Im\{v_{j,1}\}|, ..\} > B_h$ **then**
11: $\mathbf{v_j} \leftarrow \mathbf{v_j}/2$
12: $\xi_j \leftarrow 1/4 \cdot \xi_j$
13: **end if**
14: **end for**
15: $\mathbf{v_i} \leftarrow \mathbf{v_i}/|\sqrt{\xi_i}|$
16: **for** $j = i+1$ to N_t **do**
17: $s = \left(\mathbf{v_i}^H \odot \mathbf{a}^T\right)\left(\mathbf{v_j} \odot \mathbf{a}\right)$
18: $\mathbf{v_j} \leftarrow \mathbf{v_j} - s\mathbf{v_i}$
19: $\xi_j \leftarrow \xi_j - |s|^2$
20: **end for**
21: **end for**
22: $\mathbf{Q} \leftarrow [\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_{N_t}}]$

Here, ξ_i is the vector norm of column vector $\mathbf{v}_i$ and $\mathbf{a}$ is a real valued scaling vector that scales the last N_t entries down by a factor of $\sqrt{N_0}$ while leaving the prior ones unchanged. This component wise multiplication is denoted by the $\odot$-operator.

While the $\mathbf{R}$-matrix is not required for linear equalizing, it is necessary for non-linear algorithms like SIC. As mentioned in [4], the matrix is lost when applying DS, but it can be restored by keeping track of the bitshifts e_i formerly performed on the column vector i of $\bar{\mathbf{H}}_\mathbf{u}$. Combined with column sorting as proposed in [8], this leads to Algorithm 3. While the sorted MGS algorithm proposed in [7] uses the column vector norm to derive the column scaling factors, the algorithm presented here uses the absolute values in combination with DS. Hence, the resulting algorithm is computationally less complex for an SDR application.

B. Inverse Square Root Calculation

Normalization of the column vectors as required by the MGS algorithm is a computationally complex tasks, since it inhibits a square root calculation and a division. Observing Algorithms 2 and 3, one sees that but the inverse of r_{ii} is used in the QRD as well as during spatial equalizing. For that reason, it makes sense to store $(1/r_{ii})$ instead of r_{ii} in the diagonal entries of the $\mathbf{R}$-matrix.

As proposed in [9], the inverse square root calculation can be efficiently approximated by Newton's Method (NM). The

Algorithm 3 MMSE MGS-SQRD with DS and URCM

$\mathbf{V} \leftarrow \bar{\mathbf{H}}_\mathbf{u}$
$\mathbf{P} \leftarrow \mathbf{I}_{N_t}$
for $i = 1$ to N_t **do**
  $\xi_i = \left(\mathbf{H}^H\mathbf{H}\right)_{i,i} + N_0$
  $e_i = 0$
end for
for $i = 1$ to N_t **do**
  $k = \text{argmin}_{j=i,..,N_t}(\xi_j)$
  exchange columns i and k in $\mathbf{V}$, $\mathbf{R}$ and $\mathbf{P}$
  exchange elements i and k in ξ and $\mathbf{e}$
  for $j = i$ to N_t **do**
    if $\max\{|\Re\{v_{j,1}\}|, |\Im\{v_{j,1}\}|, ..\} < B_l$ **then**
      $\mathbf{v_j} \leftarrow 2\mathbf{v_j}$
      $e_j \leftarrow e_j + 1$
      $\xi_j \leftarrow 4 \cdot \xi_j$
    else if $\max\{|\Re\{v_{j,1}\}|, |\Im\{v_{j,1}\}|, ..\} > B_h$ **then**
      $\mathbf{v_j} \leftarrow \mathbf{v_j}/2$
      $e_j \leftarrow e_j - 1$
      $\xi_j \leftarrow 1/4 \cdot \xi_j$
    end if
  end for
  $r_{ii} = 1/\sqrt{\xi_i} \cdot 2^{+e_i}$
  $\mathbf{v_i} \leftarrow \mathbf{v_i}/\sqrt{\xi_i}$
  for $j = i + 1$ to N_t **do**
    $s = \left(\mathbf{v_i}^H \odot \mathbf{a}^T\right)(\mathbf{v_j} \odot \mathbf{a})$
    $r_{ij} = s \cdot 2^{-e_j}$
    $\mathbf{v_j} \leftarrow \mathbf{v_j} - s\mathbf{v_i}$
    $\xi_j \leftarrow \xi_j - |s|^2$
  end for
end for
$\mathbf{Q} \leftarrow [\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_{N_t}}]$

general form of NM for iteratively deriving the zeros of a function f is given by:

$$y_{n+1} = y_n - \frac{f(y_n)}{f'(y_n)} \tag{15}$$

To solve $y = 1/\sqrt{x}$, the function whose zeros one has to derive is $f(y) = 1/y^2 - x$, so the iteration is given by:

$$y_{n+1} = y_n \frac{3 - xy_n^2}{2} \tag{16}$$

Requiring no divisions, this term is very suitable for an SDR implementation. Additionally, the application of DS is limitting the dynamic range of the input values, so in combination with a start value selection from a few predefined values, good results can be achieved with less than 5 iterations.

C. Givens Rotation

In contrast to MGS, which is based on vector-vector projections and subtractions, the GR algorithm rotates the row vectors of the channel matrix to derive a set of orthogonal basis vectors. Consequently, the vector norms remain unchanged, so numerical stabilization like DS is not required here. The starting point for the MMSE approach is a composite matrix:

$$\mathbf{Z} = \begin{bmatrix} \mathbf{H} & \mathbf{I}_{N_r} \\ \sqrt{\frac{N_0}{E_s}}\mathbf{I}_{N_t} & 0 \end{bmatrix} \tag{17}$$

The matrix is processed from the bottom upwards, combining two adjacent rows in a rotation operation. Such a rotation is performed in two steps. First, real and imaginary part of the upper row vector are turned so that the leftmost, non-zero element of the vector becomes real. Then, lower and upper vectors are rotated so that the leftmost, non-zero element of the lower vector becomes zero. These rotation are repeated from top to bottom until the submatrix $\mathbf{H}$ has been turned into a triangular matrix. After N processing steps, the matrix $\mathbf{Z}^{(N)}$ now has the following shape:

$$\mathbf{Z}^{(N)} = \begin{bmatrix} \mathbf{R} & \mathbf{Q}_a^H \\ 0 & \mathbf{Q}_c^H \end{bmatrix} \tag{18}$$

The channel matrix can be written as the product:

$$\mathbf{H} = \mathbf{Q}_a\mathbf{R} \tag{19}$$

If sorting as described in Alg.3 is used, the equation includes a the permutation matrix $\mathbf{P}$ and changes according to:

$$\mathbf{H} = \mathbf{Q}_a\mathbf{R}\mathbf{P}^T \tag{20}$$

The algorithm for the sorted variant as also described in [7] is given in Algorithm 4. Each Givens rotation is represented by a multiplication with a rotation matrix $\mathbf{\Theta}$. A matrix $\mathbf{\Theta}_{(p,q,\theta)}$, which rotates rows p and q by an angle of θ contains trigonometric functions. These functions are often not available on SDR platforms, however they can be approximated by several iteration of the so called CORDIC algorithm. A description

Algorithm 4 MMSE GR-SQRD

$\mathbf{Z} \leftarrow \mathbf{Z}^{(0)}$
$\mathbf{P} \leftarrow \mathbf{I}_{N_t}$
for $i = 1$ to N_t **do**
  $\xi_i = \left(\mathbf{H}^H\mathbf{H}\right)_{i,i} + N_0$
end for
for $i = 1$ to N_t **do**
  $k = \text{argmin}_{j=i,..,N_t}(\xi_j)$
  exchange columns i and k in $\mathbf{P}$ and the first $N_r + i - 1$ rows of $\mathbf{Z}$
  compute a series of Givens rotations Θ_u so that the elements $\mathbf{Z}_{(i+1,i)}$ until $\mathbf{Z}_{(i+N_r,i)}$ become zero.
  $\mathbf{Z} \leftarrow \left(\prod_{u=(i-1)N_r}^{iN_r} \mathbf{\Theta}_u\right)\mathbf{Z}$
  for $j = i + 1$ to N_t **do**
    $\xi_j \leftarrow \xi_j - |z_{ij}|^2$
  end for
end for
$\mathbf{R} = \mathbf{Z}_{(1..N_t,1..N_t)}$
$\mathbf{Q}_a^H = \mathbf{Z}_{(1..N_t,N_t..N_t+N_r)}$

of this algorithm is beyond the scope of this work, but the interested reader is referred to [7].

D. SINR Calculation

Soft symbol demapping of the equalized symbol vectors is commonly performed using the max-log approximation according to the below given equation. The index k denotes the stream, from which a symbol is to be demapped. $L(b_{k,i})$ is the LLR value for the bit at position i within the constellation symbol from stream k, and A_i^0 and A_i^1 are the subsets of constellation symbols with a zero or one bit at position i respectively.

$$L(b_{k,i}) \approx \rho_k \left(\min_{s \in A_i^0} |z_k - s|^2 - \min_{s \in A_i^1} |z_k - s|^2 \right) \tag{21}$$

The signal to interference and noise ratio (SINR) of stream k is given by ρ_k [10] and can be calculated according to:

$$\rho_k \approx \frac{1}{\frac{\sigma_n^2}{E_s} \left[\left(\mathbf{H}^H \mathbf{H} + \frac{\sigma_n^2}{E_s} \mathbf{I}_{M_r} \right)^{-1} \right]_{k,k}} \tag{22}$$

Using the results of the regularized QR decomposition, the above equation can be reformulated to:

$$\rho_k \approx \frac{1}{\frac{\sigma_n^2}{E_s} \left[\mathbf{R}^{-1} \left(\mathbf{R}^{-1} \right)^H \right]_{k,k}} \tag{23}$$

This means, even though strictly speaking, demapping is not a part of spatial equalizing, it uses the results of the latter. Also note that in combination with regularized MGS-QRD and URCM, $\mathbf{R}^{-1}$ is directly available from the lower part of the $\mathbf{Q}$ matrix. As a consequence, it makes sense to include SINR calculation into MIMO preprocessing, as it will be done in the following.

IV. Execution Time

A fast execution time is crucial to keep the real time constraints of IEEE 802.11n. Even though more advanced detection algorithms like Sphere Decoding [11] are available, the limits of current SDR platforms require the application of more straight forward detectors. Table I lists the execution time of several suitable MIMO preprocessing and spatial equalizing algorithms for a 2×2 and 4×4 antenna configuration on a single xp70 core equipped with a VECx SIMD extension. Times for preprocessing are accumulated for the entire frame. Times for spatial equalizing contain the processing of one OFDM slot of $4\mu s$ duration.

Also note, that for linear equalizing, the QR decomposition has to be followed by a matrix-matrix multiplication, to derive the equalizer matrix $\mathbf{G}$. Naturally, this step is not required for SIC preprocessing, but in the non-linear equalizing, every received symbol vector has to be pre-multiplied by $\mathbf{Q}_a^H$ before starting the back substitution. For linear equalizing, on the other hand, the actual equalizing solely consists of a matrix vector multiplication.

Regarding preprocessing, this table shows that dynamic scaling as well as sorted QR decomposition come at an additional cost in terms of execution time, which has to be justified by a superior algorithmic performance. One also sees a major advantage of the regularized MGS QR decomposition over the GR variant, since $\mathbf{R}^{-1}$ is directly available from the decomposition in the first case, while the triangular matrix $\mathbf{R}$ has to be inverted first in the latter case. Thus, SINR calculation can be performed faster, if MGS-QRD is used.

System	**2x2**		**4x4**	
Operation	cycles	T (μs)	cycles	T (μs)
MIMO Preprocessing (per frame)				
mgs-mmse-qrd	22,848	38.08	55,536	92.56
mgs-mmse-ds-qrd (UMCR)	35,424	59.04	66,624	111.04
mgs-mmse-ds-sqrd (UMCR)	43,248	72.08	85,392	142.32
gr-mmse-sqrd	-	-	112,032	186.72
matrix-matrix mul.	2,496	4.16	11,472	19.12
sinr-calc-r	9,456	15.76	25,824	43.04
sinr-calc-r-inv	7,248	12.08	15,600	26.00
Spatial Equalizing (per OFDM slot)				
back substitution	1,188	1.98	2,736	4.56
matrix-vector mul.	1,968	3.28	3,312	5.52

TABLE I. Single core execution time

Apart from that, the table shows that the CORDIC based GR is significantly slower than the MGS variant. Even though [12] indicates that GR is more suitable for a high throughput ASIC solution, the regular data accesses of MGS are more suitable for a SIMD implementation. Moreover the VECx vector unit is not specifically tailored to the needs of baseband processing, so it contains no special instructions to speed up the CORDIC algorithm.

Even though the multi-core aspects of the P2012 SDR application are covered in [13], two important aspects of the application, related to the above presented execution times, shall be mentioned here. Firstly one sees that the execution time for preprocessing exceeds the real time assigned to the 2×2 or 4×4 preamble. However, it is not an OFDM slot that has to be processed in real time but the entire MIMO OFDM frame. For this reason, the additional latency introduced by preprocessing can be compensated by a faster spatial equalizing. Secondly, the results in Table I show that also the actual equalizing does not achieve real time execution with a single xp70 core. For this reason, the application inherent data level parallelism must be used, as presented in [13], to distribute each task along a parallelizable dimension to enable real time execution.

V. Algorithmic Performance

To achieve real time execution, the target application uses a 16 bit fixed point format. This format causes precision problems for $\mathbf{Q_a}$ and $\mathbf{Q_b}$ in (7), which were addressed by dynamic scaling and URCM respectively. To evaluate the effect of these measures, the error correction capabilities of the resulting implementation in terms of bit-error-rate (BER) are presented in this section. For that purpose, a channel simulation featuring AWGN and i.i.d. Rayleigh Fading was set up. As in the frequently used TGn-C, a 150ns power delay spread is assumed. The power delay profile is modelled as an exponential 20dB drop. The error correction of the fixed

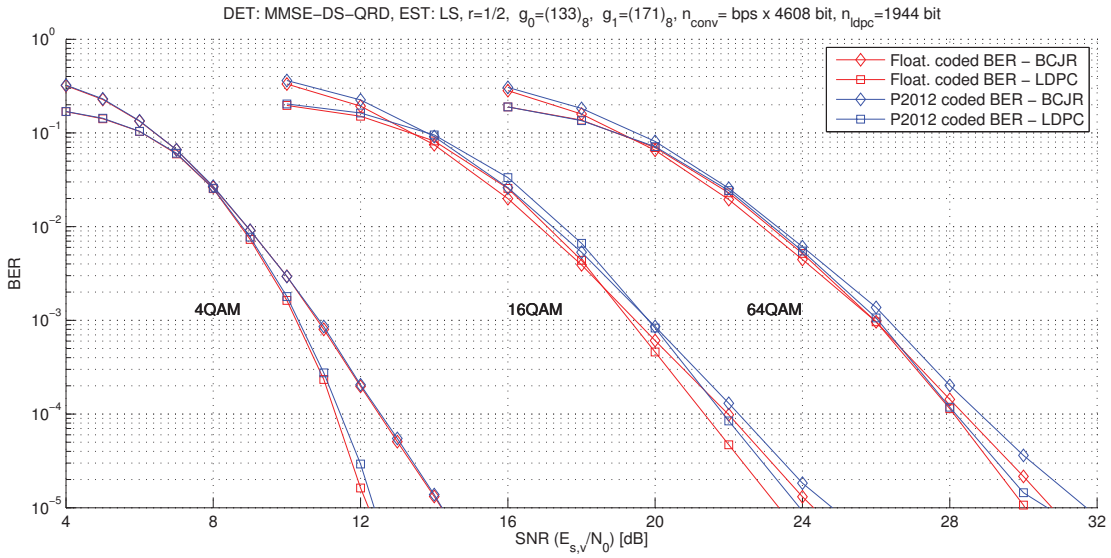


Fig. 3. Coded BER comparison of 4x4 MIMO OFDM use cases

point P2012 application is then compared to a floating point reference implementation. Note that even though this paper focuses on detection aspects, the fixed point implementation, which was benchmarked, comprises the **entire** inner modem application on the transmitter and receiver sides. Hence the resulting BERs are a realistic estimate for the actually achievable performance. Figure 3 shows the coded BER curves of the P2012 fixed-point inner modem compared to the floating point reference (UPEG), where the signal power of the SNR is given as the power per symbol of each receiver antenna. To show the capabilities of the two inner modems, they are combined with a soft-input BCJR and LDPC channel decoder. As one can see from the figure, the MGS QR decomposition extended by dynamic scaling and URCM delivers close to floating point algorithmic performance. Similar to [14], it has been observed in the scope of this work that SIC equalizing (not shown) does not outperform linear MMSE equalizing in terms of coded BER, while its computational complexity is significantly higher.

VI. Conclusion

In this paper, SDR specific aspects of non-iterative MIMO detection were discussed. The main focus was on achieving close to floating point algorithmic performance, using the limitted 16 bit precision typically available on DSP platforms. It has been seen that linear MMSE equalizing based on the Modified Gram-Schmidt QR decomposition offers the desired algorithmic performance at a reasonably low computational complexity. While the focus of this paper was on an SDR platform equipped with SIMD cores, current and future work is also investigating implementations on other types of cores like VLIW. Apart from that, it will be investigated how gains in algorithmic performance can be achieved by more advanced algorithms, while keeping the computational complexity in a region feasible for SDR solutions.

References

[1] Jim Zyren. Overview of the 3GPP Long Term Evolution Physical Layer. jul 2007.

[2] IEEE Standard for Information technology–Telecommunications and information exchange between systems–Local and metropolitan area networks–Specific requirements Part 11, Amendment 5. *IEEE Std 802.11n-2009*, pages c1 –502, oct. 2009.

[3] Kees van Berkel, Frank Heinle, Patrick P. E. Meuwissen, Kees Moerman, and Matthias Weiss. Vector Processing as an Enabler for Software-Defined Radio in Handheld Devices. *EURASIP J. Appl. Signal Process.*, 2005:2613–2625, January 2005.

[4] Hun Seok et al. Kim. A practical, hardware friendly MMSE detector for MIMO-OFDM-based systems. *EURASIP J. Adv. Signal Process*, 2008:94:1–94:14, January 2008.

[5] D. Fuin L. Benini, E. Flamand and D. Melpignano. P2012: Building an ecosystem for a scalable, modular and high-efficiency embedded computing accelerator. In *DATE'12*, 2012.

[6] D. Wuebben, R. Bohnke, V. Kuhn, and K.-D. Kammeyer. Mmse extension of v-blast based on sorted qr decomposition. In *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, volume 1, pages 508 – 512 Vol.1, oct. 2003.

[7] Peter Jan Luethi. *VLSI circuits for MIMO preprocessing*. ETH, 2009.

[8] D. Wubben, R. Bohnke, J. Rinas, V. Kuhn, and K.D. Kammeyer. Efficient algorithm for decoding layered space-time codes. *Electronics Letters*, 37(22):1348 –1350, oct 2001.

[9] Fast Inverse Square Root. Chris lomont. 2003.

[10] I.B. Collings, M.R.G. Butler, and M. McKay. Low Complexity Receiver Design for MIMO Bit-Interleaved Coded Modulation. In *International Symposium on Spread Spectrum Techniques and Applications, 2004*.

[11] E.M. Witte, F. Borlenghi, G. Ascheid, R. Leupers, and H. Meyr. A Scalable VLSI Architecture for Soft-Input Soft-Output Single Tree-Search Sphere Decoding. *Circuits and Systems II: Express Briefs, IEEE Transactions on*, 57(9):706 –710, sept. 2010.

[12] P. Luethi, C. Studer, S. Duetsch, E. Zgraggen, H. Kaeslin, N. Felber, and W. Fichtner. Gram-schmidt-based qr decomposition for mimo detection: Vlsi implementation and comparison. In *Circuits and Systems, 2008. APCCAS 2008. IEEE Asia Pacific Conference on*, pages 830 –833, 30 2008-dec. 3 2008.

[13] Torsten Kempf, Daniel Guenther, Aamir Ishaque, and Gerd Ascheid. Mimo ofdm transceiver for a many-core computing fabric - a nucleus based implementation. In *SDR'11 - The Wireless Innovation Forum Conference on Communications Technologies and Software Defined Radio*, Washington D.C., USA, dec 2011.

[14] E. Zimmermann and G. Fettweis. Adaptive vs. Hybrid Iterative MIMO Receivers Based on MMSE Linear and Soft-SIC Detection. In *Personal, Indoor and Mobile Radio Communications, 2006 IEEE 17th International Symposium on*, pages 1 –5, sept. 2006.

A COMPONENT-BASED ARCHITECTURE FOR PROTOCOL DESIGN AND DEVELOPMENT IN SDR FRAMEWORKS

Maurizio Colizza, Marco Faccio, Claudia Rinaldi, Fortunato Santucci
University of L'Aquila, Center of Excellence DEWS
L'Aquila, Italy
e-mail: colizza@westaquila.com, marco.faccio@univaq.it, claudia.rinaldi@univaq.it, fortunato.santucci@univaq.it

ABSTRACT

The increasing interest in software defined radio (SDR) as enabling technology for defining and developing advanced wireless systems, e.g. mobile ad-hoc networks (MANET) with high degree of adaptivity and ability to (re)configure in application scenarios, motivates research efforts in developing methods and tools for supporting a complete and sound design flow, that encompasses i) waveform/protocol specification, ii) thorough validation through accurate simulations and early stage testing, and then iii) rapid code development on selected target platform. This paper proposes a SW architecture which derives from the application of a new methodology fors synthesis, design and analysis, called Tissue Methodology. The proposed architecture aims at reducing the development time through the use of reconfigurable SW components and the application of automatic code generation techniques.

Key words : Architecture, Methodology, Automatic code generation.

1. INTRODUCTION

The most recent evolutions concerning telecommunication systems have presented the problem of an efficient use of spectral resources. This has pushed the research and industrial communities into the investigation of algorithms for resource dynamical access[7]. The modelling and design of a SDR requires the ability of dynamically configuring a communication system as a function of radio services offered by the environment. Furthermore, the reconfiguration covers all layers of the protocols stack. This high degree of configurability opens new possibilities in terns of services dynamical access from one side, while introducing new problems on the other side due to the dynamicity introduced in the protocol stack.

Despite those needs of high degree of configurability , it can be observed that state-of-the-art approaches offer too complex and heavy architectures resulting in underutilization of their potentialities and still lack significantly in several components, with critical drawbacks in those environments where performance estimation and assessment are particularly challenging (e.g. MANETs). A (non exhaustive) list of current weaknesses in the available tools (e.g. NSx,OPNET, OMNET) can be provided as follows,

1. simulators are not typically conceived to offer the opportunity to reuse the developed code for subsequent implementation on the target device;
2. merging of measurement code and business code is not typically addressed;
3. tools for sound and easy support of tracking projects requirements into the developed code are not available;
4. tools for supporting automatic generation of reports that rely on qualitative and quantitative performance assessment are not satisfactory;
5. high level cross verification for performance analysis (e.g. logic trigger) is not addressed.

With the motivation of bringing improvements in the depicted technical framework, our research group is involved in several projects, e.g. ARTEMIS PRESTO [8] and FP7 NoE HYCON2 [9] both co-funded by the EC. Although the main research problems are different, both projects are conceived with the problem of overcoming limitations due to the current in use technologies. Specifically, the PRESTO project is mainly focused on: 1. the improvement of test-based embedded software development and validation procedure, while considering the constraints of industrial development processes; 2. the definition of functional and performance analysis with platform optimization at an early stage of the development process. The project also intends to explicitly consider some industrial development constraints: simplified use of tools, smooth integration in current design processes, framework of tools that is flexible enough to adapt to different process methodologies, design languages and integration test, platform modeling for early comparison of results with real scenarios and fast prototyping. Through the WP6, HYCON2 also pursues research advances in developing methods and tools for analysis and design in the broad range of complex and networked embedded systems.

The present paper is intended to report on our research activity, that is focused on defining and developing a set of tools (suite) to support the sound design, appropriate verification/test and development of embedded software for SDR systems. Specifically, we are defining a workflow whose qualifying features are as follows: 1. the design of a system or subsystem in a network/protocol stack is model-based; 2. the amount of manually written code (firstly for simulation) is minimized, while the code is usually obtained from the model through a set of procedures for automatic code generation; 3. the probes for measurement may be placed in the model; they can be automatically switched off when the model is used to produce code for target devices; the model holds true independently of the target device; 4. when a probe for measure is selected, the generated track can be automatically added to a technical report. The suite is intended to provide the designer with the abilities of: 1. developing a protocol model through the composition of library components; 2. generating code for simulation and test in a network simulator starting from the model; 3. generating code for a target device, starting from the model; 4. integrating protocol models with application related models, e.g. those encountered in the context of networked control systems. The paper will report on already achieved results in terms of developed models and simulation environments.
The paper is structured as follows: in section 2 we describe the architecture proposed, in section 3 we present the requirements arising from the tissue ide (Integrated Development Environment) assumed, while section 4 is related to the application of tissue pattern for protocol designed, a case-study for IEEE 802.15.4 physical layer is investigated in section 5 and finally section 6 concludes the paper and discusses future works.

2. ARCHITECTURE DEFINITION

In order to introduce improvements with respect to limitations enumerated in the previous section, this section proposes an implementation of a new methodological approach, namely Tissue Methodology[1].
The methodology proposed in [1] emphasizes the following modeling paradigms:

1. modular programming [3], [4], [5];
2. patterns programming ;
3. events oriented programming [6];
4. fractal programming [2].

The design pattern used in this methodology was named as Tissue Pattern. The Tissue pattern has the aim of enabling the design by using a basic module which is able to :

1. receive and generate events (H);
2. process events (P);
3. storing a state space or other information (S);
4. increase their “skills” through interaction with other units or through a reconfiguration;

The growth of tissues is achieved through the repetition of basic units, as well as the fractal structures; the link between H, S and P is represented by functional calls, an access to remote resources, or any communication protocol. The basic units can be used to build macro structures that can be in turn used for growing a tissue. Following this approach, a protocol stack can be rethought as show in Figure 1,.
It show a model made up of a basic tissue pattern; each of the H modules receives events from the other layers and, at the same time, it generates events towards the others layers. Each event, in each layer, is processed from the P module. The data exchanged between two layers, or the data needed to a P module, are stored in the S module.

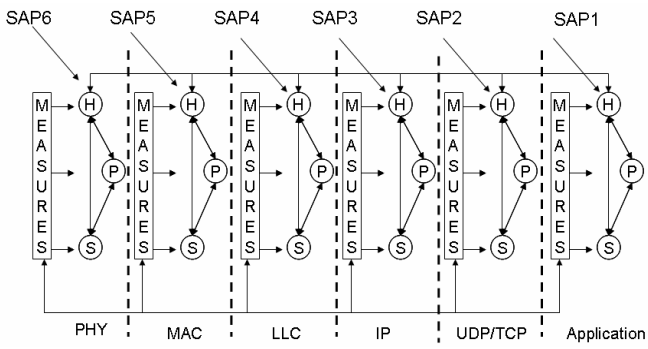


Figure 1 This figure show an example of how to use the basic tissue pattern to model a protocol stack. Specifically, the SAPs (Service Access Points) between two layers is modelled through the H module.

ID	handle
ID1	handle1
IDN	handleN

Figure 2 Data set.

It is worthwhile mentioning that each S module represents a system to store data (e.g. a bank of memories, a remote data source). The P module, or the H module, may retrieve a specific data set, through an identification code. Through this code, the module, which needs to use the data set, receives an handle; this handle enables the use of the data set in read/write mode, Figure 2. The data set could be composed by basic types (e.g. boolean, int, double) or structured data type. If the designer defines all data types needed by the project, the S module may be implemented through automatic code generation. Moreover, using standardized data structures, or rather, data structures which were obtained from predefined data structures, the measure code may be generated in automatic mode too and may be automatically switched off when the model is used to produce code for the target device. Along this vision, it is important to define :

1. a technique to exchange information between different layers, or rather, between different modules of type H;
2. a technique to index of each data set;
3. a technique to manage events generation and the processing.

In the next sections we will suggest a SW implementation for these techniques.

3. REQUIREMENTS OF THE TISSUE IDE

This section deals with the definition of requirements needed to implement a system through the exploitation of the Tissue Methodology. The first step consists in isolating the main characteristics of each paradigm that is at the basis of the Tissue Methodology. The modular programming is the first concept to be taken into account.
In order to be able to build a system through input/output functionalities and memory partitioning, the environment has to provide a support for the creation of the module, that has to be supplied with input/output ports for receiving and generating events (Req.1). Moreover, each module needs to provide an handle through which it is possible to interact with it (Req.2).
In order to exchange events, a communication protocol (Req.3) is required. One of the possible protocols to be used is the Message Passing Interface (MPI); this protocol is particularly interesting because of the existence of a version for real time systems, named Real Time MPI. This is very important for our purposes because the environment we want propose has to allow the designer to simulate the architecture that is going to be implemented on the target device (Req.4). With this design choice, starting from architectural models, it is possible to automates, through automatic code generation, the implementation of the model for simulation activities, or for prototyping activities, in order to avoid any porting activity, any redesign activity, any new testing activity (Req.5).
Basing on previously described requirements, the events simulator OMNET++ has been chosen.
The next section is devoted to describe the Sequence diagram of the design pattern which is used in Figure 1.

4. TISSUE PATTERN FOR PROTOCOLS DESIGN

The issue of designing a protocol layer using tissue methodology, can be solved by the use of the tissue pattern shown in Figure 3.
This pattern represents the basic configuration to build a system where the tasks to deal with are shared among different entities, that can be classified in logical layers. However, if a cross layer interaction is needed, the architecture depicted in Figure 1 inherently supports this interaction. This concept can be explained with the help of Figure 4.

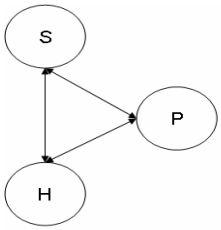


Figure 3 Basic Tissue Pattern.

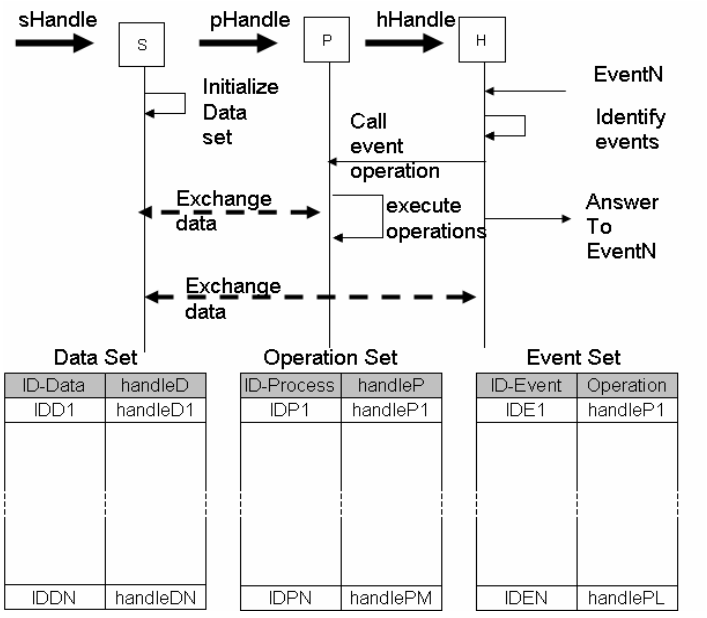


Figure 4 Sequence Diagram for a Basic Tissue Pattern.

Referring to Figure 4, it has to be observed that in addition to the Data Set there are: the operation set,containing the handles for each functionality, and the event set, containing the handles to the events. Each data set is characterized by two coordinates, the handle of the reference storage, sHandle, and the handle of the specific data, handleDk, with k referring to indices data types in range [1,N]. for the data set, [1,M] for the operation set, [1,L] for the event set.
This choice brings to the following advantages :

1. all the data set can be moved from one storage module to another;
2. any data set can be handled from any P module or H module; indeed it is possible to use the pair (sHandle,handleDk) from any module;

Moreover, the exploitation of previous advantages facilitates measurement operations; this way another requirement of our IDE (Integrated Development Environment) is fullfilled.
Before going into the development of techniques to automate the generation of the code through which implementing the tissue pattern, and hence the protocol stack, it is necessary to verify the feasibility of these techniques. Next section proposes an application of the tissue pattern to a 802.15.4 physical layer.

5. APPLICATION OF THE TISSUE PATTERN TO IMPLEMENT 802.15.4 PHYSICAL LAYER

In order to verify the feasibility of this methodology and to spotlight the differences with a traditional design, like OOP (Object Oriented Programming), we started from an existing project and we redesigned this project through the reusing the existing code in order to produce a Tissue Methodology compliant implementation. The project chosen is part of a framework developed to simulate MANET networks through the use of the OMNET++. The framework is named INETMANET; we focused on the physical layer 802.15.4 implementation. The process followed to do this conversion includes the following steps:

1. definition of data types to cover all the data managed into the phy layer;
2. association of a unique identification code to each data type;
3. association of a unique handle to each data type;

Figure 5 shows the identification codes, while Figure 6 shows the implementation of the handles for each data type, basing on the “map” data structure of C++ STL (Standard Library). The following methods have been implemented to manage data types:

1. virtual void* select802154Data(const char* data,int* typeData,wrapper_t tW): it returns the handle to specified through the typeData ID;.
2. virtual void set802154Data(const char* data,int* typeData,wrapper_t tW,void* dataMP): it adds a new data structure

in order to retrieve the handle of the storage module, the needed methods are :

1. cModule*hs802154PHY=(getParentModule()->getSubmodule("sphy"));
2. ::S802154PHY*hS802154PHY=check_and_cast<S802154PHY *>(hs802154PHY);

This is a way to satisfy the requirement Req.2; in order to meet Req.3, the basic functionalities of the H module are listed below :

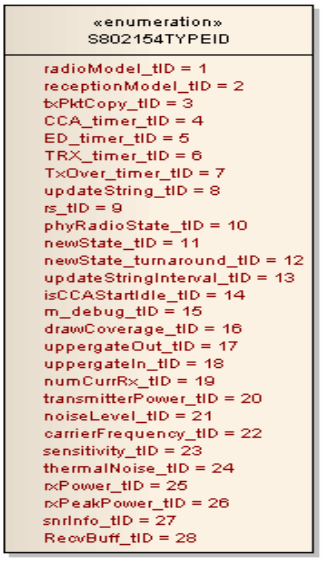


Figure 5 Data ID list.

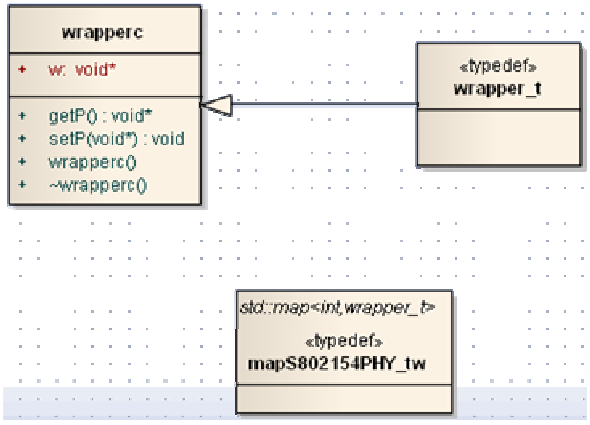


Figure 6 Class Diagram for the generic set.

1. virtual void fCSend(cMessage* msg,int idGate,int sel,simtime_t t); it is needed to control the generation of events in the H module;
2. virtual void fCSelfMsg(simtime_t t,cMessage* msg); it is needed to set internal events (e.g. Timer);
3. virtual void fCancEvent(cMessage* msg,int sel); it is needed to cancel an event which is expired or that was processed;
4. virtual void deleteSelfMsg(cMessage* msg); it is needed to cancel an internal event which is expired or that was processed;

These methods are necessary since the H module has to generate events; however, the P module, as a result of a processing, could need to generate an event.

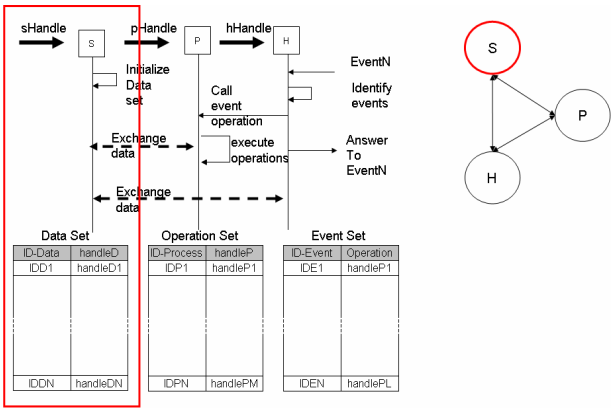


Figure 7

To do this, the P module has to be able to let the event to the H module, in order for the H module to generate the event towards the destination.
When an event is received on the H interface, the H module ask the P module for the execution of one of the following operations :

1. updateDisplayString(*drawCoverage,*sensitivity,*transmitterPower,updateString,*updateStringInterval);
2. handlePrimitive(msg->getKind(), msg) : it is useful to manage the primitives exchange between the 802.15.4 physical layer and 802.15.4 mac layer;
3. handleUpperMsg(airframe) : it is useful to manage the messages originated from the mac layer;
4. handleSelfMsg(msg) : it is useful to manage the internal messages;
5. handleLowerMsgStart(airframe);
6. bufferMsg(airframe) : it is useful to manage the queue of the air frames Protocol Data Units;

All these methods are placed in the P module. This closes the implementation of the basic tissue pattern of Figure 4. Now, basing on Software Defined Radio paradigms, the reprogramming of a device is to be taken into account. Using tissue architecture, the reprogramming could be implemented through the execution of the following steps :

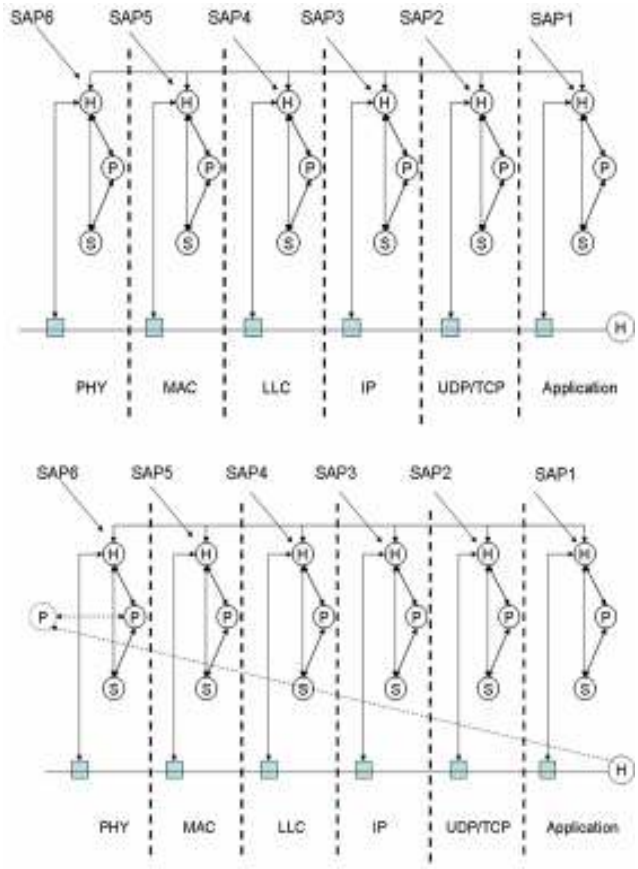


Figure 8 Example of dynamic tissue pattern reconfiguration.

1. if the generic H module is not able to identify an operation request for its P module, it sends a notification of not recognized event to another H module, devoted to recognize this type of events. This specific module, is trained to process new events through the use of an xml file, which contains the information about the type of event and the type of P module (part of a collection of P modules) which has to be instantiated.

2. The H module that has to manage unrecognized messages can instantiate a new P module, specifically designed for the new types of event, and link the new P module with the old P Module, as depicted in Figure 8. Moreover, the old H module is updated;

6. CONCLUSIONS AND FUTURE WORK

In this paper we analyzed the implementation requirements needed of a new methodology, named Tissue Methodology; moreover, an environment which satisfies these requirements was selected, specifically OMNET++. In the context of this environment, a basic tissue pattern was developed and exploited to study the applicability of the tissue methodology to implement an IEEE 802.15.4 phy layer model.

The future works, already in progress, regards the development of techniques to automate the generation of the code for S, H and P modules for the basic tissue pattern used for IEEE 802.15.4 PHY layer, and to extend the basic tissue pattern to others layers of the protocols stack, specifically 802.15.4 MAC layer.

7. REFERENCES

[1] M. Colizza, M. Faccio, C. Rinaldi, F. Santucci, " A METHODOLOGY TO DESIGN AN ADVANCED FRAMEWORK FOR EFFICIENT MODELLING AND TESTING OF MANETS", in Proc. of Wireless Telecommunication Synmposium, IEEE April 2012, to appear

[2] http://fractal.ow2.org/documentation.html

[3] K.K. Lau and Z. Wang. Software Component Models. IEEE Trans. Software Eng., vol. 33 n° 10 , October 2007.

[4] S. Sicard, F. Boyer, and N. De Palma. Using Components for Architecture-Based Management: The Self-Repair Case. in Proc. of 30th International Conference on Software Engineering (ICSE 2008). ACM, 2008, ISBN: 978-1-60558-079-1

[5] P.C. David, M. L´eger, H. Grall, T. Ledoux, and T. Coupaye. A Multi-stage Approach for Reliable Dynamic Reconfigurations of Component-Based Systems. In 8th IFIP Int. Conf. Distributed Applications and Interoperable Systems, DAIS 2008, volume 5053 of LNCS, 2008.

[6] K. Mani Chandy, Michel. Charpentier, Agostino Capponi, Towards a Theory of Events DEBS '07, June 20–22, 2007 Toronto, Ontario, Canada

[7] III Mitola, J. and Jr. Maguire, G.Q., "Cognitive radio: making software radios more personal," Personal Communications, IEEE, vol. 6, no. 4, pp. 13 –18, aug 1999.

[8] http://www.presto-embedded.eu/

[9] http://www.hycon2.eu/

COMPARISON OF CONTENTION-BASED PROTOCOLS FOR SECONDARY ACCESS IN TV WHITESPACES

Richard MacKenzie and Keith Briggs (BT Innovate & Design, Adastral Park, UK;
richard.mackenzie@bt.com, keith.briggs@bt.com)

ABSTRACT

The performance of the contention-based protocols in 802.11 and ECMA-392 are described and analyzed. Their suitability for various scenarios in TV whitespaces are evaluated. We demonstrate that by adjusting a single parameter over a limited range of values, a high throughput can be maintained. At the same time we can limit the effects of aggregate interference (where multiple secondary stations transmit on the channel at the same time), which could potentially cause interference to primary systems. We describe how the backoff behaviour of both protocols can be compared and evaluated using Markov chains. A description of an extremely fast and efficient way to solve these large and complex chains is given.

1. INTRODUCTION

TV whitespaces (TVWS) refer to one of the spectrum bands that are early candidates for dynamic spectrum access (DSA). DSA improves spectral efficiency by allowing for unlicensed/secondary users to opportunistically access spectrum while protecting licensed/primary users. In the case of TVWS the primary users can be TV broadcasters and wireless microphones.

Regulatory approval of TVWS is progressing rapidly. In the USA, the FCC has released a document which determines the final rules for the use of TVWS [1]. These new rules remove mandatory sensing requirements, thus facilitating the use of geolocation-based channel allocation. In the UK, Ofcom is consulting on a draft Statutory Instrument to make whitespace devices licence-exempt. Ofcom also plans to work with stakeholders to make information about existing licensed services that operate in the TV band available to prospective database providers, and expects that TVWS technology could be launched in the UK in 2013 [2].

While primary systems are protected from interference from secondary systems, secondary systems themselves must be able to coexist with one another; however, the rules and etiquette methods to allow for this are still being developed. Contention-based protocols using random backoff mechanisms are therefore attractive for users, especially early adopters, of TVWS due to the ease of coexistence with other systems. This can be seen from the success of contention-based access in other shared spectrum such as the 2.4GHz ISM band. The ECMA-392 standard [3] is already released and defines TVWS channel access mechanisms including prioritized channel access (PCA) for contention-based channel access. The 802.11 standard [4] can provide prioritized contention-based access using enhanced distributed channel access (EDCA). IEEE 802.11 task group AF are currently developing the modifications to this standard to allow for co-existence in TVWS. It is expected that 802.11af will include a contention-based mechanism with similar behaviour to EDCA. Test-bed implementations of ECMA-392 and 802.11 in TVWS are already available as shown in [5] and [6] respectively which suggests both could be early adopters of TVWS.

There are many potential scenarios that could benefit from opportunistic channel access; examples are provided in [7] and [8]. Currently there is a focus towards opportunistic channel access in TVWS as this band is becoming available in the near future, but also because of the favourable propagation characteristics of the TV band. This allows for relatively good coverage using relatively low transmit powers. Example scenarios that could use TVWS include home networking and indoor-to-outdoor coverage. Indoor-to-outdoor coverage describes the coverage of users on streets from access points within buildings.

In this paper the behaviour and performance of the backoff mechanisms of the 802.11 and ECMA-392 protocols are compared. Although the backoff mechanisms are similar there is a key difference in the way that contention window values are reset which can cause significantly different performance. Based on the performance results the suitability of each of these protocols is judged for different deployment scenarios. The performance of these backoff schemes is analyzed using Markov chain analysis. These sorts of Markov chains can be extremely large and difficult to solve so a description is provided in this paper of a highly efficient way to solve Markov chains which could be used for more complex systems than shown here.

Parameter adjustments are also investigated in this paper to show how high throughputs can be maintained and how aggregate interference can be kept low. Aggregate interference must not exceed the interference thresholds of any primary users. The number of simultaneous transmissions from secondary systems should therefore be limited to avoid these issues. The solution to aggregate interference may also involve further mechanisms such as power control, an example of which is described in [9], to make sure that secondary systems do not interfere with the primary users.

This is the first paper to give a detailed comparison of the behaviours of the 802.11 and ECMA-392 contention-based mechanisms. This is motivated by feedback from discussions which the authors have had following presentations which gave only a brief comparison of these issues [10], [11].

This paper is organized as follows. Section 2 provides an overview of both the 802.11 EDCA and ECMA-392 PCA access mechanisms while section 3 describes the Markov chains used and how to solve them. Results are provided and discussed in section 5 followed by conclusions in section 6.

2. PROTOCOL OVERVIEWS

Both 802.11 EDCA and ECMA-392 PCA use carrier-sense multiple access with collision avoidance (CSMA/CA) for contention-based access. First a description is provided for EDCA followed by the description for PCA. The performance of the two protocols is then compared.

When a packet arrives at an 802.11 EDCA station it is mapped into one of four access categories (AC). Each AC contends for the channel using its own channel access function (CAF). Each CAF has its own parameter set which includes CW_{min}, CW_{max}, *AIFSN* and $TXOP_{\text{limit}}$. When a packet first arrives at a CAF, the CAF will first sense the channel. If the channel is sensed idle for an arbitration interframe space (AIFS), which is one short interframe space (SIFS) plus *AIFSN* timeslots, then the packet will be transmitted. If, however, the channel is initially sensed busy, or becomes busy during AIFS, then the backoff procedure is invoked. The CAF will have a contention window (CW) with a value of *CW*. Initially CW is set to CW_{min}. A random backoff time is then selected from the range [0,*CW*]. After the CAF senses the channel to be idle for a duration of AIFS, the system will countdown its backoff for each idle timeslot. If the channel becomes busy the backoff will freeze and will continue again once the channel has been idle for a further AIFS duration. Once the backoff counter has reached zero, the CAF can transmit the packet. If more than one CAF at the same station attempt to transmit at the same time then the highest priority of those CAFs is allowed to transmit on the channel while the other CAFs assume failed transmissions due to this internal collision. If the intended recipient station successfully receives a transmitted packet then it will send an acknowledgement back to the sender station following a SIFS interval. If an acknowledgement is not received the sender station assumes that its transmission was unsuccessful. Following an unsuccessful transmission attempt the CAF increases *CW* and selects a new random delay for another backoff before attempting a retransmission. *CW* is incremented as one less than powers of two until CW reaches CW_{max}. CW remains at this value until it is reset to CW_{min}. When a transmission is successful the CAF has become the TXOP holder. This means that the TXOP holder can undergo multiple frame exchanges, separated by SIFS, so long as the total duration of the TXOP does not exceed $TXOP_{\text{limit}}$. Following a TXOP where the final transmission was successful *CW* is reset. It can also be reset once a retry limit is reached. Also, following a TXOP where the final transmission was successful the backoff procedure shall be invoked once more to reduce the probability of a packet collision.

The request to send/clear to send (RTS/CTS) mechanism is an optional feature that can be used in a frame exchange sequence. Here, the transmitting CAF first sends an RTS frame which describes the time required for the rest of the frame exchange. If this is successfully received the receiving station returns a CTS frame which also contains timing information on the frame exchange. Any neighbouring stations that hear either the RTS or CTS frame now know about the rest of the frame exchange sequence so refrain from transmitting during this time. This mechanism is particularly useful when the frame being sent is large and/or when there are hidden station issues (i.e. not all stations in the network can hear one another and so carrier-sense is not reliable).

The above description uses immediate positive acknowledgements although it is possible to send packets that do not use acknowledgements. Also block acknowledgements can be used to improve efficiency.

The ECMA-392 standard defines a superframe structure. Within this superframe is a beacon period at the start and a contention signalling window at the end. There are also two additional optional windows; a reservation-based signalling window, and a quiet period. The rest of the superframe is the data transfer period (DTP). Within the DTP both channel reservation access (CRA) and prioritized contention access (PCA) are permitted. More information on the various components of the superframe can be found in [3]. In this paper we just focus on the behaviour of PCA.

PCA behaves in a very similar way to 802.11 EDCA. Packets are mapped into ACs and have the same parameter set list, although parameter values may not be the same. TXOPs are contended for in the same manner. However, the backoff rules have some slight differences from those of 802.11 and so are explained here.

The rules for adjusting CW and invoking the backoff procedure are described in section 7.5.1.7 of the ECMA-392 standard. There is a general rule for updating CW which states that following a successful frame transmission a station will reset CW to CW_{min}. This agrees with the rules of 802.11. However there are some specific rules, A to F, which determine how CW is adjusted when invoking the backoff procedure. In the event that both the general rule and the specific rules would both modify CW, then the specific rule is applied instead of the general rule (Note: This section of ECMA-392 can be a little ambiguous but this interpretation has been confirmed from discussions with the editor of the standard). This is only the case for rules B and

C which apply following a successful transmission when that transmission is the final frame exchange in the TXOP. Rule B applies when the current CAF has no further frames in its buffer to send; in this case backoff is invoked with CW reset to CW_{min}. Rule C applies when the CAF still has frames in its buffer but the TXOP is not long enough for further frame transactions; in this case backoff is invoked with CW remaining at its current value (i.e. CW is not reset). These rules, B and C, show that ECMA-392 stations are using the current buffer status to determine whether or not to reset CW. During high loads, CW is reset less, resulting in fewer transmission attempts. In other words, during a high load these rules aim to reduce network congestion (collisions). The aim of this paper is to demonstrate how these rules can cause an ECMA-392 PCA-type system to exhibit significantly different behaviour to an 802.11 EDCA-type system.

There are ways that an ECMA-392 PCA system can behave more like an 802.11 EDCA system. First of all, as explained above, rules B and C only apply for transmissions that are the CAF's final frame transmission in a TXOP. This means that if a TXOP contains more than one frame exchange CW will already have been reset before the final frame exchange in the TXOP takes place. ECMA-392 currently states fixed channel access parameters and the TXOP values are relatively short compared to 802.11 (NB: in 802.11 the recommended value for best-effort and background ACs is actually zero but this corresponds to a single frame exchange. Also, in 802.11 the TXOP values for any AC can be updated by parameter set updates). For data transmissions, at most PHY transmission rates, this will mean each TXOP is only large enough for a single frame exchange. These relatively short TXOP values allow ECMA-392 to share bandwidth more fairly on a short timescale. However, if an ECMA-392 system wished to behave in a more aggressive manner, it could use fragmentation to allow for multiple frames per TXOP at the expense of the extra overheads associated with fragmentation. Conversely, if an ECMA-392 system wished to behave in a more conservative manner, it could try ensure that each TXOP only contains one frame transmission; this could involve frame aggregation for example. Another way to make an ECMA-392 PCA system behave more like an 802.11 EDCA system is to use the RTS/CTS mechanism. This way, even when a TXOP only has one data frame to send, more than one frame will actually be transmitted. As a result CW will be reset following the successful transmission of the RTS frame.

For this paper we compare the more conservative ECMA-392 PCA-type performance (i.e. only resetting CW following a successful TXOP when that CAF's buffer is empty) with the more aggressive 802.11 EDCA-type performance (i.e. resetting CW after every successful TXOP). For this to be the case, and for the comparison of mechanisms

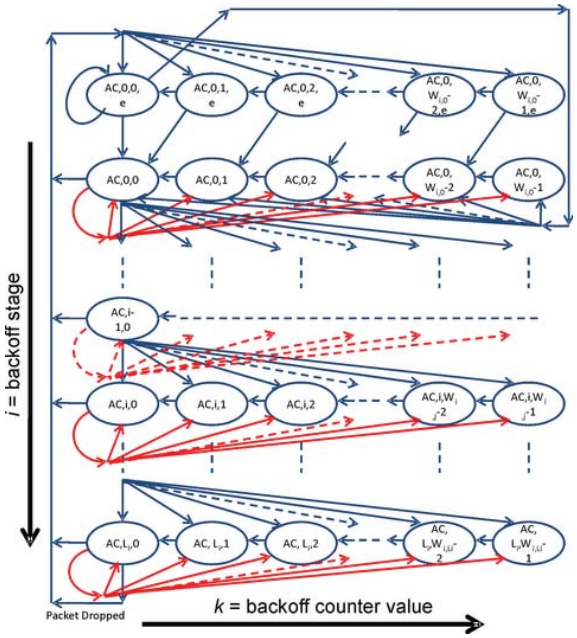


Fig. 1. Markov chain to compare the backoff behaviour of the 802.11 EDCA (blue transitions only) and ECMA-392 PCA (red and blue transitions) protocols.

to be fair, we make several assumptions:

- Each TXOP only contains one data packet. The effect is that an ECMA-392 PCA-type system will only reset CW following a successful TXOP when that AC has an empty buffer; otherwise CW will remain at its current value.
- For the ECMA-392 PCA-type system we assume PCA is used in a continuous DTP. This way overheads associated with the ECMA-392 superframe structure, where PCA would not be contending for the channel, are ignored. Likewise, we assume that an 802.11 EDCA-type system is always able to contend for the channel.
- The same physical layer is used under both MAC mechanisms so that we are comparing the MAC performance fairly.

Markov chains are one of the best ways to analyze the performance of EDCA. Figure 1 shows a Markov chain used in [12] for EDCA. This is one of the more advanced models as it considers all of the EDCA parameters, multiple ACs and non-saturated loads. The state transmission probabilities are not shown as this diagram is only used to provide a visual comparison of how the EDCA and PCA backoff mechanisms differ. The states are labelled (AC, i, k) where AC is the CAF being modelled, i is the backoff level and k is the current backoff value. The states labelled (AC, i, k, e) represent states where backoff is performed while a CAF queue is empty. States with values of $k = 0$ are the

transmitting states. The blue lines reflect the transitions available in EDCA, while PCA also includes the added red lines. Following a successful TXOP EDCA will select a new backoff value in the top level in the Markov chain ($i = 0$) whereas PCA will only do this if its transmission queue is currently empty (or the TXOP contains more than one frame exchange). Otherwise PCA will remain at the same level in the Markov chain to start the backoff procedure for the next transmission.

3. EFFECTIVE MARKOV CHAIN SOLVING

Everything we compute is derived from a *transition matrix* P, so that P_{ij} is the probability of moving from state i to state j. (In reality states have labels which are integer tuples $(i_0, i_1, \dots)$, but we ignore this complication here.) We need the *equilibrium vector* z, which is a solution of $z^{\mathrm{T}}(I-P) = 0$. Properties of P which must be considered are: the transition matrix P is large, but *sparse*; the transition matrix P is *asymmetric*; and the system of linear equations to be solved is singular and need some additional normalization condition, such as is provided by normalizing z, so that $||z|| = 1$. These properties create difficulties. There can be thousands of states, and then the full P matrix cannot be stored, and in all cases it cannot be manipulated in dense fashion. We found by experience that the best results were obtained by using the Super LU package [13]. We implemented a mapping of integer tuples to integer indices (and its inverse) using C++ hash mappings. The result is a convenient library for solving large Markov chains, with the user input in a natural form and internal re-indexing hidden from the user. The singularity problem is handled by replacing one row of $I-P$ with the equation $||z|| = 1$.

Some of the simpler cases of Markov chain models of wireless backoff have exact analytic solutions. We claim that this is of little help in practice; it is hard to check whether an analytic solution is possible, and if it is, it is easy to make mistakes deriving it. We believe that a single numerical technique, applicable to all cases, such as ours, is better in general.

A particular example of interest is the Bianchi model [14]. This has parameters: W is the minimum contention window plus one; m is the value such that $2^m W$ is the maximum contention window plus one; and p is the packet collision probability, computed from the number of users as shown below. Such a model has $(2^{m+1}-1)W$ states. The states of the system form $m+1$ downward-going "escalators" E_0, E_1, …, E_m of heights W_0, W_1, …, W_m where $W_i = 2^i W$. The states are labelled by (i,k) where $i \in \{0,1,\dots,m\}$ labels the backoff stage, and $k \in \{0,1,\dots,W_i-1\}$ is the backoff time counter (height on that escalator). The transmitting states are $(i,0)$ for $i \in \{0,1,\dots,m\}$ so the probability of a transmission, τ, is the sum of the probabilities of being in these transmitting states. The dynamics for a saturated 802.11 CAF may be summarized by these rules: From (i,k) with $k \geqslant 1$, move with probability one to state $(i,k-1)$, one step down E_i. From $(i,0)$ (the bottom of E_i) jump with probability $1-p$ to a random point on escalator E_0. From $(i,0)$ jump with probability p to a random point on escalator $E_{\min(i+1,m)}$.

In formulae,

$$\begin{aligned}
p(i,k;i,k-1) &= 1 \quad \forall i \text{ and } 1 \leqslant k \leqslant W_i-1,\\
p(i,0;0,k) &= \frac{1-p}{W_0} \quad \forall i \text{ and } 0 \leqslant k \leqslant W_0-1,\\
p(i,0;i+1,k) &= \frac{p}{W_{i+1}} \text{ for } 0 \leqslant i \leqslant m-1 \text{ and}\\
&\quad\ 0 \leqslant k \leqslant W_{i+1}-1,\\
p(m,0;m,k) &= \frac{p}{W_m} \text{ for } 0 \leqslant k \leqslant W_m-1,
\end{aligned}$$

where we use $p(i,k;j,l)$ to denote the 1-step transition probability from (i,k) to (j,l).

These rules are adjusted to the following in order to apply to a saturated CAF with ECMA-392 PCA backoff,

$$\begin{aligned}
p(i,k;i,k-1) &= 1 \quad \forall i \text{ and } 1 \leqslant k \leqslant W_i-1,\\
p(i,0;i,k) &= \frac{1-p}{W_i} \quad \forall i \text{ and } 0 \leqslant k \leqslant W_i-1,\\
p(i,0;i+1,k) &= \frac{p}{W_{i+1}} \text{ for } 0 \leqslant i \leqslant m-1 \text{ and}\\
&\quad\ 0 \leqslant k \leqslant W_{i+1}-1,\\
p(m,0;m,k) &= \frac{p}{W_m} \text{ for } 0 \leqslant k \leqslant W_m-1,
\end{aligned}$$

This results in a saturated ECMA-392 Markov chain never jumping back down to escalator E_0 and becoming stuck in escalator E_m. This explains the conservative behaviour of the backoff mechanism during high loads.

The solution is also required of a nonlinear equation. The packet collision probability p is related to the number of stations n, the probability P_{tr} of a transmission in a particular timeslot, the probability P_{s} of a transmission being successful, and throughput S by the following equations [14]:

$$\begin{aligned}
p &= 1-(1-\tau)^{n-1}\\
P_{\mathrm{tr}} &= 1-(1-\tau)^{n}\\
P_{\mathrm{s}} &= \frac{n\tau(1-\tau)^{n-1}}{1-(1-\tau)^{n}}\\
S &= \frac{P_{\mathrm{s}}P_{\mathrm{tr}}E[P]}{(1-P_{\mathrm{tr}})\sigma+P_{\mathrm{tr}}P_{\mathrm{s}}T_{\mathrm{s}}+P_{\mathrm{tr}}(1-P_{\mathrm{s}})T_{\mathrm{c}}}
\end{aligned}$$

where $E[P]$ is the average packet payload, σ the timeslot duration, T_{s} and T_{c} the duration of a successful transmission and a collision respectively. Solving the top equation requires an iterative method. An initial estimate of p is used; using this to solve the Markov chain, τ can be calculated. p must then be modified and the Markov chain re-solved until a value of τ is produced that allows for the above equation to be satisfied. For the adjustment of p at each iteration we used the multiroot solver from the GSL library [15].

TABLE 1
TEST PARAMETERS

Access Type	Data Rate (Mbps)	$E[P]$ (μs)	σ (μs)	T_s (μs)	T_c (μs)
Basic	31.65	379	9	490	490
RTS/CTS	31.65	379	9	577	106

4. AGGREGATE INTERFERENCE

During a collision more than one station will attempt to transmit at the same time. These simultaneous transmissions can cause aggregate interference. In this paper we wish to examine the aggregate interference issue by examining how many transmissions are likely to be involved in each collision. We use $\Pr[NTX = x]$ to represent the probability that when there is a transmission attempt there are x stations simultaneously attempting transmission. We can generalize the equation for P_s as the probability that when a transmission occurs only one station attempts transmission (i.e. $\Pr[NTX = 1]$). So $\Pr[NTX = x]$ can be calculated as

$$\Pr[NTX = x] = \frac{\binom{n}{x}\tau^x(1-\tau)^{n-x}}{1-(1-\tau)^n}.$$

5. TEST RESULTS

In this section results are shown to compare the behaviours of 802.11 EDCA and ECMA-392 PCA. Ways to maintain high throughputs using parameter adjustment are investigated while also evaluating the aggregate interference performance.

The above Markov chain analysis is used to provide the analytical results. Further validation is provided by a modified version of the *wlan_mac_hcf* process model from the Opnet Modeler Wireless Suite 16.0 [16]. The physical layer chosen is that of ECMA-392 for an 8 MHz channel. This offers a maximum transmission rate of 31.65Mbps. For each test all stations are saturated with packets that have 1500 byte MSDUs. As mentioned earlier, we assume that each TXOP only contains one data packet and both systems are able to contend for the channel all of the time (i.e. no overheads currently considered such as quiet periods for sensing.). We also assume that channel access parameters are adjustable, which is not currently the case for ECMA-392. So when comparing the behaviours of 802.11 EDCA and ECMA-392 PCA we refer to the behaviour as 802.11 EDCA-type and ECMA-392 PCA-type behaviours respectively. The parameters used for the Markov analysis are shown in Table 1.

For this first set of results we compare 802.11 EDCA-type and ECMA-392 PCA-type behaviours using the same parameter set values. These are CW_{min} =15 and CW_{max} =1023, which equate to W=16 and m=6 for the Markov chains used; *AIFSN*=2; *TXOP*=0 (i.e., one frame exchange per contention) and a timeslot duration, σ=9μs. The system performance is evaluated using basic access and also with the RTS/CTS mechanism. It is important to note that when RTS/CTS is used an ECMA-392 PCA system, according to the rules in the standard, would reset CW during a successful frame exchange (as the RTS frame would be seen as a successful frame transmission that is not the final frame transmission in the TXOP). For these tests this would result in the ECMA-392 PCA-type system and the 802.11 EDCA-type system behaving the same as each other when the RTS/CTS mechanism is used. For this first set of results we show the ECMA-392 PCA-type performance with RTS/CTS if CW is only ever reset when the AC has an empty buffer. The reason for this is that we wish to compare conservative and aggressive backoff mechanisms in general rather than simply limiting our choice to existing 802.11 and ECMA-392 specifications. This is also a reason for using the terminology 'ECMA-392 PCA-type' and '802.11 EDCA-type'.

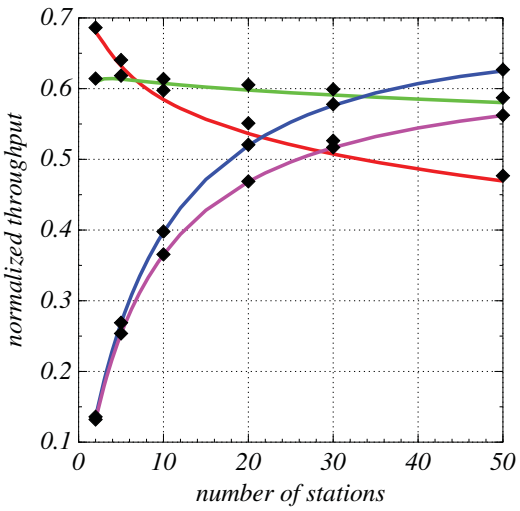


Fig. 2. System capacity results. 802.11 EDCA-type system (red: basic, green: RTS) & ECMA-392 PCA-type system (blue: basic, magenta: RTS). Black diamond=simulation.

Figure 2 shows the performance of the systems under test for networks of varying size. The first thing to note is that, for a network with few stations, an 802.11 EDCA-type system provides a high throughput while the throughput of an ECMA-392 PCA-type system is very low. As the number of stations increases the performance of an ECMA-392 PCA-type system improves while the 802.11 EDCA-type system performance gradually degrades. The best performing mechanism is the 802.11 EDCA-type system with the RTS/CTS mechanism. For the full range of network

size tested this mechanism achieves around 60% MAC layer efficiency. For a very small network the 802.11 EDCA-type system performs better without the RTS/CTS mechanism due to the increased overheads in a successful RTS/CTS frame exchange sequence. However as the contention rises for larger networks this overhead is compensated for by the much reduced time spent on collisions (It is worth noting at this point that while the RTS/CTS mechanism reduces time spent on collisions in a congested network it does not reduce the probability of collisions and therefore the aggregate interference issue is not necessarily resolved).

When the network has 50 stations, an ECMA-392 PCA-type system without RTS/CTS shows the best overall performance. Even at this point the overheads in successful RTS/CTS frame exchanges are not compensated for by the reduced overheads in collisions. This tells us that even for a large network an ECMA-392 PCA-type system is successful at avoiding collisions. This conservative approach suggests that it may be more cooperative with other secondary systems than the more aggressive 802.11 EDCA-type system.

Using these channel access parameters, the 802.11 EDCA-type system is the most suitable protocol to use when the number of active stations is low and each station has high capacity demands. Video distribution around the home is a prime example of this type of scenario. Another suitable scenario is indoor-to-outdoor coverage of the street allowing for Internet access to outdoor terminals. When there are a large number of terminals that each require only a low amount of throughput such as machine-to-machine systems, the ECMA-392 PCA-type system is most preferable; it can offer throughput that can compete with the 802.11 EDCA-type system while at the same time being more cooperative with other secondary systems.

The channel access parameters above compare system behaviour using CW_{min} and CW_{max} values that are quite common. In fact these are the recommended values for best effort traffic in 802.11 EDCA (The recommended value for CW_{min} is physical layer dependent but is typically 15 for high rate physical layers) and the fixed values for best effort traffic in ECMA-392 PCA. However, system performance can be greatly improved for each protocol by adjusting the parameter settings to match the current network scenario. Figure 3 shows how we can modify the 802.11 EDCA-type system performance simply by selecting different values for CW_{min} while fixing CW_{max} =1023. Here, only CW_{min} values of one less than powers of two are used. However, we can see that as the network size varies, different values for CW_{min} allow for the best throughput performance. If the correct value of CW_{min} is selected then a MAC efficiency of around 65% can be maintained. One of the important points here is that only one parameter is being adjusted and it only takes 6 possible values. With a reasonable estimate of the network size a system can achieve high performance using this limited parameter set.

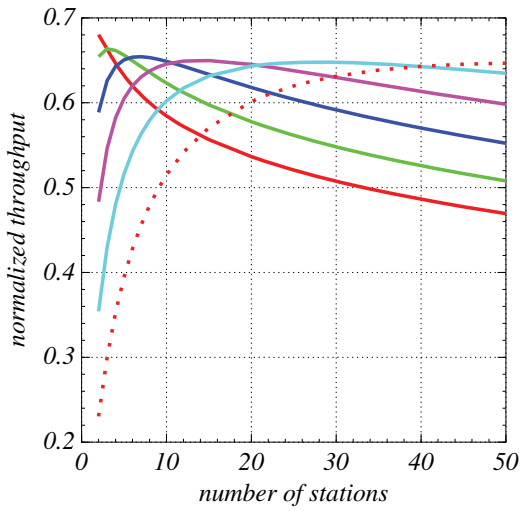


Fig. 3. Adjusting 802.11 EDCA-type system CW_{min} to maintain high throughput. (CW_{min} is red: 15; green: 31; blue: 63; magenta: 127; cyan: 255; dotted red: 511).

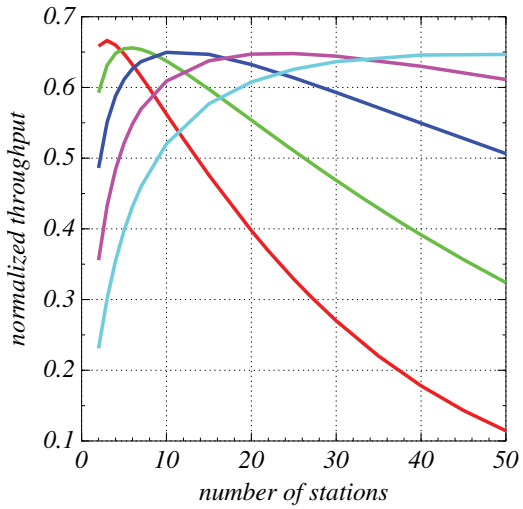


Fig. 4. Adjusting ECMA-392 CW_{max} to maintain high throughput. (CW_{max} is red: 31; green: 63; blue: 127; magenta: 255; cyan: 511).

Figure 4 shows how we can modify the ECMA-392 PCA-type performance simply by selecting different values for CW_{max}. Here CW_{min} =7 (i.e. W=8) and CW_{max} values are always one less than powers of two. As the network size varies different values for CW_{max} allow for the best throughput performance. Again, as seen with the previous set of results, if the correct parameter value is selected then a MAC efficiency of around 65% can be maintained. Here only one variable is adjusted and only takes 5 values to cover

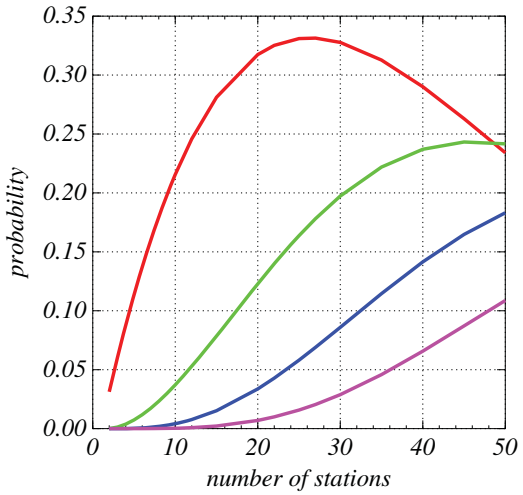


Fig. 5. Collision behaviour of ECMA-392 for CW_{min} =7 and CW_{max} =31. (red: Pr[$NTX = 2$]; green: Pr[$NTX = 3$]; blue: Pr[NTX=4]; magenta: Pr[NTX=5]).

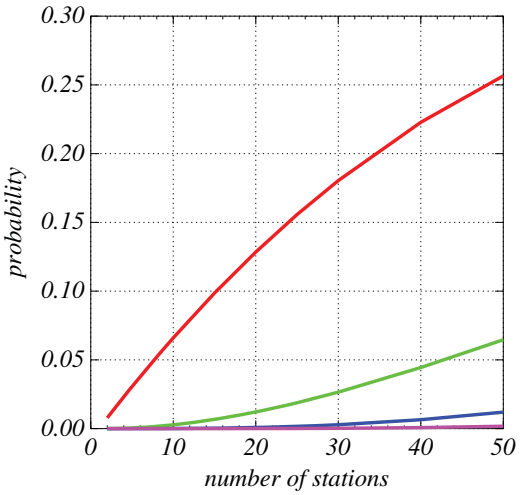


Fig. 6. Collision behaviour of ECMA-392 for CW_{min} =7 and CW_{max} =127. (red: Pr[$NTX = 2$]; green: Pr[$NTX = 3$]; blue: Pr[NTX=4]; magenta: Pr[NTX=5]).

various network sizes up to 50 stations.

Limiting aggregate interference is another reason why adjusting parameters might be preferable. When a contention-based protocol is being too aggressive there is a strong possibility of collisions. As this collision probability increases, so does the probability of more stations being involved in each collision. A high collision probability should be avoided for several reasons: (1) the system will suffer a performance degradation (e.g., drop in throughput) as a result of too many failed transmission attempts; and (2) multiple secondary systems may be using the channel. If one system transmits too often it is may not be sharing the channel very well as other secondary systems cannot gain fair access to it. Also the probability of collisions/interference among secondary systems will increase. Furthermore (3), when multiple secondary transmissions cause a collision the aggregate interference should not be able to interfere with any primary systems. For this final reason we look at the collision probability of secondary systems as the parameter settings are adjusted.

Figure 5 shows the collision performance for an ECMA-392 PCA-type system when CW_{min} =7 and CW_{max} =31 (i.e. W=8 and $m = 2$). It can be seen that when the network is small, so is the collision probability. As the network grows so does the probability of collisions. The most common form of collision is one involving just two transmissions until the network size approaches 50 stations. However, the probability of collisions with more than two transmissions becomes more of a problem as the network size increases. When the network size is around 12 stations, then about 5% of all transmission attempts involve collisions with three simultaneous transmissions. This 5% mark is reached for 4 simultaneous transmissions when the network size is about 23 stations and for 5 simultaneous transmissions when the network size is about 36 stations. When the system size is 50 stations the aggregate interference becomes a serious concern where the probability of a collision is very high and the number of transmissions involved in any collision is now likely to be more than 2.

As seen from the Figure 4, the same parameters that give these collision behaviour results only have a high throughput for a small network size. Using the example in Figure 4, these parameters would only be recommended operating parameters until the system size reaches about 5 stations. When we re-examine the results in Figure 5 for up to a network size of 5 we see that the collision behaviour of the system is quite good. As the network size reaches 5 stations, the probability of a collision with just two transmissions starts to exceed 10% while the probability of collisions with more than two stations is almost negligible. This demonstrates that the benefit of parameter adjustments can both maintain a high throughput whilst also reducing aggregate interference in a system.

To further demonstrate this point, Figure 6 shows the collision performance of the ECMA-392 PCA-type system for the operating parameters which, according to the results shown in Figure 4, would be recommended to maintain a high thoughput for a system size roughly between 9 and 17 stations (i.e. CW_{min} =7, CW_{max} =127 which means W=8, m=4). If we compare Figure 6 to Figure 5 we see that these new parameters are less aggressive and have much lower collision probabilities for higher network sizes. As these

parameters approach the largest network size that they are recommended for in order to maintain a high throughput we see again that the probability of a collision with just two transmissions starts to exceed 10% while the probability of collisions with more than two stations is very low.

6. CONCLUSIONS

In this paper the protocol behaviour of the 802.11 and ECMA-392 contention based mechanisms have been described. A comparison of the performance of these types of contention based mechanisms is provided and analyzed. Using the same parameter set 802.11 type systems are more aggresive than ECMA-392 type systems and achieve higher throughputs for small networks, whereas the more conservative ECMA-392 type systems offer better coexistence with other secondary systems using the same channel and better throughput performance for networks with many terminals. Based on these different characteristics, recommendations have been made concerning which TVWS deployment scenarios each mechanism may be better suited to.

We have also investigated the option to adjust parameter values in these mechanisms. By simply adjusting one parameter over a limited range of values a high thoughput can be maintained over a wide range of network sizes. A further benefit of this parameter adjustment has been demonstrated regarding aggregate interference. When using parameters which maintain a high throughput the collision probability of a system is kept low; when there is a collision it is unlikely to involve more than two simultaneous transmissions, which limits the issues of aggregate interference where the secondary system(s) could interfere with the channel's primary user(s).

7. ACKNOWLEDGEMENTS

The research leading to these results was derived from the European Community's Seventh Framework Programme (FP7) under Grant Agreement number 248454 (QoSMOS).

The authors would also like to give a special thanks to Jianfeng Wang, editor of ECMA-392, for clarifying some of the details about the backoff rules for ECMA-392.

REFERENCES

[1] Federal Communications Commission. (2010) Second memorandum opinion and order. [Online]. Available: http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-10-174A1.pdf

[2] Ofcom. (2011) Implementing geolocation: summary of consultation responses and next steps. [Online]. Available: http://stakeholders.ofcom.org.uk/binaries/consultations/geolocation/statement/statement.pdf

[3] *ECMA-392: MAC and PHY for Operation in TV White Space*, ECMA Std., Dec. 2009.

[4] *IEEE Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Std. 802.11-2007 (Revision of IEEE Std 802.11-1999), 2007.

[5] A. Franklin, J. Pak, H. Jung, S. Kim, S. You, J. Um, S. Lim, G. Ko, S. Hwang, B. Jeong, M. Song, and C. Kim, "Cognitive radio test-bed based on ECMA-392 international standard," in *Wireless Communication Systems (ISWCS), International Symposium on*, Sept. 2010, pp. 1026 –1030.

[6] M. Fitch, M. Nekovee, S. Kawade, K. Briggs, and R. MacKenzie, "Wireless service provision in TV whitespace with cognitive radio technology: a telecom operator's perspective and experience," *IEEE Commun. Mag.*, Mar. 2011.

[7] J. Wang, M. Ghosh, and K. Challapali, "Emerging cognitive radio applications: A survey," *IEEE Commun. Mag.*, vol. 49, no. 3, pp. 74 –81, Mar. 2011.

[8] R. MacKenzie, P. H. Lehne, U. Celentano, and M. Ariyoshi, "Identifying scenarios with high potential for future cognitive radio networks," in *Future Network Mobile Summit (FutureNetw), 2011*, June 2011, pp. 1 –8.

[9] O. Durowoju, K. Arshad, and K. Moessner, "Distributed power control for cognitive radio networks, based on incumbent outage information," in *Communications (ICC), 2011 IEEE International Conference on*, June 2011, pp. 1 –5.

[10] R. MacKenzie and K. Briggs, "MAC layer modelling using Markov chains," June 2011, talk given at Future Network Mobile Summit, 2011 (FuNeMS2011).

[11] ——, "MAC layers for white space radio systems," Nov. 2011, talk given at SDR'11-WInnComm, 2011.

[12] P. E. Engelstad and O. N. Østerbø, "Analysis of QoS in WLAN," *Telektronikk*, pp. 132–147, 2005.

[13] X. S. Li *et al.* (2011) SuperLU. [Online]. Available: http://crd.lbl.gov/~xiaoye/SuperLU/

[14] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[15] B. Gough *et al.* (2011) GNU Scientific Library. [Online]. Available: http://www.gnu.org/software/gsl/

[16] "Opnet Modeler Wireless Suite - Version 16.0," OPNET Technologies Inc., Oct. 2010. [Online]. Available: http://www.opnet.com

A COGNITIVE RADIO WITH MULTIANTENNA INTERFERENCE MITIGATION

Pertti Alapuranen (xG Technology, Inc., Sunrise, Florida, USA;
Pertti.Alapuranen@xgtechnology.com).
Richard Rotondo (xG Technology, Inc., Sunrise, Florida, USA;
Rick.Rotondo@xgtechnology.com).

ABSTRACT

Traditional cognitive radios have typically relied on sensing and dynamic spectrum access (DSA) to move away from channels with interference. However, as wireless systems proliferate and base station and user terminal density increase, the availability of "white spaces" (frequencies with low interference levels) will be difficult to come by. In some cases an opponent may be actively trying to deny access to white spaces. Therefore increasing network performance in cognitive networks will require an adaptive cognitive engine designed to work in "gray spaces" (frequencies with high levels of interference) and in some cases "black spaces" (frequencies undergoing intentional jamming).

In this paper, we propose a method for leveraging a 2X4 MIMO array to create a subspace projection that leads to significant increase in cognitive system performance. This paper also shows how the concept can be further enhanced by using the mobility of the end user terminals to increase the effective size of the antenna matrix – leading to increased improvements in interference mitigation, bit error rate and link budgets. Simulation results for both methods are presented that quantify the performance improvements of these approaches.

1. INTRODUCTION

Interference is a major impediment to reliable and efficient wireless transmissions, and can negatively impact performance and coverage in mobile systems. As more people (and their wireless smart phones, tablets, laptops, etc.) share the scarce airwaves used by today's data networks, they not only use more bandwidth, but they also increase the overall interference level in the network. New heterogeneous network architectures that reuse spectrum in an effort to increase capacity with a mix of macro, pico- and femtocells present additional challenges as self-limiting interference is introduced into the network. Thus, dealing with increasing levels of interference is a critical issue for wireless systems designers and service providers alike. Cognitive, or "smart radio" technologies combined with flexible Software Defined Radios (SDR) and Multiple input Multiple Output (MIMO) architectures offer the potential to combat interference.

By definition, a cognitive radio "provides radio resources and services most appropriate" to the user's needs [1]. However, with an expanding Digital Signal Processor (DSP) and SDR capabilities, the cognitive operation can be expanded within and beyond the original definition, and can include the use of RF and antenna systems in different modes depending upon environment and operational requirements. For example, a MIMO capable antenna system can be simultaneously used for Maximal Ratio Combining (MRC) and interference mitigating array configurations so long as the signal processing system can deliver adequate computing performance within the available size, power and thermal budgets.

This paper discusses a cognitive interference mitigating SDR MIMO radio that is initially

designed to operate in the ITU Region 2 900MHz ISM band.

2. OVERVIEW OF INTERFERENCE IN 900MHZ ISM BAND

The unlicensed 900MHZ band is extensively used by various devices from telemetry and video links to cordless phones and wireless baby monitors. To comply with FCC part15 rules in the USA the minimum bandwidth, power spectral density and transmit power are strictly limited. The minimum bandwidth is 500 kHz and maximum radiated output power is 36dBm EIRP. Large numbers of uncoordinated devices on the band also increase the noise floor to the point that traditional radio systems may not work or they deliver reduced range and data rates. Common interferers in the band include cordless phones that use frequency hopping spread spectrum modulation. Another common type of transmission used by fixed broadband systems is wide band Frequency Shift Keying (FSK) or Orthogonal Frequency Division Multiplexing (OFDM). These types of systems often occupy the entire 902-928MHz band.

To operate a mobile broadband data network in the presence of these systems, multiple methods to mitigate the interference are required.

3. A COGNITIVE APPROACH FOR OPERATING WITH INTERFERENCE

While a large amount of research is devoted to interference free secondary sharing, Preston Marshall introduces a term of Interference-Tolerant Sharing in his book of "Quantitative analysis of Cognitive Radio and Network Performance" [2]. The concept of interference tolerant sharing offers a potential for a large increase of network density. The cognitive radio described in this paper not only attempts to avoid interference; it also incorporates techniques that increase its interference tolerance while maximizing system capacity within Part 15 FCC rules.

Faced with ever increasing density of unlicensed use of ISM bands, xG Technology's xMax cognitive radio uses a multifaceted strategy for dealing with interference. The system incorporates a Dynamic Spectrum Access (DSA) concept including the ability to automatically identify and dynamically access better spectrum in response to the presence of potentially harmful interference. The system can locate another portion of spectrum with less interference in 40ms, move the communications to the new, clearer portion of the spectrum, and monitor interference levels in these new frequencies. In short, the cognitive system performs real-time sensing of interference on a channel and then rapidly changes channels when interference exceeds preset thresholds.

One of the challenges for real-time sensing is to create sensing opportunities for both the base station and user terminals. Additionally, the system has to avoid self-interference from co-located devices, specifically from base stations and access points that are collocated to serve multiple channels. The xMax system uses Time Division Duplexing (TDD) and Time Division Multiple Access (TDMA) as a channel access method. Unlike CSMA/CA, the TDMA approach allows terminals to use DSP and RF resources for scanning other frequencies during time periods when they are not communicating with their associated base station.

4. THE XMAX COGNITIVE RADIO PLATFORM

The xMax platform is targeted for commercial service, initially using the 900MHz ISM band. Delivering high system capacity while meeting low power requirements from both FCC Part 15 rules and battery powered operation presents implementation challenges when balancing functionality, range and signal processing with maximum operating time and low-cost. Additionally the size of the radio has to be within an envelope that allows for portable operation.

To allow efficient MIMO and smart antenna operation the system uses 4x2 antenna configuration, i.e. four receive antennas and two transmit antennas together with a parallel DSP processor that can dynamically execute multiple

receiver algorithms under the control of a cognitive control system. xMax cognitive radio block diagram (Figure 1) shows the main parts of the system and illustrates the number of receiver algorithms available to the cognitive engine.

Figure 2 shows the first hardware implementation of this radio. In this image the DSP subsystem is visible on a stack, where the RF subsystem is on a separate board mounted below the main digital board.

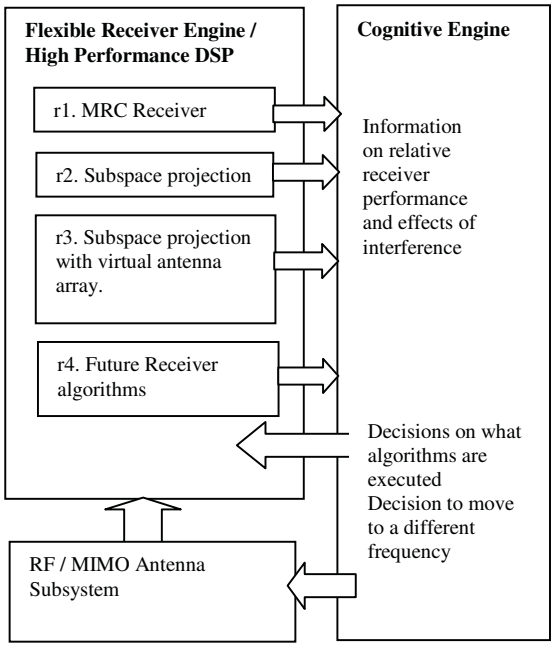


Figure 1. xMax Cognitive Radio Block Diagram

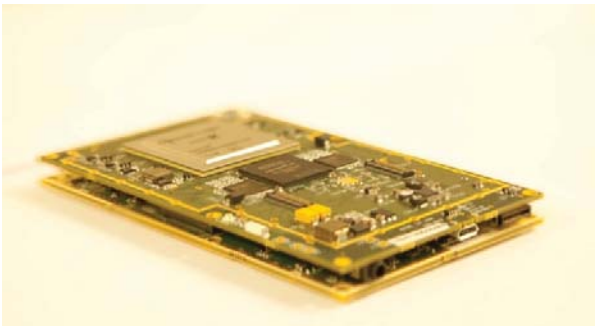


Figure 2. xMax Cognitive Radio Platform

4.1. Antenna System

Antenna system for a cognitive MIMO radio with an interference mitigation system present contradicting requirements. Small size, high performance and different modes of operation requiring four antenna elements are difficult to include into a portable, compact design.

A low envelope correlation coefficient is a key metric for good diversity performance in a fading radio channel. Typically a value of less than 0.5 is required [1]. The final antenna system used was designed to achieve 10dB diversity gain in a Rayleigh fading channel. As a result of testing multiple approaches and designs, a patch antenna solution using high permittivity low loss printed circuit board material was implemented.

Measurements confirmed that the design exceeded the goal of diversity gain by 2 dB (12dB actual). The xMax MIMO antenna (Figure 3) shows the patch antenna design. The antenna operates by exciting a patch structure through multiple ports. The volume of the antenna is extremely small and it is designed to stack on top of the cognitive radio platform. The cuts on the antenna board improve efficiency by stopping surface wave propagation.

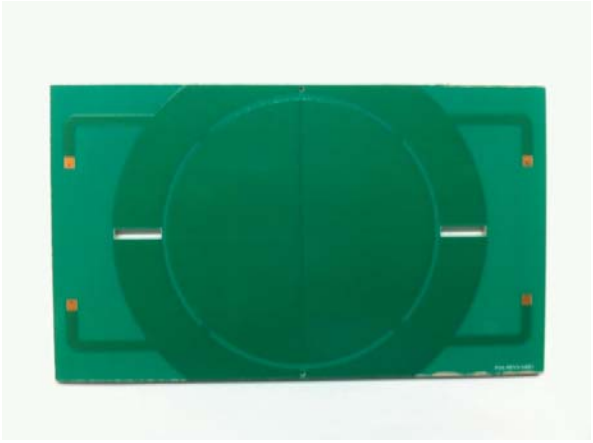


Figure 3. xMax MIMO antenna

4.2. RF Subsystem

Due to the cognitive nature of the system, the 4x2 MIMO RF subsystem design is flexible in nature. The front end is capable of 400MHz-3GHz of operation. However, to save cost in the first

implementation, the power amplifiers are designed only for 900MHz band operation. The base band signals are digitized using 14-bit converters. The xMax system allows programmable channel bandwidth up to 20MHz - where a second stage of selectivity filtering uses the DSP subsystem. Additionally, the front end band filtering can be bypassed in cases where there are no strong interfering signals present to improve performance and reduce power consumption.

4.3. Digital Signal Processing Subsystem

The DSP subsystem is able to dynamically execute multiple receiver algorithms depending on decisions from the cognitive engine. To deliver the required signal processing performance in a battery operated device a C-programmable 100 core parallel signal processor is used. The processor delivers 50GOPS and 25GFLOPS of processing performance in a single chip, and can execute 100 parallel threads while integrated Direct Memory Access (DMA) engines transfer data between the processing elements. Digitized baseband signals are fed into the DSP and routed to processor cores that execute receiver algorithms. Processed signals are compared for cyclic redundancy check (CRC) matches and if the packet is received error free the decoded data bits are delivered to the Logical Link Control (LLC) and Media Access Control (MAC) layers being executed on a Texas Instruments OMAP that contains traditional DSP and ARM processors in a single package.

4.4. The Cognitive Engine

The goal of the xMax cognitive system operating in an unlicensed band is not to protect incumbent operations, but to share spectrum fairly with other systems and to use the spectrum efficiently. However, a DSA-only system might not be able to find and operate in a low-interference channel if one is not available. For this reason the cognitive engine in xMax includes a capability to select among multiple receiver algorithms and waveforms to mitigate interference. There are several practical ways to do this; spread spectrum, adaptive antenna arrays, polarization and space diversity are few examples. The cognitive approach chosen includes the use of a flexible antenna system and signal processing system that allows the cognitive engine to dynamically decide what approach is most beneficial for mitigating encountered interference. The cognitive engine can optimize operations by instructing the physical layer to execute different algorithms, and based on real-time measurements, can choose what parameters are beneficial to implement.

It is anticipated that it will be possible to add a learning cognitive engine into the design by using a classifier in the future. For example, received signal strength, measured SNR and achieved data rate can be used to select a subset of operational modes using nearest neighbor or k-nearest neighbor clustering algorithms.

5. ADAPTIVE RECEIVER ALGORITMS

If the measurements indicate that a better channel is available, the system tunes to a new frequency within 40ms. However, if there is no better frequency available the system will use adaptive receiver algorithms to combat interference. Multiple receiver algorithms are executed under control of the cognitive engine and statistics of the interference measured by the user terminal devices and the base stations and are used to choose what receiver algorithms are most beneficial.

For example, if the receiver signal strength (RSSI) is low, MRC is effective in improving link budget and to combat fading. If RSSI is high, but packet error rate is higher than expected for the conditions, the results from multiple receiver algorithms are compared to gather information on what transmit waveform and receiver algorithm combination will be most effective.

In the case of severe interference and when there is no better frequency to move on to, the system uses the multi antenna system to orthogonalise the interference to the desired signal. Additionally the cognitive system can use

the physical motion of the user terminal to create a virtual antenna array.

The selection criteria for these modes are complex and can be tuned over time as field experience with the system is gained. For example, there may be instances where the same link performance can be reached with multiple algorithm combinations. This creates a danger of possibly selecting a local minima at a terminal where the performance goal is reached, but the solution also greatly reduces the overall system capacity. An example of this is operation in the presence of burst interference where increasing a modulation rate may improve operation by reducing the probability of data bursts being "hit" by an interferer (at high data rate bursts are shorter in time). A typical approach for non-cognitive radios in this scenario is to reduce modulation rate in an effort to improve the link budget.

5.1. Cognitive Subspace Projection Receiver

One of the adaptive antenna technologies available to the xMax cognitive engine is a subspace projection receiver. The receiver algorithm uses multiple dimensions of a receive signals from a MIMO antenna to orthogonalize the interference to the desired signal. The performance of the algorithm depends on many factors, including the number of antenna elements, element spacing and correlation to mention a few. The idea of the algorithm is to compute a projection from multi-dimensional signal space into a lower dimension so that Signal to Noise Ratio (SNR) is maximized.

Subspace projection in Figure 4 shows the idea by illustrating a 3-dimensional vector "a" projected on the 2-dimensional signal subspace "S". The point "P" is chosen so that "v" and "e" are orthogonal, where "e" is error signal, i.e. interference.

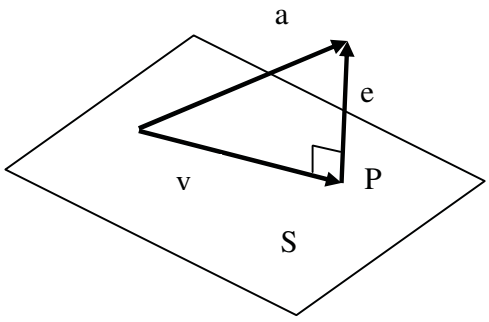


Figure 4. Subspace projection principle

For efficiency, the real-time computation of the projection can be done in the frequency domain. OFDM systems are beneficial in this case because the information in the receiver is already in the frequency domain after initial Fast Fourier Transform (FFT). However, the system performs well only when the projection is computed using carriers that are within a radio channel's coherence bandwidth.

Due to the FCC rules for ISM bands, a transmit beam forming approach is not beneficial. The FCC rules require power reduction by factor of $10*\log_{10}(N)$ where N is the number of transmit antennas. The reduction is required irrespective of the signal processing system, which in theory could reduce interference by projecting transmission into the estimated null space of a non-cognitive radio receiver [4].

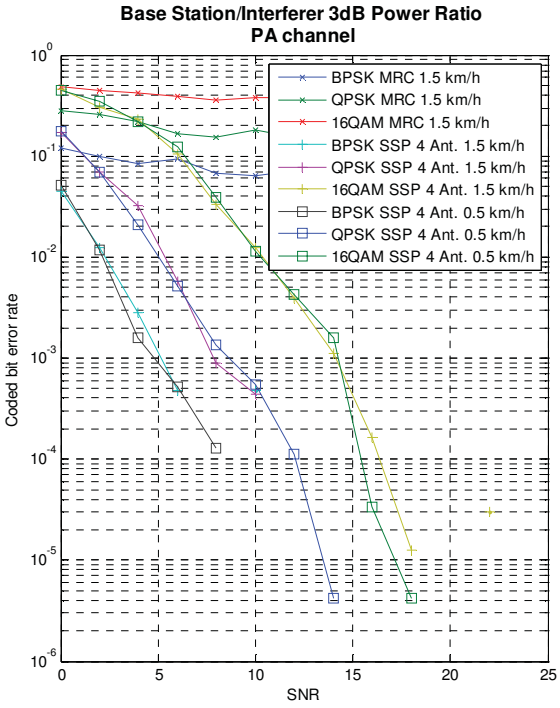


Figure 5. Simulation results of MRC and subspace projection in an interference

In a 1.4MHz 128 carrier system the simulation shows that the number of subcarriers per projection has to be in the range of 10-15 to perform in mobile radio channels. To compute the projection, a known reference signal is transmitted. For example, a Least Squares (LS) solution W can be computed as $W = (A^TA)^{-1} *A^T*d$, where the d is a known signal (local stored reference) and A contains measured complex subcarrier values. The solution is reasonably straightforward because the signal vector d is known, unlike in signal parameter estimation cases where both A and d have to be estimated [5]. In a four antenna xMax system, computations have to be done 8 times, i.e. once per each group of 13 subcarriers in a received frame.

Simulation results of MRC and subspace projection in an interference rich environment quantified in Figure 5 show the performance of a MRC receiver compared to xMax's four antenna subspace projection processing. In these scenarios, a user terminal is communicating with a picocell that is on a same channel as a nearby macro base station, resulting in 3dB higher power for macro cell signal at the receiver. In the scenario shown, the MRC receiver fails to maintain adequate link performance, while the subspace projection based receiver delivers an adequate bit error performance.

5.2. Virtual antenna system using platform motion

A cognitive radio receiver can automatically change transmission parameters and even signal formatting. In severe interference cases the motion of the end user terminal can be used to increase performance by using it to create a "virtual antenna array". When the cognitive engine decides that changing the frequency is not an optimal solution, and that MRC or subspace projection receivers are not delivering adequate performance, it is able to switch to a mode where Hybrid Automatic Repeat Request (HARQ) retransmissions are used to increase receiver signal dimensionality.

This method is an extension of the subspace projection method, except that the size of matrix A is doubled using a retransmission to a moving end user terminal. When a received packet fails, the subcarrier information is stored and HARQ retransmission is used to fill the 8x8 MIMO matrix A. The computational complexity of the solution increases roughly as the third order of the matrix A size. However, with parallel DSP processing the solution can be computed in real-time. Simulation results show that the motion and resulting increase in the size of matrix A outperforms the original four antenna projection method by more than 5dB, depending on the terminal speed. Generally, increased speed improves link performance due to the dimensions of the virtual antenna array increasing as the result of terminal motion between retransmissions.

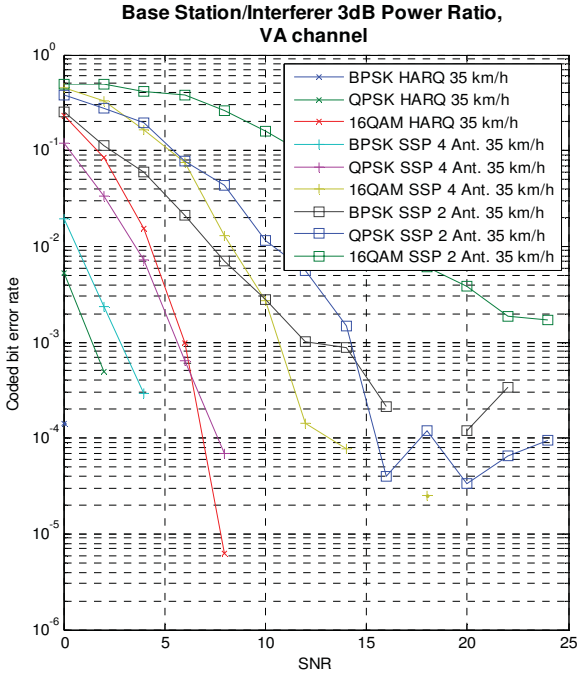


Figure 6. Coded BER of subspace projection method using platform motion to increase antenna matrix size.

Coded BER of subspace projection method using platform motion shown in Figure 6 illustrates the performance improvement using HARQ with the increase of matrix A size. It is important to note that the bit error rate improvement using HARQ comes at a cost of transmitting the data burst twice.

6. CONCLUSIONS

This paper has presented a cognitive MIMO radio with advanced interference mitigation capabilities that allows for increased network density or base stations and or user terminals. The radio system incorporates a cognitive engine that receives information on the relative performance of different receiver algorithms. This information is used to decide how best to use the antenna system and available DSP resources to maximize capacity, throughput or other network performance measures. The radio can change transmission protocols and use a virtual antenna array created by the physical motion of user terminals as a way to maintain the data link under exceptionally harsh RF conditions. Although the presented system is designed for commercial operation on unlicensed bands and does not require incumbent protection, the cognitive engine tries to share the band fairly.

The goal of this paper is to introduce an expanded view of the conventional cognitive radio concept to include a cognitive engine that can dynamically apply one or more receiver algorithms while also using multiple modes of its MIMO antenna system. This ability allows the radio to dynamically transform from a simple MRC based receiver to a relatively complex radio that uses retransmissions and terminal motion of its MIMO matrix array to increase signal dimensionality. Simulation results confirm that the concept of a cognitive radio where receiver algorithms are C-programmable applications that execute on a high performance parallel DSP under control of a cognitive engine can produce significant increases in the link budget.

7. REFERENCES

[1] J. Mitola, “Cognitive radio: An integrated agent architecture for software defined radio,” Dissertation, Doctor of Technology, Royal Institute of Technology, (KTH) , Sweden, 2000.

[2] P. Marshall , “Quantitative analysis of Cognitive Radio and Network Performance,” Norwood, MA, Artech House, pp. 215-237, 2010.

[3] R. G. Vaughan and J. B. Andersen, "Antenna Diversity in Mobile Communications," IEEE Transactions on Vehicular Technology, Vol. 36, pp. 149-172, 1987.

[4]. M. Jung, K. Hwang, S. Choi, “Interference Minimization Approach to Precoding Scheme in MIMO-Based Cognitive Radio Networks,” IEEE Communications Letters, Vol. 15, No. 8, pp. 789-791, 2011

[5] R. Roy, T. Kailath, "Esprit - Estimation Of Signal Parameters Via Rotational Invariance Techniques". IEEE Trans. Acoustics, Speech and Signal Processing, Vol. 37, No. 7, pp. 984–995, 1989.

A SIMULATION-BASED APPROACH FOR PERFORMANCE EVALUATION OF SDR BASEBAND ARCHITECTURES

Anthony Barreteau, Sébastien Le Nours, Olivier Pasquier (Lunam Université - Université de Nantes, UMR CNRS 6164 Institut d'Electronique et de Télécommunications de Rennes (IETR), Polytech'Nantes, Nantes, France; anthony.barreteau@univ-nantes.fr , sebastien.le-nours@univ-nantes.fr, olivier.pasquier@univ-nantes.fr).

ABSTRACT

Application of SDR concept for mobile device presents numerous opposing requirements including low power, low cost, high performance that create significant design challenge. In this context, system architect is supposed to design SDR baseband architecture able to support a wide variety of communication interfaces with a wide range of processing resources that fully meet these requirements. Thus, it becomes necessary to help system architect to analyze and compare the growing number of potential architectures. The creation of efficient executable models becomes a mandatory step to enable architecting of such systems. In this paper, we present a simulation-based approach for performance evaluation of SDR baseband architectures. This approach makes possible to evaluate by simulation the expected resources according to complex use-case scenarios. This approach is illustrated through the study of an adaptive multi-standard and multi-application radio communication system.

1. INTRODUCTION

Current trends in the design of radio communication systems consist in offering terminals able to easily roam among heterogeneous networks like cellular, wireless local and metropolitan area networks [1]. The purpose of these improved functionalities is to increase data rate and to propose access to a wide variety of services anytime and anywhere with a single device. As a result, the number of protocols to be supported by a single receiver, and the number of modes supported by a single protocol, calls for higher flexibility. This implies to design multi-standard and multi-application radio communication systems able to adapt to their changing environments and to make decisions about their operating behavior to fully meet expected quality of services (QoS). The inferred complexity of hardware and software resources of terminal architectures is then significantly increased in order to support such advanced services.

Currently, parallel architectures clustered by application category are adopted to implement mobile terminals [2]. These multi-core platforms consist of a set of modules like fully programmable processor cores, standard interface modules, memories and dedicated hardware blocks. The definition of such platforms is done in order to fully meet functional requirements, related to the protocols to implement, and non-functional requirements such as timing constraints, power consumption, and cost. In the process of system architecting, potential architectures are compared according to performances achieved. Performance evaluation of candidate architectures early in the development process has then become mandatory in order to avoid costly design iterations. Typically, the creation of architecture models for performance evaluation is done by mapping a model of the system application onto a model of the considered platform [3]. The resulting description is then evaluated through simulation or analytical methods. Compared to analytical methods, simulation approaches are required to investigate dynamic and non deterministic effects in the architecture model. Architecture models are then simulated to evaluate usage of resources with respect to a given set of stimuli. Simulation results are used in order to compare performances of candidate architectures. In the context of next generation communication systems, one of the challenges is to design architecture able to implement baseband processing related to various communication interfaces supported that meet power consumption requirement and costs constraints. Current trend consist in defining SDR baseband architecture made up of wide range of processing resources to meet functional and non functional requirements of these systems. SDR baseband architecture is closely linked to changes of the system and its radio environment. It then becomes mandatory to provide means to capture at a high abstraction level to improve evaluation of various operating scenarios.

This paper presents an approach for performance evaluation of SDR baseband architectures. A modeling technique based on scenario files is presented to drive the behavior of the modeled environment in order to facilitate evaluation of different use cases. Architecture and its

environment are described graphically through a specific activity diagram notation. Modeling techniques are presented to capture new requirements identified in adaptive multi-standard and multi-application radio communication systems. We also present techniques to efficiently model the various protocols supported and their related management. The behavior related to communication interfaces is described as a finite-state machine to model the use of platform resources when application executes. This model is then automatically generated in SystemC to allow simulation and to compare performances of various architectures. Simulation results are provided to analyze time properties and architecture performances according to different use-case scenarios.

The remainder of this paper is structured as follows. Section 2 analyzes the existing modeling and simulation approaches for performance evaluation of system architectures with a specific focus on the study of radio communication systems. In Section 3, the proposed modeling approach is presented and related notations are defined. In Section 4, we describe the considered case study and we detail the modeling techniques proposed to capture the different services supported. In Section 5, we present application of the proposed modeling approach to evaluate performance of a SDR baseband architecture. Finally conclusions are drawn in Section 6.

2. RELATED WORK

Performance evaluation of embedded systems has been approached in many ways at different levels of abstraction. A good survey of various methods, tools, and environments for early design space exploration is presented in [4]. Related approaches mainly differ according to the way application and platform models are created and combined. In the context of this paper, we especially focus on related works about performance evaluation for flexible terminal architectures.

In the context of next generation terminals, a multitude of radio access networks coexist and deliver different QoS according to used location, mobility, and applications. Interoperability and mobility management across these heterogeneous networks have become mandatory to enable full benefits of their complementary characteristics. This refers to concepts like Cognitive Radio (CR) and Software Defined Radio (SDR) introduced by Joseph Mitola [5]. The E²R project [6] highlighted new services and reconfiguration mechanisms required in user equipment. Innovative services were proposed to identify appropriate Radio Access Technology (RAT) and to obtain expected QoS for running application. Moreover, specific mechanisms must also be supplied to dynamically adapt the radio interface to the best configuration. It consists in providing new services to manage, deploy, and configure RAT onto the platform resources. As an example, the Software Communication Architecture (SCA) has been proposed by the Joint Tactical Radio System (JTRS) program as a common infrastructure for efficiently managing software and hardware elements [7]. In this context, works described in [8] present a modeling approach based on the combined use of UML (Unified Modeling Language) and SystemC to simulate radio communication systems build upon SCA. In [9], an approach for the design of radio communication platforms is proposed. Modeled platforms are based on a specific architecture called NoTA (Network-on-Terminal Architecture). Functional and non-functional requirements are captured in UML2.0 and SystemC is used as the simulation language to evaluate performances of virtual architectures formed. However, this work does not address the modeling of adaptation mechanisms required for multi-standard terminals. A simulation framework for performance evaluation of SDR platforms is presented in [10]. The proposed approach makes possible to model baseband functions related to different protocol and to study their allocation on a SDR platform supporting simultaneously different RATs. Similarly to our approach, baseband functions are modeled at the packet level granularity and they are described only by their processing time. In this approach, the environment is related to a specific Medium Access Control layer (MAC) which limits the number of use-cases that can be simulated. The MOPCOM design methodology presented in [11] is proposed to enable design of SDR communication systems. It defines different levels of abstraction to correctly model baseband functions of radio communication systems and to allow automatic generation of VHDL and C/C++ codes. Nevertheless, even if the creation of SystemC models is claimed, performance evaluation of architectures is not addressed. A simulation-based approach is presented in [12] to carry out design space exploration of multi-core architecture on a UMTS data link layer design case study. Functional and architectural models defined separately are then mapped together to produce a system model with performance metrics. 48 mapping are investigated to estimate execution time of 11 tasks along with the average of 1 to 11 processing element utilization. Similarly to our approach, the purpose is to allow study of different architectures with light modeling effort. However, this work is limited by the number of operating scenarios considered.

The approach presented in this paper mainly differs from the above as to the way terminal architectures are modeled. In our approach, architecture specification is done graphically through a specific activity diagram notation. The behavior related to each elementary activity is captured as a finite-state machine to express the influence of application when executed on platform resources. The resulting architecture model is then automatically generated in SystemC to allow simulation and performance assessment.

Considering the modeling of flexible radio communication systems, specific methods are proposed to favor creation of efficient architecture models. In the following, the benefits of these methods are illustrated through a specific case study.

3. CONSIDERED MODELING APPROACH FOR PERFORMANCE EVALUATION OF SYSTEM ARCHITECTURES

Further information on the considered approach presented in this section is provided in [13]. In the considered approach, the architecture model captures information about both the structural description of the system application and the description of non-functional properties relevant to considered hardware and software resources. In order to provide efficient simulation speed and light modeling effort, architecture is modeled without considering a complete description of the application. Application is represented by a workload model. Workload models are used to represent the computation and communication loads applications cause on platform resources when executed. This approach is illustrated in Figure 1.

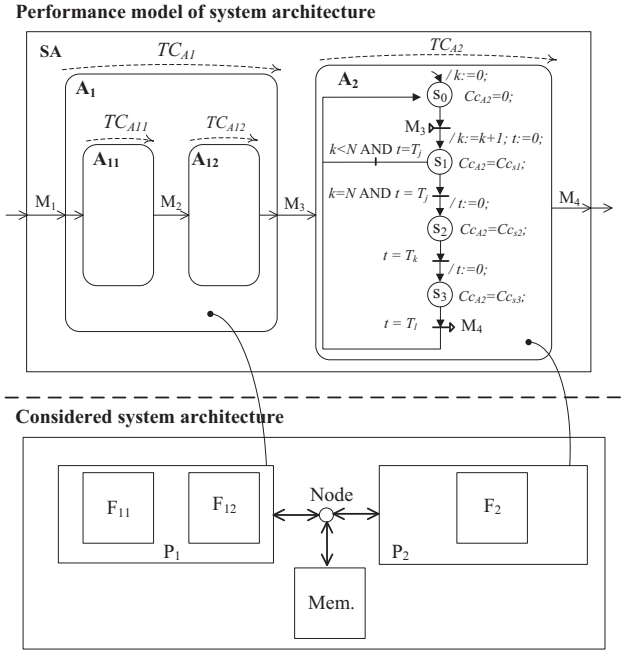


Figure 1 – Considered modeling approach for performance evaluation of system architectures.

The lower part of Figure 1 depicts a typical platform made of communication nodes, memories, and processing resources. Processing resources are classified as processors and dedicated hardware resources. In Figure 1, F_{11}, F_{12} and F_2 represent the functions of the system application. They are allocated on the processing resources P_1 and P_2 to form the system architecture. The upper part of the figure depicts the structural and the behavioral modeling of the system architecture. Each activity A_i represents a function, or a set of functions, allocated on a processing resource of the platform. As for example, activity A_{11} models the execution of function F_{11} on processor P_1. Relations M_i between activities correspond to transactions exchanged between activities. Transactions are exchanged through relations in conformity with the rendezvous protocol and they are defined as data transfer or synchronization between activities.

Behavior related to each elementary activity models the usage of resources by each function of the system application. Behavior exhibits waiting conditions on input transactions and production of output transactions. In the notation adopted, one important point is about the meaning of temporal dependencies. Here, transitions between states s_i are expressed as waiting transactions or logical conditions on internal data. A specific data value may be a time variable which evolves naturally. This data is denoted by t in Figure 1. The amount of processing and memory resources used is expressed according to the allocation of functions. In Figure 1, the use of processing resources due to the execution of function F_2 on P_2 is modeled by the evolution of the parameter denoted by Cc_{A2}. For example, Cc_{A2} can be defined as an analytical expression to give the number of operations related to the execution of function F_2. Value of Cc_{A2} can be influenced by data associated to the transaction received through relation M_3. In our approach, these properties could be provided by estimations, profiling existing codes, or source code analysis, as illustrated in [14]. The time properties used are directly influenced by the characteristics of the processing resources and by the characteristics of the communication nodes used for transaction exchange. Furthermore, the temporal behavior related to each activity is relevant to the function allocation. In case of a single processor, allocated functions are executed sequentially or according to a specific scheduling policy. In case of a multi-processor architecture, behaviors should express parallelism offered to execute functions.

Following this modeling approach, resulting model incorporates evolution of quantitative properties relevant to the use of processing resources, communication nodes, and memories. Using languages as SystemC, created architecture models can be simulated to evaluate the time evolution of performances obtained for a given set of stimuli. To facilitate creation of SystemC description we have used the framework CoFluent Studio [15]. This environment supports creation of transaction level models of system architectures. Graphical models captured and associated codes are automatically translated in a SystemC description. This description is executed to analyze models and to assess performances. Various platform configurations and function

allocations can be compared considering different descriptions of activities. In Section 5, this approach is considered to model the SDR baseband architecture of an adaptive terminal.

4. TRANSACTION LEVEL MODELING OF FUNCTIONAL PROPERTIES OF AN ADAPTIVE TERMINAL

4.1. Considered case study

In this section we describe proposed case study to illustrate modeling approach for performance evaluation of SDR baseband architectures. This case study has been used to highlight modeling techniques proposed in [16] to capture new requirements identified in the next generation of adaptive multi-standard and multi-application radio communication system. The novelty in this paper lies in considering study of non functional properties of this system.

The considered case study reflects new features of these systems. It supports two different RATs and three user applications. An appropriate management of radio interfaces should maintain required QoS to ensure end-user experience in various radio environments. This new property implies to integrate new services within radio communication systems. The system should support dynamic activation/deactivation of one or more RATs. This adaptation mechanism should be monitored by a decision making module based on delivered QoS for each application. We present in Figure 2 the considered environment and relations with the radio communication system under study in the middle. The internal structure of such a system is refined using the activity diagram notation previously presented. In the considered case study, the user is supposed to request applications to the system through relations *Application request* and *Application response*. The system delivers three kinds of information related to the three applications supported: *Voice frame*, *Web page* and *Video frame*. Expected QoS for these applications are 20 frames per second for video, one frame of 160 samples every 20 ms for speech decoding and Web browsing corresponds to reception of varying size of data at about 1s. The network environment includes the activities related to application servers and transmission of data according to Universal Terrestrial Radio Access (UTRA) [17] and 802.11 [18] specifications. These radio links are depicted by three relations: *Downlink UTRA*, *Downlink WiFi*, and *Uplink UTRA*. The data size assigned per application in each radio frame depends on the channel quality of the link. For example, UTRA Network (UTRAN) is supposed to achieve at least 384 kbps in urban outdoor radio environment and at least 144 kbps in rural outdoor. WiFi has been defined to provide up to 54 Mbps data rate in indoor environment.

The internal structure of the system is composed of two main activities, *System management* and *System data flow* depicted in the upper and the lower part of Figure 2. Activity *System data flow* is made of activities concerned by data transmission in both uplink and downlink and application processing. UTRA and WiFi communication interfaces are depicted by *UTRA reception*, *WiFi reception* and *UTRA transmission* in activity *Radio reception and transmission*, whereas applications are depicted with activities *Voice processing*, *Web processing* and *Video processing*. The

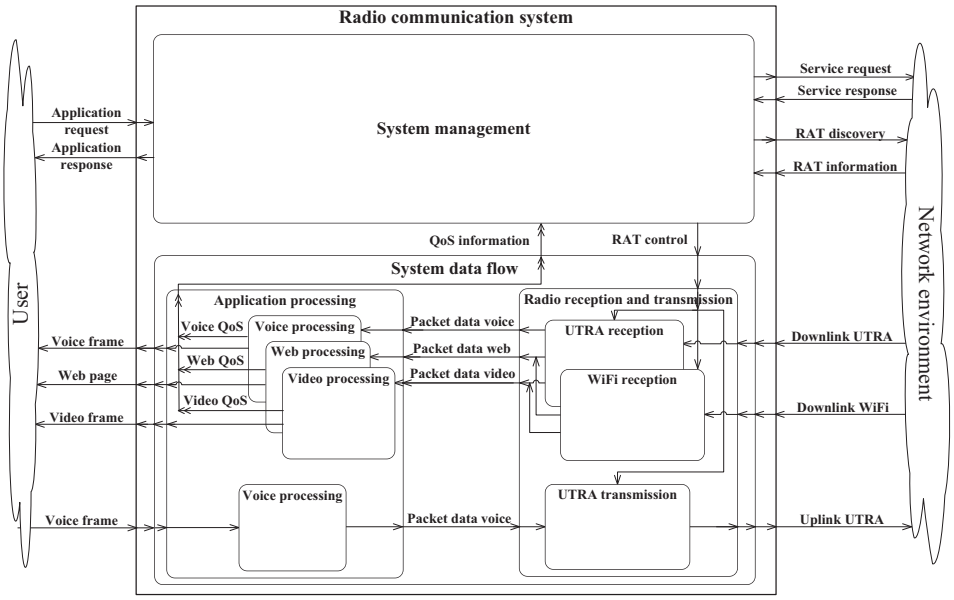


Figure 2 – Activity diagram of an adaptive multi-standard and multi-application system and its environment.

activities related to activity *System management* analyze QoS provided to user and perform request to the network environment. This activity receives information on data rate performance of the established radio link through the permanent relation *QoS Information*. In case of the QoS cannot be maintained, this activity supports a discovery service to find a RAT offering better performance. It is also able to decide and carry out a reconfiguration process of radio interface based on information provided by the *RAT control* relation. This process consists in activating the required communication interface. Besides, it communicates with the network environment according to the application request of the user and RAT to be used. Section 4.2 and 4.3 present proposed modeling techniques used to describe behavior of the system environment and communication interfaces.

4.2. Modeling the system environment

The behavior of the radio communication system is closely related to external behaviors of its environment. Indeed the system is supposed to adapt its radio interface according to the expected QoS level. This level is defined as the minimum data rate required by the user for each application. Therefore, it is necessary to model the system environment to evaluate multiple use cases and to study the dynamic behavior of the system. A generic solution consists in capturing the evolution of the environment as a scenario file containing the main time information, application declaration and associated parameters. A similar technique has been used in [9] to express sequence of service requests provided by functional requirements. Figure 3 illustrates the solution used to model the behavior of the user.

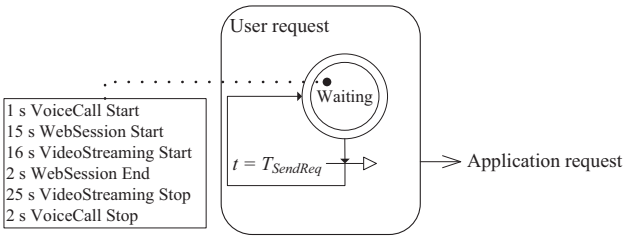


Figure 3 – Partial view of the activity diagram of the user.

The behavior related to the *User request* activity is defined to enable description of various usages of the multiple applications supported by the system. To achieve this, we propose to write in an user scenario file information describing the use case to evaluate. Information is organized as follows. Each line corresponds to an application request. The first field defines the time delay that must elapse before sending a request. The two other fields specify which application is inquired and if it must be started or stopped. This file is read in state *Waiting*. Time condition $t=T_{SendReq}$ expresses delay before sending transactions through relation *Application request*. $T_{SendReq}$ value evolves according to the time delay information indicated in each line. Transaction sent through relation *Application request* contains information related to the start or the stop of an application. The left part of Figure 3 gives an example of a user scenario file describing the successive use of the three applications proposed by the system under study.

We use a similar approach to describe the evolution of the network environment. Figure 4 presents a partial view of the activity diagram of the network environment.

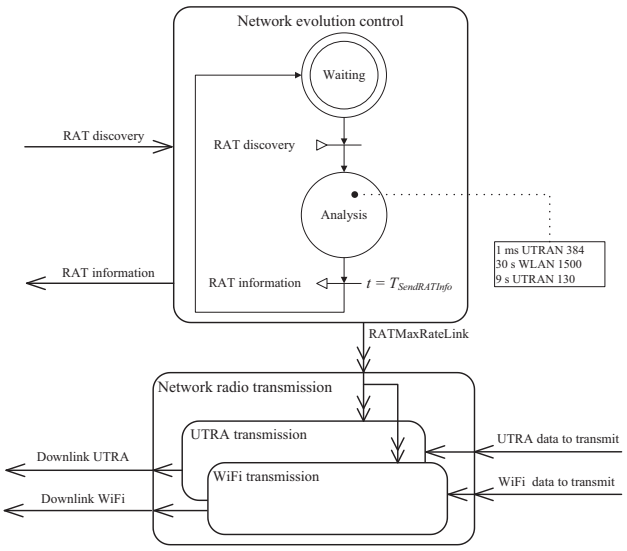


Figure 4 – Partial view of the activity diagram of the network environment.

The upper part of Figure 4 depicts the behavior of the *Network evolution control* activity. This activity is used to change the data rate offered by each RAT available in the network environment. We propose to write in a network scenario file information describing the evolution of the transmission conditions. Information is organized as follows. Each line of the scenario file defines one change in the transmission conditions associated to the network environment. A line is composed of three fields. The first one gives the time information when the transmission condition changes. The two next fields define which access network is affected among UTRAN and WLAN and what is the new data rate delivered. The state *Analysis* repeatedly reads one line in the file. After time delay, relation *RATMaxRateLink* is updating with the new data rate. Activities *UTRA transmission* and *WiFi transmission* use these information to define data rate could be used to send data on the radio links.

This approach based on scenario files makes possible to rapidly create various use cases. We can then easily assess

different system operating scenarios for several combinations of the user and the network environment behaviors.

4.3. Modeling the communication interfaces

In this section, we present a modeling technique to capture communication interfaces at high abstraction level. Then we illustrate application of the proposed modeling approach to perform evaluation of the resource usage of the communication interfaces.

In our case, transactions identified in the data flow between the network environment and the radio communication system corresponds to the amount of data transmitted by each RAT. At this level, the associated payloads of transactions are defined according to expected throughput. For UTRA, transactions periodically occur every Time Transmission Interval (TTI) set to 10 ms and the maximum payload is 480 bytes. For WiFi, the maximum payload is 2347 bytes and the transaction instants depend of the size of data to transmit. Considering this data granularity, activities can be described as finite state machines. Furthermore, the described behavior is the same for the two RATs. Figure 5 illustrates the technique used to obtain description of the activities *UTRA reception* and *WiFi reception*.

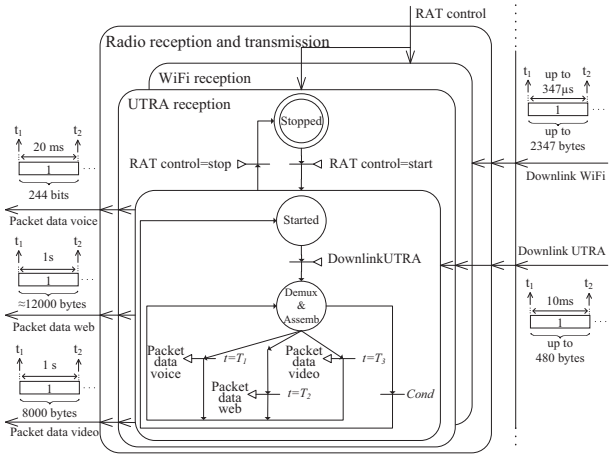


Figure 5 – Activity diagram of adaptive radio interfaces.

The behavior describes activities performed by layers involved in reception of data application (Medium Access Control layer, Radio Link Control layer for UTRA, Logical Link Control layer for WiFi). This type of description enables also to take into account the activation/deactivation of interfaces during system execution.

When *UTRA reception* or *WiFi reception* activities are started, information is exchanged with other activities in the form of transactions. As previously mentioned, transactions sent through relations *Downlink UTRA* and *Downlink WiFi* represent data transfer with the network environment. Transactions sent through the relations Packet data voice, Packet data web, and Packet data video represent the amount of data sent to related application processing activities. Packet data voice transaction must periodically occur every 20 ms with a payload of 244 bits. Packet data web transaction must be initiated 1s after the inquiring of the web page and contains quantity of data related to this web page. Packet data video transaction must periodically occur with a payload of 8000 bytes every 1s so as to display video at a rate of 20 frames/s. When *Downlink UTRA* or *Downlink WiFi* transaction is initiated by the environment, state *Demux & Assemb* in reception activities carry out both demultiplexing of data (voice, web, and video) and reassembling of previously segmented application data packets. When an application data packet is fully merged a transaction is initiated and the activity related to the application processing can start to deliver a service to the user. The behaviors of *UTRA reception* and *WiFi reception* activities evolve according to the *System management* activity. It decides to activate them through the relation RAT control. These two activities can be in two states: stopped or started. This description represents the dynamic behavior of the RAT interfaces at transaction level.

Modeling approach presented in Section 3 is applied to analyze performance obtained with different architectures to perform the chain of baseband processing related to physical layer of the two protocols. At the physical layer, synchronization, demodulation, channel equalization channel decoding, and multiple access channel extraction are linked together to form the chain of baseband processing. The behavior related to state *Demux&Assemb is* modified to allow an assessment of the related processing functions cost according architecture considered. Time interval between reception of *Downlink UTRA* or *Downlink WiFi* transactions and production of application data corresponds to time constraints to perform required baseband processing. The captured behavior enables to analyze the number of operations required to perform baseband processing related to each communication interfaces according to architecture considered.

4.4. Simulation results of the model

The CoFluent Studio tool has been used to generate and simulate an executable code of the model of our case study. It is obtained by the translation of a graphical model in a SystemC description. The code size related to this model corresponds to 4476 lines of SystemC code and 62% are automatically generated by the tool. In our case, simulation consists in analyzing how the system reacts according to various use cases. Functional performance evaluation enables to verify if properties such as timing constraints and QoS can be maintained by the system. The selected use case

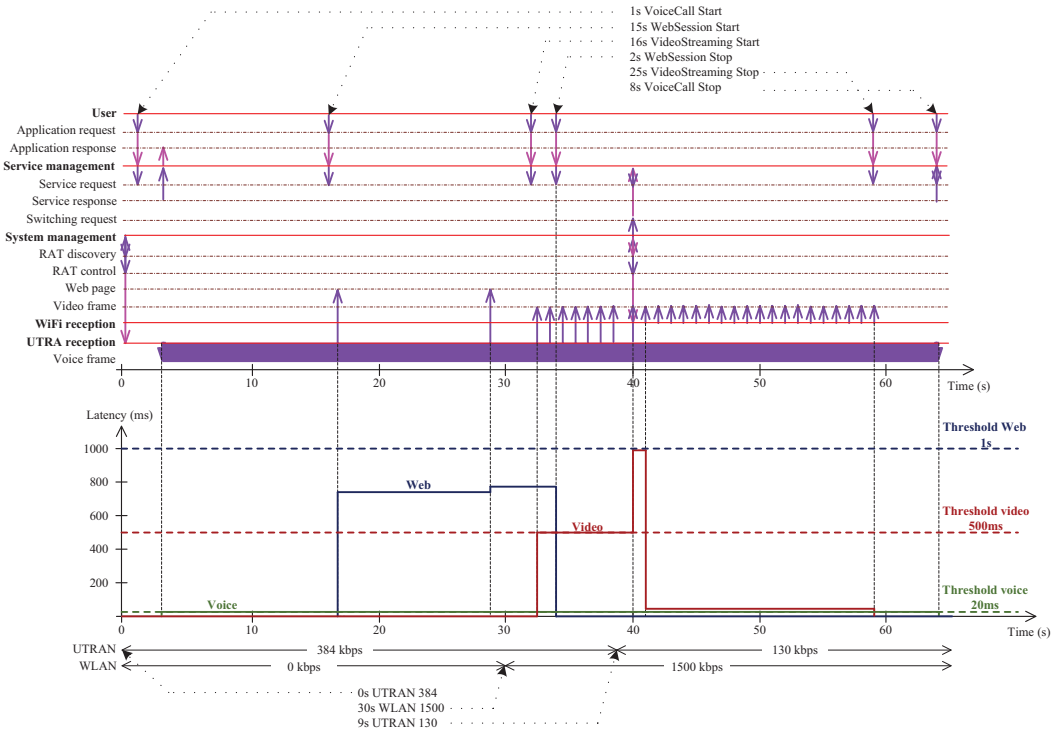


Figure 6 – Observation of transactions initiated during model simulation and related QoS evolution.

considers a user who successively requests the three applications supported by the system. Moreover the user evolves in a changing radio environment. At the beginning of the system execution, only UTRAN is available with a basic data rate set to 384 kbps to ensure data transmission required by all running applications. Then the data rate of the UTRAN is decreased to 130 kbps to model a changing environment. Therefore the WLAN is detected to support the video streaming application with the QoS expected. Figure 6 shows a way to display simulation results to observe the detailed behavior of the TLM model of the system.

The upper part of Figure 6 presents transactions exchanged during the simulation between activities. Activities and relations between activities are depicted on the left part. The lower part of Figure 6 presents evolution of the latency related to the three applications considered. This result of simulation is used to analyze the level of quality of services proposed to the user. On the right part the values of the latency threshold for each application are depicted. One can notice on the upper part of Figure 7 the successive start and stop requests of each application as expressed in the user scenario file depicted in the upper part. Initially the three applications are delivered by UTRAN. It is depicted by transactions sent by activity *UTRA reception* through the relation *Web page*, *Video frame* and *Voice frame*. After 39 s, the user is in an area where the data rate delivered by UTRAN is not sufficient to ensure both the QoS for voice and video streaming applications. Indeed, we observe on the lower part of Figure 6 that the QoS threshold related to video application is exceeded. The system performs a discovery process and identifies a new RAT in its environment. Then it decides to activate the activity *WLANreception* and the video streaming application is switched on to this RAT. This type of display is useful to validate the functional and behavioral model of the system with timing information.

5. PERFORMANCE EVALUATION OF THE FLEXIBLE BASEBAND ARCHITECTURE

5.1. Considered architecture

In this section, we present application of the proposed modeling approach described in Section 3 to perform evaluation of the resource usage of the communication interfaces. Here, we focus on the architecture required to perform UTRA and WiFi baseband processing. Figure 7 shows the studied architecture and the activities *UTRA reception* and *WiFi reception* presented in Figure 2.

The lower part of Figure 7 depicts the studied architecture. This architecture consists in implementing UTRA and WiFi baseband functions as a set of dedicated hardware resources denoted by P_1. For clarity reason, different functions related to UTRA and WiFi baseband processing are not represented. The upper part of Figure 7 depicts activities *WiFi reception* and *UTRA reception*. TC_{WiFi} and TC_{UTRA} denote the time constraints to be met by

the related activities. TC_{WiFi} and TC_{UTRA} express the time constraints to satisfy when baseband functions are executed by P_1 each time a UTRA or WiFi frame is received. The behavior of these activities depicted in Figure 5 is modified in order to express the computational complexity per time unit each function causes on the resources when executed. Simulation results of the model makes possible to observe evolution of the usage of processing resources used to implement each baseband functions according to different operating scenarios. The computational complexity per time unit metric is considered here because it directly impacts area and energy consumption of the terminal architecture.

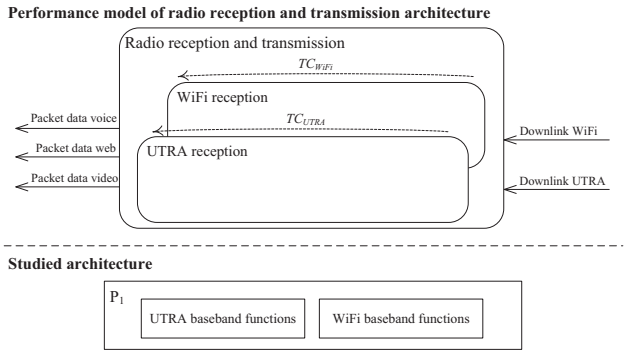


Figure 7 – Studied architecture to perform baseband processing related to activities UTRA reception and WiFi reception.

In the next section, we illustrate the simulation results could be obtained with our approach for one specific baseband function. We focus here on estimation of the required computational complexity per time unit to perform processing related to channel decoding function. Channel decoding functions for UTRA and WiFi are respectively performed with Turbo [19] and Viterbi decoder [20]. Computation duration and number of arithmetic operations are the two parameters required to estimate the resulting computational complexity per time unit. Based on a detailed analysis of resources required for channel decoding of UTRA [21] and WiFi [22], we have defined analytical expressions to give relations between functional parameters related to the different configurations of UTRA and WiFi frames and the resulting computational complexity in terms of arithmetic operations. The computation durations to perform channel decoding, denoted by $T_{ProcTurbo}$ and $T_{ProcConvol}$, are set in order to meet TC_{UTRA} and TC_{WiFi} time constraints.

5.2. Simulation results of the model

In this section, we present simulation results obtained to estimate maximal computational complexity per time unit required to perform UTRA and WiFi channel decoding with the considered architecture. Simulation of the model of flexible terminal enables to analyze the processing resource usage according to dynamic and non deterministic behavior of the system and its environment. Results presented are obtained considering the operating scenario described in Section 4.4. $T_{ProcTurbo}$ and $T_{ProcConvol}$ has been set to 2,5 ms and 4 ms to satisfy TC_{UTRA} and TC_{WiFi} time constraints.

Figure 8 presents a partial view of the evolution of the computational complexity per time unit related to the turbo decoding of data received by activity *UTRA reception*.

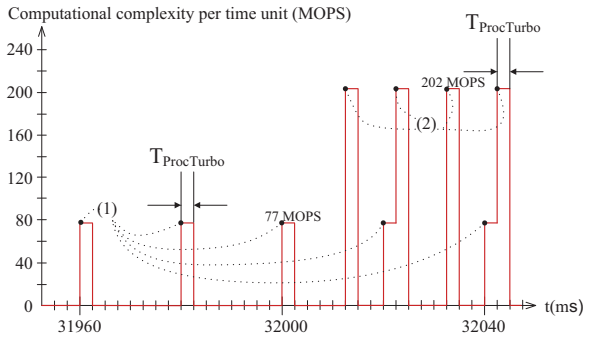


Figure 8 – Partial view of the time evolution of the computational complexity per time unit (in MOPS) for UTRA channel decoding.

Instants (1) and (2) correspond respectively to start instants of the channel decoding of data packets related to voice and video applications. We can observe resource utilization of the turbo decoder with the computation duration set. During this time interval the computational complexity per time unit varies between 77 MOPS and 202 MOPS.

Figure 9 shows the observation obtained with the simulation of the scenario presented previously.

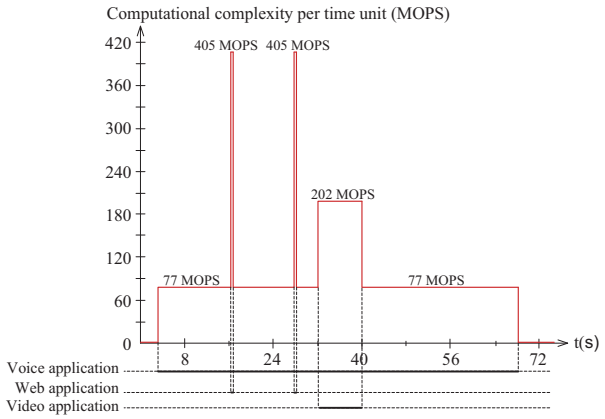


Figure 9 – Time evolution of the computational complexity per time unit (in MOPS) for UTRA channel decoding.

The lower part of Figure 9 depicts the resource utilization of P_1 according to the user applications

supported. The upper part of Figure 9 shows evolution of the computational complexity per time unit for the UTRA channel decoding. We observe that a computational complexity per time unit of 405 MOPS is required for turbo decoding of Web data packets. This value corresponds to the maximum computational complexity per time unit that should be achieved by the hardware architecture to perform UTRA channel decoding with this operating scenario.
Figure 10 shows the observation obtained with the complete simulation of WiFi channel decoder considering the scenario previously presented.

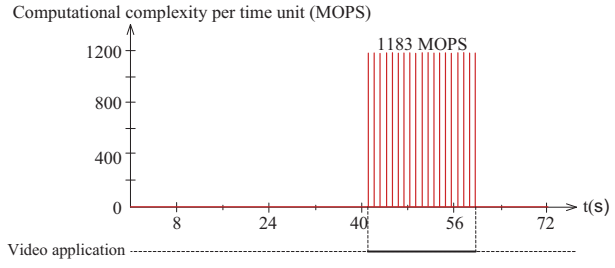


Figure 10 – Time evolution of the computational complexity per time unit (in MOPS) for WiFi channel decoding.

In this operating scenario, the WiFi reception interface is used to receive video streaming data. This radio interface is activated after 40s. We can observe that a maximal computational complexity per time unit of 1183 MOPS is also required. To estimate expected resources for this architecture, it is necessary to analyze the global computational complexity required to execute simultaneously UTRA and WiFi channel decoding.

Figure 11 presents the partial view of the global computational complexity per time unit for studied architecture with operating scenario considered.

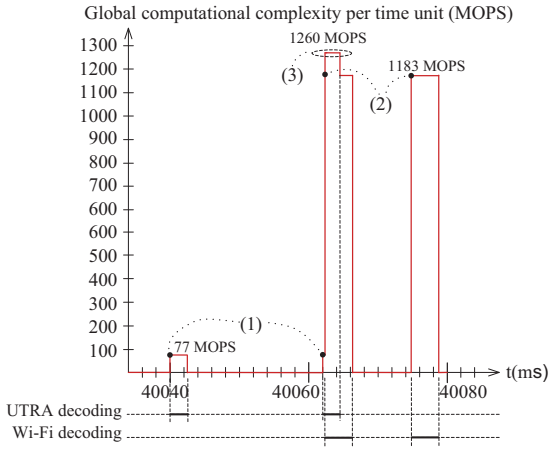


Figure 11 – Time evolution of the global computational complexity per time unit (in MOPS) for UTRA and WiFi channel decoding.

The lower part of Figure 11 depicts the resource utilization of P_1 to perform both UTRA and WiFi channel decoding. The upper part of Figure 11 shows the selected time interval where the maximal computational complexity per time unit is observed. Instants (1) and (2) correspond respectively to start instants of the channel decoding of data packets related to UTRA and WiFi protocols. We can observe resource utilization of the turbo decoder with the computation duration set. The time interval denoted (3) shows the maximal computational complexity per time unit observed with this operating scenario and architecture considered. It is estimated to 1260 MOPS. It occurs when UTRA and WiFi channel decoding are performed simultaneously. This kind of result highlights interest in simulation based approach compared to analytical approach. Indeed, the resource usage evolves dynamically because different wireless protocols with different range of performance requirements can be switched at unknown timing moments. The simulation time to execute the model for the operating scenario considered took about 500 ms on a 2.66 GHz Intel Core2 duo machine. Simulation time achieved and proposed approach based on scenario files enable system architect to iterate quickly over different operating scenarios to evaluate and compare resource usage. These simulation results can be used to study different resource utilization according to different RAT operating modes. Moreover, proposed modeling approach allows quick performance comparison with other potential architectures.

6. CONCLUSION

Radio communication systems are assumed to support new services to provide adaptation capabilities of radio functionalities according to internal and external changes. One important challenge is to provide performances evaluation of radio communication architecture highlighting these new requirements. In this paper, we have presented a simulation-based approach and modeling techniques to evaluate by simulation performances of an adaptive multi-application and multi-standard system. This approach leads to the creation of abstract representation of the architecture to efficiently deliver simulation results. Finally, the related computational complexity required to perform baseband signal processing for the various RAT supported has been evaluated to make possible platform sizing. Further research is directed towards validation of estimates provided by simulation and applying the same modeling principle to other non-functional properties such as dynamic power consumption.

10. REFERENCES

[1] J. Govil, "4G Mobile communication systems: turns, trends and transition", *In Proceedings of the International Conference on Convergence Information Technology (ICCIT'07)*, Gyeongju, Republic of Korea, November 2007.

[2] K. van Berkel, "Multi-core for mobile phones", *In Proceedings of Design, Automation and Test in Europe (DATE'09)*, 2009.

[3] D. Densmore, R. Passerone, and A. Sangiovanni-Vincentelli, "A platform-based taxonomy for ESL design", *IEEE Design and Test of Computers*, vol. 23, no. 5, pp. 359-374, September/October 2006.

[4] M. Gries, "Methods for evaluating and covering the design space during early design development", *Integration, the VLSI Journal*, vol. 38, no. 2, pp. 131-183, 2004.

[5] J. Mitola III, "Cognitive radio: an integrated agent architecture for software defined radio". Ph.D. dissertation, Computer Communication System Laboratory, Department of Teleinformatics, Royal Institute of Technology, Stockholm, Sweden, May 2000.

[6] End-to-End Reconfigurability. (2007) IST project summary. ftp://ftp.cordis.europa.eu/pub/ist/docs/directorate_d/cnt/e2rii_en.pdf

[7] J. Bard and V. J. Kovarik, *Software defined radio: the software communication architecture*, John Wiley & Sons, March 2007.

[8] G. Gailliard, E. Nicollet, M. Sarlotte and F. Verdier, "Transaction level modeling of SCA compliant software defined radio waveforms and platforms PIM/PSM", *In proceedings of the Design, Automation and Test in Europe Conference (DATE'07)*, San Diego, California, April 2007.

[9] K. Kronlöf, S. Kontinen, I. Oliver and T. Eriksson, "A method for mobile terminal platform architecture development", I*n the proceedings of the Forum on specification and Design Languages (FDL'06)*, Darmstadt, Germany, September 2006.

[10] M. Trautmann, and al., "Simulation Framework for Early Phase Exploration of SDR Platforms: A Case Study of Platform dimensionning," *in proceedings of Design, Automation and Test in Europe Conference (DATE'09)*, Nice, 2009, pp. 312-316.

[11] S. Lecomte, S. Guillouard, C. Moy, P. Leray, P. Soulard. "A co-design methodology based on model driven architecture for real time embedded systems", *Mathematical and Computer Modelling*, vol. 53, no. 3-4, pp. 471-484, 2011.

[12] D. Densmore, A Simalatsar, A. Davare, R. Passerone and A. Sangiovanni-Vincentelli, "UMTS MPSoC Design Evaluation Using a System Level Design Framework," *in proceedings of Design, Automation and Test in Europe Conference (DATE'09)*, Nice, 2009, pp. 478-484.

[13] A. Barreteau, S. Le Nours and O. Pasquier, "A State-Based Modeling Approach for Efficient Performance Evaluation of Embedded System Architecture at Transaction Level" *Journal of Electrical and Computer Engineering*, vol. 2012, Article ID 537327, 16 pages, 2012.

[14] J. Kreku, M. Hoppari, T. Kestilä, and al., "Combining UML2 application and SystemC platform modelling for performance evaluation of real-time embedded systems", *EURASIP Journal on Embedded Systems*, vol. 2008, 2008.

[15] CoFluent Design, http://www.cofluentdesign.com/

[16] A. Barreteau, S. Le Nours and O. Pasquier, "Executable models for performance assessments of adaptive mobile systems", I*n the proceedings of Software Defined Radio Technical Conference (SDR'09)*, Washington DC, USA, 2009.

[17] 3GPP. 3GPP Technical Specification Group, 25 series: Radio Aspects. Technical specification (Release 8).

[18] IEEE Standard association. IEEE 802.11-2007, IEEE standard for information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, Technical specification, June 2007.

[19] 3GPP. (2002) TS 125.201, UMTS; Physical layer - general description. http://www.3gpp.org/ftp/Specs/html-info/25201.htm

[20] IEEE Computer Society. (2007) IEEE Standard association. IEEE 802.11-2007, IEEE standard for information technology - Specific requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications. [Online]. http://www.ahltek.com/WhitePaperspdf/802.11-20%20specs/802.11-2007.pdf

[21] 3GPP. (2002, Sep.) TS 125.212, UMTS; Multiplexing and channel coding. [Online]. http://www.3gpp.org/ftp/Specs/html-info/25212.htm

[22] G. D. Forney, "The Viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268-278, Mar. 1973.

ANALOG-TO-DIGITAL CONVERSION – THE BOTTLENECK FOR SOFTWARE DEFINED RADIO FRONTENDS

Gerald Ulbricht (Fraunhofer Institute for Integrated Circuits, Erlangen, Germany; gerald.ulbricht@iis.fraunhofer.de)

ABSTRACT

Multi-standard capability of current wireless equipment requires a major degree of complexity and flexibility of classical radio frontends. Wideband architectures in combination with software defined radio (SDR) technologies can considerably reduce the demands on the analog frontends. Consequently, analog-to-digital converters (ADC) with high sampling rates and large dynamic range are necessary. Even though current ADCs are improving steadily, they are still one of the limiting components in software defined receivers.

This paper gives an overview of the state-of-the-art of analog-to-digital converters and dynamic range enhancement techniques like signal averaging or non-uniform quantization.

1. INTRODUCTION

An increasing number of radio access technologies and frequency bands for mobile communications present a major challenge for the equipment manufacturers. Future base stations have to handle GSM (Global System for Mobile Communications) with EDGE (Enhanced Data rates for GSM Evolution), UMTS (Universal Mobile Telecommunications System) with HSPA (High Speed Packet Access) and LTE (Long Term Evolution), handsets additionally Wi-Fi, near field communication (NFC), Bluetooth, ultra-wideband (UWB) or GPS (Global Positioning System) in a frequency range between roughly 800 and 3000 MHz. Therefore, software defined radio is considered the most attractive technology to meet the multi-standard and multi-band challenges for future radios.

There are several architectures of radio frequency (RF) frontends for multi-standard and multi-band radios. It is now 20 years since J. Mitola III coined the term *software radio* in 1992 and with that the vision of a radio only defined by software [1]. The idea behind is to abstain from an RF frontend and do everything in software. Of course, some RF components like antennas, low noise amplifiers (LNA), power amplifiers (PA) and anti-aliasing filters will still be necessary in most instances. However, there are no mixers or band filters which are otherwise required in conventional RF frontends.

The determining component of a software radio receiver is the ADC. Since there are no frequency selective components in that kind of frontend, all interferers reach the ADC without any attenuation. Svensson pointed out that the transmit signal of a GSM terminal can reach a power of up to -3.5 dBm at the antenna of a second terminal 2 m away [2]. A dynamic range of at least 110 dB is required in order to not decrease the sensitivity of that terminal. In addition to the dynamic range requirement, the consumed power of an ADC with more than 6 GHz sampling rate and the desired dynamic range would be not acceptable [3].

For multi-band application with carrier frequencies of up to 3 GHz, direct sampling seems not to be feasible in the near future. However, software radios are already available for HF receivers with a performance comparable to a high performance conventional architecture [4], [5]. Several technologies and techniques have been developed to increase the dynamic range of the analog-to-digital conversion in order to face the challenges of future software defined radios.

This paper reviews the process of the analog-to-digital conversion in Chapter 2, in Chapter 3 provides an overview of ADC technologies and the dynamic performance currently available in research and products. Chapter 4 provides methods and techniques to enhance the dynamic range of the analog-to-digital conversion on board level, and finally, Chapter 5 concludes this paper.

2. ANALOG-TO-DIGITAL CONVERSION

The process of the analog-to-digital conversion comprises three operations: 1) sampling, as a conversion from continuous time to discrete time; 2) quantization, as a conversion from continuous values to discrete values; and 3) coding, generating a binary representation of the sampled value.

2.1. Sampling

Ideal sampling is expressed as the multiplication of the analog signal $x(t)$ and a periodic pulse train $s_a(t)$

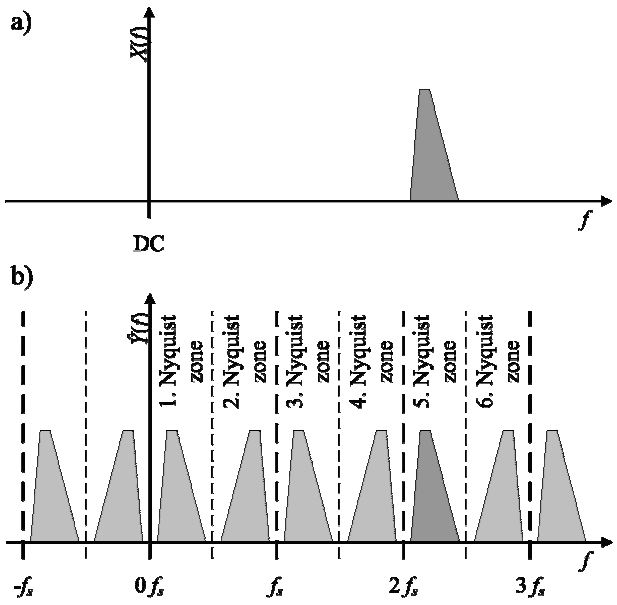


Fig. 1: Sampling process: Spectrum of a) continuous time signal, b) images created by sampling

$$\hat{y}(t) = x(t) \cdot s_a(t) = x(t) \sum_{n=-\infty}^{\infty} \delta(t - nT) = \\ = \sum_{n=-\infty}^{\infty} x(nT)\delta(t - nT). \tag{1}$$

$s_a(t)$ can also be described as a Fourier series

$$s_a(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT) = \frac{1}{T} \sum_{n=-\infty}^{\infty} e^{jn\frac{2\pi}{T}t}. \tag{2}$$

Substituting equation (2) into (1), we obtain

$$\hat{y}(t) = \frac{1}{T_S} \sum_{n=-\infty}^{\infty} x(t) e^{jn\frac{2\pi}{T_S}t}. \tag{3}$$

With the frequency shift property, we obtain for the Fourier transformation

$$\hat{Y}(f) = \frac{1}{T_S} \sum_{n=-\infty}^{\infty} X(f - nf_S), \tag{4}$$

with $f_S = 1/T_S$.

It can be seen from (4) that the spectrum of the sampled signal includes the spectrum of the original signal, but also an infinite number of images spaced at multiples of the sample frequency. The result of the sampling process of a bandpass signal is illustrated in Fig. 1. We can conclude that the analog signal has to be bandlimited to at least half of the sample frequency f_S in order to prevent aliasing. Fig. 1 demonstrates how subsampling can be used to shift a bandlimited signal to a low intermediate frequency (IF) by using a digital low-pass filter.

2.2 Quantization

After sampling, the resulting numbers still can take on an infinite amount of values, which cannot be represented by a digital word. With quantization, each sample is mapped to a discrete value causing a quantization error $e_k(x)$. Assuming $e_k(x)$ has zero mean and is uniformly distributed within a quantization step, the power of the quantization error can be calculated for an ideal ADC according to [6]

$$P_q = \sigma^2 = \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} e_k^2(x) dx = \frac{\Delta^2}{12}, \tag{5}$$

with Δ being the value of a quantization step.

If the analog input signal of the ADC is zero mean and uniformly distributed, the signal-to-quantization-noise ratio (SQNR) is

$$\begin{aligned} \text{SQNR} &= 10 \log_{10}\left(\frac{P_s}{P_q}\right) \text{dB} = 10 \log_{10}\left(\frac{V_{FS}^2}{\Delta^2}\right) \text{dB} = \\ &= 10 \log_{10}(2^{2N}) \text{ dB} = 6{,}02N \text{ dB}, \end{aligned} \tag{6}$$

with $V_{FS} = \Delta \cdot 2^N$ denoting the full scale voltage, and N the number of ADC bits. Typically, analog signals are not uniformly distributed. If the peak-to-average power η of the signal is known, the SQNR is calculated according to

$$\begin{aligned} \text{SQNR} &= 10 \log_{10}\left(\frac{P_{s,peak}/\eta}{P_q}\right) \text{dB} = \\ &= 10 \log_{10}\left(\frac{\left(V_{FS}/2\right)^2}{\eta\sigma^2}\right) \text{dB} = \\ &= 6{,}02N + 4.77 - 10 \log_{10}(\eta) \text{ dB}. \end{aligned} \tag{7}$$

The signal-to-quantization-noise ratio of a sinusoidal input signal is the well-known equation

$$\text{SQNR}_{\text{sinus}} = 6{,}02N + 1{,}76 \text{ dBFS}. \tag{8}$$

2.3 Coding

The digital output of the quantizer is coded to a specified output format. Typically, ADCs use straight binary codes with 0 as minimum and $2^N - 1$ as maximum value for single ended inputs, and two's complement with 2^{N-1} as most negative and $2^{N-1} - 1$ as maximum value for differential inputs. Other codes like Gray or BCD coding are possible but scarcely used.

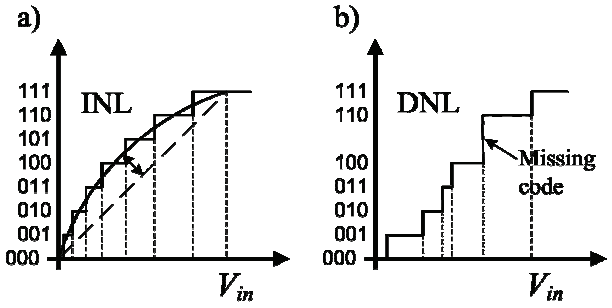


Fig. 3: 3-bit ADC with a) integral non-linearities, b) differential non-linearities

2.4 Non-ideal analog-to-digital conversion

So far, we only have considered ideal ADCs. In practice, the ADC comprises imperfections which limit the performance.

2.4.1 Static converter characteristics

Though the analog-to-digital conversion is a non-linear process, the transfer characteristic can be approximated within the input range of the converter by the linear equation $D = K + GA$, where D is the digital output, A the analog input signal, and K and G are constants. Linear errors occur if K (*offset error*) and G (*gain error*) are not ideal. Non-linear errors are separated into *integral non-linearities* (INL) and *differential non-linearities* (DNL) as illustrated in Fig. 3.

INL is defined as the maximum deviation of the transfer characteristic from a straight line. Two different methods are common to define INL with different results: 1) maximum deviation from the ideal transfer characteristic, and 2) maximum deviation from a straight line with minimum mean squared error (MSE) to the measured characteristic. Typically, the first method is used for values in datasheets.

The change of the input amplitude by $\Delta = V_{FS}/N$ is causing a step of 1 LSB at the digital output of an ideal ADC. The maximum deviation thereof is defined as the DNL. DNL can result in *missing codes* meaning that a quantization level is missing.

All four errors of the transfer characteristic are causing signal distortion and reducing the dynamic performance of the ADC. For choosing the right ADC for a communication receiver, the dynamic characteristics are of greater significance than the static ones.

2.4.2 Noise performance

In Section 2.2 we have already calculated the quantization noise and the SQNR of an ideal ADC, but we have not considered its spectral distribution. Bennett derived in [7] that the quantization noise is typically Gaussian distributed and white within the Nyquist bandwidth.

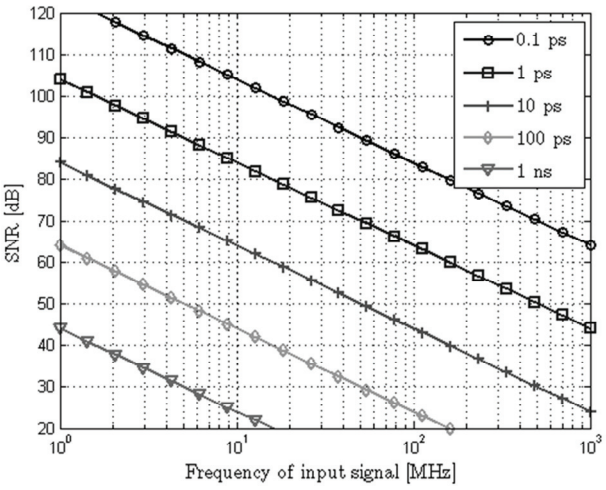


Fig. 2: SNR degradation due to jitter

Quantization noise is not the only and not necessarily the most significant noise source within an ADC. According to [8], three other noise sources can degrade the signal-to-noise ratio (SNR): 1) thermal noise, 2) jitter, and 3) comparator ambiguity. Thermal noise is generated at the analog frontend of the ADC by temperature dependent random movement of electrons in resistive components. Assuming a full scale sinusoidal input signal, the SNR affected by thermal noise in reference to the input is given by the equation

$$\mathrm{SNR}_{\mathrm{therm}} = 10\log_{10}\left(\frac{\Delta^2 \cdot 2^{2N}}{16kTR_{\mathrm{eff}}f_S}\right)\ \mathrm{dBFS}, \qquad (9)$$

with k being Boltzmann's constant $1.38 \cdot 10^{-23}$ J/K, T the temperature in Kelvin and R_{eff} the effective ohmic resistance, representing the complete thermal noise power of the ADC.

So far we assumed an ADC sampling the signal with a periodic pulse train. Due to imperfections of the sample and hold circuitry (*aperture jitter*) and phase noise of the external sample clock (*clock jitter*), the time $T_S = 1/f_S$ is constant in average but not from one sample to the next. The error caused by jitter is proportional to the slew rate of the input signal. That implies higher errors with increasing signal frequency. We assume a sinusoid full scale input signal $v(t) = V_{FS}/2\ \sin(2\pi ft)$. Using that, the derivation of a sinusoid is again sinusoidal and the RMS value of a sinusoid is the amplitude divided by $\sqrt{2}$, the SNR caused by jitter is

$$\mathrm{SNR}_{\mathrm{jitter}} = 20\log_{10}\left(\frac{1}{2\pi f\tau_a}\right)\ \mathrm{dBFS}, \qquad (10)$$

where τ_a is the RMS jitter. The resulting degradation of the SNR due to jitter is illustrated in Fig. 2.

The fourth noise source is based on the finite regeneration time constant of the comparators. According to

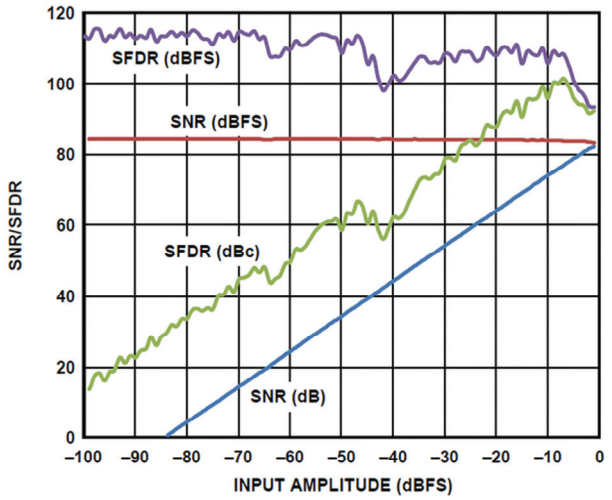


Fig. 4: SNR and SFDR of the 16-bit ADC AD9650 versus input amplitude (from [8])

[8], the effect is only significant at frequencies above several GHz and shall not be further considered.

The cumulative SNR is calculated according to

$$\mathrm{SNR} = -20\log_{10}\sqrt{\sum\left(10^{-\frac{\mathrm{SNR}_i}{20}}\right)^2}, \qquad (11)$$

with SNR_i being the signal-to-noise ratio caused by one individual noise source. Like the quantization noise, we also consider the cumulative noise as white within the Nyquist bandwidth. Since the total noise power does not depend on the sampling frequency, we can conclude that increasing the sampling frequency will improve the SNR with respect to the channel bandwidth (BW). This so-called *processing gain* is increasing the SNR to

$$\mathrm{SNR}_{\mathrm{channel}} = \mathrm{SNR} + 10\log_{10}\left(\frac{f_S}{2\cdot \mathrm{BW}}\right)\ \mathrm{dB}. \qquad (12)$$

Assuming an ADC with $f_S = 104\ \mathrm{MSample/s}$ and 82 dB SNR, sampling the complete 35 MHz wide GSM900 band we calculate the SNR with respect to the 200 kHz GSM channel to

$$\mathrm{SNR}_{\mathrm{channel}} = 82 + 10\log_{10}\left(\frac{104}{2\cdot 0.2}\right)\ \mathrm{dB} = \\ = 106.1\ \mathrm{dB}. \qquad (13)$$

2.4.3 Spurious Signals

Similar to non-linearities in the analog domain, INL and DNL add harmonic distortion to the digitized signal. *Total harmonic distortion* (THD) is one measure to determine the dynamic effect of the non-linearity of the ADC and is calculated using

$$\mathrm{THD} = 10\log_{10}\left(\frac{\sum_{i=2}^{\infty} P_i}{P_0}\right)\mathrm{dB}, \qquad (14)$$

with P_0 being the power of the fundamental and P_i the power of the i-th harmonic. Typically, only the first five harmonics are considered. The THD is a measure of how much power is lost in the harmonics, but of minor interest for communication receiver applications since the lost power can be neglected due to the high linearity of state-of-the-art ADCs.

For a software defined radio receiver, the spurious-free dynamic range (SFDR) is given particular attention, instead. The SFDR is the relation of the strongest spur within the Nyquist bandwidth to the power of the fundamental. For an SDR receiver, the power of the highest spurious signal is important, because generated by an interferer it can cover the wanted signal. Typically, second and third harmonics of the input signal are most significant. Using equation (4), we discover that harmonics, which are initially outside, convolute into the Nyquist band and become relevant. A careful selection of the IF and sampling frequency ensures that the highest spurious fall outside the frequency band of interest and improves the effective SFDR.

In contrast to non-linear distortion in analog circuits, the power of the spurious signals in state-of-the-art ADCs is almost independent of the input signal power, as depicted in Fig. 4. This example of a state-of-the-art 16-bit ADC shows clearly a linear degradation of the SFDR with decreasing input power. Only close to full scale, the behavior is opposite.

In addition to harmonics, two or more in-band interferers can cause harmful intermodulation distortion (IMD). Most attention is turned on third order IMD at $2f_2 - f_1$ and $2f_1 - f_2$, where f_1 and f_2 are the frequencies of interfering signals.

3. ADC ARCHITECTURES

There are quite a few ADC architectures known with pros and cons which shall not be considered here in detail for which comprehensive information is found in the literature (e.g. [6]). Flash converters promise the highest sampling rates, because the result is generated within one single clock cycle. The high speed is bought by a huge circuit complexity if high resolution is required. The number of comparators grows exponentially with the resolution bits of the ADC and a 12-bit converter necessitates already 4095 equal comparators. The power consumption increases also with the number of comparators and therefore, flash converters are common only for high speed applications (several GSamples/s) with resolutions up to 8 bits.

Most of the state-of-the-art ADCs for wireless communications receiver applications are folded flash, subranging or pipelined architectures. These architectures use more than one flash ADC with reduced resolution. There are several clock cycles necessary to generate the

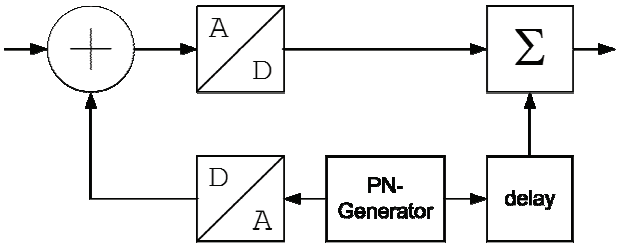


Fig. 5: Subtractive large-scale dither

digital output signal of the converter, which limits the maximum sampling speed.

Within recent years, ΣΔ-converters found their way also into wireless communication receivers. Their high dynamic range is achieved by oversampling in combination with noise shaping. The maximum oversampling ratio and stability of the noise shaping loop limit the useable bandwidth to about 10 MHz.

A good overview of published ADC performance can be found in [10]. State-of-the-art converters currently available on the market achieve an SNR around 80 dB and an SFDR of almost 100 dB at sampling rates above 100 MHz.

4. DYNAMIC RANGE ENHANCEMENT TECHNIQUES

Even though ADC performance is increasing permanently, their SNR and SFDR will be not sufficient for multi-standard, multi-band Software Radios within the near future. However, there are a few measures able to increase the dynamic range of an SDR receiver which shall be discussed within this chapter.

4.1 Automatic Gain Control

Almost every wireless receiver comprises an automatic gain control (AGC) to react on changing input levels. In classical analog frontends, the gain is controlled for maximum signal to noise and distortion ratio (SINAD) within the wanted frequency channel. Actually, the dynamic range of an AGC is only limited by the dynamic range of the controllable amplifiers or attenuators of the analog frontends which can be up to several tens of dB. With software radios, the situation is more critical. No narrowband filters suppress interfering signals before analog-to-digital conversion and controlling the gain for optimum SINAD of the wanted signal might not be possible.

The main goal of the AGC is to ensure not to overdrive the ADC. Of course, other constrains like linearity of amplifiers and mixers have to be considered, too. A strong interferer could cause the AGC to attenuate the received signal so far that the SINAD for the wanted signal is not anymore sufficient for correct demodulation. Therefore, *a priori* information about the expected waveform and the actual interference scenario can help to find the optimal AGC parameters (e.g. time constants, switching thresholds). More details available about the AGC loop dynamic behavior can be found in [11].

4.2 Dither

In literature, two different forms of dither are known: small scale and large scale dither. In Section 2.4 we assumed uniformly distributed quantization errors and therefore a white quantization noise. Depending on the input signal, this is not always the case. According to Bennett, quantization noise is only white if the input signal is random to a certain extent and exercises a high number of quantization levels of the ADC [7]. This is e.g. not true for small sinusoidal input signals. In order to get a flat frequency response of the quantization noise in any case, a so-called small-scale dither signal is added before quantization. [12] and [13] demonstrate that an additive random signal in the range of $\pm\Delta/2$ eliminates the nonlinearity and whitens the noise. The noise power added to the signal is equal to the quantization noise and the noise power increases by 3 dB, which typically can be neglected compared to thermal noise and jitter SNR degradations.

Small-scale dither is not able to whiten the distortion caused by the INL of the ADC and a higher power of the dither signal is required. This dither signal adds a significant amount of interference. There are two methods known to remove this interference from the digital signal: If band-limited noise is added in a frequency range outside the wanted frequency channel, it can easily be rejected by digital filtering. The second method is to generate a pseudo random noise (PN) sequence in the digital domain and to add it after a digital-to-analog conversion before quantization. The dither signal can be subtracted afterwards in the digital domain (see Fig. 5). If the dither level is well adjusted, a significant increase of SFDR can be achieved. Estrada showed in [14] a 13 dB improvement of the SFDR of an 8-bit high-speed ADC by injection of a band-limited dither signal. Note that dithering can only reduce spurious signals caused by the quantizer. Spurs generated e.g. by non-linearities of the frontend are not affected.

4.3 Non-uniform quantization

In order to still be able to receive a weak signal in the presence of a strong interferer, one approach is to quantize small levels with fine quantization intervals and strong levels with coarse intervals. This technique is well-known in audio signal processing. Non-uniform quantization can be performed with particularly designed ADCs, but it is more common to use an analog compressor and a uniform

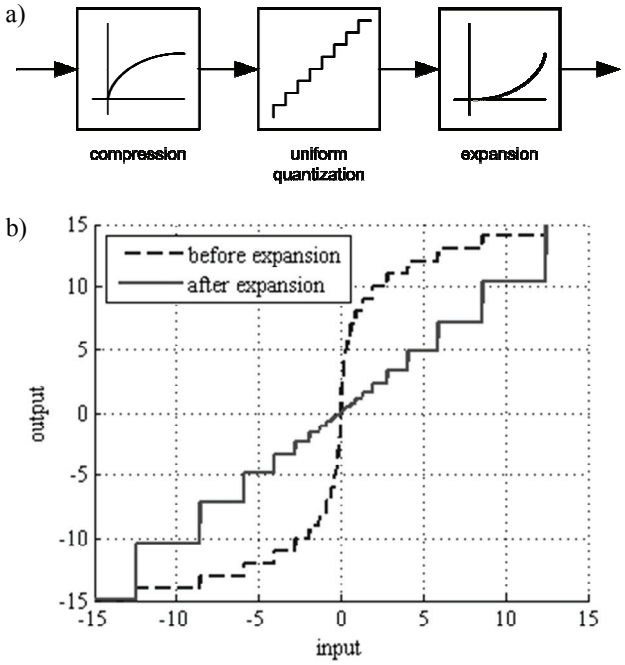


Fig. 6: Non-uniform quantization a) block diagram b) μ-law compression with 8-bit quantizer, μ = 255

quantizer. For compensation of the non-linearity, a digital expansion is necessary after analog-to-digital conversion as shown in Fig. 6a, e.g., a μ-law non-linearity can be used for a compression according to

$$y(x) = \operatorname{sgn}(x)\frac{\log(1+\mu|x|)}{\log(1+\mu)}. \qquad (15)$$

Fig. 6b shows the characteristic of the output signal of the uniform quantizer and the expander, respectively.

Fox demonstrates in [15] that non-uniform quantization does not work properly for an SDR receiver. The weak signal is superimposed on a strong interfering signal and is experiencing fine and coarse quantization depending on the current level of the interferer. Fig. 7 illustrates the difference of the power spectral density (PSD) for uniform and non-uniform quantization. In both cases, a 14-bit uniform quantizer was simulated, the non-uniform quantization was generated by a μ-law compressor with μ=255.

In a), a weak QPSK-signal was used as an input signal. Due to the finer quantization of small input levels, the non-uniform quantization shows an improved SNR. The plot shows also that the uniformly quantized PSD exhibits a non-white noise characteristic since the input signal contains only little frequency content and is utilizing only a few quantization levels.

In b), the input signal additionally contains a strong QPSK-interferer. Now, non-uniform quantization results in higher quantization noise due to reasons described above. Non-uniform quantization does not seem to give an improvement for SDR receivers, since receiving scenarios

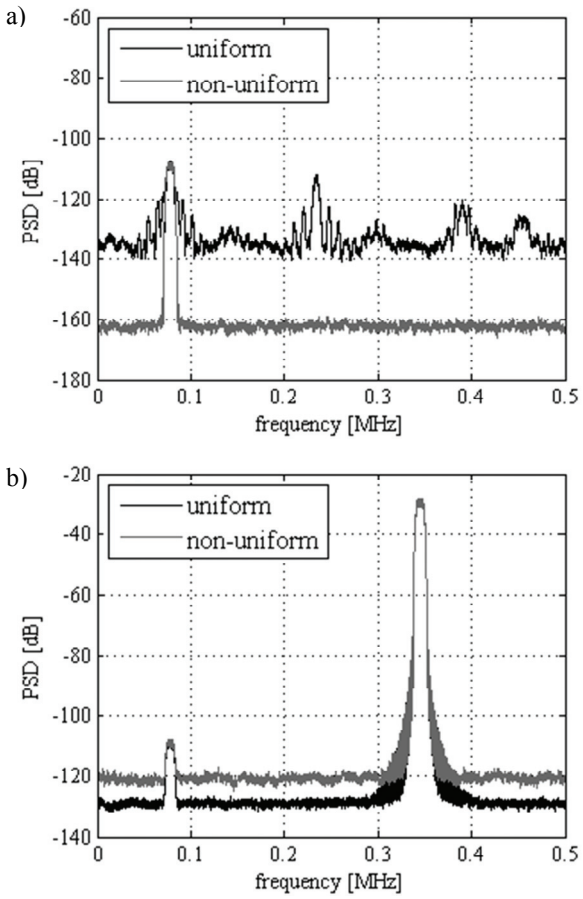


Fig. 7: Uniform and non-uniform quantization, N=14, μ=255 a) PSD of a weak signal b) PSD of a weak signal and a strong interferer

similar to these simulated in case a) can be managed by an AGC.

4.4 Oversampling and time-interleaved ADCs

The complete noise power of the analog-to-digital conversion is considered to be within one Nyquist bandwidth ($f_S/2$). Assuming white noise, the SNR improves with increasing sampling frequency f_S. Doubling of the sampling frequency increases the channel related SNR by 3 dB as shown in Equation (12). In addition, oversampling relaxes the requirements for the anti-aliasing filter, but increases the power consumption and is limited by the maximum sampling rate of the ADC.

With time-interleaving of two or more ADCs the sampling rate can be increased beyond the maximum sampling rate of the single component. The technique was first disclosed by Black and Hodge in [16]. Each ADC samples the same input signal with the same frequency f_S, but with a different phase of the sample clock. The phase of

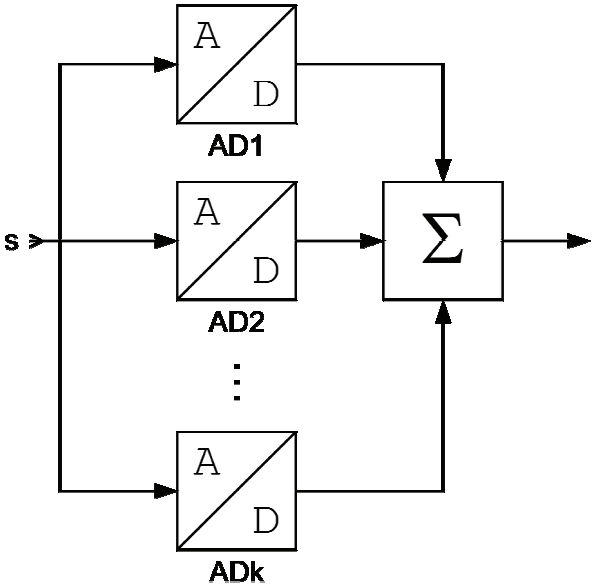


Fig. 8: Signal averaging with parallel ADCs

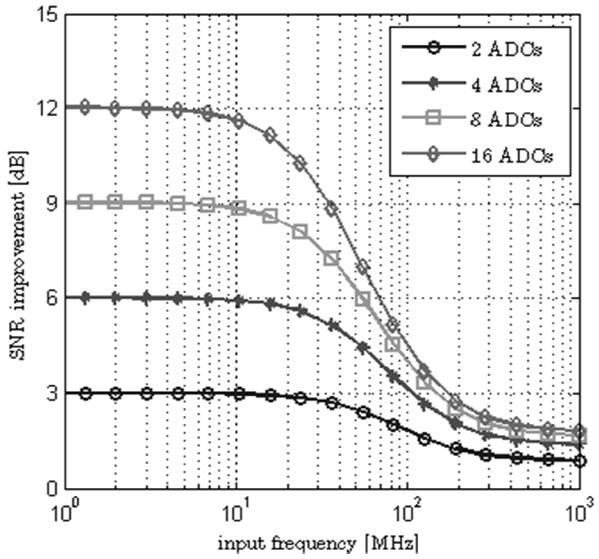


Fig. 9: SNR improvement with signal averaging; data were calculated with SNR_{therm} = 82 dB, τ_a = 75 fs, and τ_c = 100 fs

the sample clock for the individual ADC can be calculated according to

$$\varphi_k = 2\pi \cdot \left(\frac{k-1}{K}\right), \qquad (16)$$

with k being the individual ADC and K being the total number of time-interleaved ADCs. The output signals of the ADCs is multiplexed to result in a total sampling rate of $K \cdot f_S$.

The spurious performance of a time-interleaved ADC depends on the offset, gain and phase matching of the ADCs. The effect of these errors is thoroughly analyzed in [16], [17] and [18]. Spurs caused by offset errors are almost independent of the input signal and occur at multiples of f_S, while spurs due to gain or phase errors depend on the input signal and occur at $f_{in} + i \cdot f_S$, with f_{in} being the frequency of the input signals and $i = 1,2,3,...$ [17] provides the theory and the equations to calculate the resulting SFDR. [18] demonstrates that a gain and phase error of 0.02% causes an SFDR of 74 dB. Digital post-processing can improve the spurious response of time-interleaved ADCs in the presence of mismatch (see e.g. [18]).

4.5 Signal averaging

A further method to improve the SNR of the analog-to-digital conversion was first presented by Seifert and Nauda in [19], known in the literature as signal averaging with parallel ADCs. The idea behind signal averaging in general is increasing the SNR by generating and adding several replicas of the wanted signal. If the noise of the replicas is uncorrelated, it adds on an RMS basis, while the signal adds coherently.

The principle of the technique is illustrated in Fig. 8. Every ADC runs with the same sampling clock. Assuming uncorrelated noise, the SNR increases by 3 dB with every doubling of the number of parallel ADCs. In practice, only thermal noise and aperture jitter is uncorrelated, while clock jitter and quantization noise has to be assumed to be correlated. With that, the SNR is a function of the frequency of the input signal for k parallel ADCs according to [20]

$$\mathrm{SNR} = \left(\frac{1}{k \cdot \mathrm{SNR}_{\mathrm{therm}}} + \frac{(\omega \cdot \tau_a)^2}{k} + (\omega \cdot \tau_c)^2\right)^{-1}, \qquad (17)$$

with τ_a being the RMS aperture jitter, τ_c the RMS clock jitter, and ω the angular frequency of the input signal. The improvement of the SNR for $k = 2 \ldots 16$ ADCs is illustrated in Fig. 9. For SNR and aperture jitter, values from the datasheet of the AD9650 [8] were used and a clock jitter of 100 fs was assumed. Fig. 9 shows that almost no improvement is possible when the clock jitter becomes dominant over thermal SNR. Unfortunately, SFDR does not improve with averaging signal ADCs.

Analog Devices demonstrated about 5.5 dB SNR improvement within the 1st Nyquist zone with its AD10678 comprising four parallel 14-bit ADCs [21].

There has been some effort expended to decorrelate the clock jitter of the individual ADCs e.g. by phase or frequency offsets of the input signal. Lauritzen proves in [20] that 3 dB is the maximum increase of SNR in that case. A further benefit of phase or frequency offsets is the decorrelation of the spurious response. Decorrelation can also be achieved by using independent clock sources for each individual ADC.

5. CONCLUSION

A review of noise and distortion sources limiting the performance of the state-of-the-art analog-to-digital converters was assessed in this paper. There is still a gap between the requirements of a software radio for multi-standard mobile communications receivers and the dynamic range of available ADCs. The paper presented techniques which are able to enhance the dynamic range of the frontend. Automatic gain control is an appropriate method, but the ADC remains the limiting device at the presence of a strong interferer. Significant improvement of the dynamic range can be achieved with signal averaging at the expense of high hardware effort. However, the increase of the SNR is limited to 3 dB at frequencies where the jitter is the dominant noise source.

6. ACKNOWLEDGEMENT

This work has been supported by the Embedded Systems Initiative (ESI) funded by the Bavarian Ministry of Economic Affairs, Infrastructure, Transport and Technology.

7. REFERENCES

[1] J. Mitola, “Software radios-survey, critical evaluation and future directions,” *National Telesystems Conference NTC-92*, pp. 13/15–13/23, 1992.

[2] C. Svensson, “The blocker challenge when implementing software defined radio receiver RF frontends,” *Analog Integrated Circuits and Signal Processing 64*, Nr. 2, pp. 81–89, 2010.

[3] K. H. Lundberg, “High-speed analog-to-digital converter survey,” *unpublished*, 2005.

[4] N. C. Davies, “A high performance HF Software Radio,” *8th International Conference on HF Radio Systems and Techniques*, pp. 249–56, 2000.

[5] P. T. Anderson, “An (almost) all-digital HF communication receiver,” *RF Design*, May 1999, pp. 56–64

[6] J. H. Reed, “Software radio: a modern approach to radio engineering”, Prentice Hall, 2002. – ISBN 9780130811585

[7] W. R. Bennett, “Spectra of Quantized Signals,” *Bell System Technical Journal*, Vol. 27, pp. 446–471, 1948

[8] R. H. Walden, “Analog-to-digital converter survey and analysis, *IEEE Journal on Selected Areas in Communications*, Vol 17, Nr. 4, pp. 539–550, 1999

[9] Data sheet for AD9650 16-bit, 25 MSPS/65 MSPS/80 MSPS/ 105 MSPS, 1.8 V Dual Analog-to-Digital Converter, http://www.analog.com/static/imported-files/data_sheets/AD9650.pdf

[10] B. Murmann, "ADC Performance Survey 1997-2012," [Online]. Available: http://www.stanford.edu/~murmann/adcsurvey.html

[11] E. Kölble, Approach to solve the AGC API issue in the tactical SDR domain - a waveform provider perspective, *Proceedings of SDR'11-WInnComm-Europe*, pp. 48–53, 2011

[12] L. Schuchman, “Dither signals and their effect on quantization noise,” *IEEE Transactions on Communication Technology*, pp 162–165, 1965

[13] N. M. Blachman, “The effect of a nonlinearity upon signals in the presence of noise,” *IEEE Transactions on Communications*, pp 152–154, 1973

[14] Estrada, A.; , "Improving high speed analog to digital converter dynamic range by noise injection," *IEEE Autotestcon 2007*, pp.669-676, 2007

[15] B. L. Fox, “Analysis and dynamic range enhancement of the analog-to-digital interface in multimode radio receivers,” *Master Thesis*, 1997

[16] W. C. Black, D. A. Hodge, “Time interleaved converter arrays,” *IEEE Journal of Solid-State Circuits*, Vol. SC-15, No. 6, pp. 1022–1029, 1980

[17] N. Kurosawa, H. Kobayashi, K. Maruyama, H. Sugawara, K. Kobayashi, “Explicit analysis of channel mismatch effects in time-interleaved ADC systems,” *IEEE Transaction on Circuits and Systems I: Fundamental Theory and Applications*, Vol. 48, No. 3, pp. 261–271, 2001

[18] M. Looney, Advanced digital post-processing techniques enhance performance in time-interleaved ADC systems,” *Analog Dialogue 37-8*, 2003

[19] E. Seifert, A. Nauda, “Enhancing the dynamic range of analog-to-digital converters by reducing excess noise,” *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, pp. 574–576, 1989

[20] K. C. Lauritzen, S. H. Talisa, M. Peckerar, “Impact of decorrelation techniques on sampling noise in radio-freauency applications,” *IEEE Transactions on instrumentation and measurement*, Vol. 59, No. 9, pp. 2272–2279, 2010

[21] Data sheet for AD10678 16-bit, 80 MSPS A/D-Converter, http://www.analog.com/static/imported-files/data_sheets/AD10678.pdf

A SYSTEM ARCHITECTURE FOR REAL-TIME MULTI-PATH MIMO FADING CHANNEL EMULATION

Elliot Briggs, Tanja Karp, Brian Nutter * †

Texas Tech University
Electrical and Computer Engineering
Lubbock, Texas, USA 79410
elliot.s.briggs@ttu.edu,
tanja.karp@ttu.edu, brian.nutter@ttu.edu

Dan McLane

Chirp, Inc.
21900 Burbank Blvd., Suite 200
Woodland Hills, CA 91367
dmclane@chirpdata.com

ABSTRACT

Creating a flexible, programmable multi-path fading channel emulator in hardware presents many design and implementation challenges. Past work has focused on creating a stochastic Jakes process generator as well as hardware efficient implementations of arbitrary upsampling. Significant advances in the design of the arbitrary ratio upsampler have been made, which has greatly reduced the hardware complexity and increased the system's performance, allowing the design to support MIMO emulation for 2 transmit antennas in a single FPGA device. Also building on previous work, the Kronecker model is introduced, allowing the user to vary the spatial correlation features in the emulated channel. This paper will quickly summarize the theoretical aspect of MIMO channel emulation. The main focus will be given to architectural and implementation details, following the system design process from end to end. A high level of attention will be given to multi-rate signal processing design aspects that minimize computational complexity. The results section highlights metrics from both simulation and FPGA hardware integration.

Index Terms— MIMO, Channel Emulation, Jakes, Kronecker, FPGA, ASIC, Multi-Rate, Arbitrary Resampling

1. INTRODUCTION

The time-varying nature of a SISO channel can be modeled by the Jakes process [1, 2, 3], which models the temporal statistics of a rich scattering channel observed by a mobile device. Numerical and hardware implementations of Jakes process generators exist throughout the literature, often only focusing on approximating the autocorrelation properties of the Jakes process using various techniques, i.e. [4, 5]. Oftentimes in the literature, implemented architectures neglect practical, fundamental system-level details and overlook the possibility for drastic optimizations and computational complexity reduction, i.e. [6]. This work will highlight some of the possible optimizations that arise when the a full channel emulation architecture is considered.

This work will illustrate the implementation of the Jakes Doppler model used to generate the desired temporal statistics of the channel. The extension to a multi-antenna (MIMO) system requires the introduction of spatial correlation between the antenna arrays in the system, which can be applied using the Kronecker model [7, 8]. The Kronecker model is a simplistic model that is used to correlate a generated set of i.i.d. Jakes processes according to the channel's properties. By combining the Jakes and Kronecker models to achieve programmable temporal and spatial correlation, a MIMO channel emulator can be designed.

First, the generic MIMO channel will be introduced, providing a conceptual framework and establishing notation. Next, the Jakes Doppler model and the Kronecker model will be briefly introduced and discussed. Having introduced the two channel models, a system architecture is presented, from which detailed system-level implementation aspects are highlighted. The system architecture heavily leverages beneficial aspects of multi-rate signal processing techniques to produce a very computationally efficient architecture, targeted for FPGA/ASIC implementation. Finally, system-level simulation results and FPGA integration details are displayed.

2. THE GENERIC MIMO CHANNEL

To establish a conceptual framework, first a generic MIMO system is be described. The number of transmit and receive antennas in the MIMO system are indicated by M and N respectively. The $M \times 1$ transmitted symbol vector $\mathbf{x}$ passes through the channel, modeled by the multiplication of $\mathbf{x}$ with the $N \times M$ matrix of complex coefficients $\mathbf{H}$ to form the re-

*This work has been possible under continuous support of Innovative Integration (www.innovative-dsp.com)

†This material is based in part upon work supported by the Texas Higher Education Coordinating Boards Norman Hackerman Advanced Research Program under Project No. 003644-0041-2009.

ceived $N \times 1$ symbol vector $\mathbf{y}$.

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{x} \\ \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} &= \begin{bmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,M} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N,1} & h_{N,2} & \cdots & h_{N,M} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix} \end{aligned} \tag{1}$$

The multiplication between $\mathbf{H}$ and $\mathbf{x}$ model the passage of the N transmitted symbols traveling through $N \times M$ paths in the channel which are then summed accordingly at the receiver's N antennas.

3. APPLYING TEMPORAL CORRELATION: THE JAKES DOPPLER MODEL

Eq. 1 describes the generalized static MIMO system. A time index is added to Eq. 1 so that a continual transmission of symbols through the channel can be represented. Also, if an element of time is included with the channel matrix $\mathbf{H}$, the channel is allowed to become *time varying*. For simplicity, let the number of transmit and receive antennas $M = 2$ and $N = 2$ respectively and let t indicate the discrete time index in the system. Modifying Eq. 1 accordingly,

$$\begin{aligned} \mathbf{y}[t] &= \mathbf{H}[t]\,\mathbf{x}[t] \\ \begin{bmatrix} y_1[t] \\ y_2[t] \end{bmatrix} &= \begin{bmatrix} h_{1,1}[t] & h_{1,2}[t] \\ h_{2,1}[t] & h_{2,2}[t] \end{bmatrix} \begin{bmatrix} x_1[t] \\ x_2[t] \end{bmatrix}. \end{aligned} \tag{2}$$

The time-varying nature of each complex path gain in $\mathbf{H}[t]$ can be modeled by the Stochastic Jakes Doppler process [1, 2, 3]. The Jakes process is one of many stochastic processes that is used to model mobile wireless channel environments. The Jakes process requires two parameters, the carrier wavelength and the mobile's velocity. The ratio of these parameters defines the maximum Doppler frequency f_{max}.

$$f_{max} = \frac{v}{\lambda} \tag{3}$$

To define the Jakes process, the channel coefficients must be decomposed into their real and imaginary components.

$$j_{n,m}[t] = \mu_0[t] + j\mu_1[t] \tag{4}$$

In the Jakes process, the continuous-time cross correlation between the real and imaginary components of $j_{n,m}$, μ_0 and μ_1 respectively is defined by

$$r_{\mu_0\mu_1}(\tau) = 0, \forall\tau\,, \tag{5}$$

and the continuous-time autocorrelation of each real and imaginary component is defined by

$$r_{\mu\mu}(\tau) = J_0(2\pi f_{max}\tau)\,, \forall\tau\,. \tag{6}$$

Taking the Fourier transform of Eq. 6 defines the continuous-frequency Jakes' power spectral density function.

$$\begin{aligned} S_{\mu\mu}(f) &= \begin{cases} \frac{1}{\pi f_{max}\sqrt{1-(f/f_{max})^2}}, & |f| \le f_{max} \\ 0, & |f| > f_{max} \end{cases} \\ f &\in (-\infty, \infty) \end{aligned} \tag{7}$$

4. APPLYING SPATIAL CORRELATION: THE KRONECKER MODEL

The independent, time-varying complex channel coefficients $j_{m,n}[t]$ are generated to have the temporal statistical properties defined by the Jakes process. The spatial features in a MIMO channel introduce an additional dimension of correlation. Intuitively, if the antennas at the transmitter and receiver are closely spaced, the channel conditions between each antenna pair are likely to be correlated.

The spatial statistics resulting from antenna spacing, radiation pattern, polarization, and the local scattering environment can be lumped into a correlation matrix for each transmit and receive antenna arrays. The correlation matrices for the transmitter and receiver antenna arrays are defined by

$$\begin{aligned} \mathbf{R}_{TX} &= \begin{bmatrix} \rho_{1,1}^{TX} & \rho_{1,2}^{TX} & \cdots & \rho_{1,M}^{TX} \\ \rho_{2,1}^{TX} & \rho_{2,2}^{TX} & \cdots & \rho_{2,M}^{TX} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{M,1}^{TX} & \rho_{M,2}^{TX} & \cdots & \rho_{M,M}^{TX} \end{bmatrix} \\ \mathbf{R}_{RX} &= \begin{bmatrix} \rho_{1,1}^{RX} & \rho_{1,2}^{RX} & \cdots & \rho_{1,N}^{RX} \\ \rho_{2,1}^{RX} & \rho_{2,2}^{RX} & \cdots & \rho_{2,N}^{RX} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{N,1}^{RX} & \rho_{N,2}^{RX} & \cdots & \rho_{N,N}^{RX} \end{bmatrix}, \end{aligned} \tag{8}$$

where $\rho_{m,n}^{TX}$ and $\rho_{m,n}^{RX}$ indicate the correlation coefficients between the m^{th} and n^{th} antenna at the transmit and receive antenna arrays, respectively [7, 8]. In wireless standards, such as the 3GPP long-term evolution (LTE) standard, several of these matrices are defined for varying scenarios [9]. The Kronecker model attempts to model the spatial correlation between the transmitter and receiver by defining an $NM \times NM$ correlation matrix of the MIMO channel, generated by performing the Kronecker product between the two matrices defined in Eq. 8.

$$\mathbf{R}_{MIMO} = \mathbf{R}_{TX} \otimes \mathbf{R}_{RX}\,, \tag{9}$$

where $\otimes$ indicates the Kronecker product operator.

Using $M = 2$ and $N = 2$ as an example, the desired spatial correlation properties are generated using $\mathbf{R}_{MIMO}$ to obtain $\mathbf{H}[t]$ from $\mathbf{J}[t]$, the matrix of i.i.d. Jakes processes.

$$\begin{aligned} \mathrm{Vec}\{\mathbf{H}[t]\} &= \mathbf{C}\mathrm{Vec}\{\mathbf{J}[t]\} \\ \begin{bmatrix} h_{1,1}[t] \\ h_{2,1}[t] \\ h_{1,2}[t] \\ h_{2,2}[t] \end{bmatrix} &= \begin{bmatrix} c_{1,1} & 0 & 0 & 0 \\ c_{2,1} & c_{2,2} & 0 & 0 \\ c_{3,1} & c_{3,2} & c_{3,3} & 0 \\ c_{4,1} & c_{4,2} & c_{4,3} & c_{4,4} \end{bmatrix} \begin{bmatrix} j_{1,1}[t] \\ j_{2,1}[t] \\ j_{1,2}[t] \\ j_{2,2}[t] \end{bmatrix}, \end{aligned} \tag{10}$$

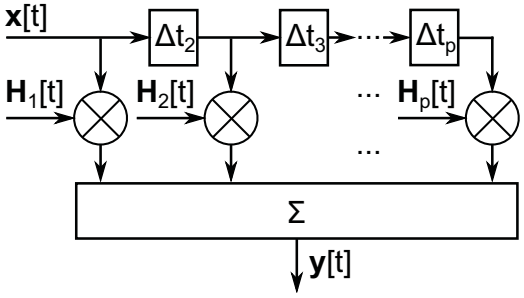


Fig. 1. Tapped Delay Line Model

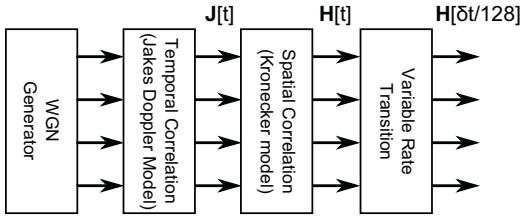


Fig. 2. Channel Matrix Generator

where the lower-diagonal $\mathbf{C}$ matrix is obtained by performing the Cholesky decomposition of $\mathbf{R}_{MIMO}$.

$$\mathbf{R}_{MIMO} = \mathbf{C}\mathbf{C}^T \tag{11}$$

5. GENERAL IMPLEMENTATION ARCHITECTURE

In addition to spatial and temporal correlation, a wireless MIMO channel typically has multiple paths or echoes. To keep the model simple, assume that each element in the channel matrix has the same $P \times 1$ delay vector $\boldsymbol{\tau} = [\tau_1 = 0, \tau_2, \ldots, \tau_P]^T$. Using the $\boldsymbol{\tau}$ vector along with P independent $\mathbf{H}$ matrices, the multi-path MIMO channel can be mathematically described by

$$\mathbf{A}(t, \tau) = \sum_{p=1}^{P} \mathbf{H}_p[t]\, \delta(\tau - \tau_p) \ . \tag{12}$$

Graphically, the system can be described by the tapped-delay line model, illustrated in Fig. 1. This indicates the overall architecture of a multi-path MIMO channel emulator. The tapped delay-line structure that utilizes the P channel matrices will be referred to as the "channel processing" component.

The graphical representation of each $\mathbf{H}$ matrix generator is shown in Fig. 2, which includes the additional "variable rate transition" component. This component performs arbitrary-ratio upsampling for each of the elements in $\mathbf{H}$ so the user can vary the Doppler by programming the desired upsampling factor [10].

6. MULTI-RATE SYSTEM DESIGN

Modern wireless standards are requiring increasingly higher sampling rates, thus wider bandwidths to be processed, specifically LTE-Advanced (release 10), requiring up to 100 MHz of occupied channel bandwidth [11]. As the carrier wavelengths and mobile velocities remain constant, the required rate of the Jakes process generation and the Kronecker spatial correlation components remains constant. To bridge the rate disparity between the high operating rate of channel processing and the low-rate Jakes and Kronecker generators, an arbitrary-rate upsampler is inserted at the end of the processing chain. Assuming the channel processing component operates at a fixed rate, changing the upsampling factor in the variable-rate component will change the rate at which the Jakes and Kronecker components operate, thus allowing the Doppler frequency to be varied!

The rate disparity between the two sections of the channel matrix generator architecture is exploited to form a very computationally efficient implementation. Previous work will be improved [10] and the necessary extensions to support MIMO will be added.

6.1. Jakes Process Generation

As in [2, 3], an FIR filter will be used to process white Gaussian noise to generate each Jakes process. The FIR coefficient set is generated by sampling the square-root Jakes power-spectral density given by Eq. 7. Note that the equality condition in Eq. 7 must be removed to avoid possible divisions by zero.

$$\begin{aligned} H(f_k) &= \sqrt{S_{\mu\mu}(f_k, f_d)}, \\ f_k &= \frac{k-1}{N_{Jakes}} - \frac{1}{2}, \\ k &= 1, 2, \ldots, N_{Jakes}\ , \\ f_d &= \frac{2f_{max}}{f_s}\ , \end{aligned} \tag{13}$$

where N_{Jakes} indicates the number of sampling points and f_s denotes the operating rate of the filter. For a running design example, $f_d = .8$ and $N_{Jakes} = 256$ will be selected. Performing the discrete Fourier transform of $H(f_k)$ yields the sampled Jakes autocorrelation function. Applying a Kaiser window to the result yields the desired impulse response of the FIR Jakes filter minimizing Gibbs phenomenon. Fig. 3 shows the generated Jakes FIR filter before and after applying the Kaiser window to show its effects.

A 2x2 system requires 4 complex Jakes processes, thus 8 real Jakes processes must be generated, requiring 8 Jakes FIR filters. If an absolute Doppler frequency $f_{max} = 100$ Hz is desired and $f_d = .8$, the required filter operating rate is $f_s = 250$ Hz, thus 250 8-element vectors per second must be produced. If the workload of each Jakes filter, defined by the

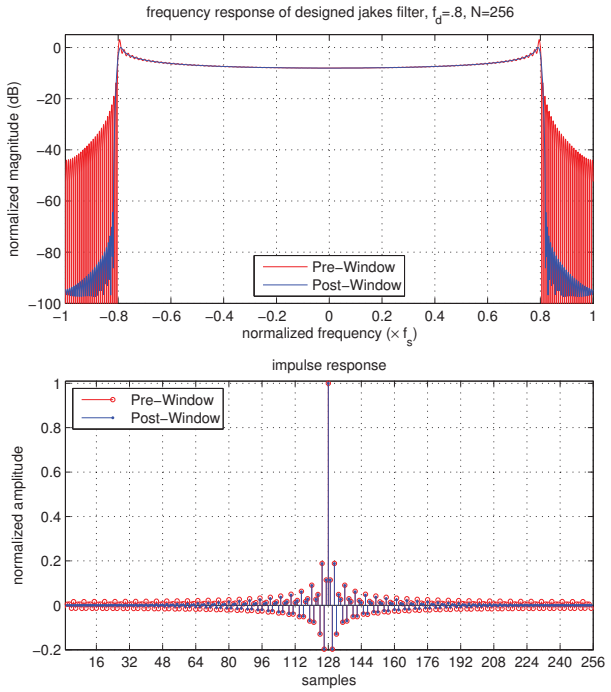


Fig. 3. Jakes FIR Filter

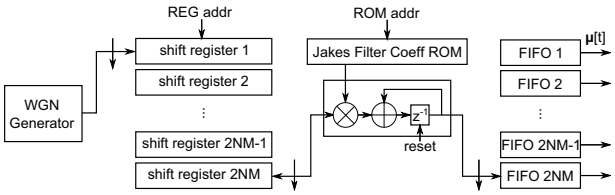


Fig. 4. A Single MACC, $2NM$ Channel Jakes Generator

number of MACC operations required for each output sample, is $W_{Jakes} = 256$, only $8W_{Jakes}f_s = 512,000$ MACCs per second must be performed, a task easily performed using a single MACC element in an FPGA/ASIC implementation.

Additional optimizations can be made when one of the fundamental properties of WGN is exploited. Each sample in a WGN sequence is independent, and uncorrelated with the others, thus a single WGN generator can be time-multiplexed for use with many Jakes process generators! Depending on the maximum possible operating rate of the WGN generator, it is possible to use a single WGN generator across all of the channel matrix generation components!

Fig. 4 depicts the system diagram of an $2NM$ channel Jakes process generator using a single MACC element. The first commutator delivers samples to each $2NM$ shift registers, each containing N_{Jakes} elements. After the shift register is loaded with a new input sample, it is selected by an output commutator, providing access to all N_{Jakes} elements to the MACC component. The register address ROM addresses

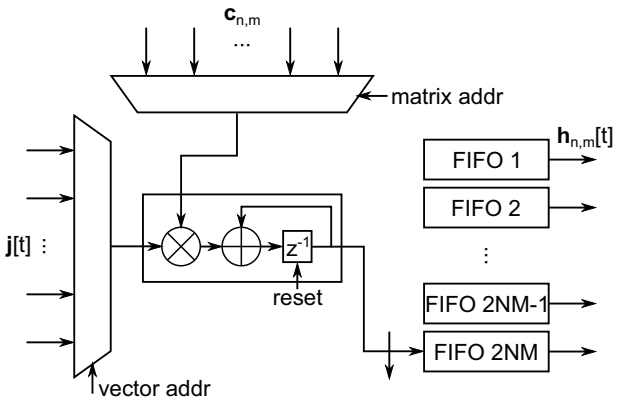


Fig. 5. Matrix-Vector Multiplication Component for the Kronecker Model

are then incremented accordingly, allowing the MACC to perform the necessary operations to generate a single output sample. The output sample is then loaded into its respective FIFO buffer, which are used to time-align each output of the MACC component, resulting in vectors of $\mu[t]$. The Jakes FIR filter is symmetric in nature, requiring only half of the coefficient set to be stored in the Jakes coefficient ROM.

6.2. Applying the Kronecker Model

After generating the $2NM$ Jakes processes, the Kronecker model must be applied to achieve the desired spatial correlation characteristics for the MIMO channel. The required mathematical operation follows from Eq. 10. For a 2x2 MIMO system, the Kronecker model's matrix operation requires 20 real multiplications and 16 real accumulate operations for each produced vector of complex Jakes processes. In the running example, this equates to 250 matrix-vector operations per second, requiring only 5,000 real multiplications and 4,000 real accumulates per second. These operations can be performed using a single MACC element as illustrated in Fig. 5.

6.3. Arbitrary-Rate Upsampling

The arbitrary-rate upsampling component is used to match the very low rate of the Jakes and Kronecker model elements with the potentially high rate channel processing component. The variable rate feature allows the Doppler frequency to be selected by the user. An attractive arbitrary resampler architecture that uses a three-staged approach, capable of bridging the rate disparity in the system has been selected [12, 13]. This architecture utilizes a pair of polyphase upsamplers working in conjunction with a linear interpolation stage to achieve arbitrary-ratio resampling while maintaining levels of dynamic range greater the self-imposed requirement of 16-bits (96 dB). With the assumption that only very large upsampling factors will be required, the mentioned resampling

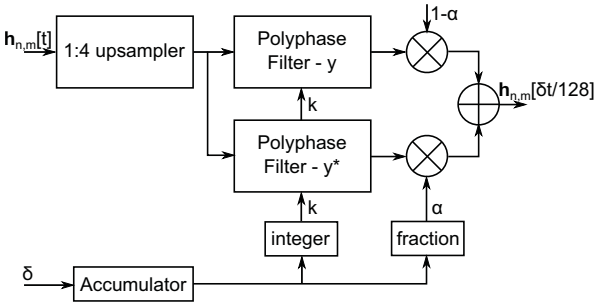


Fig. 6. 3-Stage Arbitrary Resampling Architecture

architecture can be simplified, enabling a highly resource-efficient implementation.

The architecture found in [12, 13] is shown in Fig. 6. To achieve the desired 96 dB dynamic range, the input signal must first be oversampled by a factor of 4, thus the dual polyphase structure is preceded by a 4x upsampling stage. Each of the dual polyphase filters are designed to nominally upsample by a factor of 32x, allowing the final linear interpolation stage to gain the largest jump in overall upsampling.

6.3.1. Cascaded Half-Band Polyphase Preprocessing Upsamplers

A computationally efficient (and resource efficient) 4x upsampling "preprocessing" stage can be implemented using a cascade of 2x upsampling half-band polyphase FIR filters. The half-band FIR filter is specially designed to have almost half of its set of coefficients occupied by zeros, thus the polyphase implementation has two arms, one of which collapses to a pure delay-line, eliminating the need for many MACC operations and dramatically reducing coefficient storage requirements. The overlaid frequency response of the two designed filters is shown in Fig. 7. The half-band filters in Fig. 7 require only 18 and 5 coefficients to be stored for the filters shown by the black and red curves respectively. The total workload of the filter, computed by the number of required MACC operations per output sample, is $W_{\mathrm{HB4x}} = \frac{18+5}{4} = 5.75$. The upsampler must produce 4x the number of vectors per second as the Kronecker model component, therefore in the running example, 1,000 vectors per second, each containing 8 independent filter operations. The vector generation requires 8,000 outputs per second, thus 46,000 MACC operations per second. Each 2x upsampling component in the cascade can be implemented using the structure shown in Fig. 8. One arm of the polyphase half-band filter can be implemented as a pure delay line. Advantageously, the input shift register in Fig. 8 and the delay line hold the same set of samples during operation, thus the shift register implements the storage for the polyphase arm containing non-zero coefficients, and the delay-line for the

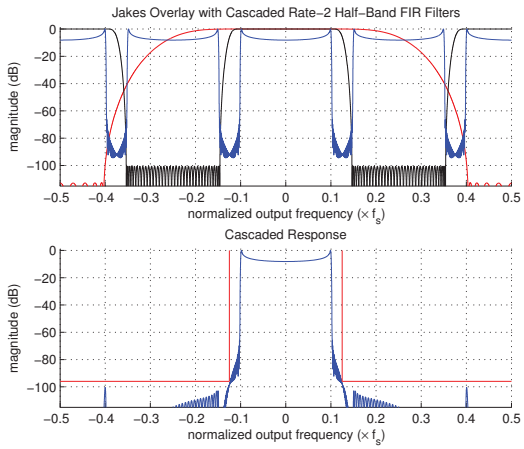


Fig. 7. Cascaded Response of Rate-2 Half-Band Polyphase Upsamplers

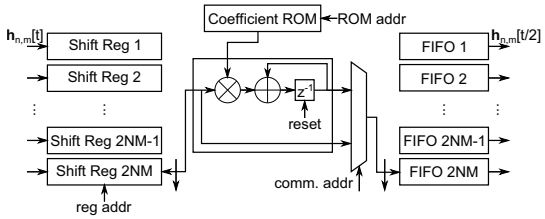


Fig. 8. Rate-2 Half-Band FIR Upsampling

other portion of the filter. A multiplexer is used as a commutator to select samples from the output of the MACC element or the last element in the input shift register.

6.3.2. Fixed-Rate 32x Polyphase Upsampler With Variable-Rate Linear Interpolator

The architecture taken from the literature and shown in Fig. 6 indicates the requirement of two polyphase filters, each of which is designed using the same prototype filter. The prototype filter can exploit the fact that the incoming signal has already been oversampled by a factor of 4, therefore it is only required to provide attenuation in the frequency regions occupied by spectral replicates generated by the upsampling process. Using the Remez algorithm, only the frequency regions that require additional attenuation are included in the stop-band constraints. The remaining regions are considered "don't care" regions, allowing the size of the coefficient set to be dramatically reduced. The response of the cascaded Jakes and 4x upsampling filter is overlaid with the 32x upsampling prototype filter in Fig. 9, magnified on along the frequency axis for clarity. The designed prototype filter is symmetric and has a 192 tap impulse response, thus the implemented filter requires the storage of 96 coefficients.

Given a low enough operating rate, the dual polyphase

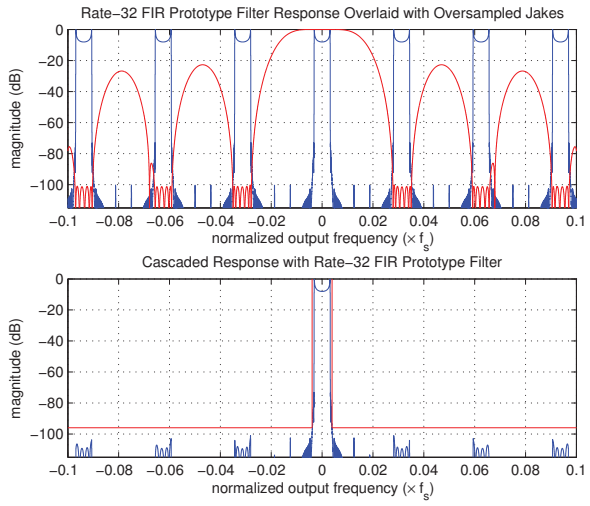


Fig. 9. Prototype Filter Design for the Rate-32 Upsampler

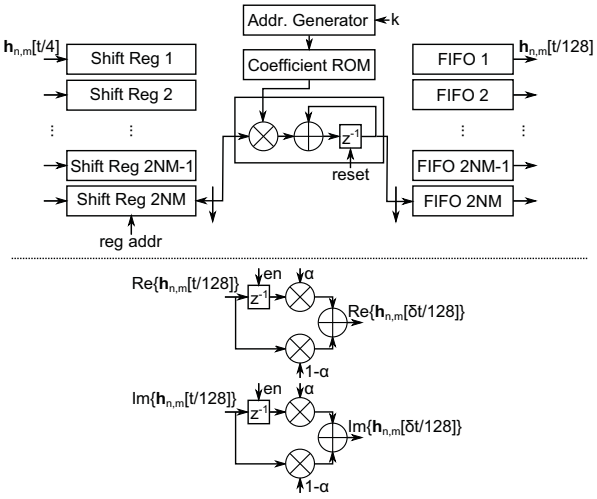


Fig. 10. Arbitrary-Ratio Upsampler: Rate-32 Polyphase Upsampling with Linear Interpolators

structure shown in Fig. 6 can again be implemented using a single MACC component. Using the running example, 1,000 vectors arrive at the 32x upsampler per second, thus 64,000 vectors per second are produced, recalling that the dual structure produces two separate output vectors for each input. Assuming the variable upsampler is always required to upsample by a factor greater than 32x, one of the produced vectors is always a delayed copy of the previous vector, thus an added memory element reduces the workload of the filter by a factor of two and the dual polyphase filter implementation collapses to the standard single polyphase upsampler structure. After this optimization, the workload of the filter is $W_{32x} = \frac{192}{32} = 6$. Using this optimization, the number of vectors produced per second is reduced to 32,000, thus requiring 256,000 outputs per second and 1.536 MMACs per second. With this low workload, it is realistic to assume this upsampler can also be implemented using a single MACC component in an FPGA/ASIC.

The top half of Fig. 10 illustrates the implementation details for the 32x polyphase upsampler. The integer portion of the accumulator k, depicted in Fig. 6 is used to control the virtual commutator address in the coefficient ROM.

The bottom half of Fig. 10 shows the variable linear interpolation components, which operate at the rate of the channel processing portion of the system (Fig 1). The fractional value of the accumulator α, is used to perform the convex combination between the newest and previous output vectors of the 32x upsampling component. The value of the accumulator input δ, determines the final upsampling factor. To determine the desired value of δ for this architecture, the following equation can be used

$$\delta = \frac{256 f_{max}}{f_d f_s} , \tag{14}$$

where f_s is the operating rate of the channel processing component. If the parameters $f_s = 200$ MHz, $f_d = .8$, $f_{max} = 100$ Hz are selected, $\delta = 1.6 \times 10^{-4}$. Given a value of δ, the upsampling factor of the system, including the 4x preprocessing stage, can be determined by

$$M = \frac{128}{\delta} . \tag{15}$$

Therefore, for $\delta = 1.6 \times 10^{-4}$, $M = 8 \times 10^5$.

The frequency response of the full cascade of resampling filters is shown in Fig. 11 using $\delta = 0.0225$ to approximate an irrational resampling ratio of $\frac{128}{0.0225} \cong 5688.8889$.

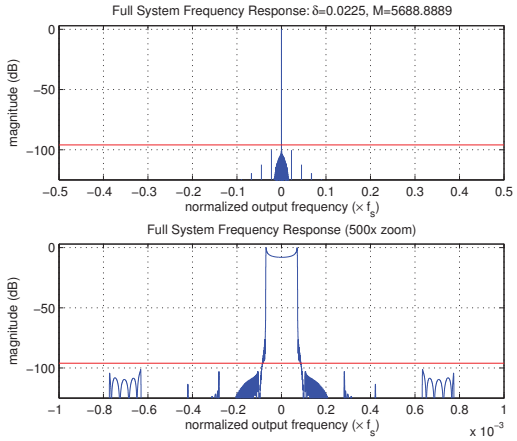


Fig. 11. Frequency Response of the Full System: $\delta = 0.0225$

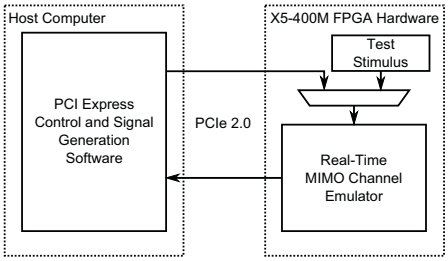


Fig. 12. Hardware Test Setup

	Used	Available	Percentage
slices	1,646	14,720	11%
BRAM	6	244	2%
DSP48E	38	640	6%

Table 1. FPGA Resource Consumption for a Single-Tap Emulator (Excluding WGN Source, $f_s = 200$ MHz)

7. RESULTS

The 2x2 MIMO channel emulator has been implemented using the X5-400M FPGA hardware designed by Innovative Integration. The X5-400M hardware features a Xilinx Virtex 5 SX95T FPGA and communicates with a host computer using PCI Express Gen 2.0. The implemented channel emulator runs at a maximum sampling rate of 200 MHz, processing 1.4901 GBytes/s. The test scenario is depicted in Fig. 12. The host computer generates LTE, or arbitrary signals in software to be processed by the channel emulator. The computer simultaneously and continuously transmits test signals and receives processed signals, which are stored on the host computer for analysis. In this test, the PCI express is capable of providing sustained data transactions of several hundred MBytes/s. In a typical channel emulation configuration, requiring the full processing rate to be sustained, the signal into the channel emulator is sourced from 4 ADCs, processed, and played by 4 DACs (when complex signals baseband signals are used).

Given $f_s = 200$ MHz, 200 MHz of processing bandwidth is achieved. Implementing the arbitrary resampler's accumulator using a width of 22-bits provides a selectable Doppler resolution of 0.149 Hz. In addition to the programmable Doppler frequency, the spatial correlation matrix can be selected by the user for a given antenna array configuration. The scaling factor of the channel path is performed by scaling the matrix used by the Kronecker model off-line.

The final FPGA resource consumption for a single channel matrix and channel processing element is given in Tbl. 1 (i.e. $P = 1$). The channel processing element multiplies the complex channel matrix with the transmitted complex signal vector (shown in Fig. 1). This matrix-vector multiplication requires 16 real multiplications and 16 real addition operations at the transmitted signal's sampling rate, accounting for

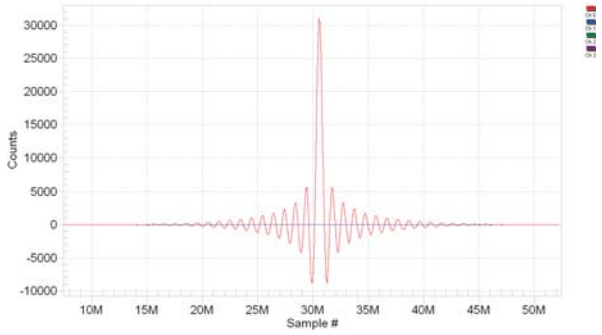


Fig. 13. Hardware-Sourced Jakes Processing Cascade Impulse Response

16 DSP48E elements in the total amount.

To test the Jakes process generation components, a test vector is sourced in hardware, taking the place of the WGN source. Fig 13 shows the impulse responses generated by the 4 complex Jakes signal processing chains by sourcing the vector $[1 - j1, 1 - j1, 1 - j1, 1 - j1]^T$ into the Jakes filter cascade. The transmitted signal vector for both antennas is set to a constant $[1 + j1, 1 + j1]$. For this test, the Kronecker matrix $\mathbf{C}$ is set to the identity matrix. Doing so reveals the cascaded impulse response of the system. The impulse response reveals itself on the real part of the received signal vector and the imaginary part is constant zero. In this test, the Doppler is set to 200 Hz, yielding an impulse response with approximately 50 million taps! The length of the impulse response is determined by the variable Doppler frequency parameter.

Next, the statistical properties of the Kronecker model are verified. To perform verification, WGN is transmitted into the channel emulator and the correlation properties of the received signals are examined. The correlation for each matrix defined by the LTE specification are shown below [9].

$$\begin{aligned} \mathbf{R}_{\mathrm{RXlow}} &= \begin{bmatrix} 1 & 0.0011 \\ 0.0011 & 1 \end{bmatrix} \\ \mathbf{R}_{\mathrm{RXmed}} &= \begin{bmatrix} 1 & 0.2874 \\ 0.2874 & 1 \end{bmatrix} \\ \mathbf{R}_{\mathrm{RXhigh}} &= \begin{bmatrix} 1 & 0.6683 \\ 0.6683 & 1 \end{bmatrix} \end{aligned} \tag{16}$$

8. CONCLUSION

A MIMO channel emulator architecture has been presented that allows the user to select the desired temporal, as well as spatial correlation properties of the channel. A theoretical overview of the Jakes Doppler model as well as the Kronecker spatial correlation model has been given. A detailed architecture has been presented that targets FPGA/ASIC implementation. The implemented design is very resource-efficient, consuming a small fraction of the Xilinx Virtex 5 SX95T

FPGA. Simulation results were given by showcasing the ability of the arbitrary resampling component to achieve very large, high-resolution upsampling factors, allowing the user to select the desired Doppler frequency. Finally, using the three LTE-specified spatial conditions, variable levels of correlation between the received signals has been displayed.

9. REFERENCES

[1] W.C. Jakes, *Microwave Mobile Communications*, John Wiley & Sons Inc, 1974.

[2] M.C. Jeruchim, P. Balaban, and K.S. Shanmugan, *Simulation of Communication Systems: Modeling, Methodology, and Techniques*, Information technology: transmission, processing, and storage. Kluwer Academic/Plenum Publishers, 2000.

[3] M. Pätzold, *Mobile Fading Channels*, J. Wiley, 2002.

[4] M. Patzold, R. Garcia, and F. Laue, "Design of High-Speed Simulation Models for Mobile Fading Channels by Using Table Look-up Techniques," *Vehicular Technology, IEEE Transactions on*, vol. 49, no. 4, pp. 1178 –1190, jul 2000.

[5] Fei Ren and Yahong Zheng, "Hardware Emulation of Wideband Correlated Multiple-Input Multiple-Output Fading Channels," *Journal of Signal Processing Systems*, vol. 66, pp. 273–284, 2012.

[6] A. Alimohammad, S.F. Fard, B.F. Cockburn, and C. Schlegel, "A Novel Technique for Efficient Hardware Simulation of Spatiotemporally Correlated MIMO Fading Channels," in *Communications, 2008. ICC '08. IEEE International Conference on*, may 2008, pp. 718 –724.

[7] K.I. Pedersen, J.B. Andersen, J.P. Kermoal, and P. Mogensen, "A Stochastic Multiple-Input-Multiple-Output Radio Channel Model for Evaluation of Space-Time Coding Algorithms," in *Vehicular Technology Conference, 2000. IEEE VTS-Fall VTC 2000. 52nd*, 2000, vol. 2, pp. 893 –897 vol.2.

[8] J.P. Kermoal, L. Schumacher, K.I. Pedersen, P.E. Mogensen, and F. Frederiksen, "A Stochastic MIMO Radio Channel Model with Experimental Validation," *Selected Areas in Communications, IEEE Journal on*, vol. 20, no. 6, pp. 1211 – 1226, aug 2002.

[9] 3GPP, "User Equipment (UE) Radio Transmission and Reception," 2010.

[10] E. Briggs, D. McLane, and B. Nutter, "A Real-Time Multi-Path Fading Channel Emulator Developed for LTE Testing," in *Wireless Innovation Forum Conference*, Dec. 2011.

[11] S. Sesia, M. Baker, and I. Toufik, *LTE, The UMTS Long Term Evolution: From Theory to Practice*, Wiley, 2009.

[12] F.J. Harris, *Multirate Signal Processing for Communication Systems*, Prentice Hall PTR, 2004.

[13] C. Dick and F. Harris, "Options for Arbitrary Resamplers in FPGA-Based Modulators," in *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on*, nov. 2004, vol. 1, pp. 777 – 781 Vol.1.

MAPPING COGNITIVE RADIO SYSTEM SCENARIOS INTO THE TVWS CONTEXT

Per H. Lehne (Telenor, N-1331 Fornebu, Norway, per-hjalmar.lehne@telenor.com)
Richard MacKenzie (BT Innovate & Design, Ipswich, UK, richard.mackenzie@bt.com)
Dominique Noguet (CEA-LETI, MINATEC, FR-38054, Grenoble cedex 9, France; dominique.noguet@cea.fr)
Vincent Berg (CEA-LETI, MINATEC, FR-38054, Grenoble cedex 9, France; vincent.berg@cea.fr)
Ole Grøndalen (Telenor, N-1331 Fornebu, Norway, ole.grondalen@telenor.com)

ABSTRACT

Cognitive Radio has been one of the key research topics in the wireless community for about 10 years. The digital switch-over in the TV bands provides opportunities for Cognitive Radio Systems (CRS) to operate in the UHF spectrum under incumbent protection restrictions. Regulation bodies, in particular the FCC and OFCOM in the UK, have specified parameters under which CRS shall operate. In this paper we analyze key scenarios for CRS stemming from the QoSMOS project. Then, we analyze how these scenarios can be mapped into the TVWS context by considering link budget computation based on FCC and OFCOM transmit power recommendations as well as statistical propagation models for the UHF band. We also consider the expected capacity which can be achived when using TVWS as a capacity extension in an LTE network. We eventually conclude on the most promising scenarios in the context of the TVWS usage.

1. INTRODUCTION

Radio spectrum is a finite resource. There are many spectrum bands which already suffer from congestion, while at the same time there are other spectrum bands that are highly underutilized. Improved spectrum utilization is essential to allow for future wireless services to satisfy the increasing user demand for wireless capacity, coverage and quality of service. In an attempt to improve the utilization of currently underutilized spectrum bands, there is a growing regulatory trend to allow for license-exempt users to gain opportunistic access to spectrum that is in underutilized licensed spectrum bands. An opportunistic user must act as a cognitive radio in order to avoid interference with primary/licensed users. It should also cooperate fairly with other opportunistic users (also known as secondary/license-exempt/cognitive users).

"White space" (WS) is a term used to describe a part of radio spectrum (this will be described temporally and spatially as well as by its frequency) that can be available for opportunistic access. An issue that can occur with white spaces is the need for fairness among other opportunistic users which can make it difficult for commercial systems to provide high enough quality of service (QoS) guarantees when using white spaces alone. This is due to the fact that the load contributed by opportunistic users can be unpredictable, yet the provision of even a minimal service level will impose a lower limit on the available bandwidth required. In some scenarios systems may be able to function using white spaces alone, whereas other systems may use white spaces in addition to some licensed spectrum, to provide congestion relief and added functionality. In this paper white spaces existing in the TV band (TV white space (TVWS)) is considered as a particular band of interest as this band is currently being opened up for opportunistic channel access in many areas of the world.

Identifying scenarios at an early stage in system development is important as this can keep further development aligned, working with a common goal in mind. The scenarios identified in this paper are being used by the QoSMOS project [1], [2] to help guide the development of tools and techniques to bring these cognitive radio concepts closer to real-world systems. It can be noted that some of these scenarios are also considered by ETSI RRS [5]. The requirements for systems that could operate in these scenarios have been produced in [3] and [4]. However these scenarios can also offer guidance for cognitive radio developments outside of the QoSMOS project.

The structure of the paper is as follows. In Section II a description is given of the three scenarios and the criteria used to select them. In Section III we analyze how these scenarios can be applicable to the TVWS context bearing in mind regulatory constraints and statistical propagation models.

2. SCENARIOS FOR COGNITIVE RADIO SYSTEMS

If a Cognitive Radio System (CRS) is going to be attractive for most actors in the wireless industry, it has to provide a significant benefit compared to what is possible with today's and tomorrow's mainstream wireless technology. Mainstream technology like 3GPP's LTE, with the

evolution towards LTE-Advanced, and Wi-Fi has a great momentum in the market, and will also provide significant improvements in performance as well as cost in the years to come.

2.1. Evaluation criteria

Three top criteria have been defined in order to select feasible deployment scenarios for a CRS providing both managed QoS and high mobility.

Benefit from CRS technology. The CRS solution should be able to provide a significantly better performance than existing (conventional) systems.

Benefit for actors. Deploying CRS for a particular scenario should provide a significant potential benefit for the actors. It should have a joint maximized benefit both for end users and industrial actors (This includes service providers and network operators, but could also include actors such as database administrators). A successful CRS should be commercially attractive. This criterion addresses the commercial side of the CRS, and the selected scenarios must be likely to provide a better business case than conventional systems.

Managed QoS and mobility. The scenario should cover a range of QoS and/or mobility demands. A scenario's QoS requirements depend on the traffic classes that it will serve and how demanding these traffic classes are.

Further, seven criteria have been used for targeting the most interesting and promising scenarios for business case studies.

Market Potential. The scenario should have a large market potential, e.g. with respect to the number of user terminals or expected revenue for the service. This potential could actually come from reduced costs, e.g. reduced spectrum costs or lower power requirements.

Best Solution. No other solution should appear as a better (w.r.t. e.g. performance, lower cost, have environmental benefits, etc.) solution for the given scenario.

Technical Feasibility. It must be probable that this system can be implemented with current state of the art technology or beyond state of the art technology achievable within a reasonable time frame.

Economic Feasibility. It must be probable that within a period of 3-10 years it will be possible to produce equipment and services to a cost that match the users' willingness to pay. The scenario must offer profitability for all major actors in its ecosystem.

Regulatory Feasibility. If the solution requires regulatory changes in order to be deployed, the changes should be such that it is reasonable to expect that they can be realized within a reasonable time frame.

Ecosystem Feasibility. The ecosystem may consist of customers, partners, suppliers, competitors and local and national authorities. If the scenario imposes great changes in the ecosystem (e.g. roles that disappear), it will be much harder to get acceptance for the solution in the industry.

Benefits for the society. Local or national authorities may be willing to support deployment of a system if the social benefits it represents are large. Political support can also make it much easier to get acceptance for regulatory changes.

2.2. Scenario descriptions and example use cases

Applying the criteria above has resulted in the scenarios described below.

2.2.1. Scenario "Cognitive femtocell"

The femtocell scenario, depicted in Figure 1, describes a user situation with low mobility, but high demands on throughput and QoS. It may also be described as a "hot spot" scenario. Femtocells are always connected to an infrastructure. Both indoor and outdoor deployment is possible.

The stakeholders in this scenario are both mobile and fixed operators as well as private and enterprise users. Examples of use cases for this scenario are:

- Private wireless access solution of the same type as Wi-Fi is used today.
- Public hot spots, where several femtocells comprise a larger coverage area.
- The use of indoor femtocells to provide outdoor coverage in e.g. urban/suburban streets.

The main benefits of using cognitive radio for femtocells are:

- Better interference control than current 3G/LTE femtocell technology which can improve capacity and coverage,

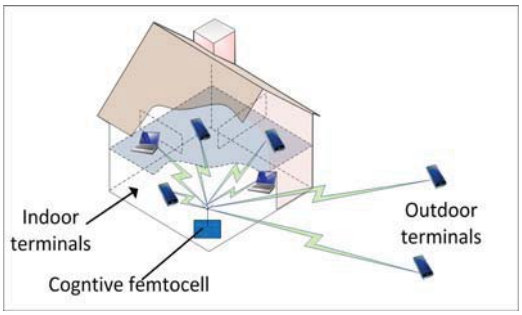


Figure 1. Overview of Cognitive Femtocell

- Better user experience due to more frequencies being available and potentially larger coverage.

2.2.2. Scenario "Cellular extension in whitespace"

Cellular extension in white space, depicted in Figure 2, is where mobile network operators (e.g. LTE-operators) will utilise white space spectrum in addition to their own licensed spectrum. The suitability of a spectrum band for this scenario depends on whether it is to be used for coverage or capacity enhancements.

The stakeholders in this scenario are mainly network operators and service providers. Examples of use cases for this scenario are:

- Increased mobile broadband coverage in rural areas with low traffic demand.
- Peak hour traffic offloading.
- Rural broadband involving the provision of wireless Internet connectivity to homes in rural locations through a base station.

The main benefits of using cognitive radio in this scenario are:

- Better user experience due to more frequencies being available and potentially larger coverage.
- Increased operational bandwidths, resulting in improved load balancing, improved link quality and more flexible services.
- The use of low frequencies increases range and the transmit power can be kept low. This reduces power consumption and reduces health risk concerns (especially for uplink transmissions).

2.2.3. Scenario "Cognitive ad hoc network"

The cognitive ad hoc network scenario, depicted in figure 3, typically includes properties of high dynamics and different nodes and terminals. Ad hoc networks are typically limited in space and time.

The stakeholders in such a scenario are, among others, end users (both private and enterprise), equipment vendors

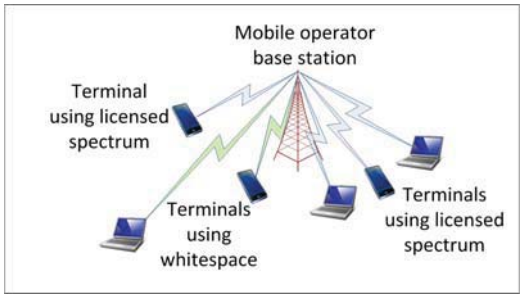


Figure 2. Overview of cellular extension in white spaces

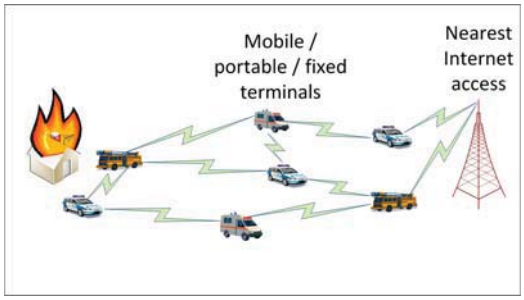


Figure 3 Overview of Cognitive ad hoc Network

and the public sector.

Examples of use cases for this scenario are:

- Emergency ad hoc networks with several actors (police, paramedics and fire fighters) who will typically have two needs: One is to communicate efficiently between one another; the other is to establish a connection to a rescue co-ordination centre.
- A network established for a business meeting to exchange documents and other information. Dependent on the type of event, such a network may be partly pre-planned before the actual event.

The benefits of using cognitive radio for ad hoc networks are:

- The capacity can be increased to serve peak demands without the need for such bandwidth to be allocated during off-peak times.
- The use of low frequency bands is beneficial especially in emergency scenarios due the improved propagation through walls.

3. SCENARIO IMPLEMENTATION IN THE TVWS

3.1. Main parameters for TVWS usage

In this section we analyse how the scenarios described in section II applies to the specific case of the TV whitespace (470-790MHz band). To this aim, the allowed transmit power and the propagation conditions are key elements to determine the link budget for each scenario. In the following, we consider a "typical" transceiver with a 6 dB Noise Figure (NF) and a 1dB insertion loss. These values are derived from consumer UHF silicon TV tuner for which noise figure is between 4 dB and 10 dB. We also assume a SNR of 8 dB, which correspond to a capacity of 2.8 b/s/Hz using Shannon's capacity theorem. From these figures, maximum range can be computed based on statistical propagation models for each scenario.

The transmit power considered in this paper come from the FCC rules [6] and OFCOM statement [7]. The key parameters used hereafter are given in table 1. A channel

bandwidth of 8 MHz is considered in the following calculation. Then, the budget for propagation loss can be computed for the transmit power values of table 1.

Table 1. Transmit power allowed by FCC and foreseen by OFCOM

Parameter	FCC	OFCOM
Power for FD in adjacent band	Not allowed	Not applicable
Power for FD in non-adjacent band with geo-location capability	30dBm (1W) (36dBm EIRP with 6dB gain antenna)	Not applicable
Power for PPD in adjacent band	16dBm (40mW) (Gain antenna not allowed)	4dBm
Power for PPD in non-adjacent band with geo-location capability	20dBm (100mW) (Gain antenna not allowed)	17dBm
Power for PPD in non-adjacent band without geo-location capability	17dBm (50mW)	

FD = Fixed Device ; PDD = Portable Personal Device

Table 2 Propagation budget

TX EIRP	36.00 dBm	20.00 dBm	17.00 dBm	4.00dBm
RX noise Power	-104.97 dBm			
RX Noise Figure	6 dB			
Required SNR	8 dB			
RX Antenna Gain	0 dBi (best case) or -7 dBi (worst case)			
Cable and Connector Loss	1 dB			
Building penetration loss	15 dB			
Minimum RX Levels:				
Best case	-94.97 dBm			
Worst case	-89.97 dBm			
Best case incl. build. pen.	-79.97 dBm			
Worst case incl. build. pen	-72.97 dBm			
Max Propagation Loss for Service:				
Best case	125.97 dB	109.97 dB	106.97 dB	93.97 dB
Worst case	118.97 dB	102.97 dB	99.97 dB	86.97 dB
Best case incl. build. pen.	110.97 dB	94.97 dB	91.95 dB	78.95 dB
Worst case incl. build. pen	103.97 dB	87.95 dB	84.95 dB	71.95 dB

The receiver (user terminal) antenna gain will vary according to terminal type and antenna solution. The typical gain is -7 dBi for a built-in antenna of a handheld terminal, which is the dimensioning case recommended by DVB-H [8]. The best case is not likely to exceed 0 dBi even for external antennas. Both cases are considered. In addition, we look at the use case of indoor-to-outdoor coverage for the cognitive femtocell scenario and model this by adding a penetration loss of 15 dB. The results for estimated maximum propagation loss for service are provided in table 2.

We intend to present average coverage figures, therefore the median path loss will always be calculated for the middle of the UHF TVWS frequency band (630 MHz) and no shadowing is accounted for. For a particular user a different carrier frequency and the presence of shadowing might significantly deteriorate or improve the link budget presented hereafter.

3.2. Range expectations for cognitive radio scenarios

Cognitive femtocell scenario, as well as the ad-hoc network scenario can be divided into two subcases. The first one corresponds to PPD to PPD communication. This link is expected to be a short to medium range indoor link in the femtocell case, with ranges similar to Wi-Fi. The second subcase is a wireless connection to the core network and corresponds to a FD to PPD case, where a long range communication is expected and where the PPD is expected to be fixed. This subcase also corresponds to a rural broadband access configuration. Because the PPD is assume to be fixed antenna gain at the receiver can be envisaged for the downlink. On the other hand, cellular extension in WS involves mobile PPD where no antenna gain at the receiver can be considered.

Thus, as far as range estimation is considered, the scenarios of section 2 can be classified into the categories of table 3.

Table 3. Mapping of QoSMOS usage scenarios to propagation scenarios

Usage scenario	Propagation scenario	Typical range
Cognitive femtocells and ad hoc	Indoor short range for PPD	1 – 100m
Cellular extension	Fixed long range access	1 – 10km
	Mobile cellular	0.1 – 10km

3.3. Range estimation for indoor PPD

These types of propagation conditions have been studied by Saleh and Valenzuela [9]. They propose to model the path loss by using the following equation:

$$PL(r) = 10\log_{10}\left(r^{-\alpha}\right) - 10\log_{10}\left(G_t G_r \left[\frac{\lambda_0}{4\pi}\right]^2\right) \quad (1)$$

Where r is the distance from transmitter, G_t, G_r are respectively transmit and receive gains of the antennas, λ_0 the wavelength of the signal in free space. α is the propagation path loss coefficient that varies from 1.5 to 6. The value of α is a function of the topology of the building where the propagation occurs. Typical values for UHF indoor propagations are between 3 and 4 for same floor propagation and 4 to 6 for propagation across multiple floors. These figures may notably be found in [10], where propagation models have been surveyed.

The distance for which quality of service is guaranteed can then be derived. We propose to use either an α of 3 or 6.

Table 4. Range for indoor PPD case

Carrier Frequency	630 MHz		
TX EIRP	20.00 dBm	17.00 dBm	4.00 dBm
Cell Range:	Best case (0 dBi Rx antenna)		
Indoor, $\alpha = 3$	524 m	416 m	154 m
Indoor, $\alpha = 6$	23 m	20 m	12.5 m
Indoor-to-outdoor, $\alpha = 3$	166 m	132 m	48 m
Cell Range:	Worst case (-7 dBi antenna)		
Indoor, $\alpha = 3$	306 m	243 m	90 m
Indoor, $\alpha = 6$	17.5 m	15.6 m	9.5 m

Under these assumptions, the propagation under indoor conditions is thus expected to range from 10m to 500m depending on the building topology and materials. Therefore it can be concluded that indoor short range communication for portable devices is a viable scenario for TVWS operation. Indoor-to-outdoor communication is also viable, assuming that the indoor only path loss is kept low (α=3) with ranges from 50 to 170 m.

3.4. Range estimation for fixed long range access

In this scenario we suggest using the Okumara-Hata model of propagation. Path loss is given by the following equation for urban environment:

$$PL_{Urban}(d) = 69.55 + 26.16 \cdot \log(f) - 13.82 \cdot \log(h_b) - C_H + [44.9 - 6.55 \cdot \log(h_B)] \log d$$
$$C_H = 0.8 + (1.1 \cdot \log(f) - 0.7) h_m - 1.56 \log(f) \quad (2)$$

where h_b is the height of the mobile antenna and h_B the height of the base station antenna. For sub-urban environment path loss is given by the following equation:

$$PL_{Suburban}(d) = PL_{Urban}(d) - 2\left(\log\left(\frac{f}{28}\right)\right)^2 - 5.4 \quad (3)$$

When considering broadband access, the receiver at the customer premises is located at a higher level (usually at the top of the roof), 4m is a reasonable average value. Secondly the receiver antenna may be highly directional, providing gains up to an extreme 20 to 24dBi. There is also an option to use a low-noise preamplifier to decrease the noise figure to 1…2 dB instead of 6 dB considered in the other scenarios. Therefore the maximum propagation loss may be typically equal to 146dB (using 20dB extra gain in the propagation path loss).

Table 5. Range for long range fixed access

Carrier Frequency	630 MHz
Mobile Terminal Height	4 m
Base Station Height	15 m
CH	5.59
TX EIRP	36.00 dBm
Max Propagation Loss	146 dB
Cell Range	
Okamura-Hata -- Urban	4.72 km
Okamura-Hata -- Suburban	8.27 km

This scenario gives relatively large propagation range up to almost 10 km, which validates the use of fixed access in the TVWS as far as a gain antenna at the receiver can be considered. It shall be noted though that in the femtocell case, the femtocell may be indoor and an additional loss for building penetration must be taken into account, reducing the range to slightly more than 2 km.

3.5. Range estimation for mobile cellular extension

This scenario corresponds to a cellular base station allowed to emit to power levels up to 36 dBm EIRP, and using the derivation from above, 119dB propagation loss budget. We also suggest for these channel conditions the use of the Okumara-Hata propagation model. The maximum expected coverage is then estimated and given in table 6.

Table 6. Range for mobile cellular extension

Carrier Frequency	630 MHz	
Mobile Terminal Height	1.5 m	
Base Station Height	15 m	
CH	0.00	
TX EIRP	36.00 dBm	
Rx antenna gain	-7 dBi	0 dBi
Max Propagation Loss	119 dB	126 dB
Cell Range		
Okamura-Hata -- Urban	0.63 km	0.97 km
Okamura-Hata -- Suburban	1.1 km	1.7 km

We are assuming a base station located 15 m above ground and antenna of the mobile terminal at 1.5m, and using the maximum allowed EIRP of 36dBm (or a maximum propagation path loss of 126 dB) as defined for fixed transmitters.

This gives maximum cell range of 600 – 1000 m for urban environments and 1.1 – 1.7 km for suburban environments. Therefore cellular extension in the TVWS can only be intended for rather small cells, like for instance in congested areas where TVWS can offload part of the cellular traffic.

For Non-Line-of-Sight (NLOS) conditions, which will likely be encountered for the mobile user scenario, the path loss model for the "Rural macrocell" scenario, defined by 3GPP in [12], can also be applied. The pass loss model defined in [12] is known to scale well down to 450 MHz, hence it is adequate to cover the TVWS. The mean path loss in dB has been found to be:

$$PL_{Urban}(d) = 161.04 - 7.1\log(W) + 7.5\log(h) - (24.37 - 3.7(h/h_B)^2)\log(h_B) + \\ + (43.42 - 3.1\log(h_B))(\log(d) - 3) + 20\log(f) - (3.2(\log(11.75h_{UT}))^2 - 4.97) \quad (4)$$

where W is the average street width, h is the average building height, h_B is the base station height and h_{UT} is the user terminal height.

The shadow fading is given as lognormal, with a standard deviation of σ = 8 dB for the NLOS case, σ = 4 dB for the LOS case within the so-called breakpoint distance, σ = 6 dB for the LOS case past the breakpoint distance.

These figures are in good agreement with the shadow fading data specified by TV broadcast recommendations and the mean path loss approximate the model of (2) and (3). For a mean propagation loss between 119 and 126 dB (worst and best case Rx antennas), the range is found to be between 600 and 900 m for the urban environment (W=20 m, h=10 m, h_B=15 m, h_{UT}=1.5 m).

4. CAPACITY CONSIDERATIONS FOR THE CELLULAR EXTENSION SCENARIO

One of the most important parameters to consider in the cellular extension scenarios is the amount of extra capacity that cognitive radio can add to a cellular network, e.g. by using one vacant 8 MHz TV channel. A simulation study was performed to estimate the extra capacity that can be provided by a cognitive LTE system adapted to operate in 8 MHz channels.

4.1 Simulation model

The SEAMCAT simulator [13] was used to estimate the achievable aggregate downlink and uplink bitrates for the cognitive LTE system. This is a tool provided by CEPT to estimate interference between networks. While primarily being a tool for evaluating interference scenarios, it also includes a module that can be used to estimate the capacity obtained in a LTE network.

The cognitive radio can be used to provide extra capacity in hot-spots or it can be used at all BS sites throughout the network to give a uniform increase in the offered capacity. Based on this, three different simulation scenarios were considered:

i) A single LTE BS with one omnidirectional sector
ii) A single LTE BS with 3 sectors
iii) An "infinite" network of LTE BSs, each having 3 sectors.

The network in the last scenario consisted of 19 identical hexagonal 3 sector cells, where the capacity was determined for one of the sectors of the centre cell. SEAMCAT uses a wrap-around technique to remove the network edge effects and thereby create a model of an "infinite" network.

All sub-carriers were used in all sectors, i.e. the re-use factor was 1. It was assumed that all the UEs had unlimited data to send and that all Resource Blocks (RBs) were used at all times, which means that the network load was 100%.

In LTE UL, power control was applied to the active users so that the UE Tx power was adjusted with respect to the path loss to the BS it was connected to. A look up table was used to map throughput in terms of spectral efficiency (bps per Hz) with respect to calculated SNIR (= C/N+I) (dB) level. The tables were taken from the 3GPP TR36.942 document [14]. The maximum spectrum efficiency was 4.4 bit/s/Hz in downlink and 2 bit/s/Hz in uplink, giving maximum bitrates of 33.484 Mbit/s and 15.22 Mbit/s respectively.

For each iteration the UEs were distributed randomly over the geographical area covered by the LTE network. Then, the path loss to all BSs in the network were calculated for each UE and put in a ranked list with the BS with the lowest path loss at the top.

Mobility and hysteresis of handover will have the effect of delaying handovers such that not all UEs will be connected to the optimum base station. This effect is taken into account in SEAMCAT by keeping only the BSs that are less than a handover margin of dBs of the minimum path loss in the BS list. The BS a UE is connected to is then chosen at random from this shortened list. A handover margin of 3 dB was used in the simulations.

The 3GPP antenna pattern from TR 36.942 [14] was used for the 3 sector cells and the antenna gain was assumed to be 6 dBi. The centre frequency was set to 630 MHz and the Hata propagation model for urban environments was used with a log-normal shadow fading of 10 dB. The wall penetrations loss was a random variable with a mean of 10 dB and a standard deviation of 5 dB. The UE and BS

receiver noise figure was set to 6 dB. As explained in [8] the antenna solution in a small hand held terminal has to be an integral part of the terminal construction and will therefore be small when compared to the wavelength. Based on this, the UE antenna gain is assumed to be -7 dBi.

The LTE network is assumed to be located in an urban environment and consist of equally sized cells. The inter-site distance is assumed to be 750 meters, which is a typical number for urban deployments.

4.2 Simulation results

Figure 4 shows the total bit rate as a function of the total site EIRP for indoor UEs.

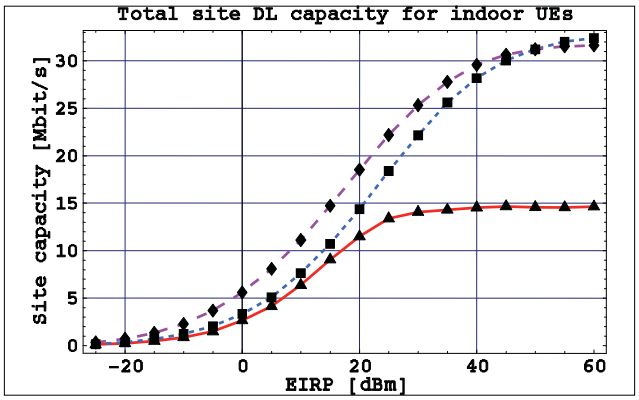


Figure 4. Total site capacity for indoor UEs for single cell with 1 sector (diamonds), single cell with 3 sectors (squares) and an "infinite" network of 3 sector cells (triangles). The Inter-site distance is 750 meters.

It can be seen that the internal interference in a multi-cell LTE network limits the site capacity to about 14.6 Mbit/s. The EIRP limits set by FCC for fixed devices (ref. table 1) are 36 dBm and 30 dBm, which corresponds to a site capacity of 14.3 Mbit/s and 14.1 Mbit/s respectively. Since this is only marginally lower than the maximum achievable capacity, it can be concluded that these EIRP limits will not limit the site capacity in this kind of network.

In the single cell (hot-spot) case the EIRP limits proposed by FCC and OFCOM will limit the capacity. For the single cell with 1 sector case, the maximum site capacity is 33.4 Mbit/s which is almost reached at an EIRP of 55 dBm. With a limit of 36 dBm for the sector EIRP, the site capacity will be reduced to 28.2 Mbit/s. The case of a single cell with 3 sectors has an even higher maximum site capacity, but a site EIRP of 36 dBm limits the site capacity to about 26.2 Mbit/s.

These calculations have been performed with the assumption that the UE antenna gain is -7 dBi which is expected to be realistic for a handheld user terminal. For terminals with larger form factors, such as laptops and tablets, larger antennas can be used and UE antennas gains approaching 0 dBi might be reached. For such terminals the margins in the infinite network case will be even higher.

In practice, there will be a mix of indoor and outdoor UEs and a mix of UEs with different form factors. Hence, the average capacities that will be achieved will be somewhat higher.

Figure 5 shows the uplink site capacity as a function of the maximum UE transmit power. The gain of the BS antenna was set to 6 dBi. All UEs was assumed to be located indoor.

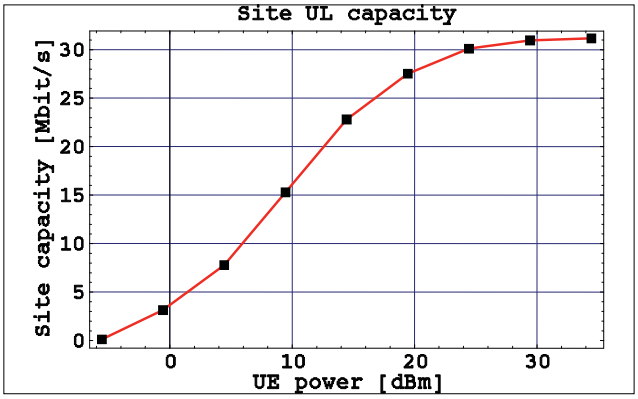


Figure 5. Total site uplink capacity as a function of the maximum transmitted UE power. All UEs are indoor.

According to Table 1, the maximum UE power when operating in a non-adjacent band is specified as 20 dBm by FCC and 17 dBm by OFCOM. This gives total site uplink capacities of 27.9 Mbit/s and 25.5 Mbit/s respectively, which is much higher than the achievable downlink capacities. For operation in adjacent bands, the maximum UE power is specified as 16 dBm and 4 dBm by FCC and OFCOM respectively. A maximum UE power of 16 dBm gives a site capacity of 24.5 Mbit/s, which also more than sufficient compared to the achievable downlink capacities. But a maximum UE power of 4 dBm, will give an uplink capacity of 7.3 Mbit/s, which is less than half the achievable downlink capacity. However, since many services require much lower uplink bitrates than downlink bitrates, even this uplink capacity should be sufficient in many realistic traffic scenarios.

5. CONCLUSIONS

Cognitive Radio concepts are applicable to many different scenarios. The TVWS secondary usage is the first opportunity where they could be deployed at a large scale. Regulation bodies are specifying the rules for white space operation in these bands, with an incumbent protection priority in mind. Therefore, transmit power is limited and

this paper aimed at analyzing which of the scenarios can realistically be foreseen in the TVWS.

From the FCC and OFCOM figures link budget calculation and capacity estimates, it can be concluded that indoor WLAN-like scenarios and fixed broadband access are the most realistic scenarios among those considered by the QoSMOS project. Extension to cellular networks is also possible, but shall focus on dense areas where cells of 1 km are viable from a market point of view. This is typically the case where cellular system offload is required.

The additional capacity offered by cognitive radio in the cellular extension scenario will be limited by the EIRP limits in a single-cell case (hotspot), but still have an acceptable performance. In the multi-cell case, the capacity limitations are dominated by co-channel interference coming from neighbour cells.

6. ACKNOWLEDGMENTS

The research leading to these results was derived from the European Community's Seventh Framework Programme (FP7) under Grant Agreement number 248454 (QoSMOS).

7. REFERENCES

[1] QoSMOS project website. Available at: http://www.ict-qosmos.eu/

[2] R. MacKenzie, P.H. Lehne, U. Celentano, M. Ariyoshi, B. Cendón, "QoSMOS consolidated scenarios", FP7-ICT-2009/248454 QoSMOS Project Deliverable D1.2, Dec. 2010

[3] P.H. Lehne, U. Celentano, J. Lehtomäki, D. Noguet, R. Datta, P. Delahaye, R. Wansch, V. Berg, P. Grønsund, "QoSMOS consolidated system requirements", FP7-ICT-2009/248454 QoSMOS Project Deliverable D1.4, March 2011

[4] D. Noguet, R. Datta, P.H. Lehne, M. Gautier, G. Fettweis, "TVWS regulation and QoSMOS requirements", *2nd International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace in Electronic Systems Technology (WirelessVITAE)*, Chennai, India, March 2011.

[5] M. Mueck, D. Noguet, "TV White Space Standardization and Regulation in Europe", *2nd International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace in Electronic Systems Technology (Wireless VITAE)*, Chennai, India, March 2011

[6] FCC final rule, "Unlicensed Operation in the TV Broadcast Bands", US Federal Register Vol. 74, No.30, pp 7314-7332, Feb. 17 2009

[7] Digital dividend: cognitive access, statement on licence-exempting cognitive devices using interleaved spectrum, OFCOM, July 2009

[8] ETSI TR 102 377 V1.2.1, Digital Video Broadcasting (DVB); DVB-H Implementation Guidelines. November 2005

[9] A. A. M. Saleh and R. L. Valenzuela, "A Statistical Model for Indoor Multipath Propagation", *IEEE Journal on Selected Areas in Comms,* vol. 5, 1987, pp 128-137

[10] T. K. Sarkar, and al., "A survey of Various Propagation Models for Mobile Communications", *IEEE Antennas and Propagation Magazine*, vol. 45, No. 3, June 2003

[11] H. Okamoto, K. Kitao, S. Ichitsubo, "Outdoor-to-Indoor Propagation Loss Prediction in 800-MHz to 8-GHz Band for an Urban Area", *IEEE Trans. Vehicular Technology*, Vol. 58, No. 3, March 2009

[12] 3GPP TR 36.814 V9.0.0, 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 9), Mar 2010

[13] SEAMCAT – Spectrum Engineering Advanced Monte Carlo Analysis Tools. http://www.seamcat.org

[14] 3GPP TR 36.942 9.2.0. 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) system scenarios (Release 9), December 2010

COGNITIVE MULTI-MODE AND MULTI-STANDARD BASE STATIONS: ARCHITECTURE AND SYSTEM ANALYSIS

C. Armani (Selex Elsag, Italy; claudio.armani@selexelsag.com); R. Giuliano (University of Rome Tor Vergata, Rome, Italy; romeo.giuliano@uniroma2.it); F. Mazzenga (University of Rome Tor Vergata, Rome, Italy; mazzenga@ing.uniroma2.it); A. Neri, (University of Rome RomaTRE, Rome, Italy; neri@uniroma3.it).

ABSTRACT

Each wireless technology/standard has been optimized to provide a specific set of services, in accordance to technical and economic aspects. The possibility of changing the communication technology allows the introduction of more flexible management of transmitted power and radio resources in accordance to the offered traffic, services and QoS. The adoption of multi-mode base stations (BSs) offers an additional degree of freedom for efficient usage of the radio resources. One or more radio access technologies can be activated in the single cell area in accordance to offered traffic load and service requests. Users can be distributed among the different technologies. Multi-mode BSs open new and very interesting scenarios for the development and deployment of innovative mobile access networks.

In this paper we consider new access network architecture based on multi-mode BSs even including possible cooperation among different service providers. Capacity improvement due to spectrum sharing is known and shall be encouraged.

It is well known that radio access systems providing spot-like coverage can be used to off-load primary mobile radio systems extending over the entire cell area. Practical examples are provided by Wi-Fi for WCDMA and femtocells. Multi-mode BSs facilitate the implementation of the off-loading concept. Achievable improvements in the case of LTE off loading UMTS are analyzed in this paper.

1. INTRODUCTION

In the last two decades, several mobile communications standards have been developed worldwide to operate within 800 MHz up to few GHz. Each technology has been optimized to provide a specific set of services, according to technical and economic aspects. Waveforms utilized in different standard are characterized by different parameters in terms of transmitted power, occupied bandwidth per channel and Quality of Service (QoS). Multi-standard Software Defined Radio (SDR) systems allow to implement every standard on a single hardware/software platform [1] and, when SDR is combined with a cognitive engine, it allows the efficient use of available bandwidth and power by adapting the modulation scheme as a function of the offered traffic and required services [2], [3]. However, from a practical point of view, current terminals are not SDR-based and implement multi-standard transceivers by using one or more ASIC modules specifically designed and optimized for each one of the considered technologies. The possibility of changing the communication technology so to (optimally) redistribute users and services allows introducing flexible management procedures of power and spectrum resources in accordance to the offered traffic. If different technologies are classified in accordance to power, bandwidth/bit rate required for each service and QoS, they can be intelligently used to manage power and bandwidth in accordance to the number of terminals in the area and their requests in terms of services and QoS.

A flexible wireless network architecture that can meet the previous requirements can be based on multi-mode BSs able to transmit/receive signals in accordance to several wireless access standards (e.g. 2G, 3G, 4G-LTE and even Wi-Fi). The design and deployment of a radio access network (RAN) using multi-mode BSs opens new research challenges mainly oriented to identification of algorithms and techniques for self-organization/configuration, self-healing and self-optimization for the best use of radio resources available in different frequency bands (i.e. from 900 MHz up to 3 GHz). The multi mode network shall also be designed to support collaborative functionalities, mainly oriented to the (fair) spectrum sharing among the different operators.

The wireless network architecture presented in this paper is able to support the above functionalities. It is assumed that each operator has its own infrastructure and can share spectrum resources with other operators. The access network owned by each operator includes multi-mode BSs. The principle architecture of the considered integrated network infrastructure is presented and the main functionalities of its components are detailed. Some practical implementation problems are analyzed even

considering the presence of legacy terminals. Practical hints for the planning and design of a cellular access network with multi-mode BSs are discussed. Some aspects related to spectrum sharing are also presented and analyzed. In particular, the advantages offered by the straightforward implementation in multi mode BSs of the off-loading concept is analyzed in the case of UMTS-WCDMA covering the entire cell area and spot coverage by UMTS-LTE.

The paper is organized as follows. In Section 2 the considered network architecture is presented. In Section 3 a discussion on the network planning and optimization issues are presented taking into account of the presence of multi-mode and collaborative BSs in the network. In Section 4 and Section 5, we analyze the cooperative dual-technology access network work and its performance, respectively, based on WCDMA and LTE in terms of gain due to the technology selection possibility respect to the capacity achievable by the WCDMA network only. Conclusions are drawn in the last Section.

2. SCENARIO AND SYSTEM ARCHITECTURE

The network architecture considered in the paper is detailed in this Section.

2.1. Multi-mode BS

In this section the main features of multi-mode BS are summarized. We consider a BS with an assigned bandwidth B over a variable, carrier frequency f_0. The BS can communicate with terminals in accordance to one or more radio standards such as GSM/GPRS, EDGE and UMTS including LTE. The specific communication technology to be used in B is not fixed a priori but can be selected by the BS in accordance to current traffic conditions and service requests. As an example the BS can decide to use a part of the bandwidth B to scale down to GPRS/EDGE services for each terminal in the area if the number of requests exceeds the practical capacity of UMTS. This means that the BS can allocate a portion of the bandwidth B to transmit GPRS/EDGE sub-carriers instead of UMTS-HSPA signals. If priority concepts are also introduced, the BS can allocate another portion of the managed bandwidth B to serve high priority users with UMTS-HSPA technology thus providing a higher bit rate. In Table 1 an example of service division among technologies has been reported mainly based on coverage and bit rate provided by the different techniques.

Furthermore, taking into account of the different sensitivities associated to each standard, energy efficient strategies can also be introduced. The wireless technology selection allows complying with energy and coverage requirements. In Figure 1 the coverage radii of GSM and WCDMA technologies are reported. The request of download Internet content from MS_2 and MS_3 is performed with different technologies by the users due to their position in the cell (i.e. based on the trade-off between the required bit rate and the allowable technology).

Table 1: Example of technology vs service division

Technology	Coverage	Supported bit rate	Service examples
GSM	High	Low	Voice, SMS
GPRS/EDGE	Medium/ high	Low/ medium	Email, browsing
UMTS-WCDMA	Medium	Medium	Gaming, content download
UMTS-HSPA	Low/ medium	High	Streaming
UMTS-LTE	Low/ medium	High with lower delay	Conference

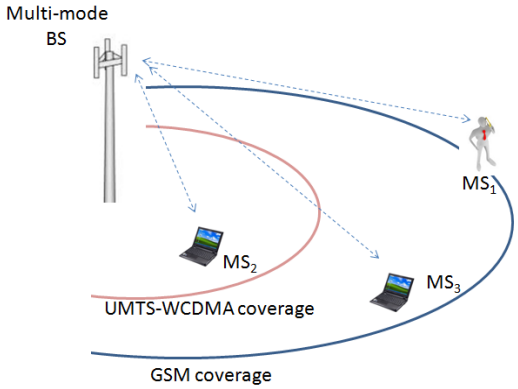


Figure 1: Coverage for different technologies.

The architecture and the main features of the considered multi-mode BS is depicted in Figure 2.

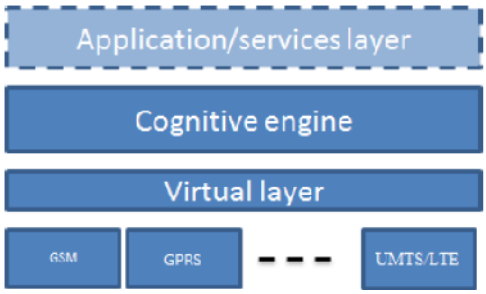


Figure 2: Architecture for the considered multi-mode cognitive BS

The multi-mode BS can receive and transmit signals according to multiple standards (e.g. GSM, GPRS, EDGE, UMTS-WCDMA, UMTS-HSPA and UMTS-LTE). To ease BS management it is necessary to introduce a virtualization

layer to abstract the available wireless technologies to the cognitive/management engine [4]. This permits a simplified "description" of available spectrum resources in terms of "bit tubes" each one characterized by its availability, QoS (i.e. bit error rate, delay and jitter) and bit rate. This abstraction of PHY layer permits to simplify the description of the cognitive and management algorithms since the detailed characteristics of each one of the several PHY layers are wiped out. Cognitive/management layer sends request to the virtual layer for channel assignment having specific characteristics expressed in terms of QoS, achievable bit rate and error rate, etc… The protocol entities inside the virtual layer manage access to the radio resources taking into account for the requests of users, QoS requirements and possibly using information describing the position of the user(s) in the area, their channel quality, the power required for transmission(s), the interference situation and the congestion status of the cell. In this sense the virtualization layer can be seen as an advanced and intelligent multi-mode MAC/DLC layer managing and interfacing to several radio technologies. Starting from connection requests from higher layers, it provides radio connectivity over a multi-mode radio access scenario by proper selection of the radio technology and then channelization. Other important functions hidden inside the virtualization layer are about mobility (e.g. handover), power control, resource scheduling, adaptive modulation and coding. Similar concepts have been already presented in modern radio access networks based on the single RAN solutions presented in [5].

In order for the single BS to accommodate service requests by terminals in the cell area, BS uses parts of the bandwidth *B* to transmit signaling channels that are specific of the considered technologies. If the BS does not wish to activate one radio technology, the corresponding signaling channels are turned off. Multi-mode terminals transmit their access request using the procedures specific for the radio standard they are using for connection and whose signaling channels are transmitted by the BS. Based on received service requests over the different signaling channels, BS can assess the overall traffic load in the cell and for each terminal it can measure the propagation characteristics. Then, BS assigns the best technology to each requesting terminal so to maximize system throughput. Multi-mode BS can receive the service request on one technology selected by the terminal. However, if BS detects favorable propagation conditions and/or the service characteristics of the required service can also be satisfied by means of another and (possibly) better performing standard, the BS can "force" the terminal to select another technology. As an example, the BS can force a multi-mode terminal to use a specific technology by forbidding the use of the others i.e. by switching off signaling channels in the bandwidth *B*, and/or explicitly indicating its choice in the return paging channel. In the latter case it is implicitly assumed that protocol procedures for multi-mode terminals to communicate with the BS have been modified to include information on the radio technology to be used. The former solution based on switching off the signaling channels can be used to permit legacy terminals to interface and communicate over a multi-mode BS. As an example, if multi-mode BS receives one service request over the GSM signaling channel and channel conditions are very favorable, it can deny the request on GSM paging channel, switch off the GSM signaling carrier and switch on the UMTS carrier. In this way, the BS implicitly forces the legacy terminal to search for another technology.

2.2. Multi-mode BS in the access network

The multi-mode BSs described in the previous section can be used to create a flexible RAN. Assuming full cooperation among different operators, a first functional architecture for the considered network is depicted in Figure 3.

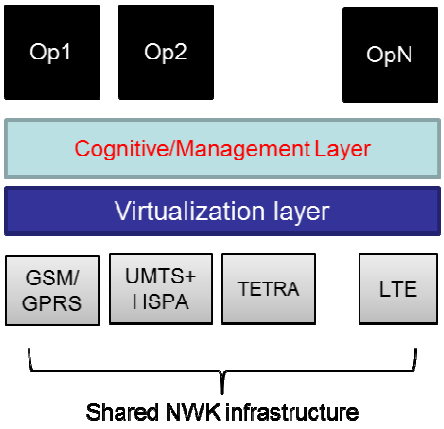


Figure 3: Functional network architecture for a full cooperative access network

To obtain the architecture in Figure 3, it has been assumed that operators share their physical infrastructures and spectrum resources (i.e. a complete cooperative scenario). Operators appear like "applications" running over the cognitive/management layer. The physical network is completely virtualized and operators contend the access to transmission resources i.e. the channels. Virtualization layer now extends over the multi-mode BSs belonging to the operator infrastructure. Contention of radio resources is mediated by the cognitive/management layer which receives the requests of services by users and operators, and it is designed to optimize the spectrum usage (at local and global level), transmission features and so on. The network architecture presented in Figure 3 may turn out to be unrealistic for many reasons. The business model for operators is not completely clear e.g. profit margin for the

different operators should be identified; it is not clear who is the owner and maintainer of the physical network infrastructure; finally legal and regulatory aspects can be very complicated and very difficult to solve.

An alternative and more realistic architecture for the multi-mode access network is presented in Figure 4. Each operator owns one physical network infrastructure and exerts full control/management over it. The physical radio infrastructure of each operator is completely virtualized and cognitive/management layer implements algorithms for the (self) optimization of radio resource utilization.

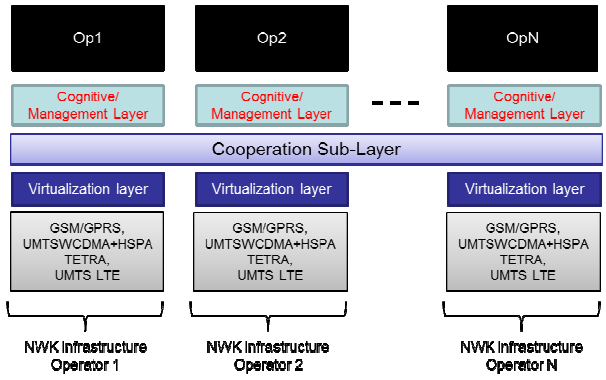


Figure 4: Considered network architecture for the cooperative multi-mode access network

In order to enable resource sharing among operators, a Cooperation sub-layer has been inserted in the network architecture of Figure 4. Protocol entities belonging to this layer reside in the network of each operator and communicate among them using intranet/Internet. Neglecting for the moment the several issues related to standardization, privacy, confidentiality, etc., the main purpose of this layer is to propagate requests for additional radio resources that can be issued by one operator to the others when its network becomes congested. Assuming that more operators cover the same area with their infrastructures (that is typical in the current practice), the request is analyzed and discussed by other operators. In the case only one operator positively responds, the price for resource borrowing has to be negotiated. In the case where more than one operator responds, other forms of negotiations/auctions can take place until one operator wins. The outcome of multiple negotiations depends on several factors such as the traffic load conditions of the networks of the different operators. Poor traffic loaded operators are more prone to borrow their resources rather than leave them unused. Obviously, operators can deny the request of resource. In general, negotiations among operators can be driven by (private) agreements. In general, algorithms for cooperation shall be designed so to guarantee an optimal usage of the spectrum resources and, at the same time, to guarantee a fair treatment for every operator; selfish operators shall be avoided since they do not permit to achieve the advantages of cooperation. For an assigned cooperation strategy, it would be of interest to study the equilibrium properties so to evidence the possible presence of dominant operators. To this aim cooperative game theory can be helpful.

As indicated in Figure 2, the cooperation layer is above the virtualization layer. This permits the adoption of a common and simplified description for the cooperation algorithms in terms of channels, QoS parameters.

3. PLANNING AND DEPLOYMENT OF MULTI-MODE ACCESS NETWORK

The design of a mobile access network with multi mode BS is a complex issue and could be organized in accordance to the following steps.

1. Network planning and initial deployment;
2. Definition of the network entry procedure;
3. Identification, description and classification of services to be supported by the network;
4. Frequency planning and band allocation;
5. Resources request and assignment. The problem of optimal resource allocation is very important and in multi-mode case the selection of technology is an additional degree of freedom to be included in the resource allocation procedures.

In the following of this section, additional details on the steps and procedures listed in the previous points are presented.

Step 1: Network design and deployment

Even though the network can be designed and deployed to be self-organizing and self-optimizing, the problem related to the initial positioning of BSs in the area needs to be addressed. For simplicity it is assumed that carrier frequency is assigned. Since multi-mode BSs are considered, to determine the distance among BSs it is necessary to select one radio technology as reference for initial network dimensioning of coverage. Technology selection can be driven by the expected offered traffic over the considered area, urban, sub-urban and rural. As an example for rural area GSM/GPRS technology can be considered for initial planning and BSs' positioning. In this case the multi-mode BSs can offer UMTS services only for users inside the coverage area that can be served by UMTS (see Figure 1). If the expected traffic is large, such as in urban areas, it may be convenient to consider WCDMA for initial planning. After initial deployment, network can gradually evolve to accommodate more traffic by adding multi-mode BSs so reducing the inter-BSs distance. Network growing can continue until all the service area can

be served by a single high capacity technology such as LTE (or a newer one).

Given the initial coverage planned in accordance to a specific technology at carrier frequency f_0, if one or more BSs can adaptively change technology and carrier frequency network topology can vary with time. In particular one or more cells can increase or decrease their coverage if carrier frequency of the BSs is varied during normal operations. If necessary, to avoid cell breathing the maximum transmitter power level could self-regulated by the BSs so to guarantee the invariance of coverage extension. The addition of new BSs and the corresponding registration and insertion in the network can be managed/regulated by self-organizing algorithms.

Step 2: Network entry

To provide basic access functionalities, BSs transmit signaling channels in accordance to the reference radio technologies used for initial planning. Network entry is then carried out using the procedures specific for these technologies.

As an alternative, the technologies supported by the largest set of terminals expected in the area could be considered as candidate. BSs receiving access request also provides the cognitive/management layer with information on channel quality, terminal capabilities (e.g. supported standards), and types of service requested. As shown in the following, this information can be used not only for channel assignment but even for the (possible) radio technology swapping.

Step 3: Classification and description of communication services

In principle it is possible to distinguish between voice and data services. The characteristics of voice services in terms of QoS are well known and they are supported by every one of the considered 2G–4G technologies. Instead, data services differs for bit rate and QoS requirements. A first classification of services vs supporting radio technology is reported in the following points:

- Voice: supported by GSM+GPRS, WCDMA, LTE;
- Data – Category 1 (up to 200 kb/s with different QoS requirements): supported by GSM+GPRS, EDGE, UMTS, LTE;
- Data – Category 2 (from 200 kb/s up to 500 kb/s with different QoS requirements): supported by WCDMA, LTE;
- Data – Category 3 (more than 500 kb/s with different QoS requirements): well supported only by LTE.

Services could also be classified in accordance to the extension of the coverage they can be provided. An example is reported in the following points:

- Voice: services offered by GSM+GPRS up to 5 km (to be conservative); WCDMA: up to 1 km; LTE: up to 500 m.
- Data Cat. 1: services offered by GSM+GPRS, EDGE up to 4 km; WCDMA: up to 1 km; LTE: up to 300 m.

Classification indicated in the previous points can be used for the smart assignment of radio resources when even the selection of radio technology is accounted for. As shown in the previous points, similar services can be offered by more than one technology with the same/similar level of QoS. However some technologies may require reduced bandwidth occupancy and can serve users far from the BS without requiring an increase in transmitter power especially if carrier frequency can also be changed.

Step 4: Frequency planning and band allocation

In order to visually describe the situation of spectrum resource allocation over each carrier and for each multi-mode BS, we can consider a table indicating on the horizontal axis the carrier frequency and on the vertical the BS. An example is provided in Figure 5.

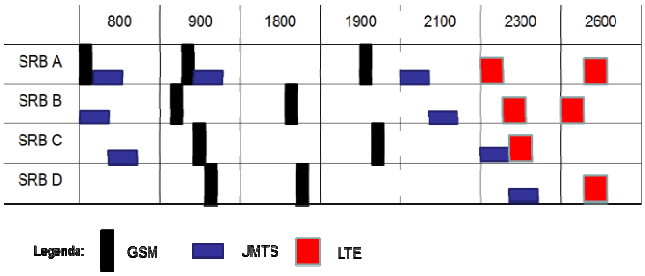


Figure 5: Example for visual representation of radio resource allocation in a multi-mode network

The width of each cell in the table is proportional to the available band around the considered carrier frequency. For each cell it is possible to draw stripes whose color indicates the radio technology and whose width is proportional to the occupied bandwidth as the example illustrated in Figure 5. The problem of radio resource optimization can be studied over a diagram similar to that in Figure 5 by finding the optimum positions of the stripes as a function of the offered traffic and of the interference constraints considered in the adjacent channel interference (ACI) parameter. The co-channel interference and ACI can be a serious concern when multi-mode BSs are considered. As an example, the proper band allocation for GSM and LTE requires to account for co-channel interference so that reuse patterns need to be introduced. In general, inter-cell interference coordination techniques [6], can be adopted. Finally, the situation of band allocation depicted in Figure 5 can change over time in accordance to the (self) spectrum optimization algorithms.

Step 5: Request for resources and admission control

As working principle the user shall be served using "the best" technology compatible with the required service and QoS requirements and guaranteeing the optimal usage of spectrum and transmission resources. Terminal contacts the network using the signaling channels of the radio technology used for initial network dimensioning i.e. the BSs transmit signaling channels for it. We assume that a user X require a data service over the WCDMA interface. The BS measures a favorable transmission channel condition and is aware user could be served (for example with an improved bit rate) using LTE technology which can adopt a higher modulation efficiency thus reducing the bandwidth occupation. Then the BS denies the access request and instructs terminal to move on LTE technology. For legacy terminals this means BS denies the request, then switches off the WCDMA signaling channel (thus forcing the terminal to search for other technologies) and switches on the LTE signaling channel. In the considered example LTE carrier has been turned on to better serve the user. Other users can directly connect to LTE if they are inside its coverage area. When the LTE carrier becomes underutilized i.e. the number of user served by LTE is small, LTE traffic can be re-directed to WCDMA and LTE carrier turned off in order to reduce energy consumption as well as to re-assign that portion of bandwidth B to another radio technology that, for example, can be helpful to guarantee a wider coverage than LTE.

As outlined previously, when a new carrier is turned on it is necessary to consider co-channel interference and ACI. Thus active coordination among BSs in the area is an important feature. Multi-mode BSs should be able to directly communicate. This concept has been already introduced in the LTE network [7] with the definition of the X2 interface. The X2 paves the way for the effective implementation of self organizing and self optimizing networks.

Before concluding this Section, it should be remarked that admission control (AC) is another important procedure for proper network operations. AC shall avoid overloading situations, in which QoS of other users deteriorates, [8]. In multi-mode BS networks any admission criteria shall account for the possibility of selecting the radio technology. As an example, if one WCDMA user cannot be admitted due to the high interference generated by it, BSs can instruct user to select another (active) technology in the cell, for example GPRS, to provide it with a similar service in terms of bit rate and QoS.

Another interesting application of multi-mode BSs consists in using one radio technology to off-load the technology considered for initial planning. To this aim off-load technology shall have a spot-like radio coverage. This concept has been already introduced for example in the integration of hot spot Wi-Fi and WCDMA. A WCDMA terminals detecting Wi-Fi hot-spot executes a vertical and seamless handover to hot spot technology. Multi-mode BSs permit to easily extend this concept it to any pair of technologies. As an example, starting from WCDMA as design technology, LTE can be used to off-load WCDMA when interference increases. In particular assuming LTE coverage is confined in the WCDMA cell area, the same LTE band can be re-used (when necessary) in the neighboring cells without causing harmful co-channel interference. Terminals connecting to UMTS that are inside the LTE coverage can be re-directed by the multi-mode BS to LTE thus reducing the overall WCDMA interference. The achievable performance improvement in this case will be analyzed in the next Section.

4. COOPERATION AMONG TECHNOLOGIES

Thanks to the presence of the virtualization and cooperation layers in Figure 4, telecommunication operators are able to share their spectral resources in order to maximize system capacity. Each licensed bandwidth assigned to one operator could be shared and properly used. In this paper we consider the overall bandwidth assigned to operators to UMTS-WCDMA system (e.g. at frequency 2 GHz). It is composed by $N_u = 9$ bandwidth of 5 MHz assigned to 4 telco operators. Radio mobile system is planned in order to cover the service area with this technology. We assume that operators have co-located BS according to an optimization criterion. Moreover we assume that a second technology could be deployed i.e. some WCDMA carriers could be turned off and UMTS-LTE carriers could be turned on. In this way operators can select the proper technology based on required service and coverage criteria. Generally LTE has a reduced coverage then a possible deployment between two considered technologies is reported in Figure 6 where RLTE is the radius of LTE covered cell, RWCDMA is the radius of WCDMA covered cell and DLTE is the inter-site distance between two LTE cells.

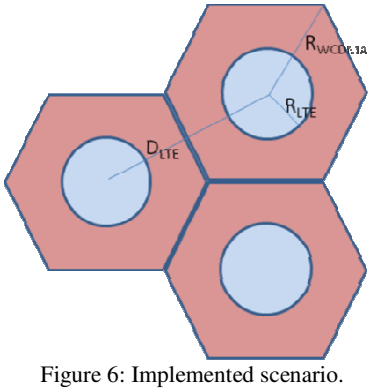


Figure 6: Implemented scenario.

According to this scenario, users can be forced to use WCDMA or LTE access in order to maximize the cell throughput i.e. according to a smart selection of the technology. The selection criterion is based on the user position. Users that are in the LTE covered area can be moved from WCDMA to LTE off loading the consumed capacity in the WCDMA system. We assume that the maximum allowable number of WCDMA user in a cell is reported in (1):

$$N_{WCDMA} = \frac{\eta_{UL}}{1+i} \cdot \left(1 + \frac{W}{R_b \cdot v \cdot \left(E_b/N_0\right)_t}\right) \quad (1)$$

where η_{UL} is the load factor, i is the ratio between the other cell interference and the own cell interference, W = 3.84 Mchip/s is the WCDMA system bandwidth, R_b is the user bit rate, v is the activity factor of user transmission and $(E_b/N_0)_t$ is the energy per bit to noise spectral density ratio target for the considered service.

The number of user off loading WCDMA and moved to LTE is $\gamma \cdot N_{WCDMA}$ in the cell, where γ is the off loading factor and it is equal to the LTE area fraction respect to WCDMA area. In formulas, approximating cells as circles, it yields:

$$\gamma = \frac{{R_{LTE}}^2}{{R_{WCDMA}}^2} \quad (2)$$

The off loaded users can be replaced by the corresponding user quantity. A part of it, in particular a fraction γ of the replaced users, falls in the LTE coverage area and then it could be moved again. Then, at the second time users moved are $\gamma^2 N_{WCDMA}$. This could be reiterated until the LTE maximum capacity is reached. If it is not reached some channel remain unused in LTE system and they could be used for other users in LTE area. Note that theoretically, we have to consider the integer part of off-loading user. Nevertheless we can negligee the integer part thanks to the rate matching box in the receiver chain that allows properly adapting the transmission rates and then users.

After L iterations the number of off loaded users is:

$$N_{off} = N_{WCDMA} \sum_{l=0}^{L} \gamma^l = N_{WCDMA} \cdot \gamma \cdot \frac{1-\gamma^L}{1-\gamma} \quad (3)$$

Note that N_{off} cannot exceed the maximum allowable users of the LTE system. Then $1 \leq N_{off} \leq N_{LTE}$, where N_{LTE} is the number of allocable users in an LTE system and its value is reported in (4):

$$N_{LTE} = \frac{N_{system\_carrier}}{N_{user\_carrier}} \cdot \frac{R_{b-LTE}}{R_b} \quad (4)$$

where $N_{system\_carrier}$ is the number of subcarrier assigned to the system (e.g. for 5 MHz $N_{system\_carrier}$ is equal to 300), $N_{user\_carrier}$ is the number of subcarrier assigned to one user, $R_{b\text{-}LTE}$ is the user bit rate and it is equal to $N_{sym} \cdot N_{bit}/1ms$ with N_{sym} the number of symbol per subframe (1 ms long) and N_{bit} is the number of bit/Hz depending on modulation scheme (i.e. QPSK = 2 bit/Hz, 16QAM = 4 bit/Hz, 64QAM = 6 bit/Hz). Note that we have not considered coding and MIMO.

Considering that the modulation scheme depends on the experienced $(E_b/N_0)_t$ by the user, we can individuate three regions corresponding to QPSK, 16QAM and 64QAM modulation. Moreover increasing the modulation scheme provides greater bandwidth efficiency or allows inserting a higher number of users in the time-frequency resource grid. Then $R_{b\text{-}LTE}$ is rearranged as in (5):

$$R_{b-LTE} = \left(3 \cdot g_{64QAM} + 2 \cdot g_{16QAM} + 1 \cdot g_{QPSK}\right) \cdot \frac{N_{sym} \cdot 2bit}{1ms} \quad (5)$$

where g_k with k={64QAM, 16 QAM, QPSK} is the fraction of covered LTE cell for 64QAM, 16 QAM and QPSK modulation scheme. If $N_{off} > N_{LTE}$, another carrier should be allocated for LTE.

5. RESULTS

The off loading capability leading to WCDMA capacity improvement is analytically evaluated in the scenario described in the previous section.

To have reuse distance equal to 1 for LTE system in Figure 6, it is necessary to evaluate the interference produced by the other cells on the reference cell. To this aim we assume to have a maximum allowable noise raise $r = (I + \eta)/\eta$ that the reference cell is able to tolerate without exceeding the bit error rate or equivalently the $(E_b/N_0)_t$. Then considering the condition for coverage we have:

$$P_{LTE} G_{TR} L_{TR} \frac{L_0}{{R_{LTE}}^{\delta}} = \left(E_b/N_0\right)_t \eta r \quad (6)$$

And considering the condition to calculate the interference I for the reference user, we have:

$$P_{LTE} G_{TR} L_{TR} \frac{L_0}{{D_{LTE}}^{\delta}} = I = \eta(r-1) \quad (7)$$

where P_{LTE} is the transmitting power, G_{TR} and L_{TR} are the join antenna gains and join internal losses for transmitter and receiver, respectively, η is the thermal noise, L_0 is the minimum coupling loss and δ is propagation pathloss exponent. Combining (6) and (7), it yields:

$$\frac{L_0}{{R_{LTE}}^{\delta}} \cdot \frac{{D_{LTE}}^{\delta}}{L_0} = \left(E_b/N_0\right)_t \cdot \frac{r}{r-1} \quad (8)$$

Then, considering that $D_{LTE} \cong 2 \cdot R_{WCDMA}$ and substituting (2) in (8), we obtain the maximum γ with reuse 1:

$$\gamma_M \equiv \gamma = 4 \cdot \left(1 + \left(E_b/N_0\right)_t \cdot \frac{r}{r-1}\right)^{\delta/2} \quad (9)$$

We assume $(E_b/N_0)_t$ = 7 dB for QPSK, $(E_b/N_0)_t$ = 13 dB for 16QAM, $(E_b/N_0)_t$ = 19 dB for 64QAM and r=2 dB. Results for γ regions are reported in Table 2.

Table 2: Values of g for several modulation schemes and constant propagation exponents

	δ=3	δ=3.5	δ=4
$\gamma_{QPSK} \equiv \gamma$	0.5959	0.7335	0.8520
γ_{16QAM}	0.2606	0.3701	0.4778
γ_{64QAM}	0.1080	0.1770	0.2546

In the considered scenario g_k with k={64 QAM, 16 QAM, QPSK} are equal to, respectively:

$$g_{64QAM} = \frac{{R_{64QAM}}^2}{{R_{LTE}}^2} = \frac{\gamma_{64QAM}}{\gamma_{QPSK}};$$

$$g_{16QAM} = \frac{{R_{16QAM}}^2 - {R_{64QAM}}^2}{{R_{LTE}}^2} = \frac{\gamma_{16QAM} - \gamma_{64QAM}}{\gamma_{QPSK}};$$

$$g_{QPSK} = \frac{{R_{LTE}}^2 - {R_{16QAM}}^2}{{R_{LTE}}^2} = \frac{\gamma_{QPSK} - \gamma_{16QAM}}{\gamma_{QPSK}}. \quad (10)$$

The improvement due to smart technology selection is evaluated respect to the case which the operators work in their assigned bandwidth separately.

The gain of technology selection, G_{TS}, is reported in (11):

$$G_{TS} = \left(1 - \frac{l}{N_u}\right) \cdot \left(1 + \gamma \cdot \frac{1-\gamma^L}{1-\gamma}\right) \quad (11)$$

where l is the number of 5 MHz bandwidth assigned to LTE technology.

In Figure 7, the technology selection gain is reported as a function of different values of γ. Several values of L and l are considered as parameters.

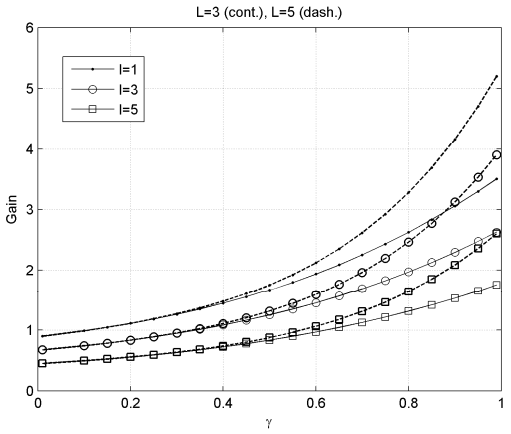


Figure 7: Gain G_{TS} vs γ values

When $G_{TS} \geq 1$, it is convenient to adopt this strategy until γ is allowed to respect the LTE reuse 1 and the number of reiteration does not exceed the maximum number allocable in LTE system (i.e. $N_{off} > N_{LTE}$).

6. CONCLUSIONS

The adoption of multi-mode BS offers an additional degree of freedom for efficient usage of the radio resources. Radio access systems providing spot-like coverage has been proposed in order to off-load primary mobile radio systems extending over the entire cell area. The proposed system architecture highlights a virtualization layer for a full cooperative access network.

Planning procedure has been described to allow the deployment of Multi-mode access network. Entry procedure has also been considered and resources request and assignment has been included in the RAN deployment.

Finally selection technology algorithm has been highlighted. Performance in terms of number of LTE users off-loading the WCDMA technology has been evaluated. The fraction g of area to be allocate to WCDMA and LTE technology in the considered scenario has been analyzed for several values of off loading iterations and number of bandwidth assigned to LTE.

7. REFERENCES

[1] F. K. Jondral, "Software-defined radio: basics and evolution to cognitive radio", *Jour. Wir. Comm. Netw.*, Vol. 2005 Is. 3, Aug 2005.

[2] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey", *Computer Networks Jour.*, Vol. 50, Is. 13, p. 2127-2159, Sept. 2006.

[3] J. Mitola III, "Cognitive Radio for Flexible Mobile Multimedia Communications", *Jour. Mob. Netw. Appl.*, Vol. 6, No. 5, p.435-441, 2001.

[4] Y. Zaki, L. Zhao, C. Goerg, A. Timm-Giel, "LTE mobile network virtualization: Exploiting multiplexing and multi-user diversity gain", *Jour. Mob. Netw. Appl.*, Is. 16, p. 424–432. June 2011.

[5] Eric Xu, *"LTE Single RAN: la parola a HUAWEI"*, TELECOM ITALIA Notiziario Tecnico, Mar 2010, p.42-45.

[6] G. Boudreau, J. Panicker, Ning Guo, Rui Chang, Neng Wang, S. Vrzic, "Interference coordination and cancellation for 4G networks", *IEEE Com Mag*, Vol. 47, Is. 4, Apr. 2009, p. 74-81.

[7] 3rd Generation Partnership Project, Technical Specification, "X2 general aspects and principles (Release 8)", 3GPP TS 36.420 V8.0.0 (2007-12).

[8] N Agoulmine, *Autonomic Network Management Principles: From Concepts to Applications*, Pub. Elsevier, Nov. 2010.

IMPROVING EFFICIENCY OF GENETIC ALGORITHM BASED OPTIMIZER FOR COGNITIVE RADIO

Ajay Sharma (Defence Electronics Applications Laboratory, DRDO Dehradun, India; contactmeajay@gmail.com); Gaurav Kapur (Defence Electronics Applications Laboratory, DRDO Dehradun, India; kapoor_gaurav25@rediffmail.com); Vipin K. Kaushik, Lal C. Mangal, Ramesh C. Agarwal (Defence Electronics Applications Laboratory, DRDO Dehradun, India)

ABSTRACT

An important function of cognitive radio is to adapt its configuration to optimize its resources and performance. Recent researches on cognitive radio have proposed Genetic Algorithm (GA) based multi-objective optimization for this purpose. As the optimization involves GA processing, which is a time consuming activity, it is difficult to get immediate response from system even using best available hardware. Also, to obtain more accurate results from GA, the population size needs to be increased which makes the algorithm take more time to give solution. There are applications that require to have small response time while optimizing radio resources as in case of tactical radio environments. So there is the need to improve efficiency of optimization process in terms of time. The paper considers this problem and proposes a methodology to modify the way GA processing results are used, so that cognitive radio responds faster.

1. INTRODUCTION

Cognitive Radio is an emerging technology in the field of wireless communication. It is useful in operational scenarios where electromagnetic spectrum and radio resources need to be used efficiently and a seamless communication is required. The cognitive radio senses its radio environment, analyses it, and takes appropriate actions to adapt its configurations suitable for current environment. The cognitive radio needs to have a reconfigurable hardware platform to meet this requirement. Software Defined Radio (SDR) provides such a platform to meet this requirement. The cognitive radios are infact intelligent SDRs that use Artificial Intelligence (AI) techniques to adapt its configurations to maximize system capacity.

For efficient spectrum usage, Dynamic Spectrum Access (DSA) is an enabling technology for cognitive radio. DSA enables cognitive radio to sense spectrum allocated to licensed users for detection of free bands, use these bands for communication, and vacate it when the licensed users appears, while causing minimal interference to licensed users.

The cognitive radio has cognitive engine as its subsystem that performs the tasks like radio resource optimization, reasoning and learning. The important radio resources that it needs to optimize are the spectrum and battery power. The radio needs to provide the required Quality of Service (QoS) specified by its user, while optimally using these resources. The radio also uses a knowledge repository for storing knowledge about itself and its radio environment having other radio nodes, to do reasoning and take decisions. In addition, the radio learns from its previous experiences stored in its knowledge-base to take actions in unprecedented scenarios [1]. The paper addresses the resource optimization aspect of cognitive engine.

Recent researches have shown that GA can be used for multi-objective optimization of radio resources [2] [3] [4]. GA is a suitable search technique where search space is very large but it may be difficult to incorporate it in systems where immediate response is required. Also, to obtain more accurate results from GA, the population size needs to be increased, that requires more time to converge. For applications that require to have small response time like in tactical environments, there is the need to improve its efficiency in terms of time. There are advanced GA techniques to enhance the performance of genetic algorithms in terms of accuracy and time. For accuracy, niching can be used to maintain population diversity throughout the GA to find global optimum [5] [6]. Parallel Genetic Algorithms can be used to exploit parallel processing for improving performance [7]. Other methods include biasing the initial population using domain knowledge [8] and using case-based initialization techniques for GA [9]. These techniques improve the performance but still not appropriate for scenarios where

immediate response is required due to underlying time requirement for GA processing dependent on population size.

Case-base decision theory can be applied in the context of cognitive engine where previously observed problems, actions taken and their utilities are stored as cases in a case-base [1] [2]. When the input problem is similar to any case in case-base, the corresponding action is taken based on the utility of action. The paper proposes a similar method for improving efficiency of optimization process. The method is similar in the sense that it stores the GA result as an entry in a lookup table acting like a case-base.

The key idea is to store the optimization results of the genetic algorithm based optimization for given environment parameters and use them subsequently even if the environment parameters change. The approach suggested exploits the observation that there is an overlap between the optimal solutions of GA when there is change in environment parameters.

The rest of paper describes the radio resource optimization in Section-2. The Section-3 discusses the observation from simulations of GA approach, proposed method and its simulation, followed by conclusion in Section-4.

2. RADIO RESOURCE OPTIMIZATION

The SDR has many parameters that it needs to optimize during its operation. These parameters are called Meters [1]. Examples are, Power Consumption, Bit Error Rate (BER), Frame Error Rate (FER), Delay and Throughput. The parameters that represent the configuration of radio to achieve these multiple objectives are called Knobs [1]. Examples are, Transmit Power, Coding Rate, Modulation Order, Modulation Type and Transmit Frequency. The radio user has a QoS requirement in terms of radio meters like minimizing BER, minimizing delay or maximizing throughput. The performance of radio is affected by the radio environment parameters like channel type and propagation conditions. The radio resource optimization problem is to find a configuration for radio in terms of its knob settings that achieves the QoS requirements specified in terms of its meters, while optimally using its radio resources.

2.1. Optimization and Pareto Front

The problem of optimizing the radio meters is a multi-objective optimization problem, because it needs to optimize on its multiple aspects simultaneously. The problem is more complex by the fact that objectives may be directly or indirectly interdependent and conflicting. i.e. improvement in one objective adversely affects other objective. As an example, the objectives BER and power consumption are interdependent and to improve BER requires to increase transmit power or use a more complex Forward Error Correction (FEC). As, it is difficult to optimize all objectives simultaneously, solutions can be found that are compromises. Such solutions are said to lie on a front, called as Pareto front. These solutions have the characteristic that the improvement in one objective is at the cost of degradation for other objectives of the problem.

The concept of non-dominance of solutions can be used for evaluation of pareto front to perform multi-objective optimization using GA [10] [11] [12]. The concept is briefly described here.

2.2. Multi-objective optimization and Non-Dominance

Multi-objective optimization problem can be defined by the basic formula as shown in (1).

$$\begin{aligned} &min/\max \boldsymbol{y} = f(\boldsymbol{x}) = \left(f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \dots, f_n(\boldsymbol{x})\right) \quad (1)\\ &subject\ to\ \boldsymbol{x} = (x_1, x_2, \dots, x_m) \in X\\ &\qquad\qquad \boldsymbol{y} = (y_1, y_2, \dots \dots, y_n) \in Y \end{aligned}$$

The input value **x** belongs to the allowed set of inputs X. In the context of SDR, **x** represents a radio configuration, where $x_1, x_2, \ldots, x_m$, are individual knob values. The output value **y** belongs to the set of outputs Y. It represents objective vector, where $y_1, y_2, \ldots, y_n$ are individual meter values. The objective vector **y** is evaluated using objective functions $f_1, f_2, \ldots, f_n$ for a given **x**.

Mathematically, for a maximization problem with two decision vectors **a**,**b** ∈ X, **a** dominates **b** (also written as **a** ≻ **b**) iff

$$\begin{aligned} &\forall i \in 1,2,\dots,n:\ f_i(\boldsymbol{a}) \geq\ f_i(\boldsymbol{b})\\ &\wedge\ \exists j \in 1,2,\dots,n:\ f_j(\boldsymbol{a}) >\ f_j(\boldsymbol{b}) \end{aligned}$$

A non-dominated individual is one, which is not dominated by any individual in set. Additionally, **a** covers **b**, iff **a** ≻ **b** or **a** = **b**. All decision vectors which are not dominated by any other decision vector are called non-dominated or Pareto-Optimal. The family of all non-dominated solutions is denoted as Pareto-Optimal Set or Pareto-optimal front.

2.3. Wireless Systems Genetic Algorithm

Genetic Algorithm can be used to perform multi-objective optimization for SDR. Wireless Systems Genetic Algorithm (WSGA) is one such algorithm that defines a genetic model

for SDR and does the radio resource optimization through genetic processing [2]. The algorithm is briefly described here.

The basic requirement of any GA based algorithm is the chromosome definition. In WSGA, the PHY and MAC layer parameters of SDR are encapsulated in the genes of a chromosome. The chromosome shown in Figure-1 represents the PHY-layer traits. The algorithm initializes with a population of such chromosomes referred to as individuals, and analyzes populations through many generations of fitness evaluation, selection, crossover and mutation, to find non-dominated individuals.

In WSGA, the decision vectors **a** and **b** are represented by individual chromosomes. The functions to evaluate objectives are called objective functions. In (1), $f_i(\mathbf{x})$ is an objective function that evaluates i^{th} objective for decision vector, **x**. These objective functions are used to determine the fitness for each individual chromosome. Different versions of same objective function can be there for the evaluation of an objective. For example, to evaluate BER there are different objective functions for different channel condition (like AWGN and Non-AWGN channel). The functions are also assigned weights to represent the relative importance the user has associated with different objectives.

It uses a relative tournament selection method, where two chromosomes are randomly chosen to select one as a parent. In this method, fitness of the winner from a comparison on single objective is scaled by the associated weight. The method does not guarantee all winners to be non-dominated but it maintains diversity within population while moving towards pareto front. Cross-over and mutation are simple one point genetic operations. An individual solution is obtained using a linear-logarithmic utility function while considering weights associated with each objective. The individual with highest utility is output as the final solution of algorithm.

Power	Modulation Order	Coding Rate	Data Rate	Frequency

Fig. 1. SDR knobs represented as GA chromosome

The accuracy of final solution lies in that the objective values it achieves is as close as possible to the objective values desired. Higher is the number of candidate solutions in pareto front, an output solution more closer to desired objective can be found. The number of candidate solutions relies on population size used in algorithm but increasing population size means more computation and a proportionately longer time required to generate pareto front. Thus it is difficult to get immediate results from system when using large population size. In the next section, an optimizer design is proposed that can give immediate solution while using pareto front generated from GA.

3. THE PROPOSED APPROACH

In order to improve the performance of optimization process, the paper proposes an optimizer design that uses GA as a one time activity to generate pareto front. The optimal solutions making this pareto front and the GA input is stored in a lookup table. In subsequent calls, these solutions are used to generate pareto front to find new configuration for radio. In the following subsections details of the approach are given.

3.1. Parameter Space and Objective Space

The knobs of SDR represented by X in (1) form an m-dimensional parameter space as shown in Figure-2. The permutations of different settings of all knobs lie in this space and represent all possible configurations of SDR. The configuration points in this space are mapped to the points in an n-dimensional objective space formed by all objectives. These are represented by Y in (1). The mapping is done by objective functions f_1, f_2, ..., f_n. For the optimization process, it is required to find the most optimal configurations from the parameter space. The optimal configurations of SDR have mapping to points lying on pareto front in objective space. These points lie on pareto front.

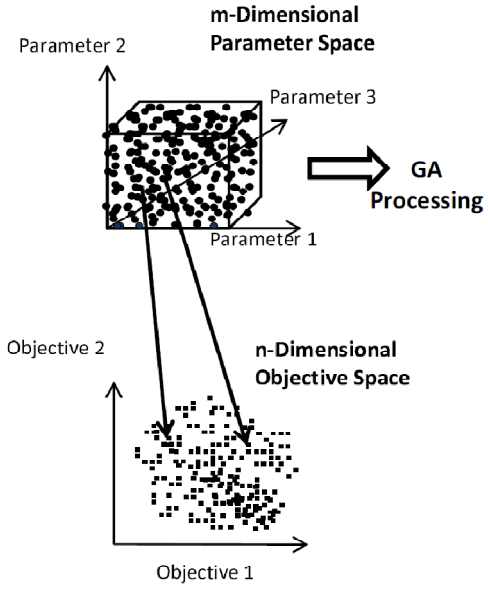


Fig. 2. The Parameter space and Objective space formed by knobs of SDR

3.2. Observation from simulations

The pareto front generation using GA has been performed with different environment parameter values in simulations. It has been observed that these pareto fronts were lying at different positions in objective space but they all have many common elements in their corresponding optimal configurations in parameter space. This has been shown in simulations in Section-III by re-mapping of two such optimal configurations to the objective space using same environment parameter values to get pareto fronts. It has been observed that these pareto fronts overlap at many positions.

This observation means that optimal configurations generated by GA are well applicable even if the environment parameters change. The objective space points or the pareto front can be re-calculated using changed environment parameter value and this front can be used to find new configurations.

The set of optimal configurations are generated for a given SDR knobs, objectives and GA parameters like population size, number of generations, crossover and mutation rate. These are collectively referred here as GA input. Thus it can be said that the set of configurations that makes pareto front in objective space remains valid for a given GA input, even if the environment parameter values change. The set of optimal configuration are required to be regenerated using GA only when there is some change in GA input, like the SDR is upgraded to support new knob settings or new objectives are introduced.

The reason behind this behavior can be understood with an analogy to a competitive examination scenario. In which the ranking of candidates remains approximately same when examination of particular subjects is taken multiple times, each time with question papers of varying difficulty. The ranking is only affected when the number of participating candidates is varied. The participating candidates represent the configurations in parameter space. The good candidates are the set of configurations that makes pareto front. The question papers of varying difficulty are the varying environment condition. The subjects are the objectives and the marks obtained are the values of objectives. The varying radio environment has no effect on pareto front, the good solution retain their rank, only the values of objectives achieved are varied.

It has been said that there is an overlap between the optimal solutions of GA when there is change in environment parameters. The positions at which it does not overlap was observed to be due to change in individuals competing for non-dominance during GA processing. The change in competing individuals happens due to occurrence of a combination of environment parameter value and knob settings value that may not be valid, and individuals having such knob settings are discarded by GA.

3.3. Proposed method

The suggested approach is based on the observation just described and divides the optimization processes in two phases. In first phase, the GA is run with a given GA input and pareto front is generated as shown in Figure-3. The configuration set in the parameter space that formed this pareto front is stored in a lookup table as shown in Figure-4. The lookup table has two fields, one having GA input and the other having configurations that formed the pareto front. In the second phase, a new configuration is found that is optimal in current radio environment for a user specified Quality of Service (QoS). The user specifies the GA input and the QoS requirement. The QoS requirement is specified

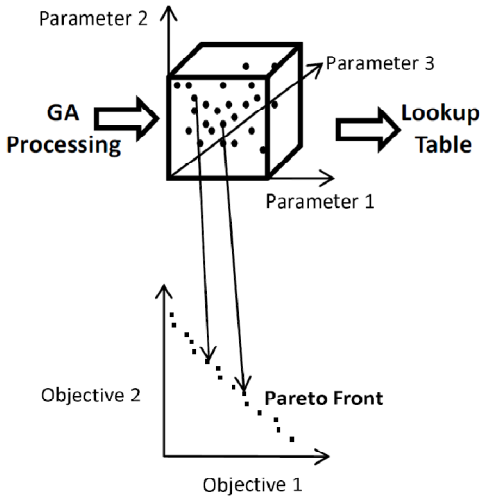


Fig. 3. Resulting pareto front after GA processing

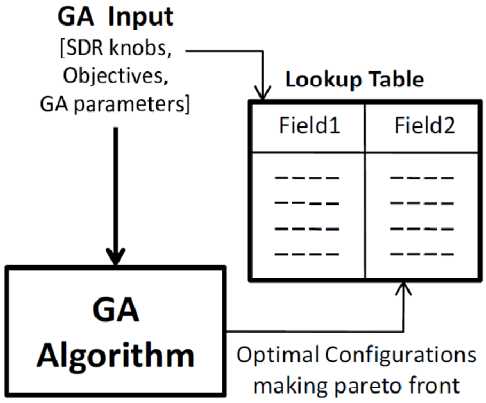


Fig. 4. Lookup table for storing Pareto front

as the objective parameters with their desired upper or lower limit depending on whether particular objective is to be minimized or maximized, respectively. To find a new configuration, the lookup table is first searched for an entry that has same GA input as specified by user. If an entry is found, the corresponding optimal configuration set is taken and its objective space points are calculated using objective functions with current environment parameters. This set of points is the pareto front for current environment. The point in this pareto front that is closest to the required QoS and satisfies the limits specified is taken, as shown in Figure-5 and the configuration in optimal solutions that mapped to this point is given as the new configuration. The second phase handles the varying radio environment conditions and is called in each call to optimizer at runtime. Figure-6, shows the flow chart for the proposed optimization process.

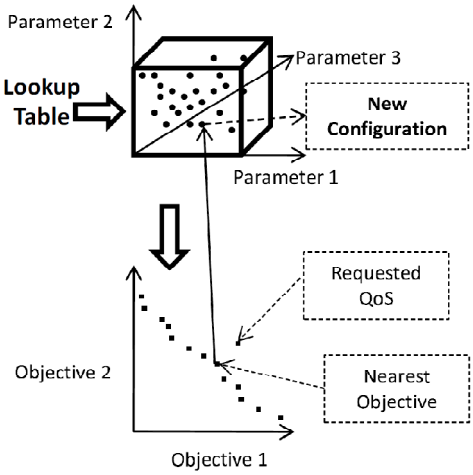


Fig. 5. Finding new configuration under current radio environment

The nearest point is found by a linear search on points in pareto front using Euclidean distance. The time taken to find a new configuration is thus limited by the order $O(l)$ in second phase, where l is the number of points in pareto front. Searching the entry for given GA input, is also done by linear search in lookup table. Thus the search time within lookup table is also linear and limited by the number of entries in the lookup table.

Whenever there is change in population size or introduction of new knob setting or change in crossover & mutation rate or different objectives are chosen, the first phase needs to be performed and a new entry is made in the lookup table. An example lookup table entry is shown in Figure-7. The entry has been stored in XML format. The <GA_INPUT> tag holds the GA input information and forms one field of entry. The other field stores the corresponding optimal configurations under the <OPTIMAL_SOLUTIONS> tag. The lookup table can be

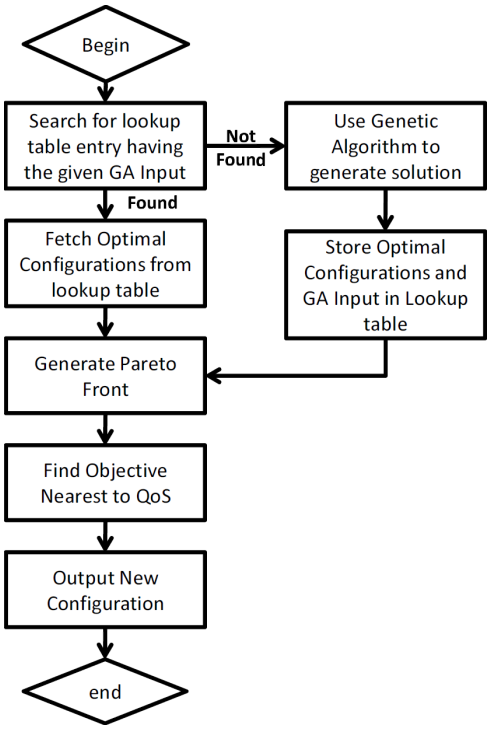


Fig. 6. Flowchart of optimization process

```
<lookup_table_entry>
  <GA_INPUT>
    <parameters>
      <p name="modulation_order_psk" ... />
      <p name="coding_rate" .../>
      <p name="data_rate" ... />
      <p name="txpowerdbm" .../>
      <p name="txfreqkhz" .../>
    </parameters>
    <objectives>
      <o name="ber" .../>
     <o name="bandwidth" ... />
      <o name="power_consumption" .../>
    </objectives>
    <ga_conf_parameters>
      <g population_size="4000", .../>
    </ga_conf_parameters>
    </GA_INPUT>
    <OPTIMAL_SOLUTIONS configurations="...." />
</lookup_table_entry>
```

Fig. 7. Example lookup table

filled with entries for all required combination of objectives and their optimal solutions offline. As the change in other part of GA input like SDR knobs happens once in a while, the required solution set can found most of times from lookup table.

3.4. Simulation

The design is run with a given GA input comprising of SDR knobs, objective parameters and the GA configuration parameters. The SDR knobs and the GA configuration parameters used in simulation are shown in Table-1 and Table-2, respectively. The objective set taken has three objectives BER, bandwidth and power consumption. The parameters in SDR knobs are modulation order for Phase Shift Keying (PSK) modulation, the coding rate, data rate, transmit power and transmit frequency. The parameters transmit power and transmit frequency have their minimum and maximum values specified. Their values are taken at the steps specified in table. The parameters modulation order, data rate and coding rate have enumeration type values.

In the first phase of simulation, genetic algorithm with strength pareto approach for multi-objective optimization is used to obtain the set optimal solutions [12]. It starts with an initial population of random configurations. The configurations have each parameters' random values taken from their domains. Objective set for each individual in population is evaluated using the objective functions. The non-dominated members in population are found and copied to an external set of non-dominated members. The non-dominated solutions that are covered by other non-dominated solutions are removed. This set keeps growing with each generation, so its size is kept in control by means of clustering of solutions. The clustering is done whenever the number of non-dominated solutions exceeds a given maximum. The clustering is done by finding the clusters with minimum average distance and making one cluster of them. A representative solution of this cluster is chosen. The new cluster is replaced by this representative solution.

Fitness of each individual in the original population and the non-dominated solution set is evaluated. For individuals in non-dominated set the fitness is calculated by dividing the number of individuals it dominates in original population with the size of original population plus one. For individual in original population the fitness is calculated by summing the fitness of all individuals in non-dominated that dominated it and adding one to it. This is to ensure that non-dominated solution will always have better fitness assigned than the individuals in original population, where the goal is to minimize fitness.

The selection is done from both the original population and non-dominated solutions. It is done by a binary tournament procedure till the mating pool is filled. The one point crossover and mutations are done with the specified probability to generate new population for next generation. The process is repeated for the given number of generations. The solutions in non-dominated set are the configurations that make pareto front in objective space and this solution set is stored in the lookup table.

Table 1. SDR Knobs

Knob	Values
Modulation Order PSK	2,4
Coding Rate	1/2, 1/3, 3/4
Data Rate	10000, 20000, 30000 bits per second
Transmit Power	-100 to 10 dBm (at 0.04 dBm steps)
Transmit Frequency	900 to 920 MHz (at 1 KHz steps)

Table 2. GA Parameters

Parameter	Values
Population Size	4000
Non-Dominated Set Size	5600
Mating Pool Size	2400
Generations	6
Crossover	0.98
Mutation	0.02

Table 3. Execution Time Comparison

Population Size	Using GA	Using Stored Results
4000	5880.6 Seconds	6.144 Seconds
400	67.22 Seconds	0.1593 Seconds
50	456.8 Milliseconds	48.12 Milliseconds

In the second phase of simulation, the QoS is specified as BER<=10^{-5}, Bandwidth<=20000 and Power Consumption<=1.5 Watt and the GA input as shown in XML fragment in previous subsection, is provided. The lookup table is searched for this GA input and its corresponding optimal solution set is taken. This set is used to generate pareto front which is plotted in the objective space as shown in Figure-8. The effect of environment parameter is taken in terms of Signal to Noise Ratio (SNR) at receiver node. The change in this parameter is effected by changing the distance between transmitter and receiver nodes. The SNR is calculated in objective functions by assuming a line of sight communication for simplicity. The distance between transmitter and receiver nodes has been taken as 2Km for the pareto front shown. The requested

QoS is shown by a point in the plot. A point is searched in this objective space that is nearest to this point. The objective set point found is BER=4.35835*10^{-6}, Bandwidth=20000 and Power Consumption=1.38 Watt. The corresponding configuration in parameter space is Modulation Order PSK=2, Coding Rate=3/4, Data Rate=10000 bits per second, Transmit Power=-9.84 dBm, Transmit Frequency=914.310 MHz. This configuration is given as the output of optimization process.

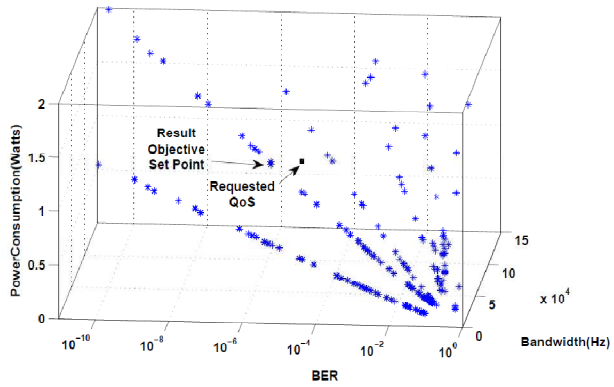


Fig. 8. Objective space showing pareto front formed by result of GA computation for transmitter and receiver separated by 2Km.

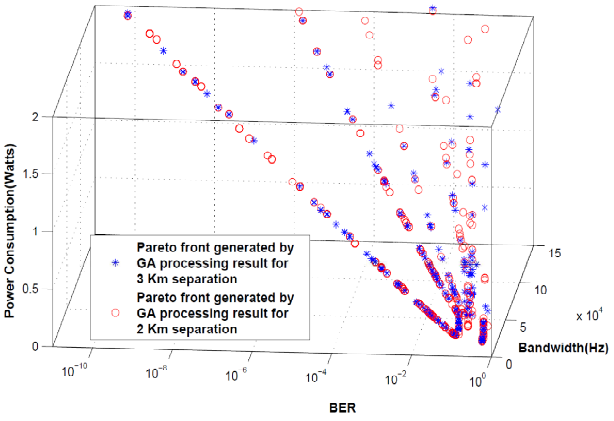


Fig. 9. Objective space showing overlap between pareto fronts of optimal solution set by GA processing for 3Km separation and the optimal solution set by GA processing for 2Km separation, both solution sets mapped using 3Km separation.

The pareto front generation using GA processing is performed with different environment parameter values as shown in Figure-8 and Figure-9. The pareto fronts with different environment parameter values lie at different positions in objective space as is obvious in two figures. The observation that both pareto fronts have many common elements in their corresponding optimal configurations in parameter space is shown by re-mapping both to the objective space using same environment parameter values, in Figure-9. One is generated by an optimal solution set for 3Km separation using GA and other is generated by re-mapping of a second optimal solution set using 3Km separation, whereas the second set was given by GA processing for 2Km separation shown in Figure-8. The two pareto fronts can be seen to be overlapping at many positions.

Table-3 shows the performance improvement when using proposed approach. It shows a comparison of execution times when genetic algorithm is run for getting solution and when stored results in lookup table are used. It is clear that proposed approach takes exceedingly lesser time for giving final output.

4. CONCLUSION

The scheme proposes to use GA as a one time activity in optimization process in Cognitive Radio so long as the knob settings of SDR are not upgraded. At runtime the lookup table is used to get optimal configuration for the current radio environment. The search time in second phase is linear and limited by the number of points in pareto front and the number of entries in the lookup table. So the scheme suggested can do great improvement in terms of time taken to get new configuration as compared to when GA is used to generate pareto front at runtime. Conversely, the scheme improves the performance in terms of effectiveness of solutions in their optimality. When pareto front is generated with high population size, the points in objective space are dense and hence, solutions are obtained that most nearly achieve the QoS, compared to when points are sparse.

5. ACKNOWLEDGMENT

Authors thank Director, Defence Electronics Applications Laboratory, for permission to publish this paper.

6. REFERENCES

[1] B. Fette, *Cognitive Radio Technology*, Elsevier, New York, 2006.

[2] T. W. Rondeau, "Application of Artificial Intelligence to Wireless Communications," *Ph.D. Dissertation*, Virginia Polytechnic Institute and State University, September, 2007.

[3] T. W. Rondeau, B. Le, D. Maldonado, D. Scaperoth, C. W. Bostian, "Cognitive Radio formulation and implementation" Center for Wireless Telecommunications, Virginia Tech, 2006.

[4] T. W. Rondeau, B. Le, C. J. Rieser, C. W. Bostian, “Cognitive Radios with Genetic Algorithms:Intelligent Control of Software Defined Radios” *Software Defined Radio Forum Technical Conference, pp. C-3-C-8*, Phoenix, 2004.

[5] C.M. Fonseca and P.J. Fleming, “Multiobjective Optimization and Multiple Constraint Handling with Evolutionary Algorithms” *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 28, pp. 26-37, 1998.

[6] J. Horn, N. Nafpliotis and D.E. Goldberg, “A Niched Pareto Genetic Algorithm for Multiobjective Optimization” *IEEE Proceedings of the World Congress on Computational Intelligence*, Vol. 1, pp. 82-87, 1994.

[7] J.P. Cohoon, W.N. Martin and D.S. Richards, ``Punctuated Equilibria: A Parallel Genetic Algorithm,'' *Proceedings of the Second International Conference on Genetic Algorithms*, Vol. 1, pp. 148-154, 1987.

[8] J. Arabas and S. Kozdrowski, ``Population Initialization in the Context of a Biased Problem-Specific Mutation,'' *IEEE Proceedings of the Evolutionary Computation World Congress on Computational Intelligence*, pp. 769-774, 1998.

[9] C.L. Ramsey and J.J. Grefenstette, ``Case-Based Initialization of Genetic Algorithms,'' *Proceedings of the Fifth International Conference on Genetic Algorithms*, Vol. 5, pp. 84-91, 1993.

[10] E. Zitzler and L. Thiele, ``An evolutionary algorithm for multiobjective optimization: The strength pareto approach'', Swiss Federal Institute of Technology (ETH), TIK-Report,No. 43, May 1998.

[11] E. Zitzler and L. Thiele, ``Multi objective evolutionary algorithms - a comparative case study and the strength pareto approach'', *IEEE Trans. Evolutionary Computation*, Vol. 3, 257 - 271, 1999.

[12] Ivo F. Sbalzarini, Sibylle Muller and Petros Koumoutsakos, ``Multiobjective optimization using evolutionary algorithms'', *Proceedings of the Summer Program,* Center for Turbulence Research, 2000.

Baseband Signal Processing Framework for the OsmocomBB GSM Protocol Stack

Harald Kröll, Christian Benkeser, Stefan Zwicky, Benjamin Weber, Qiuting Huang

Integrated Systems Laboratory
ETH Zurich, 8092 Zurich, Switzerland
e-mail: {kroell,benkeser,zwicky,huang}@iis.ee.ethz.ch

***Abstract*—The open source GSM protocol stack of the OsmocomBB project offers a versatile development environment regarding the data link and network layer. There is no solution available for developing physical layer baseband algorithms in combination with the data link and network layer. In this paper, a baseband development framework architecture with a suitable interface to the protocol stack of OsmocomBB is presented. With the proposed framework a complete GSM protocol stack can be run and baseband algorithms can be evaluated in a closed system. It closes the gap between physical layer signal processing implementations in Matlab and the upper layers of the Osmocombb GSM protocol stack. The functionality of the system is verified with a testbed comprising a base station and a receiver board with RF transceiver and FPGA.**

I. Introduction

Recently, the open source community has discovered the GSM protocol as an interesting exploration area, mainly for security aspects. Among the various successful attempts of open source implementation of several parts of the GSM network, the community behind the Open Source Mobile Communication Baseband (OsmocomBB [1]) project has implemented rather complete versions of the data link layer (L2), and the network layer (L3) of the mobile station (MS) side. The physical layer (L1) is running partially on a host computer and partially on hardware of a MS.

Unfortunately, hardware support for L1 is limited: reverse engineered legacy phones with modified firmware are mainly used for running OsmocomBB software. Access below the DSP application interface (API) can hardly be achieved, which limits the scope of new applications and implementations. Important tasks in digital baseband domain, such as channel equalization or decoding, are mostly implemented on the DSP, and therefore not accessible for further investigation. This deficiency restricts using OsmocomBB for (research) activities on the physical layer, which includes analog and digital front-end, baseband signal processing and L1 control functionality.

The signal processing, hardware development, and communication technology community has strong interest in an expandable baseband development framework with an interface to L2 and above of the GSM protocol stack. OsmocomBB's L1CTL protocol between L1 and L2 is well defined, but there is no development environment available in an ubiquitous scientific computing language such as Matlab or GNU/Octave, which can be connected to L1CTL. A framework with an interface of this type simplifies the validation of the functionality of baseband implementations towards higher layers in a closed system without expensive measurement equipment. Baseband engineers could use OsmocomBB during the design process and during testing signal processing blocks that require interaction with L2/L3.

Contribution: In this paper, a Matlab-based physical layer development framework architecture with an appropriate interface to the L2/L3 implementation of OsmocomBB is presented. The framework contains digital baseband signal processing with corresponding L1 controller and time processing unit (TPU), as required for GSM receivers. The different signal processing blocks are partitioned into so called primitive functions, which carry out essential tasks like signal filtering, symbol detection, parameter estimation, bit scrambling, and decoding. The functionality of the presented framework and interface architecture is verified on a testbed.

Outline: The paper is organized as follows. In Section II an overview on mobile phone architecture is given, and the need for crossing the boundary between L1 and L2/L3 is substantiated. An interface that connects OsmocomBB with a physical layer Matlab implementation is presented in Section III. The Matlab framework architecture is explained in Section IV. The testbed setup is described in Section V and Section VI concludes the paper.

II. The Missing Link

The protocol stack for GSM is a layered architecture that is based on the concepts of the ISO Open Systems Interconnection (OSI) model with 7 abstraction layers. The layered structure allows the distribution of work to specialists that can focus on a specific layer without having to consider the multitude of problems and issues that occur in the remaining 6 layers. In particular, baseband signal processing algorithms and architectures for the physical layer can be developed by neglecting L2/L3 procedure or operations of even higher layers. The separation of layers in the GSM standard has led to the classical partitioning of hardware in mobile phones, as depicted in Figure 1.

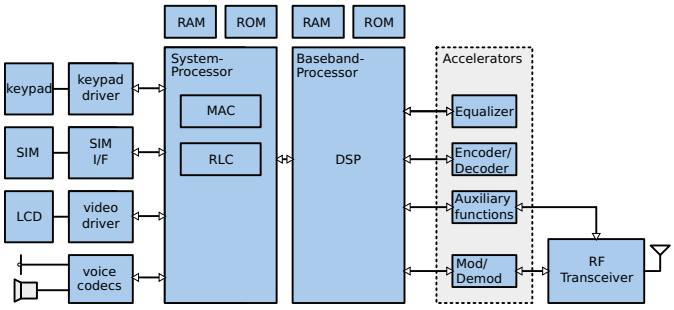


Fig. 1. A common GSM MS architecture. The PHY is distributed over a DSP and hardware accelerators [2].

Digital baseband signal processing tasks with low and medium computational complexity are typically executed on the *Baseband Processor*, a power-efficient Digital Signal Processor (DSP) optimized for mobile applications. The most complex parts of the digital baseband signal processing are usually directly mapped to dedicated hardware *accelerator blocks*, in order to achieve the required performance (e.g., bit error rate, throughput) at reasonably low power consumption. RLC/MAC layer procedures instead are suitable for an integration in software, because computational complexity is fairly low and high flexibility is required. Therefore, these tasks are typically realized on a reduced instruction set (RISC) microprocessor, the *System Processor*, which is connected to the physical layer implementation via DSP's API.

Although the strict separation of layers simplifies independent development and integration of the specific layers, it impedes optimizations and applications that require crossing the layer boundaries. For example, hybrid ARQ which is a key feature of modern mobile communication standards to enable high average throughput requires interaction between RLC/MAC layer and channel decoding in the physical layer. The hybrid ARQ technique specified for GSM/EDGE [3] is called *Incremental Redundancy* (IR). IR manages the storage of erroneously received data packets, and the combination of the stored data with re-transmissions of the same data packet. The combination of the received data packet with previously received and stored data significantly increases the probability of correct decoding[1], and therefore the average data throughput. Channel decoding is a computationally expensive baseband signal processing task in the physical layer, whereas the organization and controlling of re-transmissions, and the memory management of the stored data blocks is a procedure, that is typically controlled by RLC/MAC layer (see for example [2]). Therefore, in order to simulate the entire IR functionality, in order to evaluate average receiver throughput (with IR enabled) accurately, and in order to optimize IR implementations, being able to operate across the layer boundary between physical layer and RLC/MAC layer is desired from designer's point of view.

[1]Different (exclusive and non-exclusive) puncturing schemes are usually used, in order to increase the information gain with each re-transmission. Refer to, e.g., [4] for further details.

More than that, having access from higher layers to the physical layer and vice versa renders new applications possible. Physical layer procedures, intermediate results of baseband signal processing blocks, or simply raw baseband samples can be be monitored from higher layers, which simplifies debugging and enables new visualization opportunities of physical layer operations. New applications that require interaction between physical layer and higher layers are possible, such as *user cooperation*. In an exemplary user cooperation scenario, several mobile devices M_i support a mobile device M_0 by acting as it's remote antennas. Raw received baseband samples of the mobile devices M_i are forwarded from their physical layer to their application layer. From there, an *app* organizes the transmission of these samples to M_0 via an ad-hoc radio technology, where the samples are combined in the physical layer. Various combining schemes (e.g., [5]) can be applied in order to improve the probability of correct decoding.

We conclude that there is a need to have access to the physical layer in mobile phones, and to be able to model physical layer functions in combination with higher layers. In the following, we describe our approach of interfacing OsmocomBB with a physical layer framework.

III. The Interface between L1 and L2

The GSM specifications do not foresee a detailed protocol for the communication between L1 and L2. The GSM standard [6] defines basic messages[2] for the communication with the data link layer and the RLC/MAC layer. They are subdivided into request (REQ), confirm (CONF) and indication (IND) message types. The messages of OsmocomBB's L1CTL protocol are inspired from these message types of the GSM standard. A set of examples for L1CTL messages is given in Table I.

TABLE I
L1CTL message examples

Functionality	L1CTL messages
Reset PHY	L1CTL_RESET_REQ
	L1CTL_RESET_CONF
Synchronization	L1CTL_FBSB_REQ
	L1CTL_FBSB_CONF
Power Measurement	L1CTL_PM_REQ
	L1CTL_PM_CONF
Control Channel Mode	L1CTL_CCCH_MODE_REQ
	L1CTL_CCCH_MODE_CONF
Data indication	L1CTL_DATA_IND

The default OsmocomBB interface implementation between L1 and L2, called *osmocon*, uses a serial link with HDLC protocol [7] to load the firmware into the phone's memory. Using L1CTL messages, this firmware communicates via *osmocon* with a Unix domain Socket as implemented on the OsmocomBB side for the connection to L2/L3 running on the host computer. In order to be able to replace the firmware with a physical layer Matlab implementation, the

[2]The basic messages are called *primitives of the physical layer* in the GSM specifications [6].

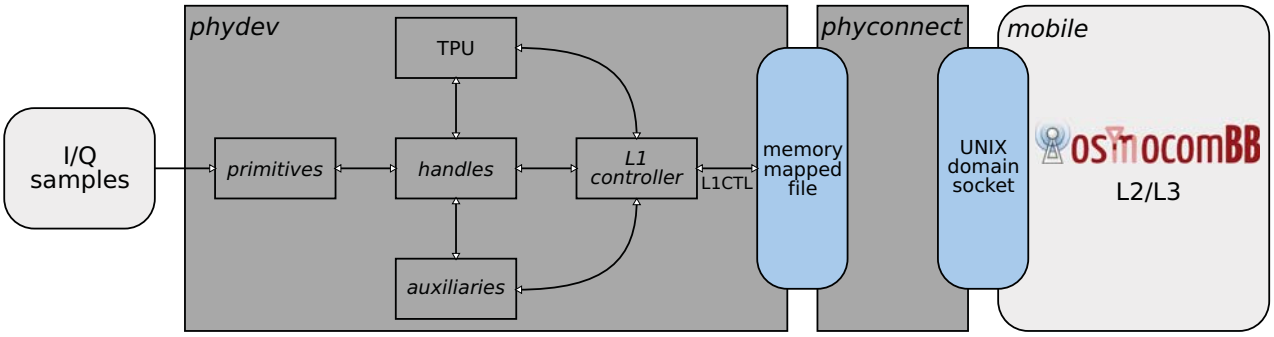


Fig. 2. Architecture of interface and framework between physical layer and *mobile*.

Unix domain socket needs to be connected with Matlab. Unfortunately, Matlab does not directly support Unix domain sockets, so an interface written in C is required for the socket communication. One solution would be to have such an interface embedded in a MEX function, such that it could be called inside a Matlab script. means of Matlab engine function calls. However, both MEX function calls and Matlab engine function calls are blocking, what prohibits parallel execution L1CTL protocol handling and baseband processing. Instead, the interface (herein referred to as *phyconnect*) for the proposed physical layer development framework connects the Unix domain socket to Matlab via memory mapped file as depicted in Figure 2.

In order to prevent accidental overwriting of data in the memory, a handshake protocol has been implemented. Thus, *phyconnect* sending data to Matlab waits first for Matlab to retrieve any data in the memory mapped file. By the same token, Matlab waits for the *phyconnect* process to retrieve data first before overwriting it. The memory mapped file has a total length of 880 bytes, which is used to build an array of 220 entries of 32 bit unsigned integers. There are entries for all information that needs to be accessible by L1, *phyconnect* and L2, such as GSM counters, L1CTL message properties, payload, and necessary information for the handshake protocol. We conclude that the proposed interface is a flexible solution to connect RLC/MAC layer of OsmocomBB with physical layer implementations in Matlab. In the following section we present our framework architecture that uses *phyconnect* to enable the simulation of our L1 Matlab realization in combination with the L2/L3 software, the *mobile* application, of OsmocomBB.

IV. Framework Architecture

Our baseband signal processing framework is shown in Figure 2. It comprises a GSM physical layer implementation, referred to as *phydev*, and the interface *phyconnect* to connect the *mobile* application of OsmocomBB, as explained in the previous section. *Phydev* is a Matlab realization of the physical baseband receiver that is typically implemented on the baseband processor assisted by accelerator blocks in dedicated hardware (cf., Figure 1). The main components of *phydev* are the TPU, the L1 controller, the *primitive functions*, and the *handles*, which will be explained in the following.

A. Layer 1 controller

The layer 1 controller builds the connection between the interface towards L2/L3 and the actual physical layer processing units. It implements the PHY finite state machine of a MS for GSM, as specified in [6] and shown in Figure 3.

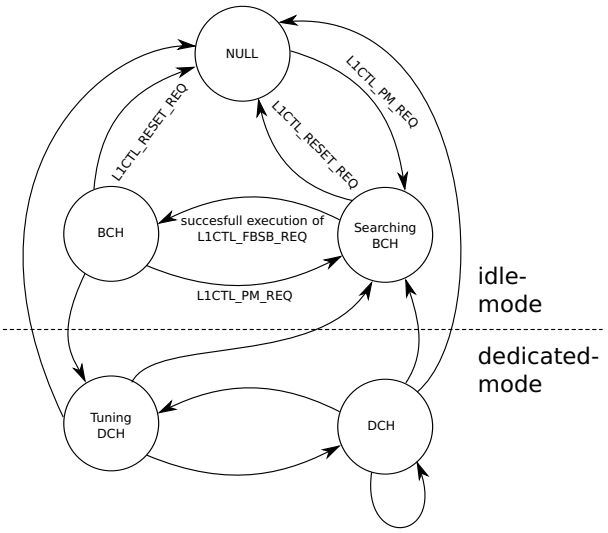


Fig. 3. PHY finite state machine with the states for the idle and dedicated mode [6].

After switching on the MS the state machine starts from the NULL state. From this, after having received a L1_PM_REQ message, the cell search procedure starts. First, the power levels of all possible GSM carriers are measured and reported to higher layers in the *searching BCH* state, in order to find the beacon carrier. Next, synchronization in time and frequency is performed after having received a L1CTL_FBSB_REQ message. In the *BCH* state the system information carried on the logical BCCH channel is extracted and reported to L2/L3. At this point the GSM state *camping on any cell* is achieved. Note that the states of dedicated mode (c.f. Figure 3) are not implemented in our framework so far.

The controller receives L1CTL messages from the interface via the memory mapped file. Subsequently, the messages get processed and multiplexed to a corresponding *handle*. The *handle* schedules the execution of primitive functions

if applicable and creates confirmation (CONF) or indication (IND) messages for the layer above.

During the *searching BCH* state the involved *handles* operate on consecutive samples, as there is no information about the GSM timing available. This streaming based operation is applied for the synchronization procedure and the power measurements (c.f., the example shown in Figure 5). After synchronization in frequency and time (in *BCH* state) the execution is frame based as depicted in Figure 4.

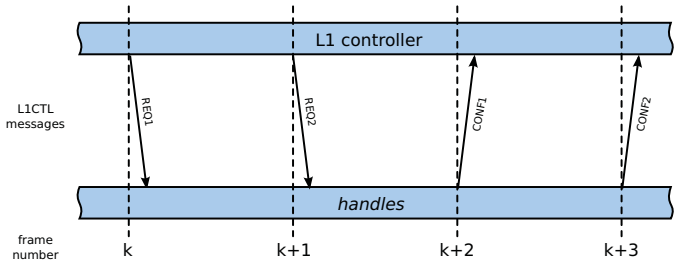


Fig. 4. Frame based execution on the physical layer in dedicated mode.

The complex signal processing algorithms do not run fast enough in Matlab to be executed in real time as required for the physical layer implementation. As timing is essential in GSM, we have implemented a TPU in *phydev* that emulates the four specified GSM timebase counters. This TPU allows the simulation of the timing between execution of primitives, L1 controller procedures, and communication with higher layers. More specifically, *handles* are called according to the TPU counter states. For each call, the counters are increased by the number of samples the corresponding primitive function processes. The corresponding L1CTL messages are generated and transmitted to OsmocomBB.

B. Primitive Functions

To ensure that the execution of operations on I/Q samples is carried out at the right time, functions which provide a result after a predefined number of samples are necessary. To this end, functions that process a fixed number of I/Q samples are implemented in *phydev*, referred to as *primitive functions*. The number of samples to be processed and other primitive-specific parameters are input arguments of the functions. The output of a primitive function is processed by the corresponding *handle* and forwarded to the L1 controller. The signal processing blocks represent entities as typically implemented as accelerator blocks or on the baseband processor, such as channel equalization, digital filtering stages or correlations. In addition to the primitive functions, auxiliary functions (*auxiliaries*) for common RF transceiver operations like RF power control or oscillator tuning (*DCXO_tune*) are provided.

In the following the main *primitives* and related signal processing blocks are explained in detail.

1) Power Measurements (PM_meas): Mobile stations for GSM have to measure the received signal power on all possible GSM carriers after power on, as well as the received signal power of the beacon carrier transmitted by surrounding base stations during operation. These tasks can be performed with the *PM_meas* primitive function. The input arguments of *PM_meas* are the number of samples to be processed and the absolute radio frequency channel number (ARFCN). The output of *PM_meas* (RMS_LEVEL) is the RMS power in dBm computed over the amount of processed samples. The L1CTL_PM_REQ-*handle* maps the result to an integer value RX_LEV and computes the running average according to the GSM specifications in [8].

2) Frequency burst detection (FB_det): The first synchronization step after the initial power measurements in the cell search procedure is the detection of the frequency burst (FB) on the beacon carrier, which is broadcast approx. every 47 ms. The FB is transmitted as a complex sinusoid, what enables a variety of detection strategies at receiver side. In this framework FB detection is performed according to [9], where the predictability of a complex sinusoid is exploited. By calling the *FB_det* primitive function, this detection algorithm is computed on the samples to be processed given as input arguments. The instance of time of the last successful FB_det primitive call gives a coarse timing estimate which is used in *SB_synch*.

3) Carrier frequency offset estimation (FB_est): Carrier frequency estimation is crucial during cell search, but also during normal operation. In our framework the carrier frequency offset is estimated from the FB's complex sinusoid by using the T&F estimator from [10], where a correlation based estimator is used. This approach has a significantly lower computational complexity when compared to, e.g., costly periodogram-based estimators. By calling the *FB_est* primitive function, this carrier frequency offset estimation algorithm is executed by processing the number of samples specified as input arguments. The output of *FB_est* is the estimated carrier frequency offset.

4) Synchronization burst detection (SB_synch): After synchronization in frequency has been achieved, timing synchronization needs to be performed. In GSM the detection of the synchronization burst (SB) which is also broadcast on the beacon carrier allows precise synchronization in time. A coarse timing estimate is already provided by the FB detection and given as input argument to the *SB_synch* primitive function. Thus, the SB detection needs only to be performed on the part of the received samples, which has been identified by the coarse synchronization as SB. Finally, the position of the 64 bit extended training sequence of the SB is detected accurately by performing a correlation between the received signal and the known training sequence.

5) Synchronization burst demodulation (SB_demod): Some of the GSM system information is transmitted on the SB and needs to be extracted during the cell search procedure. Demodulation of the symbols on the SB is performed with the primitive function *SB_demod*, which is similar to *NB_demod* for the demodulation of symbols on a normal burst (NB), which is explained in more detail in the next paragraph. Different training sequence lengths of SB and NB require two separate primitive functions.

6) Normal burst Demodulation (NB_demod): The primitive function *NB_demod* demodulates traffic and control data on normal bursts (NB). To this end, channel estimation, channel equalization and symbol demodulation have to be performed. In our implementation the channel profile is estimated in MMSE sense, as described in [11], where the midamble (training sequence) of each GSM burst is exploited. Channel equalization and symbol demodulation are performed with maximum likelihood sequence estimation (MLSE) [12], as typically used in GSM receivers. Input arguments of *NB_demod* are the number of samples to be processed, and the demodulated symbols are output.

Other tasks like interleaving, bit swapping, burst demapping or channel decoding (with a Viterbi decoder implementation) are executed by the L1 controller correspondingly.

An exemplary scenario for the processing of a L1CTL_FBSB_REQ message is depicted in Figure 5. As can be seen in Figure 3, successful processing of L1CTL_FBSB_REQ is required to reach the *BCH* state in the PHY functional state machine. After all the primitives and auxiliaries are called by the L1CTL handle, it composes a L1CTL_FBSB_CONF message and sends it back to the controller.

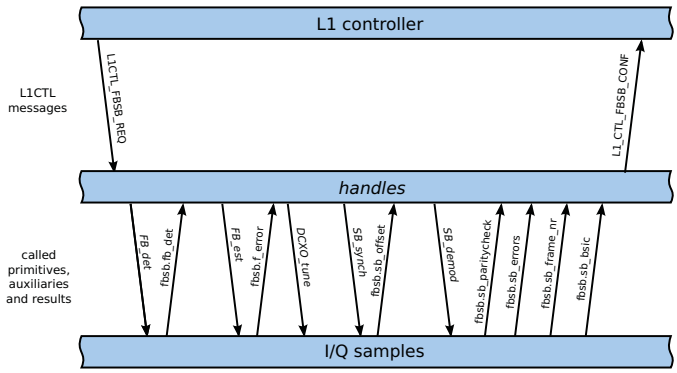


Fig. 5. Processing of L1CTL_FBSB_REQ message.

V. Testbed Setup

In order to feed our baseband signal processing framework with real-world data, a testbed setup, shown in Figure 6, has been developed. It consists of a GSM base station (BTS) emulator, a receiver board, a computer and a spectrum analyzer for debugging purposes.

The BTS emulator has been realized with OpenBTS [13] (an open source BTS software) and GNU radio software running on a computer, and a USRP[3] board with antenna to transmit the signal over the air. On the receiver board an antenna is connected to a state-of-the-art RF transceiver with ADC[4], which converts the RF signal to baseband and digital domain. The oversampled signal is fed into an FPGA on the

[3]Universal Software Radio Peripheral, from Ettus Research.

[4]IRIS305 RF Transceiver from ACP Semiconductors, Zollikon, Switzerland.

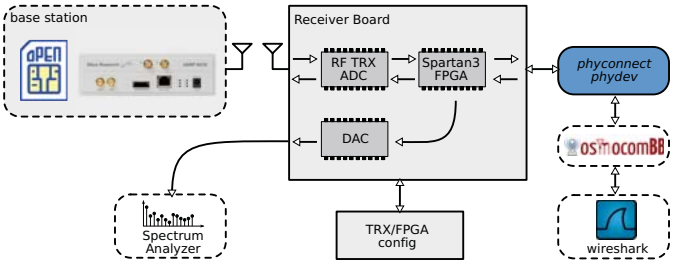


Fig. 6. Testbed Architecture with involved components.

receiver board, where decimation filtering and downsampling is performed with our digital front-end (DFE) implementation. A DAC on the receiver board allows to convert this signal back to analog domain and visualize the signal on a spectrum analyzer, which enables us to quickly identify problems on the receiver board. From the FPGA the DFE output signal is sent to a computer, where our framework processes the received samples and communicates with the *mobile* application of OsmocomBB, as described in the previous sections. The *mobile* application encapsulates the down-link data in UDP packets, which are forwarded for visualization to the Wireshark [14] protocol analyzer running on the same computer. A picture of the testbed setup is shown in Figure 7.

Fig. 7. Testbed setup with USRP N200 and receiver board.

The testbed setup has been used to verify the functionality of our framework in combination with L2/L3 of OsmocomBB by performing the initial cell search procedure in GSM (c.f. Section IV-A). To this end, the BTS emulator transmits a standard-compliant GSM beacon channel. The signal is received and processed on the receiver board as previously described, and corresponding I/Q samples are loaded from the receiver board into *phydev*, where the beacon carrier is found, and synchronization in time and frequency is achieved. The GSM system information (SI) broadcast on the beacon carrier is correctly extracted and propagated through OsmocomBB to Wireshark, where the SI messages are displayed.

VI. Conclusion

The OsmocomBB project provides an excellent open-source software implementation of the data link layer and the network layer of GSM. Unfortunately, the physical layer is not covered by OsmocomBB and hardware support is limited. Our proposed baseband signal processing framework shows how the physical layer can be realized in Matlab and connected to OsmocomBB with a dedicated interface. The verification of the framework and interface with the developed testbed setup proofs the feasibility of the proposed approach.

References

[1] "OsmocomBB, http://bb.osmocom.org, April 2012."

[2] L. Chang and Y. Wang, "EDGE incremental redundancy memory structure and memory management," Jul. 23 2009, uS Patent App. 12/507,835.

[3] *3GPP TR 44.004: General Packet Radio Service (GPRS); Mobile Station (MS) - Base Station System (BSS) interface; Radio Link Control / Medium Access Control (RLC/MAC) protocol*, TS 45.060, Rev. 9.0.0, Dec. 2009.

[4] E. Seurre, P. Savelli, and P. Pietri, *EDGE for mobile Internet*. Artech House Publishers, 2003.

[5] B. Djeumou, S. Lasaulce, and A. Klein, "Practical quantize-and-forward schemes for the frequency division relay channel," *EURASIP Journal on Wireless Communications and Networking*, p. 2, 2007.

[6] *3GPP TR 44.004: GSM/EDGE Layer 1; General requirements*, TS 45.004, Rev. 9.0.0, Dec. 2009.

[7] *ISO/IEC 13239: Information technology Telecommunications and information exchange between systems High-level data link control (HDLC) procedures*, TS 132 394, Rev. 9.0.0, Jul. 2002.

[8] *3GPP TR 45.008: GSM/EDGE Radio Access Network; Radio subsystem link control*, TS 45.008, Rev. 9.0.0, Nov. 2009.

[9] U. Jha, "Acquisition of frequency synchronization for GSM and its evolution systems," in *Proceedings of IEEE International Conference on Personal Wireless Communications*, 2000, pp. 558–562.

[10] D. Tufts and P. Fiore, "Simple, effective estimation of frequency based on prony's method," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 5, may 1996, pp. 2801–2804.

[11] E. Yakhnich, "Channel estimation for EGPRS modems," in *Proceedings of Vehicular Technology Conference, VTC-Spring*, vol. 1. IEEE, 2001, pp. 419–422.

[12] J. Proakis, *Digital communications*. McGraw-hill, 1987.

[13] "Openbts, http://openbts.sourceforge.net, april 2012."

[14] A. Orebaugh, G. Ramirez, and J. Burke, *Wireshark & Ethereal network protocol analyzer toolkit*. Syngress Media Inc, 2007.

SOFTWARE IMPLEMENTATION OF THE IEEE 802.11A/P PHYSICAL LAYER

Teo Cupaiuolo, Daniele Lo Iacono, Massimiliano Siti and Marco Odoni
Advanced System Technology
STMicroelectronics Italy
{teo.cupaiuolo,daniele.loiacono,massimiliano.siti,marco.odoni}@st.com

ABSTRACT

Software defined physical layer modems represent one of the main trends in communications and computing emerged in recent years. This is due on one hand to the need of supporting the requirements of modern communications systems in terms of seamless integration between different wireless technologies and multimedia convergence. On the other hand, programmable platforms are also beneficial as they allow consolidating methodologies, shortening development time and costs, extending products life-time. The drawback is that the complexity and power overhead of a pure computing fabric compared to a dedicated hardware can still represent a cost product developers are not willing to pay. This is particularly true for mobile terminals or in general for battery-powered devices. To become a concrete opportunity, baseband platforms should provide almost the same performance of custom designs while maintaining a certain degree of programmability. A good trade-off is represented by specific architectures integrating a proper mix of fine-grain general-purpose instructions and dedicated coarse-grain instructions wrapping custom hardware modules.

This paper presents a software implementation of a dual-mode IEEE 802.11a/p receiver on the Block Processing Engine (BPE), a proprietary template platform specifically designed for baseband processing. The combination of a novel extended instruction set with multi-thread processing support allows satisfying the most demanding requirements of the 802.11a and 802.11p standards.

1. INTRODUCTION

The automotive segment is now entering into the wireless arena with applications requiring reliable high data rate communications and including among the others vehicles safety, traffic control, remote tolling. To address vehicular environments, the IEEE recently introduced a new amendment to the 802.11 set of standards for wireless local area networks (WLAN) [1], namely IEEE 802.11p [2] for Wireless Access in Vehicular Environment (WAVE). Like several other standards developed for wideband digital communication, standards [1] are based on Orthogonal Frequency-Division Multiplexing (OFDM) modulation: among the others, a significant advantage of OFDM over single-carrier schemes is its ability to cope with severe channel multi-path conditions without the need for complex equalization filters.

The physical (PHY) layer [2] is mostly derived from the popular standard for indoor WLAN, IEEE 802.11a [3], with slight modifications required by the vehicular environment. The differences are designed to sustain robust vehicle-to-vehicle (V2V) communications as well as between vehicles and infrastructures located roadside (V2I). With respect to the indoor scenario, V2V and V2I require faster access time (<50 ms), increased range, robustness, reliability, security and indeed mobility. Being used also for critical safety applications, it will operate on the Dedicated Short Range Communications (DSRC) licensed spectrum at 5.9 GHz, exclusively reserved free-of-charge by the US FCC to vehicular communications.

These modifications come at the cost of a different and more complex digital baseband processing. Even with the recent proliferation of top performing programmable platforms, the traditional approach of implementing the PHY layer in hardware is still considered by many the best approach to fully satisfy the crescent needs of modern communications: area footprint and power consumption are a growing concern especially when targeting embedded portable devices. On the other hand, the process of developing an Application Specific Integrated Circuit (ASIC) is recognized as being long and costly, and can seriously compromise the ability to track a continuously evolving market, as typical of modern telecommunication systems. From this perspective, Software Defined Radio (SDR) represents a good opportunity strongly supported by research community: in its widest meaning, it aims to entirely replace the operations previously performed in hardware by purely software programmable resources.

Several advanced techniques have been introduced in the last years in order to fill the gap between the software and the hardware approach: the former has the undoubted advantage of reducing the development time, but it comes at cost of lower data rate, higher power consumption and area

TABLE I
THE 802 11A AND 11P STANDARDS
(MANDATORY VERSION)

	802.11a [3]	**802.11p [2]**
data rate	6–54Mbps	3–27Mbps
OFDM symbol duration	4µs	8µs
nb. of subcarriers	64 (48 data, 4 pilots, 12 virtual)	
bandwidth	20 MHz	5, 10 MHz
modulations orders	BPSK, QPSK, 16-QAM, 64-QAM	

requirement. One of the most widely used technique to effectively increase the throughput is the parallel processing enabled by Single Instruction Multiple Data (SIMD), which takes advantage of the high degree of data parallelism usually found in most telecommunication algorithms. To be effective, SIMD machines adopt parallel data paths of 128 bit or more. Another family is the Very Long Instruction Words (VLIW), wherein different types of functional units work in parallel. To further increase the performance, recent vector processors typically combine both of them [4][5]. Another recently emerged technique is the Single-Instruction Multiple-Task (SIMT), introduced by Coresonic to solve some of the major SIMD/VLIW drawbacks, like control overhead and improved memory utilization [6].

By today, various articles can be found about software implementation of the PHY of most common OFDM based standards, like [7][8][9], just to mention some of them. All these share the common approach of proposing a solution having some parts in software and some others in hardware. Typically software is used where there is some need for flexibility as can be the case of multi-standard support or real-time PHY reconfiguration; another case is represented by scalability requirements, like variable FFT sizes. Intensive data processing (as FIR filtering, FFT computation and Viterbi decoding) is usually left in hardware. Despite the commonalities with the 11a, indeed few works have been published about the 11p PHY: it can be recalled that [10] and [11] both developed a prototyping board reporting some implementation results, but these are still entirely hardware based.

This paper deals with a software implementation of a dual mode 802.11a/p receiver over the BPE, a mixed-grain template architecture designed for telecommunications and specifically customized for OFDM systems. We will show that a well balanced set of fine and coarse grain instructions allows to efficiently map the entire digital baseband receiver on a reconfigurable core. The distinctive flexibility of vectors processing of the BPE allows dealing with very different algorithms, ranging from filtering and synchronization down to the bit processing of de-puncturing and de-interleaving. To reduce the computational load while introducing no significant penalty in the programmability,

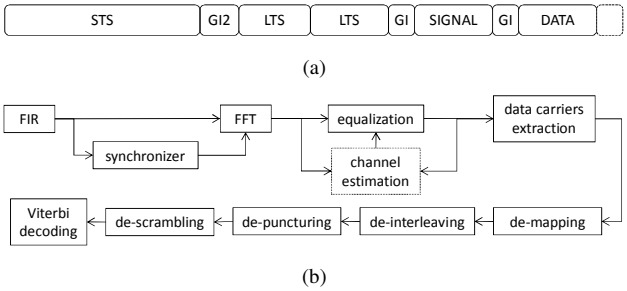


Fig. 1. (a) IEEE WLAN 802.11a/p frame structure and (b) receiver block diagram

Viterbi decoder has been implemented in hardware. Further, a new introduced multi-thread support allows executing several functions in parallel, thus effectively reducing the processing time compared to the single-thread version: the outcome is a software PHY that satisfies the most demanding data rate of 54 Mbit/s and 27 Mbit/s required by [3] and [2], respectively.

The paper is organized as follows: Sec. 2 recaps the typical algorithms involved in receivers complying with [2][3]; Sec. 3 gives an overview of the BPE processor; Sec. 4 details the algorithms mapping and profiling over the BPE; lastly, Sec. 5 concludes this paper.

2. THE IEEE 802.11A/P RECEIVER

2.1. Overview

Fig. 1a shows the structure of the WLAN frame [2][3], while the block diagram of the dual-mode receiver is shown in Fig. 1b. Referring to the frame structure depicted in Fig. 1a, the Short and Long Training Sequence (STS and LTS respectively) are known preambles used to perform time synchronization and channel estimation (CE). Details on both processes are provided later in this section.

The frame payload is composed by a first OFDM symbol (SIGNAL) holding, among the others, modulation parameters needed to demodulate the subsequent symbols (DATA): being crucial for the correct reception of the frame, the SIGNAL field is BPSK modulated and coded with maximum protection. Each symbol includes a guard interval (GI) of 16 samples to eliminate inter-symbol interference (ISI). Every OFDM symbol consists of 64 subcarriers composed by 48 data carriers (N_{SD}), 4 pilot tones (N_{SP}) and 12 null carriers (switched off to comply with spectrum emission mask). The OFDM symbol duration is 4 µs in case of 20 MHz bandwidth and 8 µs for 10 MHz.

TABLE I lists the main features of the two standards. Receivers able to demodulate both waveforms can share most of the processing, the main differences being filtering and CE functions. First of all, the larger delay spread makes the length of the 11a GI inadequate to prevent inter-symbol

interference (ISI). To face this problem while maintaining the same OFDM symbol structure (and hence same PHY processing), GI duration has been doubled by simply halving the channel bandwidth from 20 to 10 MHz. Another effect of the vehicular channel is the Doppler shift which must be compensated at receiver side with a more sophisticated estimate of the channel.

For both systems the processing performed at the receiver side can be divided in four phases: 1) *coarse synchronization*, i.e. packet detection and coarse carrier frequency offset (CFO) estimation based on the STS; 2) *fine synchronization*, i.e. frame synchronization and fine CFO estimation based on LTS; LTS is also used for the CE; 3) afterwards, processing of the signal field allows to identify the Modulation and Coding Scheme (MCS) chosen at the transmission side, i.e. the modulation order and the code rate associated to the next incoming data and other parameters, like the packet length; 4) lastly, the data field is demodulated to generate the bit stream to be propagated to upper layers.

2.2. Receiver Blocks

In this paragraph we shortly recap the typical algorithms performed at the receiver to enable the demodulation of the OFDM symbols, i.e. synchronization and CE.

2.2.1. The Short and Long Training Sequence

The STS is used to detect the presence of a frame, i.e. to perform the coarse timing synchronization. The algorithm takes advantage of the STS autocorrelation properties [12].

The coarse time synchronization can be performed by computing the M_n timing metric of the n-th sample:

$$M_n = \left|P_n\right|^2 \Big/ Q_n^2 \,, \tag{1}$$

where P_n is the autocorrelation function of $L = 16$ received samples r_n and Q_n their energy:

$$P_n = \sum_{k=0}^{L-1} r_{n+k+L} \cdot r_{n+k+L}^{*} \,, \tag{2}$$

$$Q_n = \sum_{k=0}^{L-1} \left|r_{n+k+L}\right|^2 \,. \tag{3}$$

Given the periodic structure of the STS, the timing metric takes a plateau form that begins with the first STS symbol and gradually decreases with the subsequent symbols: the frame is detected if M_n is between a (chosen) lower and an upper threshold for a given number of consecutive samples.

Then, CFO compensation takes place, using the following phase error coarse estimation:

$$\phi = \angle P_n \big/ N_{SP} \,, \tag{4}$$

where the phase of the autocorrelation function is computed over a variable number of samples (N_{SP}) depending on the estimated noise variance value.

After frame detection and CFO estimation and compensation, frame synchronization takes place in order to find the frame start. It is based on the cross-correlation between the incoming signal and a local replica of the LTS. At this time, coarse CFO estimation is refined using the autocorrelation properties of the LTS, carried (as for the STS processing) over a variable number of samples (N_{LP}). After CFO compensation, the GI of each OFDM symbol is removed and the 64 samples are transformed from the time domain to the frequency domain by 64-points FFT.

The next samples belonging to the LTS are used to compute the CE required during the processing of the OFDM symbols for equalization and soft-output de-mapping of the received bits. Afterwards, the GI2 (32 samples) is removed and the frequency domain per-tone CE H_m (for the m-th OFDM subcarrier) is computed. The basic CE corresponds to compute the mean of the two equalized received LTS $Y_{LTS1,m}$, $Y_{LTS2,m}$:

$$H_m = \frac{1}{2} \frac{Y_{LTS1,m} + Y_{LTS2,m}}{L_m} \,, \tag{5}$$

where the known LTS is denoted as L_m.

2.2.2. Signal and Data Field Processing

The signal and data fields are processed by the same algorithms, wherein the signal samples are BPSK modulated and define the MCS of the subsequent data samples. For every incoming OFDM symbol, the GI is removed and a group of 64 samples are Fourier transformed and equalized: using the previously computed CE, the symbols are (soft) de-mapped, de-interleaved and de-punctured and finally de-coded by a Viterbi decoder to produce the bit stream to be sent to the upper layers.

3. THE BLOCK PROCESSING ENGINE

In this section, we will recall only the main features of the BPE; a throughout description of the BPE can be found in [13][14]. Compared to the previous versions, the Instruction Set Architecture (ISA) has been extended withtrigonometric functions, coarse-grain FIR filtering, code generation and bit-level manipulation. From the core architecture perspective, in order to fully exploit data pipelining, it has been added the support for multi-thread function calls.

3.1. Vector Processing

The template architecture of the BPE is given in Fig. 2. The ISA implements two types of instruction: 1) basic scalar instructions (b-instructions) mainly devoted to flow control

and data access configuration; 2) dedicated vector instructions (d-instructions) performing intensive data processing. While b-instructions are locally executed, d-instructions are executed on the customizable dedicated unit (d-unit) bank. Depending on data dependencies and resources availability, units can be scheduled by the controller to run in parallel. Vectors are allocated on the dedicated memory (d-memory) bank, a set of static embedded memories allowing fast and parallel access to data.

The interconnection between d-memory bank and d-unit bank is guaranteed by the routing mesh, which is run-time configured by the controller on an instruction-by-instruction basis. To further optimize the data exchange between units, the routing mesh supports instruction pipelining through direct connection between units, thus avoiding the typical register pressure drawback of VLIW architectures: a group of pipelined d-instructions is called *macro* (instruction). Pipelined processing is the key enabler for high computational efficiency, since it allows propagating data from unit to unit without needing to store intermediate results for subsequent processing.

The controller fetches and schedules instructions one after another until one of them requires resources that have been already allocated, as can be the case for a memory or another unit. It then waits until the execution of the instructions using those resources has been completed. A side benefit of such policy is that b-instructions executed right after the scheduling of d-instructions do not cause additional delay. The latter consideration inherently suggests that maximum efficiency can be reached only using vectors long enough to absorb b-instruction execution.

3.2. The d-Instruction Set

The type of processing required by a generic PHY might be quite heterogeneous when moving down the receiver chain. The level of granularity changes according to the processing stage: whereas some blocks perform processing on a carrier by carrier basis, some others are characterized by bit-level granularity, which can cause a processing time penalty in a purely SIMD based processor. Furthermore, the type of operations performed changes significantly, ranging from typical signal processing, as can be the case of filtering, cross-correlation, equalization, to pure memory addresses computation as required for instance by the de-interleaver. Another interesting example might be the scrambler, which performs fast and with small complexity overhead only if supported by a dedicated unit (and related instruction) implementing a generic programmable linear feedback shift register (LFSR).

For the above reasons, the ISA has been extended with new instructions, like scrambling, Finite Impulse Response (FIR) filtering as well as more advanced operators like

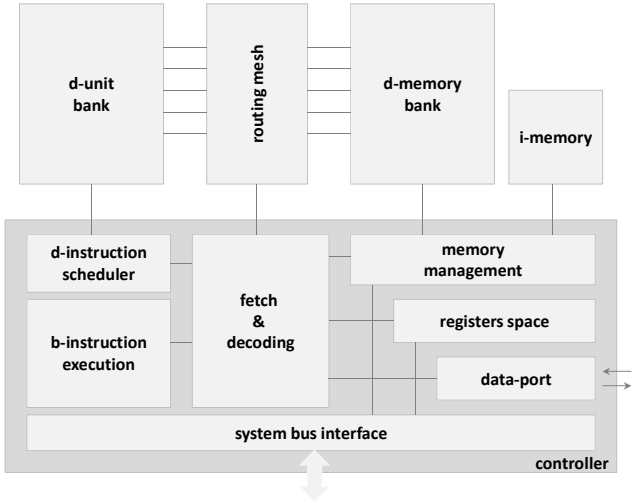


Fig. 2. The BPE template architecture

division, square root, hyperbolic and trigonometric functions all based on a low-latency fully pipelined CORDIC operator [15].

3.3. Multi-Thread Support

In [14] the technique of macro instructions pipelining has been introduced, wherein the macros themselves are defined as a group of pipelined d-instructions. Once the data dependencies of a given algorithm have been identified, macros can be connected to form a pipelined processing, where the pipeline stages are decoupled by memories implementing a ping-pong mechanism.

The macro-pipeline technique performs well within the processing of a given algorithm: the combinatorial path from the data source to the data output can be chopped into smaller pipelines until the required throughput is reached. Instead, a complete digital receiver involves several algorithms whose processing time can be effectively reduced by executing several functions in parallel. It will be shown that this enhancement is a key enabler for high data rate software based implementation.

4. MAPPING OF THE 802.11A/P PHY ON THE BPE

4.1. Short and Long Preamble Processing

The most computationally intensive kernels during the STS and LTS processing are the synchronization and the CE algorithms.

4.1.1. Synchronization and CFO computation

Provided that synchronization happens quite before data demodulation, a function implementing the synchronization algorithm is allowed to use all the available resources.

For an efficient mapping, eq. (2) and (3) can be rewritten in an iterative way: for example, the autocorrelation function P_n can be rewritten in vector form, like

$$P_{n+1} = P_n + r^*_{n+L} \cdot r_{n+2L} - r^*_n \cdot r_{n+L}. \quad (6)$$

The phase error (eq. (4)) is computed using CORDIC dedicated instructions. Once the CFO estimate is available, blocks of incoming samples are rotated executing a vector instruction from the d-instruction set.

TABLE II lists the clock cycles required for the synchronization function: the value has to be interpreted as latency, since the synchronization algorithm is not deterministic in nature. When clocking the BPE at 250 MHz, this latency translates into the need of buffering 2 OFDM symbols.

4.1.2. TD-LS channel estimation

The channel is first estimated using LTS (see eq. (5)) and further improved through reduced rank Time Domain Least Square (TD-LS) CE [17]: the incoming data are converted to the time domain (via inverse FFT), smoothed by a pre-computed matrix (related to the LTS itself) and converted back to the frequency domain. The main computational load is due to the IFFT and FFT operations: these are discussed in the subsequent section.

4.2. DATA Field Processing

4.2.1. The Fast Fourier Transform

The Fourier Transform is computed based on the Fast Fourier Transform (FFT) which is an efficient reformulation of the original algorithm that considerably reduces the computational load. In order to achieve fast FFT computation, the BPE embeds radix-2 and radix-4 butterflies as well as address calculation functions allowing data re-ordering among stages (e.g. bit-reverse). TABLE II lists the processing time required by the FFT software implementation.

4.2.2. Soft de-mapping

The simplified soft-out demapper algorithm [16] has been adopted: it avoids using costly (in terms of cycles) conditional constructs, without noticeable performance degradation compared to the optimal.

4.2.3. Intra-vector permutations: data carriers extraction, de-interleaving, de-puncturing and de-scrambling

Down the receiver chain, after de-mapping and prior to decoding, some specific intra-vector permutations need to be carried out: such operations do not perform any computation, but rather concern a change of position of the data within the vector.

TABLE II
PHY ALGORITHMS PROFILING

standard	algorithm	cycles	time [µs] @ 250 MHz
802.11a/p	synchronization [a]	1536	6.14
@DATA FIELD (MCS-7)			
802.11a/p	filtering (FIR)	162	0.65
	FFT (radix-4)	200	0.80
	equalization	64	0.26
	carriers re-ordering	56	0.22
	de-mapping	348	1.39
	de-interleaving [b]	408	1.63
802.11p	HDD DA CE	96	0.38
	TD-LS CE [c]	759	3.04
Total 802.11a[d]		1432	5.7
Total 802.11p[d]		2150	8.6

[a] it includes frame detection, CFO estimation and compensation
[b] it includes de-puncturing and de-scrambling
[c] based on the radix-4 FFT
[d] including software overhead

When performing this type of processing, SIMD based architectures turn to be heavily underutilized thus incurring in a waste of resources and a considerable increase in the required processing time. Power overhead is another side effect of underutilizing the vector. In turn, the BPE has a flexible vector management support, as described in Sec. 3.1: a fast execution time is achieved as it can be seen from TABLE II.

4.2.4. The Time Varying Channel Tracking

For a reliable communications in a mobility scenario, the time-varying channel has to be tracked, i.e. the initial CE based on LTS (Sec 4.1.2) must be updated for every OFDM symbol. This is especially true when the relative speed between two moving objects (as for two vehicles) becomes large and the Doppler effect has an impact on the system performance. In this work, the algorithm described in [18] has been implemented, specifically the TD-LS Hard Decision Directed (HDD) Data Aided (DA) CE, briefly referred to as TD-LS HDD CE.

During data processing, the CE is performed in a two steps process, as shown in Fig. 3: 1) data detection of the k-th received OFDM symbol using the CE corresponding to the (k–1)-th OFDM symbol; 2) the channel corresponding to the k-th OFDM symbol is estimated by using the newly detected symbol information based on HDD and followed by the TD-LS refinement.

The HDD CE is computed as follows: 1) frequency domain equalization of the m-th subcarrier of the k-th OFDM received symbol $Y_{k,m}$ with the available channel

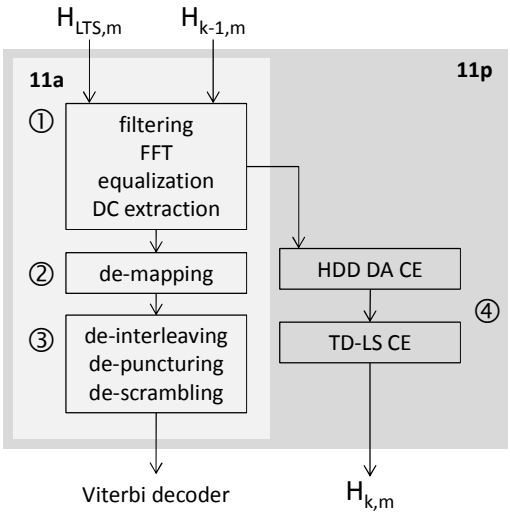


Fig. 3. DATA field processing (circled numbers refer to threads): ① filtering, FFT, equalization and data carriers extraction; ② de-mapping; ③ de-interleaving and de-puncturing; ④ channel estimation update

frequency response $H_{k-1,m}$ as $X_{k,m} = Y_{k,m}/H_{k-1,m}$; 2) hard detection of estimated data symbols $A_{k,m}$ to obtain the symbol sequence $\tilde{A}_k$: the hard detection implies hard de-mapping of the received symbol by associating the closest QAM symbol $A_{k,m}$ to the received equalized symbol $X_{k,m}$ (in terms of Euclidean distance); 3) estimation of H_k performed using as input the received symbol Y_k and the reconstructed $\tilde{A}_k$, used as an a-priori known sequence.

The CE is a computationally intensive task: a significant portion of time is spent on the FFTs and the filtering operation; refer to TABLE II for the overall processing time.

4.3. Single-Thread vs Multi-Thread Parallelism

As discussed in Sec. 3.3, the BPE versions [13][14] were based on a single-thread processor which implies that the overall processing time of the DATA field is given by the sum of the execution time of the single functions.

The first two rows of TABLE III list the clock cycles and the processing time required by a single thread to demodulate one DATA field, that is 5.7 µs for the 11a and 8.6 µs for the 11p PHY (including the overhead due to the software implementation, i.e. instructions scheduling and execution): both exceed their respectively OFDM symbol duration of 4 µs and 8 µs. The only way to increase the throughput is by taking advantage of the multi-thread support of the BPE.

Focusing first on [3], the idea is to pipeline the functions implementing the different stages of the processing, so as to demodulate different OFDM symbols in parallel: the parallelism degree is a combination of target data rate, single-stage pipe duration and availability of resources (memories and units). The multi-thread processing

TABLE III
DATA FIELD PROCESSING (MCS-7)

parallel threads	standard	clock cycles	time [µs] @ 250 MHz
1	802.11a	1432	5.7
1	802.11p	2150	8.6
3	802.11a	596	2.4
2	802.11p	1490	5.9

for the DATA payload is shown in Fig. 4a, wherein the algorithms are split among three parallel threads (①,② and ③) according to the diagram block of Fig. 3 (light-gray area): within each pipeline stage, the functions are processed serially, while the stages itself are executed in parallel. Once the overall pipeline has been filled, three OFDM symbols are elaborated concurrently, resulting in an overall elaboration time of 2.4 µs.

Regarding [2], the (TD-LS HDD) CE is the main bottleneck: the channel tracking is a recursive algorithm, thus the DATA processing cannot be pipelined as for [3]. Still, as explained in Sec. 4.2.4, the CE works independently from the demodulation thread and thus it can be run in parallel along the detection path, as shown in Fig. 3 (dark grey area): the previous three demodulation threads (①,② and ③) are executed one after each other, but in parallel with the CE thread (④). The corresponding (2-stage) multi-thread processing is shown in Fig. 4b: the overall demodulation time is 5.9 µs, which is well below than the target 8 µs symbol duration complying with [2] (see TABLE III).

5. CONCLUSIONS

In this paper we investigated the feasibility of a software implementation of the IEEE 802.11a/p [2][3] PHY on the BPE processor. The two standards share most of the receiver processing, but target two different scenarios: [3] targets indoor wireless communications, whereas [2] deals with outdoor vehicle to vehicle communications.

By properly exploiting the instruction set and the features of the BPE (instructions pipelining combined with parallel execution and multi-thread support), software real-time implementation has been developed.

We have shown that a balanced mixture of fine- and coarse-grain instructions allows to cover efficiently all the algorithms of the receiver chain. All the most computationally intensive kernels, including synchronization and FFT are performed in software; the only block left in hardware is the Viterbi channel decoder.

The data rate required by [3] can be achieved by adopting a three stage multi-thread processing, due to the absence of data dependencies (i.e. feed-back) between the computational blocks.

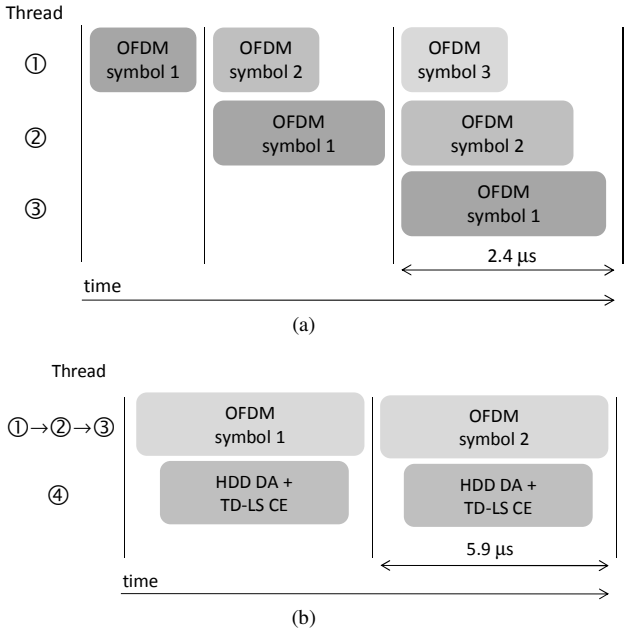


Fig. 4. Data field timing diagrams (MCS-7) based on multi-thread processing (timing not in scale): (a) 802.11a (b) 802.11p. The figure is labeled according to Fig. 3. Viterbi decoding is performed by a dedicated hardware design.

The requirements of [2] are satisfied by 2-stage multi-thread processing, but leaves little margin for optimization due to the need of recursively updating the channel tracking: the optional 20 MHz bandwidth mode, more challenging in terms of timing requirements, currently cannot be supported. It should be finally noted that a more advanced channel tracking algorithm, like Soft Decision Directed DA (based on Viterbi decoding) [19], would introduce an even higher latency, thus reducing the throughput.

In order to tackle these aspects, several further architectural enhancements are in place, like extension of the SIMD instruction width or the more future looking and challenging concept of cluster of processors.

6. REFERENCES

[1] IEEE Std 802.11 (ISO/IEC 8802-11: 1999): IEEE Standards for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area Network – Specific Requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications

[2] IEEE P802.11p-2010, "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications", Amendment 6: Wireless Access in Vehicular Environments, 2010.

[3] IEEE 802.11a-1999 Standard, "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-speed Physical Layer in the 5 GHz Band", 1999.

[4] Kees van Berkel, Frank Heinle, Patrick P. E. Meuwissen, Kees Moerman, and Matthias Weiss, "Vector Processing as an Enabler for Software-Defined Radio in Handheld Devices," EURASIP Journal on Applied Signal Processing, vol. 2005, no. 16, pp. 2613-2625, 2005. doi:10.1155/ASP.2005.2613

[5] B. Bougard, B. De Sutter, S. Rabou, D. Novo, O. Allam, S. Dupont, and L. Van der Perre, "A coarse-grained array based baseband processor for 100Mbps+ software defined radio". In Proceedings of the conference on Design, automation and test in Europe (DATE '08). ACM, New York, NY, USA, 716-721.

[6] D. Liu, A. Nilsson, E. Tell, D. Wu, and J. Eilert, "Bridging dream and reality: programmable baseband processors for software-defined radio". Comm. Mag. 47, 9 (September 2009), 134-140.

[7] V. Ramadurai, S. Jinturkar, S. Agarwal, M. Moudgill, and J. Glossner, "Software Implementation of 802.11a blocks on Sandblaster DSP", Proceedings of Software Defined Radio Technical Forum (SDR Forum '06), Orlando Florida, November, 2006.

[8] S. Eberli, A. Burg, T. Bösch and W. Fichtner, "An IEEE 802.11a baseband receiver implementation on an Application Specific Processor", Proceedings of IEEE Midwest Symposium on circuit & Systems, Montreal, Quebec, Canada, pp. 1324-1327, Aug 2007.

[9] A. Nilsson, E. Tell, D. Liu, "An 11 mm^2, 70mW Fully Programmable Baseband Processor for Mobile WiMAX and DVB-T/H in 0.12 µm CMOS", IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, USA, 90-97, 2009.

[10] D. Carona, A. Serrador, P. Mar, R. Abreu, N. Ferreira, T. Meireles, J. Matos and J. Lopes, "A 802.11p prototype implementation," Intelligent Vehicles Symposium (IV), 2010 IEEE , vol., no., pp.1116-1121, 21-24 June 2010.

[11] H. Harada, R. Funada, K. Sato, K. Iigusa and K. Li, "Research and development on UHF band inter-vehicle communication systems," Intelligent Transport Systems Telecommunications, (ITST), 2009 9th International Conference on , vol., no., pp.279-284, 20-22 Oct. 2009.

[12] T.M. Schmidl and D.C. Cox, "Robust frequency and timing synchronization for OFDM", IEEE Transactions on Communications, vol. 45, pp. 1613-1621, Dec. 1997.

[13] D. Lo Iacono, J. Zory, E. Messina, N. Piazzese, G. Saia, and A. Bettinelli. 2006. "ASIP architecture for multi-standard wireless terminals". In Proceedings of the conference on Design, automation and test in Europe: Designers' forum (DATE '06). European Design and Automation Association, 3001 Leuven, Belgium, Belgium, 118-123.

[14] T. Cupaiuolo and D. Lo Iacono, "Software Implementation of Near-ML Soft-Output MIMO Detection," Washington, DC, USA, 30 November - 3 December, 2010, Software Defined Radio Forum 2010 (SDR'10).

[15] J. Volder, "The CORDIC Trigonometric Computing Technique", IRE Transactions on Electronic Computers, 1959

[16] F. Tosato, P. Bisaglia, "Simplified soft-output demapper for binary interleaved COFDM with application to HIPERLAN/2", in: Proceedings of IEEE International Conference on Commun., New York, April/May 2002, pp. 664-668.

[17] E. Dall'Anese, A. Assalini, and S. Pupolin, "On reduced rank channel estimation and prediction for OFDM-based systems," in Proc. Int. Symp. on Wireless Pers. Multimedia Commun., Jaipur, India, Dec. 2007.

[18] M. Siti, A. Assalini, E. Dall'Anese and S. Pupolin, "Low Complexity Decision-Directed Channel Estimation based on a Reliable-Symbol Selection Strategy for OFDM Systems" (2010) IEEE International Conference on Communications (ICC'09) - Workshop on Vehicular Connectivity - Cape Town, South Africa, 23-27 May 20.

[19] L. Jarbot, "Combined decoding and channel estimation of OFDM systems in mobile radio networks," in Proc. IEEE Vehicular Tech. Conf., vol. 3, May 1997, pp. 1601–1604.