COCORA 2014

The Fourth International Conference on Advances in Cognitive Radio

ISBN: 978-1-61208-323-0

February 23 - 27, 2014

Nice, France

COCORA 2014 Editors

Pascal Lorenz, University of Haute Alsace, France

COCORA 2014

Foreword

The Fourth International Conference on Advances in Cognitive Radio (COCORA 2014), held between February 23rd-27th, 2014 in Nice, France, continued a series of events dealing with various aspects, advanced solutions and challenges in cognitive (and collaborative) radio networks. It covered fundamentals on cognitive and collaborative radio, specific mechanism and protocols, signal processing and dedicated devices, measurements and applications.

Most of the national and cross-national boards (FCC, European Commission) had/have a series of activities in the technical, economic, and regulatory domains in searching for better spectrum management policies and techniques, due to spectrum scarcity and spectrum underutilization issues. Therefore, dynamic spectrum management via cognition capability can make opportunistic spectrum access possible (either by knowledge management mechanisms or by spectrum sensing functionality). The main challenge for a cognitive radio is to detect the existence of primary users reliably in order to minimize the interference to licensed communications. Optimized collaborative spectrum sensing schemes give better spectrum sensing performance. Effects as hidden node, shadowing, fading lead to uncertainties in a channel; collaboration has been proposed as a solution. However, traffic overhead and other management aspects require enhanced collaboration techniques and mechanisms for a more realistic cognitive radio networking.

We take here the opportunity to warmly thank all the members of the COCORA 2014 Technical Program Committee. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to COCORA 2014. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the COCORA 2014 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that COCORA 2014 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the field of cognitive radio.

We are convinced that the participants found the event useful and communications very open. We also hope the attendees enjoyed the charm of Nice, France.

COCORA Advisory Committee:

Tomohiko Taniguchi, Fujitsu Laboratories Limited, Japan
Adrian Popescu, Blekinge Institute of Technology - Karlskrona, Sweden
Alan A. Varghese, Coherent Logix, USA

COCORA 2014

Committee

COCORA Advisory Committee

Tomohiko Taniguchi, Fujitsu Laboratories Limited, Japan
Adrian Popescu, Blekinge Institute of Technology - Karlskrona, Sweden
Alan A. Varghese, Coherent Logix, USA

COCORA 2014 Technical Program Committee

Anwer Al-Dulaimi, Brunel University, UK
Annamalai Annamalai, Prairie View A&M University, USA
Cornelia Badoi, Microsoft Romania, Romania
Ilham Benyahia, Université du Québec en Outaouais, Canada
Derya Çavdar, Bogazici University, Istanbul, Turkey
Malgorzata Gajewska, Gdansk University of Technology, Poland
Slawomir Gajewski, Gdansk University of Technology, Poland
Matthieu Gautier, Université de Rennes 1, IRISA, INRIA, France
Liljana Gavrilovska, Ss. Cyril and Methodius University -Skopje, Macedonia
Tien-Ke Huang, National Tsing Hua University, Taiwan
Abubakar Sadiq Hussaini, Instituto de Telecomunicações - Aveiro, Portugal / Modibbo Adama University of Technology Yola, Nigeria
Muhammad Umar Khan, Center for Advanced Studies in Telecommunication (CAST), Pakistan
Insoo Koo, University of Ulsan, South Korea
Thomas D. Lagkas, International Faculty of the University of Sheffield, CITY College, Greece
Jia-Chin Lin, National Central University, Taiwan
Marco Listanti, University of Rome La Sapienza, Italy
Ivana Maric, Ericsson Research in San Jose, USA
Jean-Claude Moissinac, TELECOM ParisTech, France
Homayoun Nikookar, Delft University of Technology, The Netherlands
Sema Oktug, Istanbul Technical University, Turkey
Adrian Popescu, Blekinge Institute of Technology - Karlskrona, Sweden
Arnd-Ragnar Rhiemeier, Cassidian Electronics - Ulm, Germany
Usman Shahid, NDSU - Fargo, USA
Boyan Soubachov, University of Cape Town, South Africa
Silvian Spiridon, Broadcom - Bunnik, The Netherlands
Tomohiko Taniguchi, Fujitsu Laboratories Limited, Japan
Ba Lam To, Phare - LIP6 - UPMC (Paris 6), France
Alan A. Varghese, Coherent Logix, USA
Liaoyuan Zeng, University of Electronic Science and Technology of China, China

Copyright Information

Table of Contents

Tactical Applications of Heterogeneous Ad Hoc Networks – Cognitive Radios, Wireless Sensor Networks and COTS in Networked Mobile Operations

Marko Suojanen
Electronics and Information Technology Division
Defence Forces Technical Research Centre
Riihimäki, Finland
marko.suojanen@mil.fi

Jari Nurmi
Department of Electronics and Communications Engineering
Tampere University of Technology
Tampere, Finland
jari.nurmi@tut.fi

Abstract— **Military tactical communications systems will be facing great challenges in the near future. Performance of commercial-off-the-shelf (COTS) technologies is improving at fast pace while at the same time life cycles of military equipment seem to remain long. The goal of this paper is to highlight a few interesting applications where development efforts of cognitive radio (CR), wireless sensor networks (WSNs), mobile ad hoc networks (MANETs) and commercial communications technologies would provide practical solutions from military point of view. The further work will focus on evaluation of CR as a part of heterogeneous military network. Importance of situational awareness on tactics and different technologies and their correspondence to tactical applications are discussed. Cognitive radio will find many applications in tactical networks but technical readiness is still lacking in some areas. A key issue in understanding the requirements of tactical networks is that the latest technology that is the most capable is not necessarily the most appropriate alternative. The focus should be on technologies whose effects on own actions and systems are understood thoroughly. This paper sets the groundwork for the research that aims at answering the question of what kind of military communications system should be built in the future if there are several requirements that limit the solution alternatives. The requirements to be considered are: 1) cost limitation, 2) limitation on number of units, 3) interoperability requirement, 4) mobility requirement and 5) performance requirement.**

Keywords—MANET; tactical networks; cognitive radio; wireless sensor networks; COTS; situational awareness

I. INTRODUCTION

Military tactical communications systems will be facing great challenges in the near future. Performance and capability of commercial civilian technologies are improving from generation to generation. Life cycles of military-grade equipment are typically much longer than life cycles of commercial technologies. On one hand, performance and high miniaturization level are tempting characteristics also in military applications. On the other hand, robustness of commercial technologies and security issues are not necessarily at the required level for military use. [1-3]

For the past few decades several ad hoc research projects have been carried out and quite many interesting results have been achieved [4-10]. Still, there are many obstacles for widespread use of mobile ad hoc networks (MANETs) [11-13]. Ad hoc networks have not been widely deployed due to better applicability and management of hierarchical and fixed infrastructure based networks. Ad hoc networks at their best are self-organizable and communication is based on short hops between nodes. Every node acts as a transceiver and a router that delivers internal and external traffic to other nodes at one-hop distance.

If we look very far to the future, it is likely that every node in the network actually is a cognitive MANET node due to technological advances and reduced unit costs. The goal of this paper is to highlight a few interesting areas where the development efforts of cognitive radio (CR), wireless sensor networks (WSNs), MANETs and commercial communications technologies would provide practical solutions from military point of view. There have been many research efforts concerning cognitive MANETs. Cognitive radio networks and WSN-assisted CR have been considered in [4], [21], [23], [25]. Many COTS-related considerations have not focused on cognitive radio but tactical radio issues. Dual use of VHF and UHF bands in tactical networks is considered in [24]. Military use of CR has gained more attention in recent years. Authors have not discovered similar approaches that consider the use of CR, WSNs, and COTS in MANET from the performance and cost point of view with the main goal of optimal solution.

The approach taken in this paper strives to understand the performance requirements, geographical coverage, mobility, connectivity, robustness and cost of the network. This paper sets the groundwork for a PhD research that aims at answering the question of what kind of military communications system should be built in the future if there are several requirements that limit the solution alternatives. The requirements to be considered are: 1) cost limitation, 2) limitation on number of units, 3) interoperability requirement, 4) mobility requirement and 5) performance requirement.

The following sections are organized as described here. First, importance of situational awareness on the battlefield and tactics are discussed in section 2. Section 3 looks to the future and describes some military applications that might use CR and MANET technologies. In section 4, different technologies and their correspondence to tactical applications

 ISBN: 978-1-61208-323-0

are discussed. Section 5 concludes the paper and presents some research questions for further study.

II. COGNITIVE AD HOC RADIO NETWORKS AND COTS IN TACTICAL APPLICATIONS

One of the most important factors in tactical operations is knowledge of operating environment and clever methods to benefit from awareness of surroundings. On modern battlefield, it is very important to preserve the communications channels between units in every situation. Due to limited resources and lack of development in radio research all units cannot be equipped with state-of-the-art radios. Therefore, awareness of operating area and understanding of radio propagation in different terrain and conditions is important. Own operation can be fitted to current situation if the location of units, terrain models, mobility of units and radio parameters are known.

In ideal circumstances, situational picture is received timely for the current and next phase of operation. Delivery of the situational picture could be handled as lightly as possible from the transmission point of view. Every node knows its own location and the nodes have a mechanism to report their own location to other nodes. Situational picture application shows the locations of own troops on the map and situation in the surrounding area based on the role of the user. The closer to the front line the user gets, the more time-critical situational information from the area is needed. Situational picture should be seen as a whole without every detail at the strategic level but there should be alarms of abrupt changes that might change the strategic situation, e.g., related to the capability of certain battalion at certain location. The level of detail depends on the operative status of the person, i.e., he/she receives the most urgent information for the current situation. The required information depends on location, order of battle and the next planned mission among other things.

Situational picture is formed by using as many sensors as possible. In the future, inexpensive nodes could run a situational awareness application as a background task without any effect on the user of the node. Since the continuous operation consumes a lot of energy in battery-operated nodes there should be a system-level mechanism that selects certain nodes to participate in the fusion of situational picture and lets other nodes sleep. This mechanism could be implemented as a part of the topology control messages. Coarse situational picture could be formed by a small number of capable sensors, and cheap sensors could be activated to create situational picture more fine-grained in areas of interest.

One should not underestimate the role of usability that has been taken more seriously in the commercial products. Most soldiers are used to use high-end technology products in off-duty so if these same methods would be used in military equipment, a great deal of resources could be saved in somewhere else than training the user interface of military equipment. Military-grade equipment has traditionally been heavy, large and hardly mobile. Large antenna structures and limited frequency bands create a challenge for modern mobile warfare. Networking capability of military units is also limited. The goal is to have a network that connects different communications systems more tightly together and guarantees the information flow to different units. Interoperability with some legacy systems could be provided by cognitive radios but interoperability with all previous military radio systems is unrealistic.

Deployment of CR on the battlefield forms a great threat for blue tactical communications so it is vital to take these issues into consideration in own concept planning. Development of a communications system should not be solely based on technological points of view but combination of defense tactics, strategy and technology should be examined in a way that allows flow of information and operational procedures to be optimized. This could be achieved by doing a thorough requirement analysis in co-operation of joint strategy and technology perspectives. Self-organization and autonomy should be applied at lower levels of hierarchy. Future systems are expected to be semi-autonomous but they require man-in-the-loop to confirm decisions, e.g., in offensive actions.

There are contradictory challenges to implementation of future military communications systems; on one hand, there is a need to form clusters that cover large geographical area, on the other hand there is a need for systems that are highly mobile and have a great computing capability. Due to cost pressures, one has to balance between the number of units and the level of performance of each unit. It is probable that a cost-fitted system consists of a large number of low-cost units and lesser number of very capable units that process data from low-cost units and connect those to faster networks. These requirements could be fulfilled by wireless ad hoc networks that are based on CR or software defined radios (SDRs). The use of SDR/CR based nodes as capable units allows the flexible collection of data from different units that operate at different frequency bands.

III. FUTURE TACTICAL APPLICATIONS OF COGNITIVE RADIO AND WIRELESS SENSOR NETWORKS

There is an ongoing progress towards software-defined technologies both in commercial and military applications. Software-defined functionality brings increased computing power and flexibility on the field but there are also evident challenges. The short-term challenge is to have versatile radios interoperable with each other and with new ones. In long-term, the situation may not be any better due to high unit costs of new multifunctional SDR/CR units.

This section takes a view to the future by considering few tactical applications that might benefit from SDR/CR technology used in ad hoc fashion. The detailed descriptions are given in the following subsections.

A. *Geospectral awareness*

There are several simulation tools that estimate the coverage of radio transmissions of, e.g., cellular base stations. These software programs use geographic data that contains height of objects, terrain types and land cover classifications. In the future, SDR/CR nodes will have a very high processing capacity so it may be realistic to assume that radio propagation calculation based on the environment could be done on the

 ISBN: 978-1-61208-323-0

field. Requirements regarding power consumption and real-time operation will remain challenging issues in the long-term.

In this concept, SDR/CR performs at specific intervals calculation of radio environment that is based on accurate geographical data, height information, classification of terrain and radio wave propagation models depending on the selected frequency band. It is important to note that there is no need to have capability to real-time calculations. Calculation could also be done in a PC that is part of the backbone network and thus only visualization of the results is done at each node. Still, there are advantages of running the calculation in the node to limit the calculation load to communications link.

As a result of calculation, SDR/CR forms a frequency-selective awareness of the radio environment. Decisions can then be done based on that awareness. If the terminal is about to send a message to another terminal at known GPS location, based on the calculation the terminal knows where to move to have the best connection to other terminal. If the movement is for some reason impossible, the terminal can activate more appropriate waveform (automatically or by user action) to maximize the quality for that link. If the expected direction of interference is known, it can be used as a boundary condition. Then the user can move to a direction that minimizes antenna gain to threat direction and/or that maximizes the quality of the friendly communications link. Calculation is done at the selected frequency, since it would be very time-consuming to simulate wide frequency bands. Decisions, e.g., on frequency change could be made based on the spectrum awareness. Results of calculation are shown on the screen clearly with instructions on navigation to the best location in terms of radio environment.

B. Sensor network gateway and high-capability sensor node

SDR/CR node could act as a gateway or master node in a deployed WSN. Ad hoc sensor network collects sensor data and routes it towards SDR/CR that has high computing capability. In this type of scenario, backbone network and ad hoc sensor network do not have to be interoperable since SDR/CR acts as a gateway and provides transparent access to sensor data. SDR/CR has two waveforms onboard: both sensor network waveform and waveform for communication with the backbone network. Typically wireless sensors have limited resources so SDR/CR could also have some high-level sensors that require more processing power and higher data rate, e.g. video surveillance system.

C. Distributed interference detection and mitigation

Distributed interference detection would be an interesting military application. Dozens of nodes are distributed on an area and they are networked in ad hoc fashion. Every node knows its location. Nodes are routing the communications traffic in ad hoc fashion. Distant high power interference source is activated and the nearest pair of nodes detects that the quality of the link is too low for transmission. Instead of increasing the transmit power, the message is routed via other route. Nodes are aware of each other's location, since the location data is shared for every node in the ad hoc network. The message is routed via second shortest path to destination node. If it is also interfered, new routes are tried until the edge of the network is reached. Based on the signal level, ad hoc network can estimate the direction of interference source and focus the traffic to the least affected route.

IV. DISCUSSION

In previous section, a few interesting tactical applications were introduced where CRs, COTS devices, WSNs and ad hoc networks could provide improved performance by working in collaboration. Since the application requires the use of CR technology in every node, consideration of that application is not realistic in short- or mid-term perspective due to high unit costs of CR units.

A. Node identities in tactical MANETs

If the cost of the wireless military network is an issue, then other solutions than using CR technology in every node, should be found. In the typical military scenario, there are different network islands that should communicate seamlessly with each other. These islands may represent a variable-sized group of soldiers, WSNs or other surveillance networks. Some of the nodes are fixed; some of those are mobile moving at different speeds. A characteristic feature of this type of network is intermittently changing topology and traffic load at different times.

Interesting topic to be considered in the research is the use of MANET functionality in mobile terminals. These terminals could be regular smart phones equipped with MANET technology that has a capability to connect and relay between other smart phones [21]. It is quite clear that MANET functionality between smart phones shows really important communications method after fixed infrastructure is not available or is destroyed. Base stations should be equipped with CR technology.

B. Frequency issues and vulnerability of MANETs

Typically the nodes in ad hoc network use the same frequency band that may be divided into separate frequency channels. The use of a common frequency band is possible due to limited transmission range and collision-detection and collision-avoidance methods. Frequency resources will be very limited in the future. The use of common frequency band enables the use of wide frequency bands that enable greater data rates. An application of direct-sequence spread spectrum and/or frequency-hopping in MANET is an interesting topic. Frequency-hopping provides robustness for the network but implementation is complex [14], [15]. Transmissions of other nodes and mobility in the same area cause interference and delays. These factors together lower the available transmission data rate. Due to low transmission power and radio transmission range, low probability of intercept (LPI) and low probability of detection (LPD) properties of ad hoc networks are satisfactory. The nodes are vulnerable under jamming but nodes have a possibility to deliver messages via optional routes unless the whole network is blocked.

 ISBN: 978-1-61208-323-0

C. COTS and cognitive radio in tactical applications

Commercial communication technologies are more and more appropriate for military applications. There are many differences between a commercial CR and military CR. Requirements for military CR are e.g. robustness, long-term upgrade support from manufacturers, fast scanning speeds, very high frequency resolution, size, weight, power, interoperability with other communications equipment and the last but not least, security. Solutions should also be found to the question of how to counter hostile CR activities. CR traditionally preserves the capacity of the licensed primary users and tries to avoid interference to primary users at any cost. In tactical applications the same type of behavior is shown for neutral civilian bands but for the enemy bands the situation is totally different. Military CR tries to capture every signal that is possible from the enemy communications or tries to block every frequency that enemy tries to use for communications [16-20].

Many researchers have published results of communications simulation in recent years. The goal of those simulations has typically been optimization of network topology, performance of routing or MAC protocols, analysis of performance degradation due to mobility, calculation of radio transmission coverage or research of effects of external or mutual interference on communications [22]. The further study will examine these characteristics but the main focus of the research will be on evaluation of tradeoffs between performance of the network and cost of the network. Defense forces in many countries are facing the challenges of limited defense budgets, increased capability requirements, upgrading the military equipment in use and increasing prices of new military equipment. These challenges raise an important question: what level of capability and performance can be achieved at certain level of funding? The best capability and performance cannot be achieved with limited resources.

A key issue in understanding the requirements of tactical networks is that the latest technology that is the most capable is not necessarily the most appropriate alternative to military applications. The focus should be on technologies whose effects on own actions and systems are understood thoroughly. If the lacking points are recognized and mitigated then the technology selection could be successful. It is essential to consider the special requirements that military environment poses on technology. New technology can be taken into use by applying old tactics. Other alternative is to develop both technology and tactics in parallel. The latter method will probably lead to a better solution in a tactical environment. Life cycle management is an important factor in defense technology procurement planning. The performance of new technological capability, related costs, improvement of technologies, interoperability and upgradability of technologies make in the end technology selection hit or miss.

V. CONCLUSIONS AND FURTHER RESEARCH

This paper delved into issues raised by CR, MANETs, WSNs and COTS technologies in tactical wireless networks. The goal of this paper was to highlight a few interesting areas where the development efforts of these technologies would provide practical solutions from military point of view. Recognized challenges should be mitigated when applying those methods to heterogeneous tactical networks. CR would provide better use of spectrum that is limited for military users. Commercial cellular networks, e.g. LTE and femto-cells seem promising and direct access between handsets might provide more capability to ad hoc connectivity on the battlefield. Frequency-hopping in tactical communications networks evidently provides more protection for handsets that operate at long ranges. LPI/LPD features are not so important for WSNs that operate over small range, e.g. tens of meters.

The main research question is how a military communications network should be implemented in the future in order to reach severe performance requirements in different conditions but cost-efficiently. Cost-effectiveness has become a top priority lately with global economy recession, shrinking defense budgets and ever-increasing prices of military technologies. The focus should be on technologies whose effects on own actions and systems are understood thoroughly. Importance of situational awareness on tactics and different technologies and their correspondence to tactical applications were discussed. CR will find many applications from tactical networks but technical readiness is still lacking in some areas.

This paper set the groundwork for a PhD research that focuses on the following topics. Further study aims at answering the question of what kind of military communications system should be built in the future if there are several requirements that limit the solution alternatives. The requirements to be considered are: 1) cost limitation, 2) limitation on number of units, 3) interoperability requirement, 4) mobility requirement and 5) performance requirement. In particular, the following issues need to be taken into account.

- How is the cost limitation defined? There should be dependence between cost limitation and performance.
- Limitation on number of units comes from the fact that certain level of performance can be guaranteed to some percentage of units. It is not feasible to use many radios for one soldier. The cumulative performance may be better if every unit is co-operating than the performance of more expensive system where only few units have a high performance.
- Performance should not be created by degrading interoperability of the systems but all systems should be interoperable.
- Mobility requirement calls for high performance systems that can change position intermittently and can be functional even when moving at fast speed.
- Performance requirement is provision of certain level of communications capability in every condition, e.g., robustness, interference tolerance, geographical coverage and appropriate transmission range.

 ISBN: 978-1-61208-323-0

References

[1] V. Kaul, C. Makaya, S. Das, D. Shur, and S. Samtani, "On the adaptation of commercial smartphones to tactical environments," Military Communications Conference (MILCOM 2011), 2011, pp. 2205–2210.

[2] A. Hansson, J. Nilsson, and M. Sköld, "Scenario based comparison of cellular and ad-hoc tactical radio networks," Military Communications Conference (MILCOM 2001), 2001, pp. 545–549.

[3] G. F. Elmasry, "A comparative review of commercial vs. tactical wireless networks," IEEE Communications Magazine, October 2010, pp. 54–59.

[4] D. R. Vaman and L. Qian, "Cognitive radio mixed sensor and mobile ad hoc networks (SMANET) for dual use applications," 2008 IEEE International Conference on Systems, Man and Cybernetics (SMC 2008), pp. 3304–3310.

[5] H-L. Koo, B-h. Roh, Y-K. Park, S-E. Kim, and Y-S. Kim, "Data delivery with end-to-end QoS guarantee in tactical network interconnected with wireless ad-hoc network and heterogeneous networks," 11th International Conference on Advanced Communication Technology (ICACT 2009), 2009, pp. 804–809.

[6] H. Wang, B. Crilly, W. Zhao, C. Autry, and S. Swank, "Implementing mobile ad hoc networking (MANET) over legacy tactical radio links," Military Communications Conference (MILCOM 2007), 2007, pp. 1–7.

[7] O. Younis, L. Kant, K. Chang, and K. Young, "Cognitive MANET design for mission-critical networks," IEEE Communications Magazine, October 2009, pp. 64–71.

[8] N. Suri et al., "Peer-to-peer communications for tactical environments: observations, requirements, and experiences," IEEE Communications Magazine, October 2010, pp. 60–68.

[9] L. Kant et al., "Network science based approaches to design and analyze MANETs for military applications," IEEE Communications Magazine, November 2008, pp. 55–61.

[10] J. Marsden, J. Treglia, and L. McKnight "Dynamic emergency response communication: the intelligent deployable augmented wireless gateway (iDAWG)," 2012 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support, 2012, pp. 279–286.

[11] G. Mao, "Research on wireless multi-hop networks: current state and challenges," International Conference on Computing, Networking and Communications 2012, 2012, pp. 593–598.

[12] L. Li, P. Vigneron, C. Brown, and M. Shi, "On designing bandwidth constrained mobile tactical networks for complex terrains," IEEE Communications Magazine, February 2012, pp. 188–194.

[13] J. L. Burbank, P. F. Chimento, B. K. Haberman, and W. T. Kasch, "Key challenges of military tactical networking and the elusive promise of MANET technology," IEEE Communications Magazine, November 2006, pp. 39–45.

[14] T. Vanninen, M. Raustia, H. Saarnisaari, and J. Iinatti, "Frequency hopping mobile ad hoc and sensor network synchronization," Military Communications Conference (MILCOM 2008), 2008, pp. 1–7.

[15] J. G. Andrews, S. Weber, and M. Haenggi, "Ad hoc networks: to spread or not to spread?," IEEE Communications Magazine, December 2007, pp. 84–91.

[16] R. Janka, "Applying cognitive radio concepts to next generation electronic warfare," Software Defined Radio Conference 2010 (SDR'10), 2010, pp. 265–270.

[17] G. Thomas, "Situation awareness issues in tactical cognitive radio," 2012 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support, 2012, pp. 287–293.

[18] Y-C. Liang, K-C. Chen, G. Y. Li, and P. Mähönen, "Cognitive radio networking and communications: an overview," IEEE Transactions on Vehicular Technology, Vol. 60 , No 7, September 2011, pp. 3386–3407.

[19] R. Umar and A. U. H. Sheikh, "Cognitive radio oriented wireless networks: challenges and solutions," International Conference on Multimedia Computing and Systems (ICMCS 2012), 2012, pp. 992–997.

[20] B. Wang and K. J. R. Liu, "Advances in cognitive radio networks: a survey," IEEE Journal of Selected Topics in Signal Processing, Vol. 5, No. 1, February 2011, pp. 5–23.

[21] J. Lopatka, "Cognitive wireless ad hoc networks in urban environment", 19th International Conference on Microwaves, Radar and Wireless Communications (MIKON 2012), pp. 810–813.

[22] N. VanderHorn, B. Haan, M. Carvalho, and C. Perez, "Distributed policy learning for the cognitive network management system", The 2010 Military Communications Conference (MILCOM 2010), pp. 435–440.

[23] S. S. Kashi and M. Sharifi, "Connectivity weakness impacts on coordination in wireless sensor and actor networks", IEEE Communications Surveys & Tutorials, Vol. 15, No. 1, 2013.

[24] J. Whitbeck et al., "Using UHF connectivity to off-load VHF messaging in tactical MANETs", The 2011 Military Communications Conference (MILCOM 2011), pp. 961–966.

[25] J. Yu et al., "Cognitive radio network as wireless sensor network (I): architecture, testbed, and experiment", Proceedings of the 2011 IEEE National Aerospace and Electronics Conference (NAECON), 2011, pp. 111–117.

 ISBN: 978-1-61208-323-0

A Survey of Cognitive Radio Management Functions

Ola Ashour Mohammed
Department of Electrical, Computer and Software Engineering, UOIT
Oshawa, Canada
ola.ashoor@gmail.com

Khalil El-Khatib, Miguel Vargas Martin
Faculty of Business and Information Technology, UOIT
Oshawa, Canada
{khalil.el-khatib, miguel.vargasmartin}@uoit.ca

***Abstract*— Due to the fast growth in wireless communication services, the need for radio spectrum increased. However, most of the suitable radio spectrum has already been allocated using long term licenses. A considerable part of the allocated spectrum is underutilized over time and space. Cognitive Radio (CR) technology has arisen to solve the spectrum scarcity problem by allowing cognitive radio devices to opportunistically make use of the unused frequency bands in the allocated spectrum, which are termed white spaces or spectrum holes. Four CR functions have to be performed to allow CR devices to efficiently utilize the available spectrum holes without interfering with licensed devices already operating in the allocated spectrum. This paper presents a survey of the CR technology, its architecture and operation, a detailed description of the four CR functions, and the techniques and processes used in each function.**

Keywords-Cognitive Radio; Spectrum holes; Spectrum decision; Spectrum Sharing; Spectrum Mobility.

I. Introduction

The radio spectrum is a natural resource managed by governments, which have assigned fixed portions of this spectrum to various operators using long term licenses. With the trend of using wireless devices continue to increase, spectrum usage increases on a daily bases, and it is becoming certain that there is a real need for more spectrum bands to facilitate the implementation of new wireless services. However, it is hard to find free bands, as most of the suitable spectrum bands have already been assigned.

Recent measurements [1] showed that a considerable part of the allocated radio spectrum is underutilized due to temporal and geographic disparities in how the allocated spectrum is used. The unused frequency bands in time or space are usually termed spectrum holes or white spaces. One way to make efficient use of these spectrum holes is to use Dynamic Spectrum Access (DSA) techniques, which enable secondary (unlicensed) users to make use of the spectrum when primary (licensed) users are not using it.

Primary Users (PUs) have rights to access a certain part of the available spectrum and hence have a higher priority in accessing the spectrum. On the other hand, Secondary Users (SUs) can utilize the spectrum under the condition of not interfering with PUs. Thus, SUs need devices that have the ability to determine whether the spectrum is being utilized at a specific location and at a certain time [2].

Cognitive Radio (CR) is an important enabling technology for DSA which helps SUs make efficient use of the radio spectrum. CR is a wireless communication technology based on Software Defined Radio (SDR), where each device is capable of determining its location, sensing its environment and learning about its radio resources [3]. The device can dynamically adjust its operational parameters, such as transmission frequency and power, to opportunistically utilize the empty frequency bands without disturbing PUs [2], [4]. A CR device has two main characteristics which are cognitive capability and re-configurability [2], [4]. The cognitive capability allows the device to sense the medium and determine the available spectrum bands. Re-configurability enables the CR device to adjust its operating frequency, modulation technique and transmission power without the need for hardware modification.

After the transition from analog TV to digital TV transmissions, large amounts of frequencies in Very High Frequency (VHF) and Ultra High Frequency (UHF) bands have been freed up. These unused frequency portions on the TV broadcasting (UHF and VHF bands) are referred to as TV White Spaces (TVWS). Frequencies in the TV broadcast bands benefit from high bandwidth, long transmission ranges and better building penetration.

A White Space Device (WSD) is a device that can make use of the available white spaces when not being used by incumbent transmitters (TV transmitters or Wireless microphones). This device should not interfere with any of the PUs operating on that band.

This paper will provide an overview on CR architectures, its operation in both licensed and unlicensed bands, and a detailed description of the CR functions. A classification for each cognitive radio function will be also discussed. The rest of the paper is organized as follows: Section II provides a background of the regulation for utilizing white spaces and the countries that started making use of the available white spaces. The architecture of the cognitive radio network is then discussed in Section III. Section IV illustrates the operation of a CR device in the licensed and unlicensed bands. In Section V, CR spectrum management functions are presented and then a detailed description of the techniques used in each function is discussed in Sections VI, VII, VIII and IX. A conclusion is presented in Section X.

II. Background

While the USA was not the first country to switch to digital TV, they did become leaders in making the decision to utilize white spaces. The Federal Communications Commission (FCC) allowed unlicensed radio transmitters to operate in the unused frequency portions of the broadcast television spectrum [5]. These unlicensed radio transmitters (secondary devices) have to make sure that they will not interfere with licensed users [6]. The FCC ruling stated that secondary devices must both consult a database which, given a certain a certain location, would be able to provide a list of the available channels at that location, and must also perform a real-time spectrum sensing every minute to ensure that no licensed devices such as wireless microphones exist.

On September 23, 2010 a Memorandum Opinion and Order was released by the FCC [7], in which the final rules for utilizing white spaces for unlicensed wireless devices was determined. The mandatory requirement for using spectrum sensing was eliminated, paving the way for geo-location based channel allocation. Spectrum sensing had been removed from FCC rules due to many reasons: 1) it is time-consuming, 2) it increases development cost, and 3) it inappropriately protect other unlicensed devices, which should not be protected from interference, as it cannot differentiate between licensed and unlicensed devices.

In early 2011, the FCC released an order [9] conditionally designating nine TV white space database operators, including for example, Google, Motorola, and HP. Microsoft was later added to the list of approved operators. The FCC stated that a trial period of 45 days is required for all database operators before being able to announce the public availability of each database. FCC stated also that, using multiple database operators will create a healthy competitive basis between the operators Spectrum Bridge's database was the first trial that began on the 19th of September, 2011.

The second country to make a decision was the UK, when Ofcom published a document on September 2011, which expressed their intention to support commercial utilization of white spaces [10]. In that document, Ofcom mentioned that their approach is based on using geo-location databases rather than alternative approaches like sensing or beacons. Ofcom also explained the UK's preference for a harmonized approach to WSDs across Europe, as it believes that this harmonized European approach would deliver greatest benefits for citizens and consumers. Ofcom said that it will continue the development of new business opportunities while waiting European-wide regulatory decisions.

III. COGNITIVE RADIO NETWORK ARCHITECTURE

Cognitive Radio Networks (CRNs) or secondary networks do not have licenses to access the spectrum and can be referred to as un-licensed networks. CRNs can be classified based on their architecture [8], [11] as infrastructure-based or ad hoc networks. An example of infrastructure-based network is shown in Fig. 1 [8]. These networks have secondary Base Stations (BSs) or secondary Access Points (APs) that can coordinate the communication between secondary devices in their coverage areas. The secondary APs can be connected together through a wired network (Core network) to allow communication between SUs in different coverage areas.

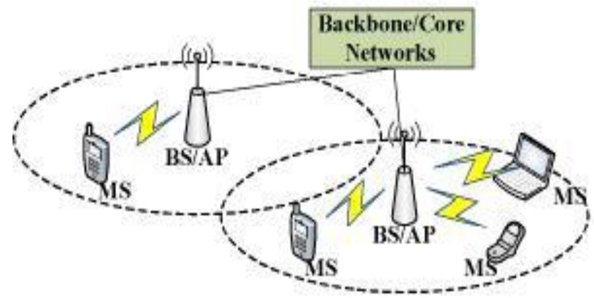


Fig. 1. Infrastructure-based CRN [8].

Ad hoc or distributed CRNs do not require an infrastructure. SUs or CR users can directly communicate with each other using certain communication protocol (Wi-Fi or Bluetooth) or they can utilize the available spectrum bands for their communications as shown in Fig. 2 [8].

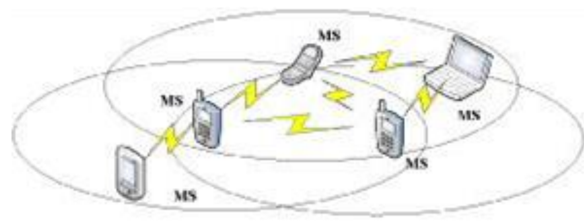


Fig. 2. Ad hoc CRN [8].

CRNs either infrastructure-based or ad hoc networks usually work inside the transmission range of primary networks as shown in Fig. 3 [12]. Primary networks or licensed networks refer to already existing infrastructure-based wireless networks like mobile networks that are allocated certain frequency bands for their operation. The infrastructure of these primary networks consists of base-stations that can control the activities of PUs.

A CR device always has to determine the available white spaces in its location, to avoid interfering with users operating in the licensed band. In infrastructure-based CRNs, CR devices have two different ways to determine the available spectrum holes. The CR devices can sense the medium and send the sensed information to the base station, which performs the spectrum decision, spectrum sharing and spectrum mobility functions. The other way of determining the available spectrum holes is through the base station itself, which can obtain a list of the available spectrum holes by contacting a database of incumbents. This approach will be covered in detail in Section VI-B.

In ad hoc CRNs, each device should have a cognitive ability, which allows the device to sense the medium and determine the available white spaces in its location. CR devices can cooperate in determining the available spectrum holes by sharing the sensed information with each other (cooperative sensing). A CR device may also get a list of the available spectrum holes in its location by connecting to a database of incumbents.

IV. USAGE PRIORITY IN LIENSED BAND

A CR user can operate in licensed and unlicensed band as well [2]. In licensed band operation, the priority is for PUs operations. SUs can use the spectrum when not being

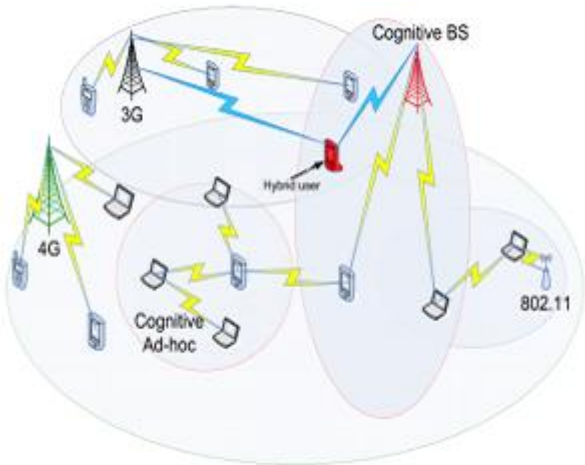


Fig. 3. CRNs operate inside the coverage area of primary networks [12].

used by PUs. SUs have to vacate the channel whenever a PU appears and move to another available channel. In unlicensed band operation, all users have equal opportunity for spectrum access (no priorities).

V. COGNITIVE RADIO SPECTRUM MANAGEMENT FUNCTIONS

A CR device has to perform four basic functions to be able to manage the available spectrum holes in its location [2], [12]. These functions, illustrated in Fig. 4, are: White Space Determination, Spectrum Decision, Spectrum Sharing, and Spectrum mobility.

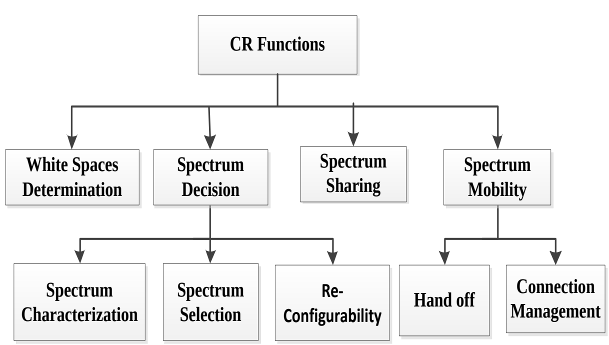


Fig. 4. Cognitive Radio Functions.

The main function of a CR device is to determine the available white spaces at a certain place and at a specific time. After determining the available spectrum holes, the spectrum decision is performed by selecting the best channel for the operation of a CR. Channel selection is usually done based on specific criteria, which could be the policy, Quality of Service (QoS) or avoiding interference to other CR devices. As the spectrum is shared among multiple SUs, Spectrum sharing is required to coordinate how SUs can coexist and access the same spectrum without interfering or colliding with each other. In Spectrum mobility, a CR device has to vacate the channel and move to another available channel if a PU appears on that channel.

VI. DETERMINING AVAILABLE SPECTRUM HOLES

The basic function of any CR device is to be able to determine the available spectrum holes which vary in time and space. A CR device should have the capability to determine its location as the available spectrum differs from one place to another. Also, the device has to repeat the calculation periodically as the available spectrum varies with time. Fig. 5 shows a hierarchal description for the approaches used in determining the available spectrum holes.

A. Spectrum Sensing

Spectrum sensing allows a CR user to periodically sense the spectrum and determine its availability for use by SUs. There are two main categories of spectrum sensing techniques [2], namely, Primary transmitter detection and Primary receiver detection (Fig. 5).

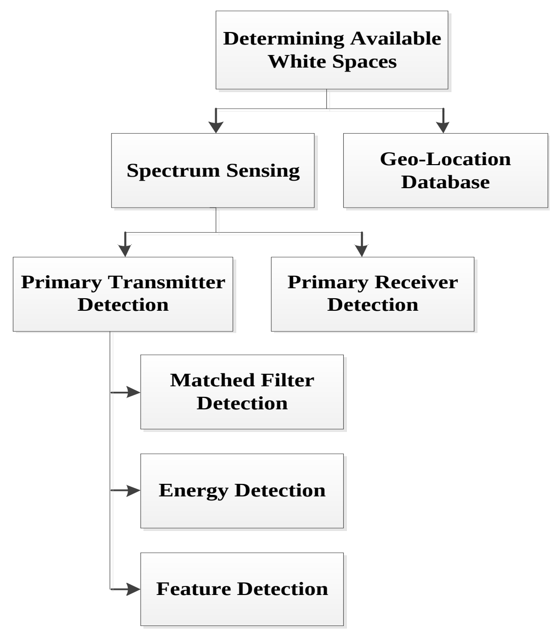


Fig. 5. Methods used for white spaces determination.

In the case of primary transmitter detection, the medium is considered available if the CR device cannot hear the signal sent from any primary transmitter. Three methods [2], [4], [13] can be used for primary transmitter detection. These methods are matched filter detection, energy detection, and feature detection.

 ISBN: 978-1-61208-323-0

The matched filter detection is the optimal detection in the presence of stationary Gaussian noise. This method maximizes the received Signal to Noise Ratio (SNR), but it requires prior information about the characteristics of the PU's signal. The matched filter operates by correlating the pattern that needs to be detected (known information about the signal) with the received signal. If the magnitude of the resulting signal is above a certain threshold; the medium is considered busy otherwise the medium is free. The matched filter is a fast detection technique, but it requires previous knowledge about the signal to be detected.

In the energy detection, no prior information about the primary transmitter signal is required. In this technique, a CR device measures the energy in a certain frequency band if it is above a predefined threshold; the medium is considered busy. If the measured value is below the threshold the medium is considered free and can be used by SUs. One of the concerns of using energy detection is that it just detects the presence or absence of a signal, but it cannot differentiate whether the detected signal is from a primary transmitter or from a secondary transmitter. Another concern is adjusting the threshold value of the detector as this value is affected by the noise level.

The feature detection, also called cyclostationary detection, depends on detecting the cyclostationary (built-in periodicity) feature of the modulated signal for detecting the presence of a signal. This kind of detector is better than energy detector as it is more robust against the uncertainty in noise power, but requires more observation time and is computationally complex.

A main problem in the primary transmitter detection is the hidden node problem. Where, a CR user may be shadowed from detecting the signal of a primary transmitter, due to the presence of an object that block the transmitter signal, as shown in Fig. 6 [2]. The hidden node problem can be solved by using cooperative spectrum sensing or cooperative detection, which allows the CR devices to share the sensed information with each other. This results in a higher detection capability, but comes with the cost of additional overhead

The primary receiver detection technique is the most efficient technique in determining spectrum holes. In this technique, the CR user needs to detect primary receivers in its communication range and avoid interfering with them. Primary receivers detecting is not an easy process. Usually, primary receivers, such as televisions or cellular phones, are passive, which makes it hard for the CR devices to detect them or determine their location. One way to allow a CR device to detect a primary receiver [14] is by utilizing the leakage power of the Local Oscillator (LO), which is emitted by the RF front end of the primary receivers. This method is currently feasible only in TV receiver detection.

Interference temperature is a model introduced by the FCC [15] to accurately measure and limit the amount of interference at the receiver. Interference temperature dictates the cumulative amount of interference from all the undesired RF energy sources that exist at a receiver at any point of time. Interference temperature provides a higher protection for the receiver against harmful interference. A CR user can use the channel while it does not surpass the limit on interference temperature.

B. Geo-Location Database

In the Geo-location database approach, a CR device does not use spectrum sensing to determine the free spectrum;

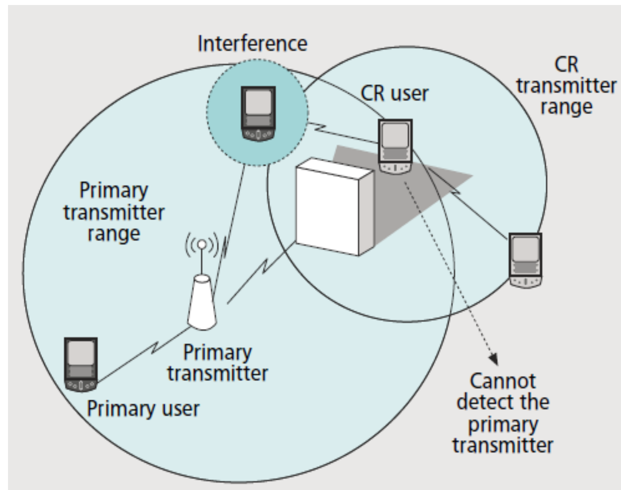


Fig. 6. Shadowing uncertainty [2].

instead it depends on an up-to-data database of incumbents. The database stores information about all primary transmitters and their locations. It also stores terrain information. The data base uses the information it has to calculate available whitespaces at the CR user`s location. According to FCC regulation [6], a CR device should have the capability to determine its location, and a way to connect to the internet to be able to access the database. The process of determining the available whitespaces is done as follows. First, the CR device provides identifying information to register with the database. Then the CR device calculates its location and send it to the database, which uses some propagation models to calculate the available white spaces at the user`s location. After that, the database will send a list of the available white spaces to the user. The database may also inform the device with the maximum allowable transmit power for its operation. In this case, the device can use its adaptable power control to ensure that the transmitting power does not exceed the maximum allowed value. Using a database to calculate spectrum holes overcomes the problem of false alarm that can happen with spectrum sensing and thus provides more efficient use of the spectrum.

Microsoft presented an approach for a geo-location database called "SenseLess: A Database Driven White Spaces Network" [16]. The SenseLess architecture consisted of a logically centralized entity called SenseLess Service. Base stations and CR devices are connected to this central entity which is responsible for determining the available white spaces for any given location. Two components mainly constitute the SenseLess service, the back-end store and the SenseLess engine. The back-end store consisted of a terrain server and a database of incumbents, such as TV transmitters (their location, antenna height, transmission power) and wireless microphones. The database is also used to cache the computed white spaces for different locations. The terrain server store high resolution terrain elevation data which can be obtained from one of the publicly available terrain maps. Sophisticated signal propagation modeling is used by the SenseLess engine to compute the available white spaces at any given location. Results showed that the Longley-Rice (L-R) with terrain propagation model [17] gave accurate results when determining the available white spaces for any given location.

In [18], a White Space Database (WSDB) was used to control the transmit power levels of the White Space Devices (WSDs). WSDs use one of the geo-location capabilities to determine their location and send it to the database. The WSDB sends a list of the available channels to the WSD to ensure that the device will not cause interference to primary incumbents. The database will also inform the device with the transmit power level that it should not exceed. The sum of all WSDs transmit powers should be kept below a certain level to avoid making interference to PUs. Considering these limitations an optimization problem was formulated to control the transmit power while maximizing the total throughput of the system uplink. Solution to the optimization problem becomes the same like the water-filling algorithm problem. Simulation results showed that in the case of using co-channel, increasing the number of users cause the spectrum efficiency to increase. While in case of adjacent channel the spectrum efficiency decrease by increasing the number of users.

VII. SPECTRUM DECISION

Spectrum decision is the capability of a CR device to choose the most appropriate channel for its operation. Channel selection should satisfy Quality of Service (QoS) requirements of SUs and at the same time ensures that they do not cause interference to PUs. Spectrum decision consists of three functions, which are spectrum characterization, spectrum selection and CR re-configurability [19], [11].

The first step after determining the available spectrum holes is to characterize them based on PUs activities and conditions of the radio environment. As the CR user opportunistically utilizes the channel; channel availability can’t be guaranteed during the whole period of its transmission because a PU may appear at any time. Modeling PUs activities can be used to predict future usage of the spectrum based on the historical information of previous spectrum usage. The condition of the radio environment is another factor that is used to characterize the channel based on interference, the number of users utilizing the same channel, and the strength of the received signal. Once the channel characterization is done, a channel that satisfies QoS requirements of the SU is selected. The last step is to adjust the transceiver parameters of the CR device to be able to communicate on the selected spectrum band.

VIII. SPECTRUM SHARING

Spectrum sharing is the most challenging function among CR functions. It addresses the problem of coordinating the transmission of CR devices to allow them to coexist and share the medium without causing interference to each other. Spectrum sharing can be categorized according to architecture, scope, spectrum allocation behavior, spectrum sharing models and spectrum access techniques [2], as shown in Fig. 7.

The spectrum sharing architecture can be centralized or distributed. In centralized spectrum sharing, a central unit is responsible for allocating the spectrum and controlling access to it. In distributed spectrum sharing, spectrum allocation and access is done by each node according to a certain policy specified by the node itself.

The spectrum allocation behaviour can be cooperative or non-cooperative spectrum sharing. In the cooperative spectrum sharing, the CR devices cooperate together to avoid interference with each other. Each CR device adjusts its transmission power taking into account other devices transmission. In the non-cooperative spectrum sharing, each CR device behaves in a selfish manner. A CR device will transmit without considering if its transmission will affect other devices transmission. Thus, in the non-cooperative case there will be high interference between CR devices, which in turn will reduce the spectrum utilization.

The spectrum access techniques can be classified as overlay spectrum sharing and underlay spectrum sharing. In the overlay spectrum sharing, SUs can opportunistically make use of the spectrum when not being used by PUs to avoid causing interference to PUs. In the underlay spectrum sharing, SUs can transmit at the same time with PUs as long as their transmission is below the noise floor of PUs. In this case, SUs use spread spectrum techniques and can only transmit over short range.

The spectrum sharing scope in infrastructure-based CRNs, can be classified to intra-cell spectrum sharing and inter-cell spectrum sharing [20]. The intra-cell spectrum sharing is related to spectrum sharing between CR users in the same cell. The inter-cell spectrum sharing is related to spectrum sharing between different cells.

The two spectrum sharing models are exclusive allocation model and common use model [20]. In the exclusive

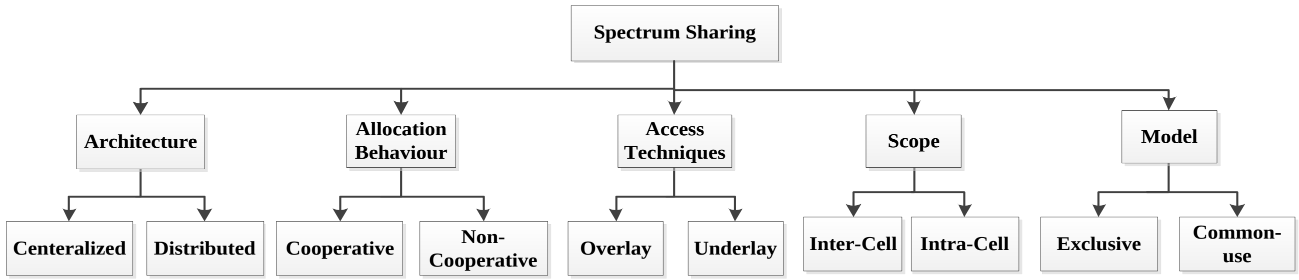


Fig. 7. Spectrum Sharing Categorization.

allocation model, each CR user is allocated a different channel to mitigate the interference between CR users. This model is optimum in maximizing the capacity of the network, but it provides unfair resource allocation in networks that have limited spectrum availability. Although the focus of this approach is on spectrum allocation, power allocation has to be considered to avoid interfering with PUs.

In the common use model, several CR users can simultaneously use the same channel by adjusting their transmission power to minimize interference. This model is preferred in networks with limited spectrum holes; as it can provide fairness in the allocation of the available spectrum. However, the achieved capacity is lower than that of the exclusive model.

Several research papers have been proposed to address the spectrum sharing problem in CRNs. The main differences between papers addressing spectrum sharing techniques are in the model they used for spectrum sharing and the objective of the network.

In [21], a combined power/channel allocation method was applied in a WiFi-like spectrum sharing scenario in TV white spaces. Three types of Secondary Users (SUs) were defined which are protected, interfered and out of range SUs. The network goal was to increase the number of supported SUs while reducing the interference secondary devices cause to each other. They used the NBS to allocate transmit power for Secondary Access Points (SAPs). SAPs cooperate by exchanging the information through relaying nodes. These relaying nodes are the interfered SUs that can hear from more than one SAP. SAPs compete on their transmission powers to maximize the number of supported SUs. SAPs have to decrease their power until there is no overlapping between their coverage areas. Two stage of cooperation were made. In the first stage, only neighboring SAPs cooperate and compete for power control, while in the second stage, all the next hop neighboring SAPs can cooperate. When the SAPs are highly overlapped, the algorithm can switch to channel allocation instead of power allocation to enhance the network performance. The switch is done if the number of interfered secondary users was above a certain number. Simulation results showed that the number of iteration required to reach optimality is decreased by SAPs cooperation. The number of supported users increased while the average SAP throughput decreased.

A downlink channel assignment and power control for an infrastructure-based cognitive radio network was implemented in [22]. The opportunistic spectrum access problem was formulated as a non-cooperative game in which the game players are the base stations. Each base station bargain to increase the number of supported CRs. Channel allocation was done by the base stations, which randomly assign channels to users. A distributed power allocation is then applied using the Iterative Water Filling algorithm. Results showed that the pure non cooperative game might have multiple Nash equilibrium points [23] that are undesirable and may lead to non-convergence. To obtain better results, the Nash bargaining solution was applied in which the cooperation of base stations was required. Simulation results showed that a unique optimal solution was achieved by using the Nash bargaining solution.

IX. SPECTRUM MOBILITY

Spectrum mobility is the process of performing a seamless transition from one channel to another available channel. After a CR user selects the channel and starts transmitting on a certain frequency band, a PU may appear on the same channel; in this case the CR user has to move to another empty channel and vacate this channel to PU to avoid causing interference to the PU. The SU may also change its channel to access another spectrum hole with better QoS. Spectrum mobility consists of two processes [24], spectrum handoff and connection management.

In the spectrum handoff process, the SU transmission is transferred from its current channel to another empty channel. Three events can trigger the spectrum handoff process. The first event is the arrival of a PU in a channel occupied by a SU. The second is the spatial movement of SUs to a place where their coverage overlap with PUs already utilizing the channel. The third is the degradation of the link quality.

The spectrum handoff process consists of two phases, evaluation phase and link maintenance phase. In the evaluation phase, the SU keep monitoring the environment to determine if an event that trigger spectrum handoff occurred, then the SU moves to the next phase (Link maintenance). In the Link maintenance phase, the SU pauses its transmission

 ISBN: 978-1-61208-323-0

on the current channel and continues the transmission on another available channel.

Connection management process is used to compensate for the unavoidable handoff delay, which happens when the SU transmission is transferred from a channel to another, by adjusting the parameters of the protocol stack according to the existing situation.

X. CONCLUSION

In this paper, we have presented a comprehensive overview of the CRN architecture and operation in licensed and unlicensed bands. The paper focused on the four CR spectrum management functions, white space determination, spectrum decision, spectrum sharing, and spectrum mobility. The white spaces can be calculated using either spectrum sensing or geo-location database. Most of the papers in the literature use spectrum sensing to calculate white spaces. However, the geo-location database approach is more accurate and efficient. Also, it overcomes spectrum sensing problems such as false alarms. The decision to select one of the white spaces depends on characterizing all the available white spaces based on the PUs activities and the conditions of the radio environment. The CR device transceiver is then adjusted to operate on the selected band. Sharing the spectrum is a very interesting and challenging function throughout the CR functions. Spectrum sharing is concerned with the way that enables all the CRNs to coexist and share the same spectrum without interfering with PUs or with each other. The spectrum sharing function was comprehensively covered according to different aspects, architecture, scope, allocation behavior, sharing models, and access techniques. The spectrum mobility was used to avoid interfering with PUs by transferring CR users to another available channel if a PU appears on that channel.

REFERENCES

[1] "Report of the spectrum efficiency working group," FCC Spectrum Policy Task Force, Nov 2002.

[2] I. F. Akyildiz, W. Y. Lee, M. C. Vuran, and S. Mohanty, "A Survey on Spectrum Management in Cognitive Radio Networks," IEEE Communications Magazine, April 2008 pp. 40–48.

[3] "Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio," Federal Communications Commission, 2005.

[4] A. Garhwal and P. P. Bhattacharya, "A Survey on Dynamic Spectrum Access Techniques for Cognitive Radio," International Journal of Next-Generation Networks (IJNGN), vol. 3, Dec 2011 pp. 15-32.

[5] "Unlicensed Operation in the TV Broadcast Bands," Federal Communications Commission, Notice of proposed rule making, May, 2004.

[6] "Unlicensed Operation in the TV Broadcast Bands," Federal Communications Commission, Nov 2008.

[7] "Unlicensed Operation in the TV Broadcast Bands," Federal Communications Commission, Sept 2010.

[8] K. C. Chen, Y. J. Peng, N. Prasad, Y. C. Liang, and S. Sun, "Cognitive Radio Network Architecture: Part I - General," in Proceedings of the 2nd International Conference on Ubiquitous Information Management and Communication, ACM, 2008 pp. 114-119.

[9] "FCC,White Space Database Administrators Guide," [Online]. http://www.fcc.gov/encyclopedia/white-space-database-administrators-guide [retrieved: Jan, 2014]

[10] "Ofcom,Implementing Geolocation," [Online]. http://stakeholders.ofcom.org.uk/consultations/geolocation/statement/ [retrieved: Jan, 2014]

[11] M. T. Masonta, M. Mzyece, and N. Ntlatlapa, "Spectrum Decision in Cognitive Radio Networks: A Survey," Communications Surveys & Tutorials, IEEE, vol. 15, Issue: 3, Nov 2012 pp. 1 - 20.

[12] E. Z. Tragos, S. Zeadally, A. G. Fragkiadakis, and V. A. Siris, "Spectrum Assignment in Cognitive Radio Networks: A Comprehensive Survey," IEEE Communications Surveys & Tutorials, vol. 15, Issue: 3, Jan 2013 pp. 1 - 28.

[13] "E3 White Paper Sensing," Nov. 2011. [Online]. https://ict-e3.eu/project/white_papers/E3_White_Paper_Sensing.pdf [retrieved: Jan, 2014]

[14] B. Wild and K. Ramchandran, "Detecting Primary Receivers for Cognitive Radio Applications," in New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on , Nov 2005 pp. 124 - 130.

[15] "Establishment of an Interference Temperature Metric to Quantify and Manage Interference and to Expand Available Unlicensed Operation in Certain Fixed, Mobile and Satellite Frequency Bands," Federal Communications Commission, Nov 2003.

[16] R. Murty, R. Chandra, T. Moscibroda, and P. Bahl, "SenseLess: A Database-Driven White Spaces Network," in IEEE Transations on mobile computing, Feb 2012 pp. 189 - 203.

[17] "Longley-Rice Methodology for Evaluating TV Coverage and Interference," OET BULLETIN No. 69, Feb 2004.

[18] S. Y. Lee, M. K. Kwon, and S. H. Lee, "Transmit power control scheme for TV White space wireless system," in Advanced Communication Technology (ICACT), 2011 13th International Conference on. IEEE , Feb 2011 pp. 1025 - 1029.

[19] I. F. Akyildiz, W.-Y. Lee, and M. C. Vuran and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey," Computer Networks, vol. 50, May 2006 pp. 2127-2159.

[20] W. Y. Lee and I. F. Akyildiz, "Joint Spectrum and Power Allocation for Inter-Cell Spectrum Sharing in Cognitive Radio Networks," in New Frontiers in Dynamic Spectrum Access Networks, 2008. DySPAN 2008. 3rd IEEE Symposium on, Chicago, Oct 2008 pp. 1- 12.

[21] B. Jankuloska, V. Atanasovski, and L. Gavrilovska, "combined power/channel allocation method for efficient spectrum sharing in TV white space scenario," in Proceedings of the 4th International Conference on Cognitive Radio and Advanced Spectrum Management, ACM , 2011 pp. 59:1 – 59:5.

 ISBN: 978-1-61208-323-0

[22] G. Hosseinabadi, "Downlink Channel Assignment and Power Control for Cognitive Radio Networks Using Game Theory".

[23] M. Felegyhazi and J. P. Hubaux, “Game theory in wireless networks: a tutorial”, EPFL technical report, LCA-REPORT-2006-002, Feb, 2006.

[24] I. Christian, S. Moh, I. Chung, and J. Lee, "Spectrum Mobility in Cognitive Radio Networks," Communications Magazine, IEEE, June 2012 pp. 114 - 121.

 ISBN: 978-1-61208-323-0

Spectrum Leasing Game for Underlay Cognitive Ratio Network with Primary System Using Adaptive Rate-Based Pricing Strategy

Wenson Chang
Department of Electrical Engineering
National Cheng Kung University
Tainan, Taiwan 701
Email: wenson@ee.ncku.edu.tw

Yu-Ping Lin
ASML Co., Ltd.
Taipei, Taiwan, 300.
Email: meekhoney@hotmail.com

Szu-Lin Su
Department of Electrical Engineering
National Cheng Kung University
Tainan, Taiwan 701
Email: ssl@ee.ncku.edu.tw

***Abstract*—In this paper, we propose a low-complexity spectrum leasing game for the underlay CR network with PS using adaptive rate-based pricing strategy. In the proposed scenario, an SU can make a request for sharing a channel with multiple PUs and pay for the spectrum lease in proportional to his transmission power. In the meantime, the PUs can determine the leasing price to maximize his own revenue at the cost of ignorable throughput degradation; In other words, one can say that the PU can actively protect himself from using the lower transmission rate by adaptively rising the leasing prize, rather than passively imposing interference-limit rule as in the conventional methods. The simulation results show that the proposed spectrum leasing game can grant SUs transmission opportunities without causing PUs throughput degradation. Moreover, the convergency of the proposed scheme is numerically proved, which accounts for the existence of Nash Equilibrium.**

Keywords-Game theory, auction, cognitive radio, pricing, spectrum leasing.

I. Introduction

Due to the scarceness of spectrum and the increasing demands for radio resources, the issue of radio resource management (RRM) has attracted much attention for decades, especially when the truth of extremely low spectrum efficiency had been revealed by the US Federal Communications Commission (FCC) [1]. To improve the spectrum efficiency, the technology of cognitive radio (CR) has been proposed to opportunistically and temporarily utilize the so-called spectrum holes [2]. Accordingly, many important CR-related topics have been extensively explored, including spectrum leasing (trading), spectrum sensing, sub-channel allocation, power and interference control, cooperative communications etc [3]–[8].

Speaking of spectrum leasing (trading), the auction theory, one of the important applications of Game theory, can be utilized to design effective protocols for managing the aforementioned interactions between primary system (PS) and secondary system (SS) as well as the competitions between secondary users (SUs) [9], [10]. Moreover, according to CR's operation modes, different spectrum leasing protocols can be designed for the purpose of protecting PS and improving the overall spectrum efficiency (more discussions can be found in the following literature survey). One should note that the key to the success of CR networks lies in the harmless interactions between the PS and SS. Therefore, from the literature, one can find that whatever the spectrum leasing protocols are applied, the well-known interference-limit (IT) rule is still pivotal to regulate SS's behaviors in the underlay CR networks.

In this paper, PUs can actively play the leading characters in the spectrum leasing game in stead of passively imposing the IT rule on SS. That is to say, rather than the IT rule, each PU can adaptively set and announce the price for spectrum sharing based on the degradation of his transmission rate. Note that the acceptable degradation level depends on each PU's willingness and tolerance. Then, each SU can play a one-shot game to decide the amount of power he can purchased from each of PU. Afterwards, the PU may adjust the price based his attainable transmission rate. The bidding process between PU and SUs may go back and forth for several times until Nash equilibrium (NE) is reached.

It should be noticed that, in the proposed scheme, all the SUs who share the same channel (owned by a particular PU) do not need to know whether the aggregated amount of interference has gone beyond the IT threshold or not, which may cost some considerable amount of message exchanges. Via the simulation results, the proposed spectrum leasing game can grant SUs transmission opportunities without causing PUs significant throughput degradation. Moreover, the existence of a unique NE can also be proved numerically, which accounts for the convergence of the proposed scheme.

The rest of this paper is organized as follows. In Section II, some important related works are reviewed. Section III describes the system model and proposed spectrum leasing game. Simulation results are provided in Section IV. Section V concludes this work.

II. Literature Survey

Here, we review several auction-based spectrum leasing (trading) schemes in the underlay and interweave CR networks.

A. Underlay CR Operation Mode

In the underlay CR networks, PS can be protected by imposing the IT rule on the SS's transmissions. Thus, in the following auction games, the SS's transmission power is regulated by this rule. In [11], two auction mechanisms were

 ISBN: 978-1-61208-323-0

designed to allocating power for SUs, of which an SU can be charged according to the received signal-to-interference-and-noise ratio (SINR) or transmission power. The charging policies, in other words bidding strategies, were to reach the pre-defined balance, i.e. NE, between SUs. In [12], the pricing method was included into the design of channel and power allocations for the CR networks. Then, the distributed price-based iterative water-filling (PIWF) algorithm as well as the corresponding media access control (MAC) protocol were proposed to reach NE. In [13], the authors allowed the primary users (PUs) to join the auction game by actively adjusting the tolerable amount of interference (the interference cap). Then, a dynamic spectrum leasing strategy was designed to control the transmission power for both PUs and SUs such that the utility functions of SUs can be maximized under the limit of the so-called interference cap.

In [14], the authors proposed a sub-optimal pricing strategy for PS to own better revenue, and SUs can also adjust their uplink transmission power to maximize their utility functions. The price was set according to the amount of each SU's transmission power. Via numerical proof, the proposed method was claimed to achieve fairness of power allocation between SUs. In [15], the Stackelberg model was applied to deal with the spectrum leasing problem between PS and SS. In additional to the payment by SUs (i.e. revenue of PUs), a shutdown mechanism was developed to prevent SUs from using unacceptable transmission power such that the IT rule can be satisfied. In [16], the problem of power control and relay selection in the multi-hop CR relay-network was investigated. An SU can pay prices to PUs who share the spectrum and SUs who relays for him to obtain optimal performance. The effects of difference pricing functions of PUs were investigated and several distributed power control algorithms were developed for the scenarios with single and multiple CR transmission pairs.

B. Interweave CR Operation Mode

In the interweave CR networks, PS can be protected by separating the access time slots or spectrum bands from those of SS, and (or) by the helpful cooperative transmissions from SS. In [17], for the purpose of maximizing quality-of-service (QoS), a PU can allure SUs for cooperative transmissions by offering a fracture of time slots in return. To own the offered fracture of time, SUs compete between each others following the distributed power control mechanism. The Stackelberg game was applied to model leader-follower relationship between PU and SUs. In [18], a PU can set a price for time sharing with a properly selected SUs and each SU can decide the amount of time to purchase according to his own QoS requirement. Note that the more time purchased, the more power should be used to forward packets for the PU. Moreover, an admission control mechanism was also develop to protect the SUs who participate in the cooperation transmissions. In [19], similar scenario was extended to multiple PUs, i.e. multiple sellers, by adopting the framework of generalized NE, of which these PUs competed with each other for the cooperative transmissions from SUs. And SUs can become willing to cooperate when the their QoS requirements can be achieved.

In [20], during the spectrum leasing period, each PU can preserve a required bandwidth for himself to satisfy his own QoS. Then, SUs can bid for the extra bandwidth, rather than the aforementioned a fracture of time, to maximize their utility function and achieve the fair allocations, i.e. NE. Similar scenario was also extended to multiple PUs. In [21], a generalized Branco's mechanism was proposed to tackle with the spectrum trading between primary service providers (PSPs) and secondary service providers. In the proposed model, the PSPs (seller) can cooperate to gain maximal profits and further share these profits. In [22], the bandwidth-sharing problem in the multi-hop relaying cellular network (MCN) was modeled by the reversed Stackelberg game. With the aid of a trust model, the base station (BS) can encourage relay station (RS) to cooperate and also discourage their misbehavior. Owing the allocated bandwidth, the RSs can serve the nearby mobile terminals. Via simulation results, the well-developed MCN cooperation can maximize the overall network performance.

III. System Model and Spectrum Leasing Game

A. System Model

Now, we consider an underlay CR network (SS) imbedded in the incumbent cellular system (PS), which is assumed to be in the uplink transmissions using adaptive modulation and coding scheme (ACM) with fixed transmission power. Each PU exclusively own a subchannel and he may share this subchannel with multiple SUs. A CR transmitting end and a CR receiving end form a CR transmission pair, which means the CR network operates in the ad hoc mode. For simplicity, a CR transmission pair is named SU in the following context.

In the proposed spectrum auction game, SUs can issue spectrum sharing requests to the neighbor PUs. After receiving these requests, each of PU can decide whether or not to join this auction and then announce the price per unit of generated interference if the requests are accepted. Note that a higher price can resist SU from producing intolerable amount of interference. According to the prices, each SU can decide the amount of transmission power he can afford to allocate to each of the subchannels. Several rounds of the bidding process (i.e., the price setting and power allocations) can give the balanced portfolio (the so-called NE). In the following context, the whole period of bidding process is named bidding phase.

B. Power Allocation of SS

The utility function of the ith SU can be defined as

$$\begin{aligned} U_i^S &= \sum_{k\in\Omega_M} u_i^{Sk} \\ &= \sum_{k\in\Omega_M}\Bigg[\alpha_i \log_2\left(1+\frac{P_i^{Sk}G_{ii}^{Sk}}{\sum_{j\neq i, j\in\Omega_N} P_j^{Sk}G_{ji}^{Sk}+Q_{ki}^{PS}+N_o}\right) \\ &\quad -\lambda_k(P_i^{Sk}G_{ik}^{SP})\Bigg] \ , \end{aligned} \tag{1}$$

where u_i^{sk} represents the utility of the ith SU sharing kth subchannel; $\Omega_M = \{1, \cdots, M\}$ and $\Omega_N = \{1, \cdots, N\}$ respectively stand for the sets of PUs and SUs who join this auction; P_i^{Sk} is the ith SU's allocated power to the kth subchannel; G_{ij}^{Sk} means the channel gain between the ith SU's transmitter and jth SU's receiver over the kth subchannel, while G_{ik}^{SP} is that between the ith SU's transmitter and kth

PU's receiver, i.e. the base station (BS); Q_{ki}^{PS} denotes the ith SU's received interference from the kth PU; N_o accounts for the additive white Gaussian noise (AWGN); λ_k is the announced price for sharing the kth subchannel. Note that α_i is the adjusting weight factor which can personalize ith SU's characteristic. Moreover, observing (1), the first term and second term account for the capacity reward and spectrum sharing cost (in other words, interference penalty). Also, the summation explains the spectrum sharing property, i.e. an SU can share multiple subchannels with PUs.

The goal of SU is to maximize the utility function via a proper power allocation. Thus, we can now form an optimization problem as what follows.

$$\max_{\mathbf{P}_i^S} \; U_i^S(\mathbf{P}_i^S, \mathbf{P}_{-i}^S)$$
$$subject\ to$$
$$0 \le P_i^{Sk} \le P_{max}$$
$$0 \le \sum_{k \in \Omega_M} P_i^{Sk} \le P_{max} \ , \quad (2)$$

where $\mathbf{P}_i^S = \{P_i^{S1}, ..., P_i^{SM}\}$ is the ith SU's power allocation vector, while $\mathbf{P}_{-i}^S$ describes that of all the other SUs belonging to Ω_N. Prior to solving this optimization problem, it should be noted that the concavity of $U_i^S(\mathbf{P}_i^S, \mathbf{P}_{-i}^S)$ with respect to $\mathbf{P}_i^S$ can be proved by calculating $\partial^2 U_i^S / \partial P_i^{Sk^2}$, of which the negativity renders the proofs of its concavity and existence of the maximum.

First, we relax the constraint of $0 \le P_i^{Sk} \le P_{max}$ and then solve the following KKT problem, which gives a power allocation scheme to satisfy this constraint. The KKT problem can be defined as

$$\max_{\mathbf{P}_i^S} \; L(\mathbf{P}_i^S, \mu_i)$$
$$subject\ to$$
$$0 \le \sum_{k \in \Omega_M} P_i^{Sk} \le P_{max} \ , \quad (3)$$

where

$$L(\mathbf{P}_i^S, \mu_i) = \sum_{k \in \Omega_M} \left[\alpha_i \log_2 \left(1 + \frac{P_i^{Sk} G_{ii}^{Sk}}{\sum_{j \ne i, j \in \Omega_N} P_j^{Sk} G_{ji}^{Sk} + Q_{ki}^{PS} + N_o} \right) - \lambda_k (P_i^{Sk} G_{ik}^{SP}) \right] + \mu_i \left[(P_{max} - \left(\sum_{k \in \Omega_M} P_i^{Sk} \right) \right] \ ; \quad (4)$$

And μ_i is the Lagrangian multiplier. Solving $\partial L(\mathbf{P}_i^S, \mu_i) / \partial P_i^{Sk} = 0$ gives

$$P_i^{Sk^*} = \left[\frac{\alpha_i}{\ln 2 (\mu_i + \lambda_k G_{ik}^{SP})} - \frac{\Gamma_i^{Sk}}{G_{ii}^{Sk}} \right]_0^{P_{max}} \ , \quad (5)$$

where $\Gamma_i^{Sk} = \sum_{j \ne i, j \in \Omega_N} P_j^{Sk} G_{ji}^{Sk} + Q_{ki}^{PS} + N_o$. Note that μ_i can be adjusted to satisfy $\sum_{k \in \Omega_M} P_i^{Sk} = P_{max}$ such that higher capacity can be achieved.

C. Pricing Strategy of PS

Here, we define the utility function for PU.

$$U_k^P = \left[\lambda_k - C_k \cdot e^{\beta_k (1 - R_k^P / R_k^o)} \right] I_k^P \ , \quad (6)$$

where C_k is kth PU's reserved price (or the cost in other words) for spectrum leasing; β_k (similar to α_i) is the price weighting factor, which can adjust the sensitivity to the change of transmission rate; R_k^o and R_k^P denotes the original and current transmission rate of kth PU, respectively; $I_k^P = \sum_{i \in \Omega_N} (P_i^{Sk} G_{ik}^{SP})$ is the amount of experienced interference. Similar to the SS's case, calculating $\partial^2 U_k^P / \partial \lambda_k^2$ can prove the concavity of U_k^P with respect to λ_k as well as the existence of maximum. Also, observing (6), one can find that the lower the current transmission rate R_k^P, the higher the cost of spectrum sharing, which can result in a higher price (as proved by the following equation).

Solving $\partial U_k^P / \partial \lambda_k = 0$ gives the optimal pricing strategy of PU as

$$\lambda_k = C_k \cdot e^{\beta_k (1 - R_k^P / R_k^o)} - \frac{I_k^P}{(\partial I_k^P / \partial \lambda_k)} \ . \quad (7)$$

Note that whenever the price is changed, the variation of experienced amount of interference with respect to price, i.e. $\partial I_k^P / \partial \lambda_k$, can be estimated by the kth PU itself.

D. Discussions on Information Exchanges

To put the proposed spectrum auction game into practice, only two additional information exchanges over one control channel are needed: one for SU to broadcast the spectrum sharing request and one for PU to announce the spectrum sharing price (λ_k). Observing (5), Γ_i^{Sk} and G_{ii}^{Sk} can be estimated by some existing techniques (which are beyond the scope of this paper). Also, thanks to the channel reciprocality, G_{ik}^{SP} can be attained by listening to λ_k over the kth subchannel in the time-division duplex (TDD) mode. Furthermore, observing (7), I_k^P and $\partial I_k^P / \partial \lambda_k$ can surely be estimated by PU itself.

It should be noticed that conventionally, an IT rule should be imposed on SS in the underlay CR network. However, in the considered scenario that a PU can share his subchannel with multiple SUs, some massive information exchanges are required. For example, each of SUs should know the percentage of the amount of interference he has produced. Then the SU can actively control the transmission power such that the IT rule is not violated. Or, alternatively, a group of SUs can sequentially adjust their transmission power, which may additionally cause some difficult problems, e.g. how to form the group of SUs and inform each of them, the fairness issue, scheduling protocol design etc. Certainly, PU can also inform each of SUs to adjust transmission power. However, in either case, massive information exchanges as well as more additional control signals are required.

IV. Simulation Results

The simulation environment is built up based on the system model described in Section III(A). In addition, PUs are uniformly distributed over the primary incumbent cell of radius 2000 (m). The transmitting end of each SU is randomly located at where the distance to BS is uniformly distributed over the

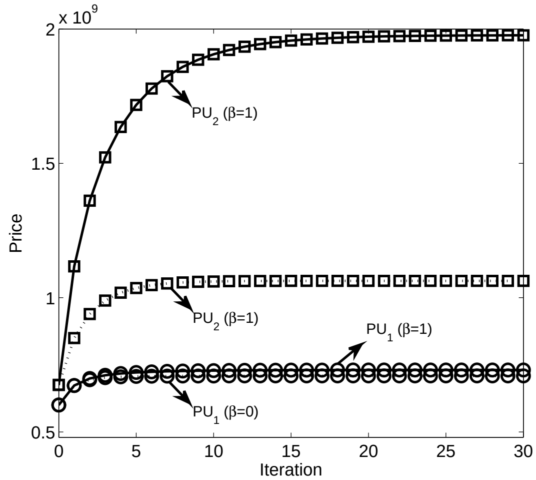


Figure 1: Convergence of PUs' prices (λ_1 and λ_2) with $\beta_1 = \beta_2 = 1$.

TABLE I: Adaptive modulation scheme of PU

Required SINR (dB)	6	10	18	24
Throughput (bps/Hz)	1	2	4	6

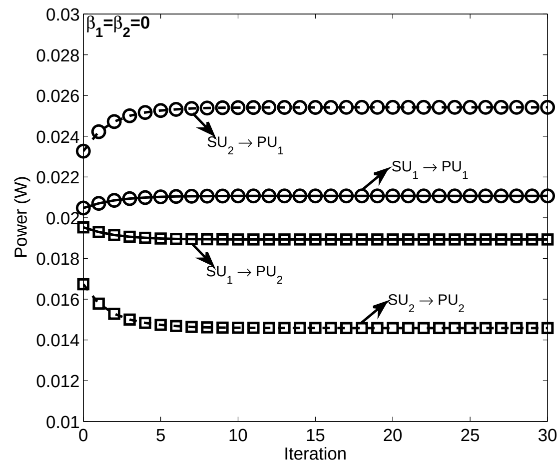
interval of [500 1000] (m); And, its corresponding receiving end is positioned randomly at where it is 50-150 meters away. The transmission power of PU and SU's maximum transmission power P_{max} are 1 and 0.1 watt, respectively. The power spectrum density of AWGN is -174 dBm/Hz. The log-distance path-loss model with exponent of three and flat Rayleigh fading are assumed. Furthermore, the stairwise effective throughput of ACM for PU are listed in Table I [23]. In the simulations, it is assumed that there are two PUs providing spectrum sharing opportunities to two SUs; And the results are averaged over 2×10^7 simulation rounds.

A. Convergency

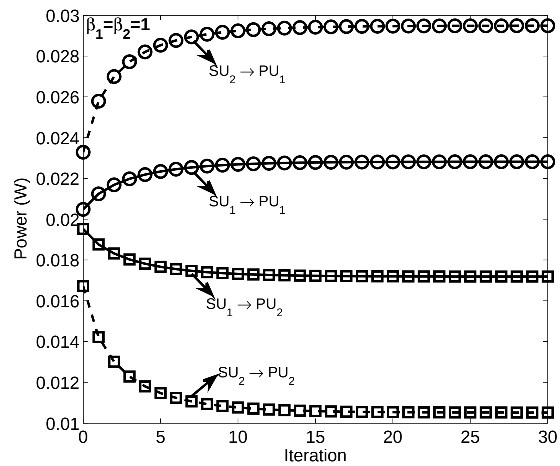
Here, we prove the convergence of the spectrum leasing game by showing the convergence of the PUs' price (λ_1 and λ_2) and SUs' transmission power. Figures 1 and 2 respectively show the snapshots of PU's price for $\beta_1 = \beta_2 = 1$, and SU's transmission power for $\beta_1 = \beta_2 = 0$ and 1. As shown in the figures, the bidding phase can be completed within ten iterations.

B. Impacts on PUs

In advance of showing the benefit SU can obtain through the spectrum auction, we first present its impact on PU. Figure 3 shows the impact of SU's activity on PU's transmission mode, i.e. ACM mode with (a) $\beta_1 = \beta_2 = 1$, and (b) $\beta_1 = \beta_2 = 4$, respectively. It can be observed that when the price weighting factor β equals to zero, PU can be severely affected, and consequently the probability of using lower ACM mode significantly rises. One should note that the case of $\beta = 0$ can be regarded as conventional method, i.e. simply maximizing the utility function without adaptively adjusting



(a) $\beta_1 = \beta_2 = 0$



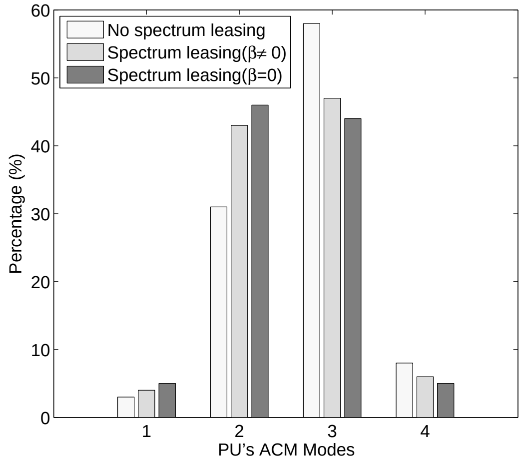
(b) $\beta_1 = \beta_2 = 1$

Figure 2: Convergence of SUs' transmission power with (a) $\beta_1 = \beta_2 = 0$, and (b) with $\beta_1 = \beta_2 = 1$.

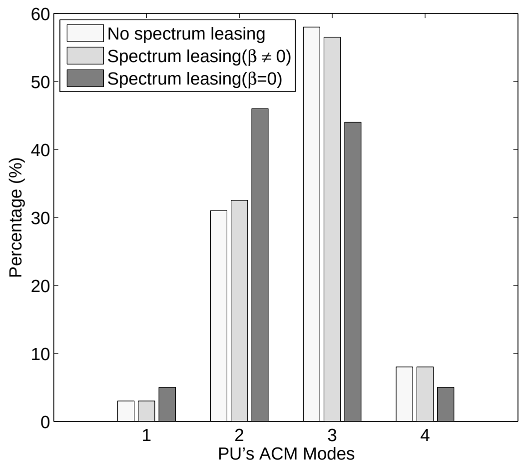
the cost C_k. In this situation, P_{max} in (2) can be regarded as the IT rule to regulate SU's transmission power. Fortunately, this unfavorable situation can be avoided by increasing the weighting factor β, i.e. sensitivity to the change of throughput. As shown in Figure 3(b), when β increases from zero to four, PU can almost maintain the same ACM mode.

C. Benefit to SUs

Figure 4 shows the average throughput of SU and PU with respect to various price weighting factors (β). One can find that higher throughput can be reached by SU with lower value of β. However, in this situation, it can cause larger loss of throughput to PU. As aforementioned, rising the β value can solve this dilemma. Therefore, using $\beta = 4$, both SU and PU can maintain at high throughput of 3.84 and 3.44 bps/Hz, respectively.

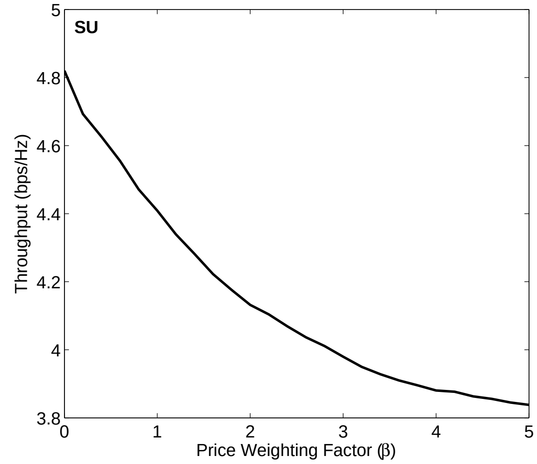


(a) $\beta_1 = \beta_2 = 1$

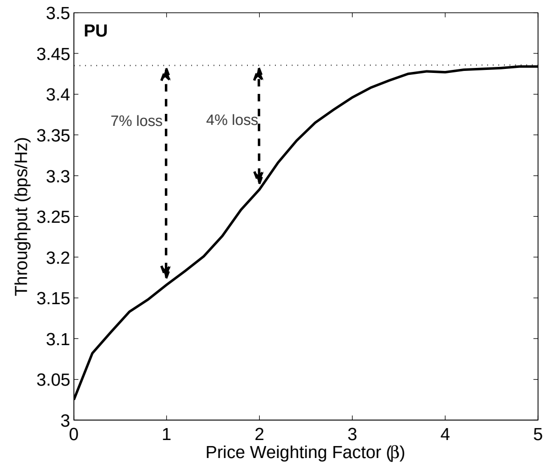


(b) $\beta_1 = \beta_2 = 4$

Figure 3: The impacts of SU's activity on PU's transmission mode, i.e. ACM mode, with (a) $\beta_1 = \beta_2 = 1$, and (b) $\beta_1 = \beta_2 = 4$.



(a) SU



(b) PU

Figure 4: The average throughput of SU and PU with respect to various price weighting factor (β).

V. Conclusion

In this paper, we have proposed a novel spectrum leasing game for the underlay CR network using rate-based pricing strategy for PU rather than the conventional interference-based method. By using the rate-based pricing strategy, several important advantages can be obtained. First, signaling overhead can be significantly decreased. Only two additional information exchanges (one for issuing SU's request and one for broadcasting the price of spectrum sharing) over one control channel are required. Second, PUs can actively join the auction and protect himself from using lower transmission rate by rising the price, instead of passively imposing the IT rule on SUs. Third, both PU and SU can simultaneously maintain at high transmission rates such that the overall spectrum efficiency can be largely improved. Many interesting future works are worth exploring, which could have potential impacts on the area of CR networks. For example: (1) apply the Stackelberg model to refine the spectrum auction game; (2) mathematically prove the existence and uniqueness of NE.

Acknowledgment

This work was jointly supported by the National Science Council, Taiwan, under the contracts 101-2221-E-006-176-MY2, 101-2221-E-006-177-MY2 and 101-2923-E-194-003-MY3.

References

[1] Federal Communications Commission, "Spectrum policy task force," ET Docket, no. 02-135, Nov. 2002.

[2] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," IEEE Journal on Selected Areas in Communications, vol. 23, no. 2, Feb. 2005, pp. 201–220.

[3] H. Sun, A. Nallanathan, C.-X. Wang, and Y. Chen, "Wideband spectrum sensing for cognitive radio networks: A survey," IEEE Wireless Communications, vol. 20, no. 2, April 2013, pp. 74–81.

 ISBN: 978-1-61208-323-0

[4] L. Huang, G. Zhu, and X. Du, "Cognitive femtocell networks: An opportunistic spectrum access for future indoor wireless coverage," IEEE Wireless Communications, vol. 20, no. 2, April 2013, pp. 44–51.

[5] M. Song, C. Xin, Y. Zhao, and X. Cheng, "Dynamic spectrum access: From cognitive radio to network radio," IEEE Wireless Communications, vol. 19, no. 1, Feb 2012, pp. 23–29.

[6] S. Sengupta and K. P. Subbalakshmi, "Open research issues in multi-hop cognitive radio networks," IEEE Communications Magazine, vol. 51, no. 4, April 2013, pp. 168–176.

[7] Y. Zhang, D. Niyato, P. Wang, and E. Hossain, "Auction-based resource allocation in cognitive radio systems," IEEE Communications Magazine, vol. 50, no. 11, Nov. 2012, pp. 108–122.

[8] Y. Zou, Y.-D. Yao, and B. Zheng, "Cooperative relay techniques for cognitive radio systems: Spectrum sensing and secondary user transmissions," IEEE Communications Magazine, vol. 50, no. 4, April 2012, pp. 98–103.

[9] D. Niyato and E. Hossain, "Spectrum trading in cognitive radio networks: A market-equilibrium-based approach," IEEE Wireless Communications, vol. 15, no. 6, Dec. 2008, pp. 71–80.

[10] K. Akkarajitsakul, E. Hossain, D. Niyato, and D. I. Kim, "Game theoretic approaches for multiple access in wireless networks: A survey," IEEE Communications Surveys & Tutorials, vol. 13, no. 3, Third Quarter 2011, pp. 372–395.

[11] J. Huang, R. A. Berry, and M. L. Honig, "Auction-based spectrum sharing," Mobile Networks and Applications, vol. 11, no. 3, June 2006, pp. 405–418.

[12] J. Huang, Z. Han, M. Chiang, and H. V. Poor, "Auction-based resource allocation for cooperative communications," IEEE Journal on Selected Areas in Communications, vol. 26, no. 7, Sept. 2008, pp. 1226–1237.

[13] S. K. Jayaweera and T. Li, "Dynamic spectrum leasing in cognitive radio networks via primary-secondary user power control games," IEEE Transactions on Wireless Communications, vol. 8, no. 6, June 2009, pp. 3300–3310.

[14] H. Yu, L. Gao, Z. Li, X. Wang, and E. Hossain, "Pricing for uplink power control in cognitive radio networks," IEEE Trans. on Vehicular Technology, vol. 59, no. 4, May 2010, pp. 1769–1778.

[15] R. D. Taranto, P. Popovski, O. Simeone, and H. Yomo, "Efficient spectrum leasing via randomized silencing of secondary users," IEEE Transaction on Wireless Communications, vol. 9, no. 12, Dec. 2010, pp. 3739–3749.

[16] Y. Xiao, G. Bi, and D. Niyato, "Game theoretic analysis for spectrum sharing with multi-hop relaying," IEEE Transaction on Wireless Communications, vol. 10, no. 5, May 2011, pp. 1527–1537.

[17] O. Simeone, I. Stanojev, S. Savazzi, Y. B. Ness, U. Spagnolini, and R. Pickholtz, "Spectrum leasing to cooperating secondary ad hoc networks," IEEE Journal on Selected Areas in Communications, vol. 26, no. 1, Jan. 2008, pp. 203–213.

[18] X. Wang, K. Ma, Q. Han, Z. Liu, and X. Guan, "Pricing-based spectrum leasing in cognitive radio networks," IET Networks, vol. 1, no. 3, Sept. 2012, pp. 116–125.

[19] T. Elkourdi and O. Simeone, "Spectrum leasing via cooperation with multiple primary users," IEEE Trans. on Vehicular Technology, vol. 61, no. 2, Feb. 2012, pp. 820–825.

[20] X. Wang, Z. Li, P. Xu, Y. Xu, X. Gao, S. Member, and H.-H. Chen, "Spectrum sharing in cognitive radio networks-an auction-based approach," IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics, vol. 40, no. 3, June 2010, pp. 587–596.

[21] S. H. Chun and R. J. La, "Secondary spectrum tradingxauction-based framework for spectrum allocation and profit sharing," IEEE/ACM Transactions on Networking, vol. 21, no. 1, Feb. 2013, pp. 176–189.

[22] S. Kim, "Reversed stackelberg bandwidth-sharing game for cognitive multi-hop cellular networks," IET Communications, vol. 6, no. 17, Nov. 2012, pp. 2907–2913.

[23] X. Qiu and K. Chawla, "On the performance of adaptive modulation in cellular systems," IEEE Transactions on Communications, vol. 47, no. 6, June 1999, pp. 884–895.

Hybrid Approach Analysis of Energy Detection and Eigenvalue Based Spectrum Sensing Algorithms with Noise Power Estimation

Pawan Dhakal
Kathmandu University
Dhulikhel, Nepal
pawan.dhakal@student.ku.edu.np

Daniel Riviello, Roberto Garello
Politecnico di Torino
Torino, Italy
{daniel.riviello, roberto.garello}@polito.it

Federico Penna
Samsung MSL
San Diego, California, USA
f.penna@samsung.com

***Abstract*—Two particular semi-blind spectrum sensing algorithms are taken into account in this paper: Energy Detection (ED) and Roy's Largest Root Test (RLRT). Both algorithms require the knowledge of the noise power in order to achieve optimal performance. Since by its nature the noise power is unpredictable, noise variance estimation is needed in order to cope with the absence of prior knowledge of the noise power: this leads to a new hybrid approach for both considered detectors. Probability of detection and false alarm with this new approach are derived in closed-form expressions. The impact of noise estimation accuracy for ED and RLRT is evaluated in terms of Receiver Operating Characteristic (ROC) curves and performance curves, i.e., detection/misdetection probability as a function of the Signal to Noise Ratio (SNR). Analytical results have been confirmed by numerical simulations under a flat-fading channel scenario. It is concluded that both hybrid approaches tend to their ideal cases when a large number of slots is used for noise variance estimation and that the impairment due to noise uncertainty is reduced on RLRT w.r.t. ED.**

Keywords-Cognitive radio; spectrum sensing; hybrid detectors; noise estimation

I. INTRODUCTION

Among the functionality provided by Cognitive Radio [1], Opportunistic Spectrum Access (OSA) is devised as a dynamic method to increase the overall spectrum efficiency by allowing Secondary Users to utilize unused licensed spectrum. For this Cognitive Radio System requires the implementation of a spectrum sensing unit in order to gain awareness of available transmission opportunities. This unit must indicate whether a transmission is taking place in the considered channel. Among several spectrum sensing methods put forward for Cognitive Radio applications, techniques based on eigenvalues of the received covariance matrix evolved as a promising solution for spectrum sensing outperforming classical ED.

Semi-blind spectrum sensing algorithms, i.e., ED and RLRT, are the optimum spectrum sensing techniques in a known noise power level scenario. However, in real systems the detector does not have a prior knowledge of the noise level. In recent years, variation and unpredictability of the precise noise level at the sensing device came as a critical issue, which is also known as noise uncertainty. With the goal of reducing the impact of noise uncertainty on the signal detection performance of ED and Eigenvalue Based Detection (EBD), several research has been proposed including [2], [3], [4] for ED and [7], [8] for EBD. Hybrid spectrum sensing algorithms based on the combination of ED and Feature Detection techniques have been put forward for the reduction of the effect of noise variance uncertainty [5], [6]. Similar hybrid spectrum sensing approach was discussed in [9] using the positive points of ED and Covariance Absolute Value detection methods while Sequeira et al. [10] used Akaike Information Criterion (AIC), Minimum Description Length (MDL) and Rank Order Filtering (ROF) methods for estimation of noise power in presence of signal for energy based sensing. In [7], the importance of accurate noise estimation has been shown for better performance of the EBD algorithms.

This paper presents an idea of auxiliary noise variance estimation and focuses on the performance evaluation of Hybrid Approach of semi-blind detection algorithms, namely ED and RLRT, using the same estimated noise variance. The rest of this paper is organized as follows: the system model is developed in Section II, test statistics of detection algorithms is noted in Section III, noise estimation approaches in relation to ED and RLRT are discussed in Section IV, Hybrid Energy Detection and Hybrid Roy's Largest Root Test schemes based on the noise estimation approaches are discussed in Section V and Section VI respectively, the simulation results and the effect of noise variance estimation on considered detection algorithms are discussed in Section VII and finally, Section VIII concludes the paper.

II. SYSTEM MODEL

We consider K sensors (receivers or antennas) for the ED / EBD detector, which senses and decides the presence or absence of the single primary signal within a defined spectrum band $\mathbf{W}$. In a given sensing time interval T, the detector calculates its detection statistic T_D by collecting N samples from each one of the K sensors. The received samples are stored by the detector in the $K \times N$ matrix $\boldsymbol{Y}$.

Let us introduce the $1 \times N$ signal matrix $\boldsymbol{S} \triangleq [s(1) \cdots s(n) \cdots s(N)]$ and the $K \times N$ noise matrix $\boldsymbol{V} \triangleq [\boldsymbol{v}(1) \cdots \boldsymbol{v}(n) \cdots \boldsymbol{v}(N)]$ where,

- $s(n)$ is the transmitted signal sample at time n, modeled as Gaussian with zero mean and variance σ_s^2 : $s(n) \sim$

 ISBN: 978-1-61208-323-0

$\mathcal{N}_{\mathbb{C}}(0, \sigma_s^2)$

- $\boldsymbol{v}(n)$ is a noise vector at time n, modeled as Gaussian with mean zero and variance σ_v^2 : $\boldsymbol{v}(n) \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}_{K\times 1}, \sigma_v^2 \boldsymbol{I}_{K\times K})$

As all the signal samples $s(n)$ of $\boldsymbol{S}$ and the noise vectors $\boldsymbol{v}(n)$ of $\boldsymbol{V}$ are assumed statistically independent, the detector must distinguish between Null and Alternate Hypothesis given by $\boldsymbol{Y}|_{H_0} = \boldsymbol{V}$ and $\boldsymbol{Y}|_{H_1} = \boldsymbol{hS} + \boldsymbol{V}$ where, $\boldsymbol{h}$ is the complex channel vector $\boldsymbol{h} = [h_1 \cdots h_K]^T$ assumed to be constant and memory-less during the sampling window.

Under $\mathcal{H}_1$, the average SNR at the receiver is defined as,

$$\rho \triangleq \frac{\mathcal{E}\|\boldsymbol{x}(n)\|^2}{\mathcal{E}\|\boldsymbol{v}(n)\|^2} = \frac{\sigma_s^2 \|\boldsymbol{h}\|^2}{K\sigma_v^2} \tag{1}$$

where, $||.||$ denotes the Euclidean norm and $\mathcal{E}$ the mean operator. The sample covariance matrix is given by

$$\boldsymbol{R} \triangleq \frac{1}{N} \boldsymbol{Y}\boldsymbol{Y}^H \tag{2}$$

and $\lambda_1 \geq \cdots \geq \lambda_K$ its eigenvalues sorted in decreasing order.

III. Test statistics

The test statistic of ED and RLRT algorithms based on the scenario developed in Section II can be noted in the following subsections.

A. Energy Detection (ED)

ED computes the average energy of the received signal matrix $\boldsymbol{Y}$ normalized by the noise variance σ_v^2 and compares it against a predefined threshold t_{ed}.

$$T_{ED} = \frac{1}{KN\sigma_v^2} \sum_{k=1}^{K} \sum_{n=1}^{N} |y_k(n)|^2. \tag{3}$$

If $T_{ED} < t_{ed}$ it decides in favor of Null Hypothesis $\mathcal{H}_0$ otherwise in favor of Alternate Hypothesis $\mathcal{H}_1$. The detection probability $P_d = Prob\{T_{ED} > t_{ed}|_{H_1}\}$ and false alarm $P_{fa} = Prob\{T_{ED} > t_{ed}|_{H_0}\}$ probabilities for this detector are well-known in the literature (e.g., [11]).

B. Roy's Largest Root Test

Using the information of the received signal matrix $\boldsymbol{Y}$ and assuming a perfect knowledge of the noise variance σ_v^2 and the channel parameter $\boldsymbol{h}$, test statistic for RLRT [12] is given by

$$T_{RLRT} = \frac{\lambda_1}{\sigma_v^2}. \tag{4}$$

RLRT is the optimum test algorithm under the "semi-blind" class of EBD algorithms, which is considered as the reference test in this class whose Detection and False Alarm Probability could be noted as [7],

$$P_{fa} = 1 - F_{TW2}\left(\frac{t_{rlrt} - \mu}{\xi}\right) \qquad P_d = Q\left(\frac{t_{rlrt} - \mu_x}{\sigma_x}\right) \tag{5}$$

where, $F_{TW2}(.)$ is the CDF of Tracy Widom Distribution of order 2. μ and ξ are centering and scaling parameter of a Tracy Widom Distribution given by,

$$\mu = \left[\left(\frac{K}{N}\right)^{\frac{1}{2}} + 1\right]^2 \tag{6}$$

$$\xi = N^{-2/3} \left[\left(\frac{K}{N}\right)^{\frac{1}{2}} + 1\right] \left[\left(\frac{K}{N}\right)^{-\frac{1}{2}} + 1\right]^{1/3} \tag{7}$$

and finally, μ_x and σ_x^2 are mean and variance parameters of a Normal Distribution given by expression,

$$\mu_x = (1 + K\rho)\left(1 + \frac{K-1}{NK\rho}\right) \tag{8}$$

$$\sigma_x^2 = \frac{1}{N}(K\rho + 1)^2 \left(1 - \frac{K-1}{NK^2\rho^2}\right) \tag{9}$$

IV. Noise Estimation

It is evident that the knowledge of the noise power is imperative for the optimum performance of both ED and RLRT. Unfortunately, the variation and the unpredictability of noise power is unavoidable. Thus, the knowledge of the noise power is one of the critical limitations especially of semi-blind detection algorithms for their operation in low SNR.

A. Offline noise estimation: Hybrid approach 1

In the first type of hybrid approaches (HED1 and HRLRT1), noise variance is estimated from S auxiliary noise-only slots in which we are sure that the primary signal is absent.

Consider a sampling window of length M prior and adjacent to the detection window containing noise-only samples for sure. Then, the estimated noise variance from the noise-only samples using a Maximum Likelihood noise power estimation can be written as,

$$\hat{\sigma}_{v1}^2 = \frac{1}{KM} \sum_{k=1}^{K} \sum_{m=1}^{M} |v_k(m)|^2 \tag{10}$$

If the noise variance is constant, the estimation can be averaged over S successive noise-only slots and (10) can be modified by averaging over S successive noise-only slots as,

$$\hat{\sigma}_{v_1}^2(S) = \frac{1}{KSM} \sum_{s=1}^{S} \sum_{k=1}^{K} \sum_{m=1}^{M} |v_k(m)|^2 \tag{11}$$

A possible scheme of RLRT/ED detection algorithm using offline noise estimation approach is shown in Fig. 1: t_{tot} represents a periodic time interval divided into a training phase (noise estimation) and a runtime phase (detection). The runtime interval can be much longer than the training one, however the noise estimation needs to be updated after t_{tot}.

B. Online noise estimation: Hybrid approach 2

In a real time scenario, it is difficult to guarantee the availability of signal free samples so as to estimate the noise variance. Some literature analyzed the performance of ED using estimated noise variance setting aside a separate frequency channel for the measurement of the noise power [13]. However, it is not always suitable to assume uniformly distributed noise in all the frequency bands of concern.

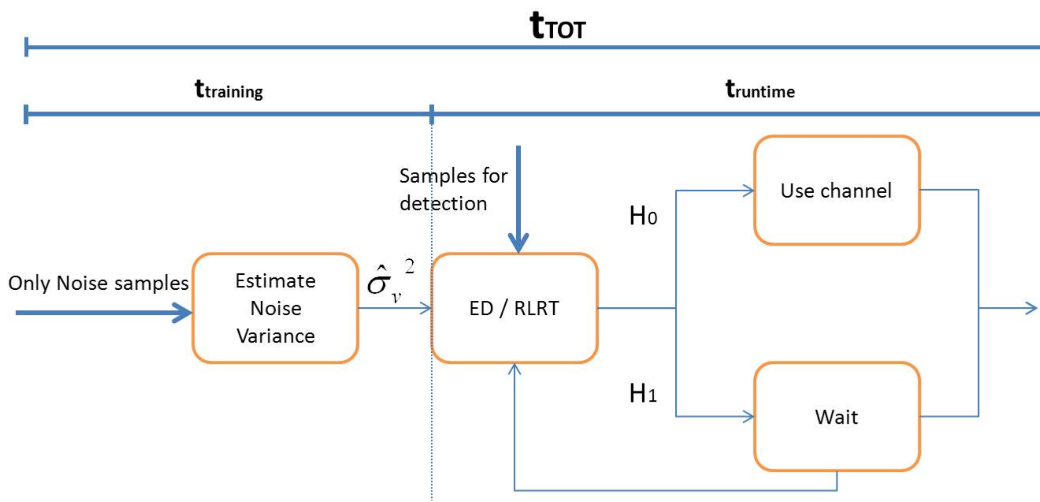


Fig. 1. HED1 / HRLRT1 with offline noise estimation approach

The second hybrid approach does not resort to the existence of auxiliary noise-only slots, but estimates the noise variance information from the previous slots declared as $\mathcal{H}_0$ by the algorithm. Now, the noise variance estimated from those S auxiliary noise-only slots (previously declared $\mathcal{H}_0$) is used in the following detection interval to get the decision about the presence or absence of the primary signal.

Given P_S the probability of receiving primary signal plus noise, P_d is probability of detection, and S is the number of slots, the Maximum Likelihood noise variance estimate $\hat{\sigma}^2_{v_2}(S)$ using M received signal samples declared noise samples by the detector from K receivers is given by,

$$\frac{\left[\sum_{s=1}^{S_s}\sum_{k=1}^{K}\sum_{m=1}^{M}|h_k s(m)+v(m)|^2+\sum_{s=1}^{S_N}\sum_{k=1}^{K}\sum_{m=1}^{M}|v_k(m)|^2\right]}{KMS} \tag{12}$$

where, $S_S = SP_S(1-P_d)$ is the number of primary signal slots missed by the detector and $S_N = S - S_S$ is the number of noise samples successfully detected.

Fig. 2 shows a possible scheme of RLRT/ED detection algorithm using online noise estimation approach; after a transient stage (offline noise estimation), the detector automatically updates the noise estimation after S slots declared $\mathcal{H}_0$ (sliding window). Unlike the first approach, no further training offline phases are required.

V. Hybrid Energy Detection

Incorporating the *offline noise estimation* and *online noise estimation* described in Section IV in ED, hybrid approaches of ED are developed and their performance parameters are derived in the following subsections.

A. Hybrid ED approach 1 (HED1)

The Energy Detection Test Statistic in (3) can be modified to HED1 test statistic using (11) as,

$$T_{HED1} = \frac{1}{KN\hat{\sigma}^2_{v_1}(S)}\sum_{k=1}^{K}\sum_{n=1}^{N}|y_k(n)|^2 \tag{13}$$

Moreover, (13) can be considered as the parametric likelihood ratio test when the signal to be detected is assumed to be Gaussian with zero mean and variance σ_s^2.

Under Null Hypothesis, after rigorous simplification, the test statistic in (13) could be approximated with a Normal Random Variable whose Probability of False Alarm P_{fa}^{HED1} for number of sensors K, number of samples N, number of auxiliary slots S and threshold t_{hed1} is given by,

$$P_{fa}^{HED1} = Q\left[\frac{t_{hed1}-1}{\sqrt{\frac{MS+Nt_{hed1}^2}{KMNS}}}\right] \tag{14}$$

Similarly, under Alternate Hypothesis, the test statistic in (13) also approximates to Normal Random Variable with different mean and variance parameters whose Probability of Detection P_d^{HED1} could be written as,

$$P_d^{HED1} = Q\left[\frac{(t_{hed1}-1-\rho)}{\sqrt{\frac{t_{hed1}^2}{KMS}+\frac{K\rho^2+2\rho+1}{KN}}}\right] \tag{15}$$

B. Hybrid ED approach 2 (HED2)

Using (12), decision statistic of HED2 can be written as,

$$T_{HED2} = \frac{1}{KN\hat{\sigma}^2_{v_2}(S)}\sum_{k=1}^{K}\sum_{n=1}^{N}|y_k(n)|^2 \tag{16}$$

Under Null Hypothesis, after rigorous simplification, the test statistic in (16) could be approximated with a Normal Random Variable whose False Alarm Probability P_{fa}^{HED2} for number of sensors K, number of samples N, number of auxiliary slots S for noise estimation using (12) and threshold t_{hed2} is given by,

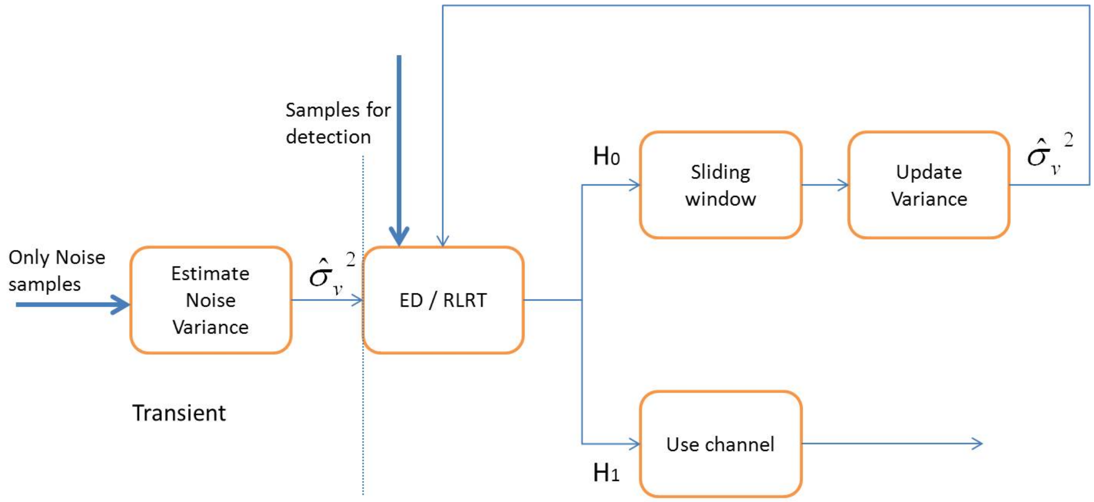


Fig. 2. HED2 / HRLRT2 with online noise estimation approach

$$P_{fa}^{HED2} = Q\left[\frac{t_{hed2} - \dfrac{S}{S+\rho S_S}}{\sqrt{\dfrac{t_{hed2}^2 NC + MS^2}{KMN(S+\rho S_S)^2}}}\right] \quad (17)$$

where, $C = (S_S K \rho^2 + \rho S_S + S)$.

Similarly, under Alternate Hypothesis, the test statistic in (16) also approximates to Normal Random Variable with different mean and variance parameters whose Probability of Detection P_d^{HED2} in a similar scenario could be written as,

$$P_d^{HED2} = Q\left[\frac{t_{hed2} - \dfrac{S(\rho+1)}{S+\rho S_S}}{\sqrt{\dfrac{t_{hed2}^2 NC + MS^2(K\rho^2+2\rho+1)}{KMN(S+\rho S_S)^2}}}\right] \quad (18)$$

VI. Hybrid Roy's Largest Root Test

In a similar way as for ED, if we incorporate *offline noise estimation* and *online noise estimation* in RLRT, hybrid approaches of RLRT are developed and their performance parameters are derived in the following subsections.

A. Hybrid RLRT approach 1 (HRLRT1)

HRLRT1 is a similar approach as HED1, which deals with the study of detection performance of the RLRT algorithm using estimated noise variance. Noise variance is estimated from S auxiliary noise-only slots where we are sure that the primary signal is absent. Using the ML estimate of the noise variance (11), the decision statistic of HRLRT1 can be expressed as,

$$T_{HRLRT1} = \frac{\lambda_1}{\hat{\sigma}_{v_1}^2(S)} \quad (19)$$

Under Null Hypothesis, after rigorous simplification, the test statistic in (19) could be approximated to the ratio of a Tracy Widom Random Variable of order 2 and a Normal Random Variable. Hence, the False Alarm Probability P_{fa}^{HRLRT1} for number of sensors K, number of samples N, number of auxiliary slots S for noise estimation using (11) and threshold t_1 is given by,

$$P_{fa}^{HRLRT1} = 1 - F_0^{H1}(t_1) \quad (20)$$

where $F_0^{H1}(t_1)$ is the Cumulative Density Function *CDF* of the Probability Density Function shown below,

$$f_0^{H1}(t_1) = C_1 \int_{-\infty}^{+\infty} |x| f_{TW2}\left(\frac{xt_1 - \mu}{\xi}\right) e^{\frac{-D(x-1)^2}{4}} \, dx \quad (21)$$

with $f_{TW2}(.)$ being the *pdf* of Tracy Widom Distribution and $C_1 = \frac{1}{2\xi}\sqrt{\frac{D}{\pi}}$.

Similarly, under Alternate Hypothesis, the test statistic in (19) approximates to Normal Random Variable whose Probability of Detection P_d^{HRLRT1} under a similar scenario is given by,

$$P_d^{HRLRT1} = Q\left(\frac{t_1 - \mu_x}{\sqrt{\frac{2t_1^2}{D} + \sigma_x^2}}\right) \quad (22)$$

where, μ_x (8) and σ_x^2 (9) are mean and variance of a Normal Random Variable.

B. Hybrid RLRT approach 2 (HRLRT2)

HRLRT2 is an alternate hybrid approach of RLRT where noise variance given by (12) is estimated from the previously received signal slots declared as $\mathcal{H}_0$ by the algorithm. The decision statistic of HRLRT2 can be written as,

$$T_{HRLRT2} = \frac{\lambda_1}{\hat{\sigma}_{v_2}^2(S)} \quad (23)$$

Under Null Hypothesis, after rigorous simplification, the test statistic in (23) could be approximated to the ratio of a Tracy

Widom Random Variable of order 2 and a Normal Random Variable. Hence, the False Alarm Probability P_{fa}^{HRLRT2} for number of sensors K, number of samples N, number of auxiliary slots S for noise estimation using (12) and threshold t_2 is given by,

$$P_{fa}^{HRLRT2} = 1 - F_0^{H2}(t_2) \quad (24)$$

where, $F_0^{H2}(t_2)$ is the Cumulative Density Function *CDF* of the Probability Density Functions shown below,

$$f_0^{H2}(t_2) = C_2 \int_{-\infty}^{+\infty} |x| f_{TW2}\left(\frac{xt_2 - \mu}{\xi}\right) e^{\frac{-(x-\mu_1)^2}{2\sigma_1^2}} \mathrm{d}x \quad (25)$$

with $C_2 = \frac{1}{\xi\sigma_1^2\sqrt{2\pi}}$.

Similarly, under Alternate Hypothesis, the test statistic in (23) approximates to Normal Random Variable whose Probability of Detection P_d^{HRLRT2} under a similar scenario is given by,

$$P_d^{HRLRT2} = Q\left(\frac{t_2 - \mu_x/\mu_1}{\sqrt{\frac{t_2^2\sigma_1^2 + \sigma_x^2}{\mu_1^2}}}\right) \quad (26)$$

where, μ_x (8), μ_1 (27) and σ_x^2 (9), σ_1^2 (28) are mean and variance parameters with,

$$\mu_1 = \frac{S + S_S}{S} \quad (27)$$

$$\sigma_1^2 = \frac{S + 2\rho S_S + \rho^2 K S_S}{KMS^2} \quad (28)$$

VII. Simulation Results

This section shows the simulation of the ROC curves and performance curves of hybrid approaches of ED and RLRT spectrum sensing algorithms. The accuracy of the the closed-form expressions is confirmed by the results presented in Fig. 3 and Fig. 4, respectively, where the theoretical formulas are compared against the simulated detection performance over S auxiliary noise-only slots (S ranges from 1 to 8). Perfect match of the theoretical and the numerical curve validates the considered model. As it can be noticed, with the increase in the number of auxiliary slots used for the estimation of the noise variance, the probability of detection increases for both hybrid approaches.

Fig. 5 illustrates the comparison of ED, HED1 and HED2 performance as a function of the SNR. Performance of HED1 and HED2 varies typically around 0 dB SNR but no visible difference can be noted in extreme high or low SNR values. Since there is a chance of mis-interpretation of noise plus primary signal as only-noise samples (used to estimate the noise variance) by ED in case of HED2, performance of HED2 is slightly lower than HED1 near 0 dB of SNR. By increasing the number of slots used for the estimation of the noise variance, the gap between HED1 and HED2 decreases and both approaches approximate the known-variance ED curve.

The convergence of the hybrid approach of RLRT to an ideal RLRT (known variance) is illustrated in Fig. 6. By increasing

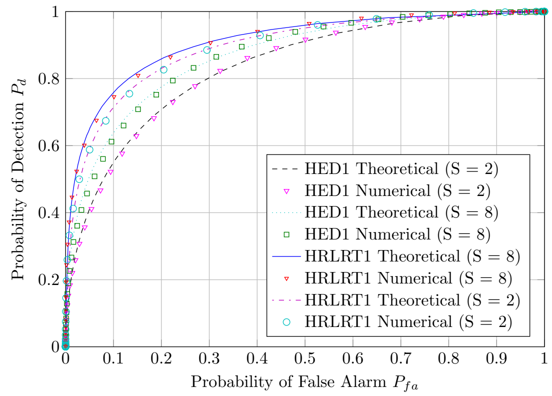


Fig. 3. Theoretical and numerical ROC plot of hybrid approach 1 of ED/RLRT. Parameters: $N = 80, M = 80, K = 4$ and $SNR = -10dB$

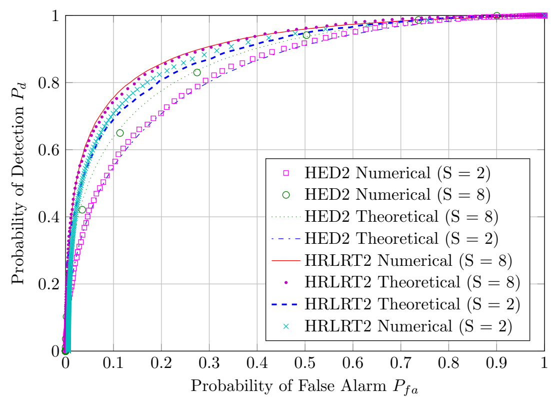


Fig. 4. Theoretical and numerical ROC plot of hybrid approach 2 of ED/RLRT. Parameters: $N = 80, M = 80, K = 4$ and $SNR = -10dB$

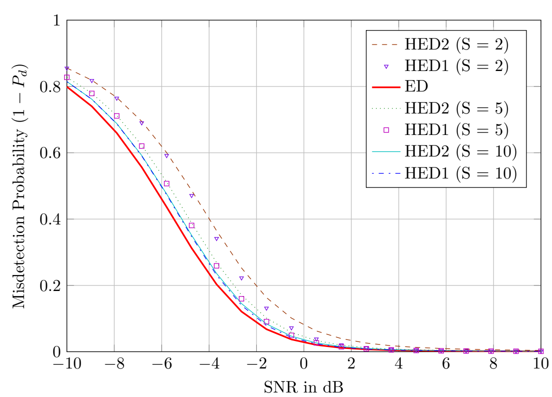


Fig. 5. Performance curves of ED and its hybrid approaches. Parameters: $N = 10, M = 10, K = 5$ and $P_{fa} = 0.05$

the number of auxiliary slots used for the estimation of noise variance, the performance of HRLRT1 and HRLRT2 converge at the ideal RLRT performance.

The performance of HED1 and HRLRT1 is compared in Fig. 7. The noise variance is estimated using (11) from S auxiliary sure noise-only slots. The curves approach the ideal ED and RLRT curves by increasing the number of auxiliary slots S, but the rate of convergence of HED1 is slower.

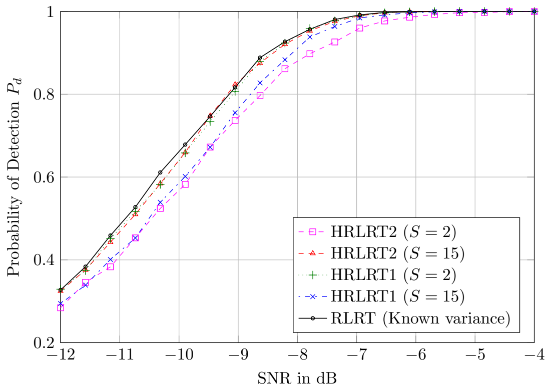


Fig. 6. Performance curves of RLRT and its hybrid approaches. Parameters: $N = 80, M = 80, K = 4$ and $P_{fa} = 0.05$

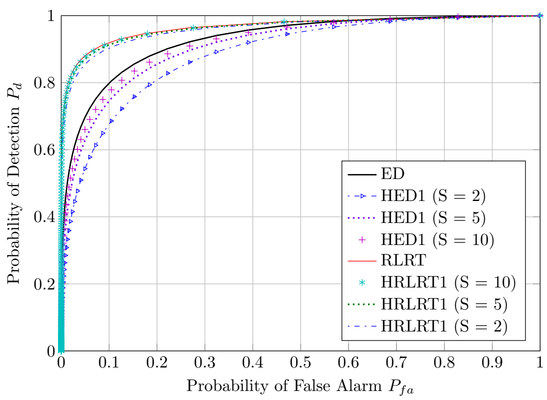


Fig. 7. ROC curves of ED/RLRT and their hybrid approach 1 (HED1/HRLRT1). Parameters: $N = 80, M = 80, K = 4$ and $SNR = -10dB$

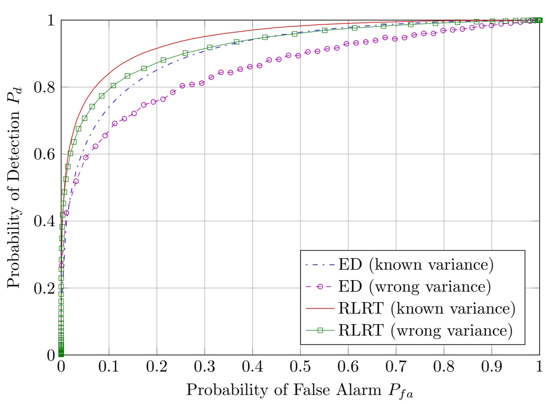


Fig. 8. Effect of Noise Variance fluctuation on ED and RLRT. Parameters: $N = 100$, $K = 4$, $var(\hat{\sigma}_v^2) = 0.0032(-25dB)$ given nominal noise variance $\sigma_v^2 = 1$.

The effect of the noise variance estimation uncertainty on ED and RLRT algorithms is considered in Fig. 8. Assuming the Gaussian distribution of the noise variance estimate with mean equal to nominal value, the ROC for ED and RLRT is plotted, setting $var(\hat{\sigma}_v^2) = 0.0032(-25dB)$. The result shows that, for the same uncertainty of the noise variance estimate, the performance gap between the ideal curve and the curve with wrong variance is larger for ED as compared to RLRT. Thus, it can be easily noticed that RLRT is more robust to noise variance uncertainty as compared to ED algorithm.

VIII. CONCLUSION

In this paper, the analysis of two semi-blind spectrum sensing algorithms, ED and RLRT, is extended to hybrid approaches. Analytical expressions for the performance parameters, P_d and P_{fa}, are derived for each algorithm. Analytical results are verified by Monte Carlo simulation and by numerical methods. In addition, the impact of noise variance estimation on ED and RLRT was carried out based on ROC curves. The results showed that the fluctuation of noise variance estimate from nominal value is severe in case of small number of auxiliary slots used for the estimation of the noise variance. Moreover, for the same uncertainty on the noise variance estimate, the performance gap between the ideal curve and the curve with wrong variance is larger for ED as compared to RLRT.

REFERENCES

[1] J. Mitola, III, Cognitive radio, "An integrated agent architecture for software defined radio", Ph.D. dissertation, Royal Institute of Technology (KTH), May 2000.

[2] R. Tandra and A. Sahai, "SNR Walls for Signal Detection," Selected Topics in Signal Processing, IEEE Journal of, Feb. 2008, vol. 2, no. 1, pp. 4-17.

[3] A. Mariani, A. Giorgetti, and M. Chiani, "Effects of Noise Power Estimation on Energy Detection for Cognitive Radio Applications," Communications, IEEE Transactions on, Dec. 2011, vol. 59, no. 12, pp. 3410-3420.

[4] A. Mariani, A. Giorgetti, and M. Chiani, "SNR Wall for Energy Detection with Noise Power Estimation," Communications (ICC), 2011 IEEE International Conference on, 5-9 Jun. 2011, pp. 1-6.

[5] Z. Khalaf, A. Nafkha, and J. Palicot, "Enhanced hybrid spectrum sensing architecture for cognitive radio equipment," General Assembly and Scientific Symposium, 2011 XXXth URSI, 13-20 Aug. 2011, pp. 1-4.

[6] F. Moghimi, R. Schober, and R.K. Mallik, "Hybrid Coherent/Energy Detection for Cognitive Radio Networks," Wireless Communications, IEEE Transactions on, May 2011, vol. 10, no. 5, pp. 1594-1605.

[7] B. Nadler, F. Penna, and R. Garello, "Performance of Eigenvalue-Based Signal Detectors with Known and Unknown Noise Level," Communications (ICC), 2011 IEEE International Conference on, 5-9 Jun. 2011, pp. 1-5.

[8] F. Penna, R. Garello, and M. Spirito, "Probability of Missed Detection in Eigenvalue Ratio Spectrum Sensing," Wireless and Mobile Computing, Networking and Communications, 2009. WIMOB 2009. IEEE International Conference on, 12-14 Oct. 2009, pp. 117-122.

[9] S. Badrinath and V.U. Reddy, "A hybrid energy detection approach to spectrum sensing," Cognitive Wireless Systems (UKIWCWS), 2009 First UK-India International Workshop on, 10-12 Dec. 2009, pp. 1-6.

[10] S. Sequeira, R.R. Mahajan, and P. Spasojevic, "On the noise power estimation in the presence of the signal for energy-based sensing," Sarnoff Symposium (SARNOFF), 2012 35th IEEE, 21-22 May 2012, pp. 1-5.

[11] H. Urkowitz, "Energy detection of unknown deterministic signals," Proceedings of the IEEE, Apr. 1967, pp. 523-531.

[12] S. N. Roy, "On a Heuristic Method of Test Construction and its use in Multivariate Analysis," The Annals of Mathematical Statistics, Jan. 1953, vol. 24, no. 2, pp. 220-238.

[13] Y. Zhuan, G. Memik, and J. Grosspietsch, "Energy Detection Using Estimated Noise Variance for Spectrum Sensing in Cognitive Radio Networks", Wireless Communications and Networking Conference, 2008. WCNC 2008. IEEE, Apr. 2008, pp. 711-716.

Software-Defined Radio for Spectrum Sensing Using Independent Component Analysis

Paulo Ixtânio Leite Ferreira
Post-Grad. Prog. in Elec. Engineering, PPgEE – COPELE
Federal University of Campina Grande, UFCG
Campina Grande, Paraíba, Brazil
paulo.ferreira@ee.ufcg.edu.br

G. Fontgalland, Bruno B. Albert, and Edmar C. Gurjão
Department of Electrical Engineering, DEE
Federal University of Campina Grande, UFCG
Campina Grande, Paraíba, Brazil
{fontgalland, albert, ecandeia}@dee.ufcg.edu.br

***Abstract*—With the increment in data rates and new wireless services, the necessity for more radio frequency (RF) bandwidth for data transmission increases. Frequency spectrum is a scarce resource, and there is a necessity to optimize its use; in cognitive radio, spectrum sensing is the first task toward this optimization. It consists of detecting unused spectrum portions, allowing its use by a secondary user. In this paper, we apply the blind source separation method called Independent Component Analysis (ICA) for spectrum sensing, by detecting the presence and bandwidth of transmitted signals and by a complementary analysis to obtain non-used spectrum portions. A measurement setup with three uncorrelated sources and predefined bands is proposed. A software-defined radio implementing three independent transmitters and a broadband receiver antenna used to capture all signals. The results show the feasibility of using ICA for spectrum sensing.**

Keywords-Spectrum sensing; ICA; blind sources; software-defined radio; cognitive radio.

I. INTRODUCTION

Cognitive radios are communication systems with capabilities of monitoring the environment, to analyze the obtained parameters, to decide about possible adjusts, and to adapt itself according to this decision [1]. In this scenario, cognitive radio can adjust the transmission frequency to an unused frequency band to optimize the spectrum use. Spectrum sensing is the determination of empty frequency bands, realized in two steps: sensing the channels and using the obtained information to decide about what channels are empty.

Spectrum sensing methods are in three broad categories, namely energy detection, stochastic methods, and analysis of signal characteristics [2], [3]. Energy detection methods due their simplicity have low performance, mainly in a noisy environment. The stochastic methods are implemented by analyzing some statistical characteristics of the signal. These methods have good performance, even in noisy environment, but they have high computational complexity. Whereas signal features analysis is also very accurate, it implies in a previous knowledge of the signal characteristics, what may be difficult to detect in transmission of other secondary users, using their own systems.

Independent component analysis (ICA) is a blind method for source separation that has been very efficient in various scenarios. It has been applied in communications, image processing, audio separation [4], determination of direction-of-arrival (DoA) [5]. In this work, ICA is applied to identify signal sources in a broad band of frequencies and to use this information for spectrum sensing. Besides its low computation complexity [4], the ICA algorithm has good performance in presence of low signal-to-noise ratio (SNR), and it does not require prior information about the sources that occupy the monitored spectrum [6].

The spectrum sensing by ICA was implemented in a software-defined radio. An experimental setup was used to validate the proposed method.

The paper is organized as follows: the basic principle of ICA method is described in Section II. The software-defined radio is presented in Section III. We describe the experiment setup in Section IV. The results and some comments about them are shown in Section V. Finally, the conclusions are drawn in Section VI.

II. INDEPENDENT COMPONENT ANALYSIS

The ICA method uses statistical assumptions for blind source separation, and it allows recovering statistically independent signals from compositions of this signal, called mixture signals [7]. A linear model relates independent signals and mixed signals. This is model let us consider a vector of n signals $\mathbf{s} = [s_1, s_2, \ldots, s_n]^T$, and a vector of m measured signals $\mathbf{x} = [x_1, x_2, \ldots, x_m]^T$, where each signal x_i ($i = 1, \ldots, m$) is a linear combination of the n source signals. We consider $n < m$ to obtain the best ICA performance, as in [7].

Fig. 1 represents the basic principles of ICA, where the measurement vector **x** is formed by combining elements of vector sources **s,** via a matrix **A** (**x** = **As**). Even **s** and **A** are unknown, ICA can find a separation matrix **W**, such that the output vector **y** (**y** = **Wx**) is the optimal approximation of **s**.

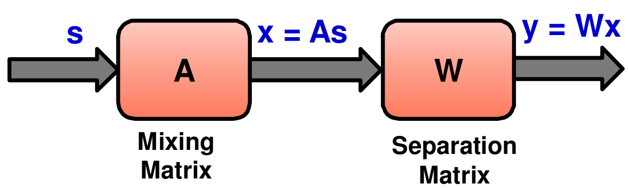


Figure 1. The ICA block diagram.

In this work, we use the FastICA algorithm. Fast ICA is an iterative fixed-point algorithm that minimize the mutual information of the estimated components from contrast function ($g = y^3$, for example) [9]. Assuming a whitened data vector and $\mathbf{w}^T$ is one of the rows of the separating matrix $\mathbf{W}$ (Fig. 1). Estimation of $\mathbf{w}^T$ is done iteratively optimizing the nongaussianity of the contrast functions and symmetric orthogonalization of $\mathbf{W}$ until a convergence is achieved. The convergence means that the old and new value of $\mathbf{w}$ point in the same direction, i.e. their dot-product is (almost) equal to 1[3].

III. Software-Defined Radio

Software-Defined Radio (SDR) proposed by Joseph Mitola III [10] as the implementation of flexible and reconfigurable radio based on software. Compared to traditional hardware implementations, SDR gives the possibility of implementing various radios on the same hardware, or changing the configuration by adjusting the software parameters. In addition, with the increase of processing power it becomes possible the use of sophisticated signal processing in the implemented radios.

GNU Radio [11] is an open source framework for development of SDR. Each SDR in GNU Radio is composed by a set of independent interconnected signal processing blocks, obtained from the built-in library or created by the user. The SDR developed using GNU Radio can run in a general-purpose processor, as a personal computer, and using a Radio Frequency (RF) interface, it is possible to transmit or to receive real signals.

The Universal Software Radio Peripheral (USRP) [12] is a RF front-end composed by a motherboard and a set of daughterboard. In the motherboard, there are analog-to-digital converters in the reception path (antenna to computer), digital-to-analog converters in the transmission path (computer to antenna), and a Field-Programmable Gate Array (FPGA) to multiplex the data from the reception daughterboard to computer and vice-versa.

The daughterboard performs the down conversion (reception) or up conversion (transmission). Each daughterboard is projected to a range of frequencies, and in a typical configuration have four daughterboard: two to transmission and two to reception.

IV. Experimental Setup

To validate the use of ICA for spectrum sensing, an experimental setup is implemented, where three SDRs, as a signal sources, and another one as receiver of the mixed signals. Each source is considered as primary user (channel owner) to transmit a signal occupying a bandwidth in a specific frequency. Care was taken for signal bandwidths to not overlap. Transmitting antennas were arranged side-by-side, and they are positioned $d = \lambda_0/2$ (λ_0 is wavelength in free space at 1.252 GHz) a part. The SDR receiver is composed by an antenna and a spectrum analyzer, it receives the mixture signal. Once captured, the data are recorded and used in an offline application of ICA.

Fig. 2 shows the measurement setup with three primary users, and up to four measurements positions. The distances (l_i, i = 1, 2, 3, 4) between the transmitters and receiver antennas can be chosen randomly, assured the far field condition. For obtained results, the chosen distances are: l_1 = 2.0 m, l_2 = 2.06 m, l_3 = 2.06 m, and l_4 = 2.24 m.

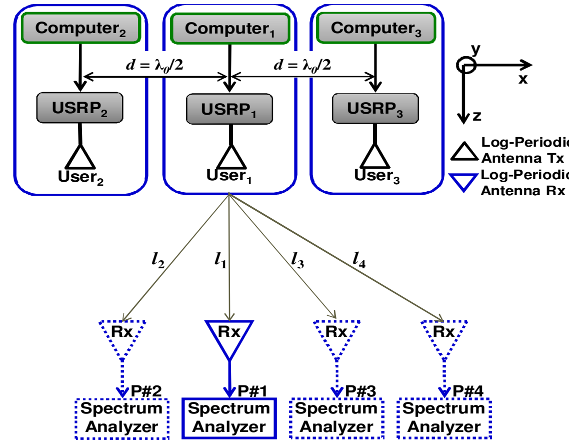


Figure 2. Measurement setup with three sources implemented in SDR.

As depicted in Fig. 2, each signal source is composed by one USRP connected to GNU Radio (computer). The SDR permits to configure the transmitted signal parameters like bandwidth, transmission frequency and amplitude. Dotted parts in Fig. 2 indicate the measurement points P#1, P#2, P#3 and P#4, sequentially.

Compared to other available setups, namely the ones that use signal generator as transmitters, the SDR provides more flexibility. In the former, waveforms are pre-programmed, and in the later the user has the capability of to configure the waveforms according to the application, e.g., by choosing the probability distribution of the signals, or any other parameter of interest.

A. *Description of Measurement Setup*

In the setup, each transmitter uses a log-periodic antenna; model WA5JVB, 0.9 – 2.6 GHz, and in the receiver was used a high gain broadband antenna (R_X) (log-periodic A.H Systems, SAS 510-7, 0.29 – 7.0 GHz). USRP and computer are connected via a Universal Serial Bus (USB) cable; and USRPs and antenna by coaxial cables of 1.10 m long. Measurements in positions P#1, P#2, P#3, and P#4 were done using a spectrum analyzer, R&S FSL6 (9 kHz – 6 GHz). The R_X positions (P#1 to P#4) were randomly chosen.

B. *Measurement Procedure*

Each primary user transmits a chirp signal generated by the SDR with 6 MHz bandwidth. Primary user frequency transmissions are $User_1$ = 1.240 GHz, $User_2$ = 1.252 GHz, and $User_3$ = 1.264 GHz. The receiver was configured to operate from 1.235 GHz to 1.270 GHz, with RBW 30 kHz , VBW 100 kHz, trace mode: max hold, and sweep point: 10,000.

The system was initially calibrated for each user individually. Firstly, the $User_2$ and $User_3$ are OFF, and measurement taken from $User_1$ as a single canal; secondly, by setting $User_1$ and $User_3$ in OFF, $User_2$ is measured; and finally, by setting $User_1$ and $User_2$ in OFF, the $User_3$ is measured.

V. Results and Discussion

the spectra of primary users are show in Fig. 3. They were obtained for measurements in position P#1. These measurements were done for comparison purposes with the recovered spectrum by ICA.

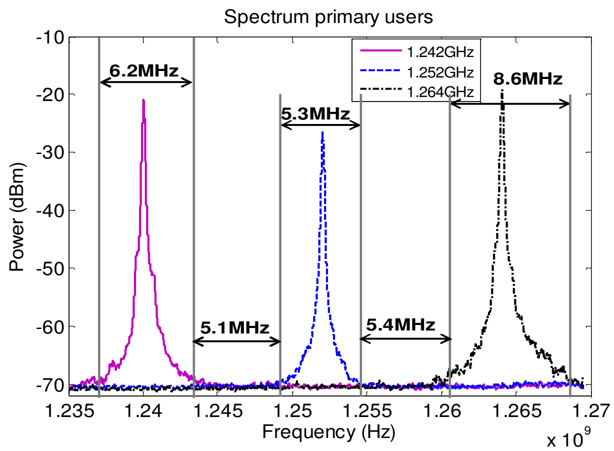


Figure 3. Spectrum original of the users, measured in position P#1.

In Fig. 3, it can be observed that the bandwidth between $User_1$ and $User_2$ is 5.1 MHz, and between users $User_2$ and $User_3$ is 5.4 MHz. The main goal in cognitive radio is to detect these empty spaces. Therefore, using ICA to detect the SDR transmission the empty band can be identified. All graphs presented here have undergone a smoothing of 0.005% through the method of moving average.

Details about signals of each user are given in Table I.

TABLE I. Parameters of Users

Primary users	**Parameters**		
	Carrier Frequency (MHz)	***BW (MHz)***	***Maximum Power (dBm)***
$User_1$	1.240	6.2	– 21.01
$User_2$	1.252	5.3	– 26.51
$User_3$	1.264	8.6	– 19.34

After measured each individual signal, the next step is to measure the three users simultaneously ($User_1$, $User_2$ and $User_3$ in ON) transmitting. The transmitters (Users) are positioned near to each other. The measurements were conducted in free space, covering the band from 1.235 to 1.27 GHz (BW = 35 MHz). This bandwidth is six times larger than the one used in [13]. The spectrum of the received signals ($User_1$, $User_2$ and $User_3$) in four different positions is shown in Fig. 4.

At the top of Fig. 4 is shown the spectrum of the received signal that is composed by three users, measured at positions P#1 and P#2 ($User_1$ + $User_2$ + $User_3$); in the bottom it is shown the spectrum measured in positions P#3 and P#4. Note that the spectrum has bandwidths available between the one allocated by primary users. In a practical case, such space must be found by the spectral detection method. In this work ICA was used.

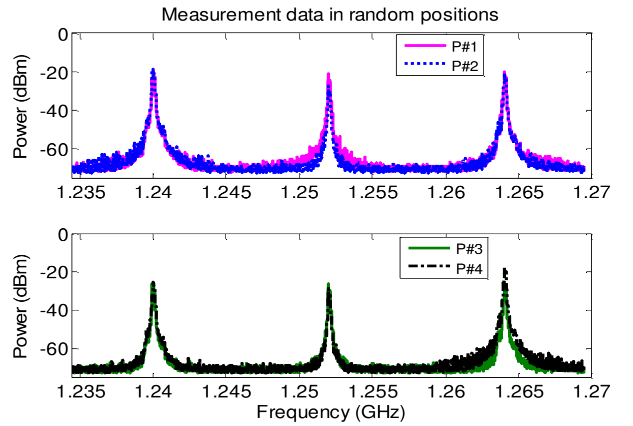


Figure 4. Composition of the spectrum received in positions P#1 to P#4.

The obtained measurements were used as the input of the ICA method. No additional information is needed in the receiver, as in [13] for instance, where it is assumed that the receiver already knows the bandwidth of the channels and the carrier frequency.

ICA was applied to the data measured in positions P#1 to P#4. They are the elements of the **x** vector (see Fig. 1). Initially, only three measured positions of the spectrum presented in Fig. 4 were considered. The results for the estimated spectrum using FastICA algorithm are shown in Fig. 5.

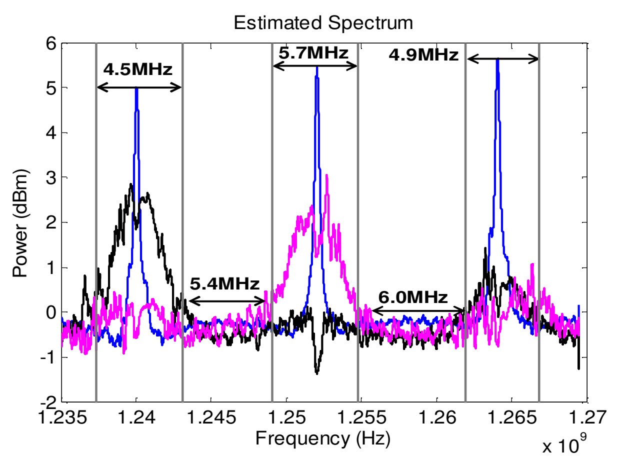


Figure 5. Estimated spectrum by ICA, from data shown in Fig. 4, in positions P#1 to P#3.

The spectrum obtained by ICA (Fig. 5) shows the frequency bands occupied by the primary users (4.5 MHz, 5.7 MHz and 4.9 MHz). The detected empty bandwidths are (2.2431 - 2.2485 GHz, and 2.2556 - 2.2665 GHz). The information about empty bands is the objective of this work, since it indicates where secondary users can be allocated to transmit. To determine the bandwidths, we consider points where the signal reaches 0 dBm. We can observe that the bandwidths are not the same as the original (Fig. 4), but are close to, except for the primary $User_3$. To overcome this problem, another measurement was considered in the

 ISBN: 978-1-61208-323-0

estimation. Fig. 6 shows the spectrum estimated to four measurements in positions P#1 to P#4.

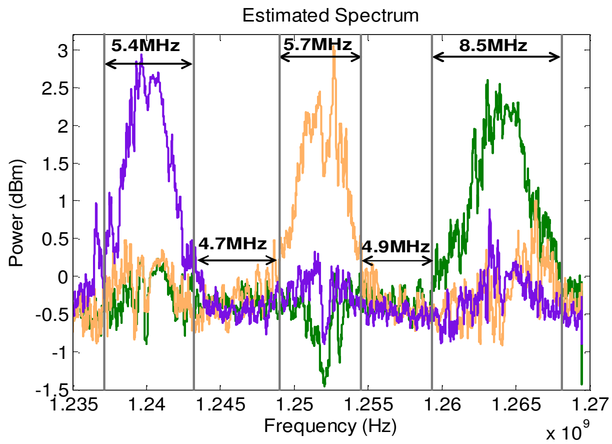


Figure 6. Estimated spectrum by ICA, from data shown in Fig. 4, in positions P#1 to P#4.

In the new estimated spectrum, shown in Fig. 6, it can be observed that the frequency bandwidths occupied by primary users are closer to the original spectrum, particularly for $User_3$. Estimated and occupied bandwidths by primary users are: 5.45 MHz, 5.75 MHz and 8.5 MHz. Once these bands are detected, such channels can be used to identify the empty bands, which are: 2.2438 - 2.2485 GHz (4.7 MHz) and 2.2546 - 2.2595 GHz (4.9 MHz). As expected, one can conclude that with more input data the ICA method was able to identify more precisely the primary users and the free frequency bands, which can be used for cognitive radio applications.

The necessity of more input data to make a better estimation does not compromise the computational cost of the FastICA algorithm. Although the signals used in the experiment do not have a defined bandwidth, the obtained results show that ICA can be applied to spectrum sensing in cognitive radio.

VI. Conclusion and Future Work

In this paper, it was shown the application ICA methods to spectrum sensing. Results showed that ICA has advantages over traditional methods for spectrum detection, since it does not requires prior information about the channels to be detected. This fact allows it to detect the existence of other users that are not regulated (primary users). ICA can also sweep a wide frequency range without compromising the complexity and speed of the algorithm.

As a future work, we propose to compare ICA methods with traditional spectrum sensing methods in same conditions, to analyze the response of ICA in three scenarios: when the user bandwidth is not constant; in the low SNR regime (around 2 dB) that is different from was shown here; and when users are occupy open TV channels.

We also intend to use SDR in the receiver, instead of the spectrum analyzer, for implementing real time spectrum estimation.

Acknowledgment

The authors wish to thank the Applied Microwave and Electromagnetics Laboratory (LEMA), Signal Processing Laboratory (LAPS), Federal University of Campina Grande (UFCG), the Federal Institute of Science and Technology of Paraíba (IFPB) – *Campus* Princesa Isabel, and the Brazilian Agency (CAPES).

References

[1] S. Haykin, “Cognitive radio: brain-empowered wireless communications”. IEEE Journal on Selected Areas on Communications, vol. 23, No 2, February, 2005, pp. 201–220.

[2] S. G. Mohapatra, A. G. Mohapatra, and Dr. S. K. Lenka, “Cyclo-stationarity Based Windowing Method for Spectrum Sensing in Cognitive Radio Networks”, in International Conference Information Communication and Embedded Systems (ICICES 2013), Tamilnadu, India, 2013, pp.603 – 608.

[3] R. S. Babu and M. Suganthi, “Review of energy detection for spectrum sensing in various channels and its performance for cognitive radio applications”, American Journal of Engineering and Applied Sciences (AJEAS), ©2012 Science Publication, 2012, pp. 151 – 156.

[4] A. Hyvärinen and E. Oja, “Independent component analysis: algorithms and applications,” in Neural Network, vol. 13, Iss. 4-5, May – Jun 2000, pp. 411 – 430.

[5] A. K. Khan, T. Onoue, K. Hashiodani, Y. Fukumizu, and H. Yamauchi, “Signal and noise separation in medical diagnostic system based on independent component analysis,” in IEEE Asia Pacific Conf. on Circuits and Systems (APCCAS 2010), Kuala Lumpur, Malaysia, 2010, pp. 812 – 815.

[6] P. I. L. Ferreira, G. Fontgalland, and B. B. Albert, “Spread-out overlapping sources by independent component analysis for location positioning,” in IEEE Transactions on Magnetics, vol. 49, No 5, May 2013, pp.1805 – 1808.

[7] P. I. L. Ferreira, G. Fontgalland, S. E. Barbin, and G. F. Aragão, “Separation of electromagnetic sources by the method of independent component analysis,” in IEEE International Instrumentation and Measurement Technology Conference (I2MTC 2013), May 6-9, 2013, Minneapolis, MN, EUA, pp. 476 – 479.

[8] P. I. L. Ferreira and G. Fontgalland, “Blind separation of multiple electromagnetic radiation sources,” Antenna Measurement Techniques Association, (AMTA 2013), in press.

[9] P. I. L. Ferreira, G. Fontgalland, and B. B. Albert, “Identification of sources with small phase difference by the method of independent component analysis,” IEEE CEFC 2012, Oita, Japan, Nov. 11 – Nov. 14, 2012 [Digests of the 15th Biennial IEEE Conference on Electromagnetic Field Computations].

[10] J. Mitola III, Software Radio Architecture: Object-Oriented Approaches to Wireless Systems Engineering. New York: Third John Wiley & Sons, 1996.

[11] GNU Radio home page, http://gnuradio.org/trac, accessed May 2013.

[12] Universal Software Radio Peripheral website, http://www.ettus.com/, accessed May 2013.

[13] N. Han, S. H. Shon, J. H. Chung, and J. Moung. Kim, “Spectral correlation based signal detection method for spectrum sensing in IEEE 802.22 WRAN systems”, the 8th International Conference Advanced Communication Technology (ICACT 2006), Gangwon-Do, Republic of Korea, 2006, pp. 1765 – 1770.

FBMC/COQAM: An Enabler for Cognitive Radio

Hao Lin and Pierre Siohan
Orange Labs
Cesson Sévigné, France
Email: hao.lin,pierre.siohan@orange.com

***Abstract*—To solve the issue of spectrum scarcity, new paradigms of spectrum access must be investigated. The Cognitive Radio (CR) is an efficient solution to response to the requirement. But it needs the support from physical layer design, especially the signal modulation. In the literature, Filter-Bank-based Multi-Carrier with Offset QAM (FBMC/OQAM) and Generalized Frequency Division Multiplexing (GFDM) were proposed as suitable enablers for CR. However, both of them have some drawbacks that eventually prevent them from a practical usage. In this paper, we present an improved modulation scheme, which successfully combines these two schemes into one named FBMC/Circular OQAM (FBMC/COQAM). With this scheme, we are able to overcome all the drawbacks of its predecessors, while still keeping their advantages.**

***Keywords*-Cognitive Radio; FBMC; GFDM; OQAM.**

I. INTRODUCTION

In the vision towards future radio systems, the spectrum scarcity tends to be an inevitable issue, which urgently calls for some new spectrum access paradigms. The CR as an emerging application effectively addresses the requirement for efficient spectrum usage. Nevertheless, the implementation of CR solutions needs the support from a suitable modulation. Indeed, one important requirement is that the modulation scheme must provide a good spectrum localization. Orthogonal Frequency Division Multiplexing (OFDM) has been widely used in today's radio communications. However, the main drawback of the OFDM systems is the rectangular pulse shaping, which leads to an unsatisfactory energy localization in frequency domain. This drawback can cause several consequences in CR applications because the high out-of-band radiation may severely pollute the neighbors in the adjacent bands. To overcome this problem or, otherwise said, to enable the CR concept, more advanced modulation scheme is necessarily required. During the past years, two main alternatives have been proposed in the CR field, i.e., FBMC/OQAM and GFDM.

The FBMC/OQAM scheme shifts the conventional OFDM paradigm [1]. The key idea is that the modulated data at each subcarrier is shaped by a well-designed prototype filter, which is different from the rectangular pulse in OFDM, so that it indeed provides a large number of degrees of freedom to use different waveform shapes that can be intentionally optimized towards the localized frequency content, straightforwardly enabling the CR applications [2]. Since then, many research works have been conducted towards the CR applications, in particular for FBMC/OQAM [3]–[7]. However, since the FBMC/OQAM is a continuous transform based scheme, it is not easy to employ a Cyclic Prefix (CP), which makes it less robust against the frequency selective fading. Another alternative, which is the GFDM, extends the traditional OFDM paradigm to a general framework by introducing a circular filtering at each subcarrier with an improved pulse shape. By this way, it trades the orthogonality with the possibility of using a non-rectangular pulse. This is the reason why the GFDM is recognized as a non-orthogonal system [8]. The GFDM was first presented for the communication over TV white space. Later, due to its good out-of-band energy attenuation, it rapidly captured a lot of attention in the CR field [9]–[12]. Nevertheless, its non-orthogonality could be a vital drawback in a practical usage. The in-band interference severely causes a performance degradation, which further was proved to be pulse shape dependent [13]. To mitigate the performance degradation, one possible solution is to use an iterative interference cancellation [11]. Thus, the receiver complexity gets largely increased and might even get exploded when considering it together with Multiple-Input-Multiple-Output (MIMO) transmission. Although, it is absolutely necessary to seek for an advanced Multi-Carrier Modulation (MCM) scheme at a price of increased complexity, when it exceeds an affordable limit, the practical implementation of such system cannot be envisaged in a near future.

In this paper, we propose a novel MCM concept that successfully combines FBMC/OQAM with GFDM. To be specific, we adopt the circular filtering for the FBMC/OQAM, which makes it a block transform scheme. Such that the CP can be easily inserted. We call this scheme FBMC/COQAM. Moreover, the core part of this scheme remains the FBMC/OQAM structure, which indeed allows to guarantee a true orthogonality system. Thus, the non-orthogonality issue of GFDM can be solved and in the meantime it enhances the robustness against the fading channels. Ultimately, it shows a good suitability for CR applications. The rest of the paper is organized as follows: in Section II, we give a brief recall of FBMC/OQAM and GFDM schemes. In Section III, we present the concept of the FBMC/COQAM and its motivation behind. In Section IV, we detail the FBMC/COQAM transmitter design in radio transmission. In Section V, we evaluate the FBMC/COQAM efficiency. Some conclusions are drawn in Section VI. For simplicity, in the following, we ignore the term FBMC for FBMC/OQAM and FBMC/COQAM, respectively.

II. OQAM AND GFDM BACKGROUND

In this section, we give a brief introduction of the OQAM and GFDM schemes which clearly shows the difference be-

 ISBN: 978-1-61208-323-0

tween these two schemes, paving the way for an introduction of our proposal.

A. OQAM modulation

The remarkable contribution of the OQAM concept is that it introduced a staggered transmission structure, which allows it to escape from the Balian-Low Theorem (BLT) [14]. So that the OQAM scheme can simultaneously employ an improved pulse shape; keep full orthogonality; and transmit at the Nquist rate. Contrary to the OFDM scheme that transmits complex-valued at subcarriers, in the staggered structure, the real and imaginary parts of the complex-valued symbols are transmitted separately with a delay of half OFDM symbol duration. More details for the OQAM concept can be found in [1] and the references therein. The baseband OQAM modulated signal writes as [15]

$$s_{\text{OQAM}}[k] = \sum_{m=0}^{M-1} \sum_{n \in \mathcal{Z}} a_{m,n} \underbrace{g[k - nN] e^{j\frac{2\pi}{M}m\left(k - \frac{D}{2}\right)} e^{j\phi_{m,n}}}_{g_{m,n}[k]}, \quad (1)$$

where M is the number of carriers; g is the prototype filter with a length of L_g and $D = L_g - 1$ (here, g is assumed to be real-valued and symmetrical); $N = M/2$ is the discrete-time offset; $\phi_{m,n}$ is an additional phase term at subcarrier m and symbol index n which can be expressed as $\frac{\pi}{2}(n + m)$. The transmitted symbols $a_{m,n}$ are real-valued. They are obtained from a QAM constellation, taking the real and imaginary parts of these complex-valued symbols. To address a perfect reconstruction of real symbols, the prototype filter must satisfy the orthogonality condition:

$$\Re\left\{\sum_{k \in \mathbf{Z}} g_{m,n}[k] g_{p,q}^*[k]\right\} = \delta_{m,p}\delta_{n,q}, \quad (2)$$

where * denotes the complex conjugation, $\delta_{m,p} = 1$ if $m = p$ and $\delta_{m,p} = 0$ if $m \neq p$.

B. GFDM modulation

The idea of GFDM is to group a set of complex-valued symbols from time-frequency lattice into one block. Then, for each block, a subcarrier-wise processing is carried out, which includes the up-sampling, pulse shaping, tail biting and finally is followed by a modulated operation to a set of subcarrier frequencies (cf. [8, Fig. 1]). The baseband GFDM modulated signal of one block, i.e., for $k \in [0, MK - 1]$, is expressed as

$$s_{\text{GFDM}}[k] = \sum_{m=0}^{M-1} \sum_{n=0}^{K-1} c_m[n] \tilde{h}[k - nM] e^{\frac{j2\pi km}{N}}, \quad (3)$$

with M the subcarrier number; K the number of symbol slots considered in one block; $c_m[n]$ the complex-valued symbols at m-th carrier and n-th symbol slot. The pulse shape $\tilde{h}[k]$ indicates a periodic repetition of the prototype filter $h[k]$ with a period of MK, i.e.,

$$\tilde{h}[k] = h[\text{mod}(k, MK)]. \quad (4)$$

The periodic filter is used to realize the circular convolution at the transmitter, which is equivalent to the tail biting process [8]. Note that there does not exist any orthogonality condition for the filter design because the GFDM itself is a non-orthogonal system.

III. COQAM: MOTIVATION AND CONCEPT

As previously stated, the OQAM and GFDM schemes inherit some weak points. For the OQAM scheme, due to its continuous transform nature, it is not straightforward to use a CP. Thus, in the radio environment, the frequency selective fading usually ruins the orthogonality of the OQAM system, leading to a performance degradation. Therefore, more complex multi-tap equalizers are needed to warrant the quality of the transmission [16]–[18]. Similar to the OQAM, the GFDM itself is a non-orthogonal scheme. Thus, the performance must be guaranteed with a more powerful receiver design. The authors in [13] have shown that even in a distortion free channel the matched filter cannot address a satisfactory performance. Although the degradation can be relieved by an appropriate filter, with an example of a raised cosine filter given in [13], the filter must be driven close to its ideal limit, which increases the complexity. Alternatively, the GFDM can employ an iterative interference cancellation method [11], resulting, nevertheless, in a high complex receiver, in particular when it is in combination with MIMO transmission.

The drawback of OQAM is due to the fact that the OQAM is not a block processing scheme, while the drawback of GFDM is due to that its non-orthogonality is restricted by the BLT. Knowing the problematic for these schemes, an intuitive question is whether we can find an improved modulation that keeps all the benefits of OQAM and GFDM and at the same time gets rid of their drawbacks. With this motivation in mind, we investigated a new MCM scheme called COQAM, whose idea is to replace the linear convolution inherited in the OQAM with a circular convolution used in the GFDM. By this way we get a modulation scheme which, as CP-OFDM and GFDM, corresponds to a block transform. The baseband COQAM modulation structure is depicted in Fig. 1. For a discrete-time signal $s[k]$ defined in a block interval such

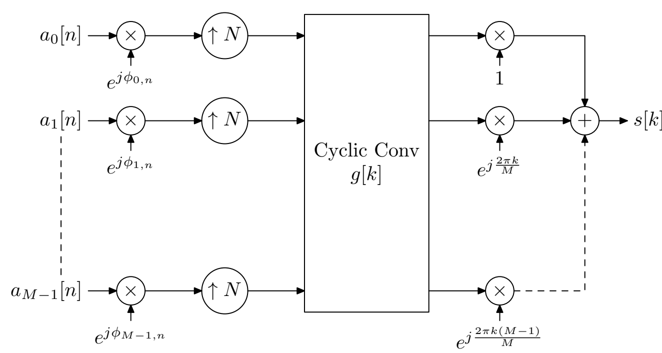


Figure 1. FBMC/COQAM baseband modulation.

that $k \in [0, MK - 1]$, the baseband COQAM modulation is expressed as

$$s_{\text{COQAM}}[k] = \sum_{m=0}^{M-1} \sum_{n=0}^{K-1} a_m[n] \tilde{g}[k - nN] e^{j\frac{2\pi}{M}m\left(k - \frac{D}{2}\right)} e^{j\phi_{m,n}}, \quad (5)$$

with K the number of real symbol slots per each block. Like GFDM, to implement a circular convolution with a prototype filter g of length $L = KM$, we introduce a pulse shaping filter denoted $\tilde{g}$, obtained by the periodic repetition of duration KM

 ISBN: 978-1-61208-323-0

of the prototype filter g, i.e.,

$$\tilde{g}[k] = g[\text{mod}(k, MK)]. \quad (6)$$

The orthogonality condition for the filter design is in line with (2). This can be readily proven using the symmetrical property of the filter.

IV. COQAM TRANSMITTER DESIGN IN A RADIO SYSTEM

In order to maintain a perfect orthogonality after a transmission through a multi-path channel, we introduce a CP to cancel the inter block interference. Moreover, as we want to prevent an alteration of the Power Spectral Density (PSD), resulting from a spectral leakage due to the block processing, a windowing is applied after CP insertion. The transmission system resulting from these two operations is named windowed CP-FBMC/COQAM (WCP-COQAM). Denoting the CP length by L_{CP}, we get $L_{\text{CP}} = L_{\text{GI}} + L_{\text{RI}}$, where L_{GI} is the CP part used to fight against the multi-path channel interference and L_{RI} is the portion devoted to windowing.

The l-th block of the WCP-COQAM signal $s_{\text{WCP-COQAM}}[k]$, for $k = 0, \cdots, KM + L_{\text{CP}} - 1$, can be obtained from the l-th block of the COQAM signal, for $k = 0, \cdots, KM - 1$, by

$$\begin{aligned} s_{\text{WCP-COQAM}}[k] &= \sum_{r=l-1}^{l+1} s_{\text{COQAM}}[\text{mod}(k - L_{\text{CP}}, KM)] \\ &\times \ w[k - rQ], \end{aligned} \quad (7)$$

where $Q = KM + L_{\text{GI}}$ and $w[k]$, defined in the $k = 0, \cdots, KM + L_{\text{CP}} - 1$ interval, is the window function computed as follows

$$w[k] = \begin{cases} \text{window coeffs.} & k \in [0, L_{\text{RI}} - 1] \\ 1 & k \in [L_{\text{RI}}, KM + L_{\text{GI}} - 1] \\ w[KM + L_{\text{CP}} - 1 - k] & \text{otherwise.} \end{cases}$$

The WCP-COQAM transmitter structure is depicted in Fig. 2, where $s_1[k]$ denotes $s_{\text{COQAM}}[k]$, which is obtained from Fig. 1 and $s_2[k]$ stands for $s_{\text{WCP-COQAM}}[k]$.

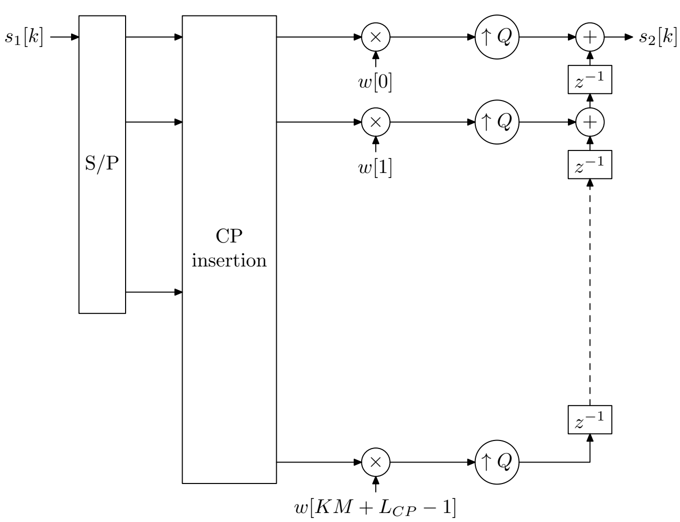


Figure 2. Transmitter of the windowed CP-FBMC/COQAM system.

There are extensive works on the window design [19]. In this paper, we simply use the Hamming window for this windowing process. Particular investigation on the window design in the context of COQAM will be envisaged in the future step. Furthermore, it is worth noting that the additional part L_{RI} does not reduce the spectral efficiency as it falls only in the overlapped samples between two successive blocks.

V. SIMULATION EVALUATION

To illustrate the efficiency of the proposed WCP-COQAM scheme, we first evaluate the PSD of the afore-mentioned MCM candidates. In our simulation, the PSD is estimated using the Welch method [20]. The parameters used in our simulation are: number of subcarriers is fixed to $M = 128$ for all MCM schemes. Regarding the prototype filter for OQAM, WCP-COQAM and GFDM, in our simulation a Square Root Raised Cosine (SRRC) filter is used with the roll-off factor 0.5 and length up to $6M$. The CP contains 32 samples for CP-OFDM, WCP-COQAM and GFDM. In particular, the windowing interval for WCP-COQAM is $L_{\text{RI}} = 16$ samples. The result of the PSD comparison is given in Fig. 3. It clearly

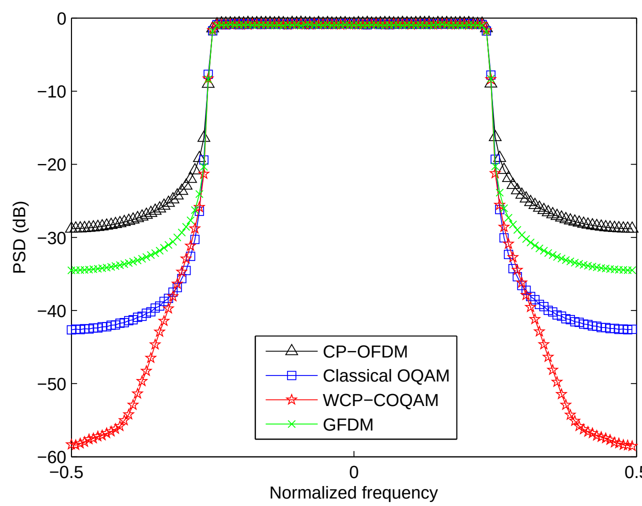


Figure 3. PSD simulation, $M = 128$, SRRC with roll-off 0.5.

shows that the WCP-COQAM can remain a satisfactory low level of the out-of-band spectral leakage. In fact, for some CR applications, non-contiguous bands are sometimes considered for addressing dynamic spectrum sharing. In this sense, the in-band notch spectrum shape becomes extremely important, as it directly reflects the spectrum management granularity, which can be used to drive its dynamics to the ultimate limit. A good MCM scheme is expected to provide a low spectral radiation in the notch band, in order not to create severe interference to the neighboring bands. As shown in Fig. 4, the WCP-COQAM is able to keep a low power leakage in the notch band. On the other hand, we note that although the PSD of GFDM is improved compared with CP-OFDM, it cannot provide a similar PSD as for the OQAM and WCP-COQAM schemes, because the block processing of GFDM still generates the spectral leakage due to the discontinuity on the edge of each block [19]. Hence, the windowing process can alternatively be employed to the GFDM for further spectral leakage mitigation. However, this cannot relieve its non-orthogonality issue.

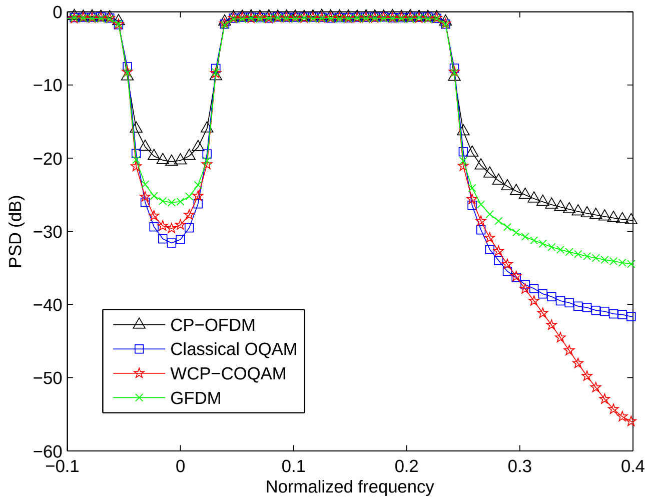


Figure 4. PSD simulation with notch, $M = 128$, SRRC with roll-off 0.5..

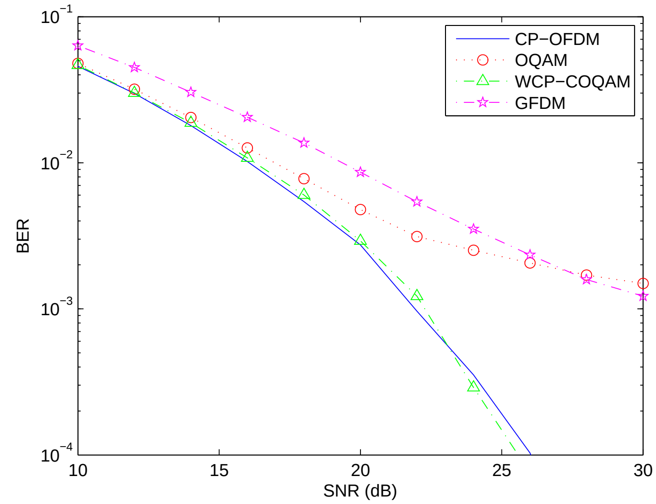


Figure 6. BER for uncoded QPSK in dispersive channel with ZF equalizer.

Next, we provide an orthogonality evaluation to show that the WCP-COQAM perfectly retains the orthogonality. Moreover, it is quite robust against the frequency selective fading. Our evaluation follows two steps. In the first step, which is similar to [13], we evaluate the Bit Error Rate (BER) performance in an Additive White Gaussian Noise (AWGN) channel. We keep the same parameter setting as in the PSD evaluation and the symbol constellation is QPSK. The receiver does not employ any equalizer but a matched filter. The simulation result is reported in Fig. 5, where we use the CP-OFDM curve as a reference because it reflects the QPSK Matched-Filter (MF) bound in AWGN. It is clearly shown that

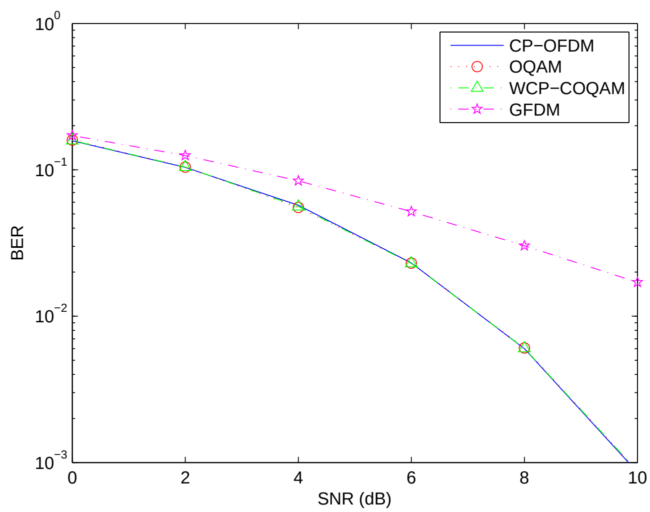


Figure 5. Matched filter BER for uncoded QPSK in AWGN channel.

the conventional OQAM and WCP-COQAM systems are fully orthogonal because their BER curves are overlapped with the CP-OFDM curve. On the other hand, the GFDM curve cannot merge with the other ones due to its non-orthogonality. A performance degradation can be seen, confirming that the matched filter cannot cancel the interference brought by the neighboring symbols. Thus, it needs a more advanced equalizer. In the second step, we show that the WCP-COQAM is still quite robust against the frequency selective fading. In our simulation, we consider a wireless dispersive fading channel, similar to [21], with impulse response $h_{\text{ch}}[k] = \sum_p \alpha_p \delta(k-p)$, where α_p are complex i.i.d. Gaussian variables with the power profile following $\mathbf{E}[|\alpha_p|^2] = e^{-p/r}$ and, r, the normalized delay spread. In our simulation, we choose $r = 4$, reflecting a severe frequency selective fading. The channel is normalized and is truncated at -20 dB. The receiver equalization technique uses a one-tap Zero-Forcing (ZF) equalization for all the MCM schemes. The simulation result, presented in Fig. 6, shows that even under a severe fading environment, the WCP-COQAM curve is still merged with the CP-OFDM. While for the OQAM and GFDM, a degradation by several decibels is displayed. Hence, the result proves that the WCP-COQAM is very robust against the frequency selective fading. With our simulations, we can conclude that the WCP-COQAM can keep all the benefits of the proposed MCM schemes and get rid of their drawbacks. Thus, this scheme can be seen as a good physical layer enabler for the CR applications.

VI. Conclusion and Future Work

In this paper, we presented a novel multi-carrier modulation scheme, which combines the conventional OQAM and GFDM concepts and successfully keeps all their benefits and at the same time overcomes their drawbacks. The simulation confirmed the achievements of this new scheme and further proved its efficiency as a suitable physical layer modulation for cognitive radio applications. In the next step, we will investigate a dedicated window design, as well as an analysis on the Peak-to-Average-Power-Ratio (PAPR).

References

[1] B. Le Floch, M. Alard, and C. Berrou, "Coded Orthogonal Frequency Division Multiplex," Proceedings of the IEEE, vol. 83, Jun. 1995, pp. 982–996.

[2] M. Bellanger, T. Ihalainen, and M. Renfors, "Filter bank based cognitive radio physical layer," in Proc. of ICT-Mobile Summit, Santander, Jun. 2009, ISBN: 978-1-905824-12-0.

[3] M. Shaat and F. Bader, "Computationally Efficient Power Allocation Algorithm in Multicarrier Based Cognitive Radio Networks: OFDM and FBMC systems," EURASIP Journal on Advanced Signal Processing (ASP), vol. 2010, 2010:528378, 13 pages.

[4] H. Zhang, D. Le Ruyet, D. Roviras, Y. Medjahdi and H. Sun, "Spectral Efficiency Comparison of OFDM/FBMC for Uplink Cognitive Radio Networks," EURASIP Journal on Advanced Signal Processing (ASP), vol. 2010, 2010:621808, 14 pages.

[5] H. Zhang, D. Le Ruyet, and M. Terré, "Spectral Correlation of Multicarrier Modulated Signals and its Application for Signal Detection," EURASIP Journal on Advanced Signal Processing (ASP), vol. 2010, 2010:794246, 13 pages.

[6] M. Renfors, "Spectrum monitoring schemes for FBMC cognitive radios," in Proc. Future Network and MobileSummit (FNMS 2010), Florence, Italy, Jun. 2010, pp. 1–9.

[7] T. Ihalainen, A. Viholainen, T.H. Stitz and M. Renfors, "Reappearing primary user detection in FBMC/OQAM cognitive radios," in CROWNCOM conference, Cannes, France, Jun. 2010, pp. 1–5.

[8] G. Fettweis, M. Krondorf, and S. Bittner, "GFDM-Generalized Frequency Division Multiplexing," in IEEE Vehicular Technology Conference (VTC Spring'09), Barcelona, Spain, Apr. 2009, pp. 1–4.

[9] N. Michailow, M. Lentmaier, P. Rost, and G. Fettweis, "Integration of a GFDM Secondary System in an OFDM Primary System," in Future Network and Mobile Summit (FNMS'11), Warsaw, Jan. 2011, pp. 1–8.

[10] R. Datta, N. Michailow, S. Krone, M. Lentmaier, and G. Fettweis, "Generalized Frequency Division Multiplexing in Cognitive Radio," in European Signal Processing Conference (EUSIPCO'12), Bucharest, Romania, Aug. 2012, pp. 2679 – 2683.

[11] R. Datta, N.Michailow, M. Lentmaier, and G. Fettweis, "GFDM Interference Cancellation for Flexible Cognitive Radio PHY Design," in IEEE Vehicular Technology Conference (VTC Fall'12), Quebec city, Canada, Sept. 2012, pp. 1–5.

[12] R. Datta, D. Panaitopol, and G. Fettweis, "Analysis of Cyclostationary GFDM Signal Properties in Flexible Cognitive Radio," in International Symposium on Communication and Information Technologies (ISCIT'12), Gold Coast, Australia, Oct. 2012, pp. 663 – 667.

[13] N. Michailow, S. Krone, M. Lentmaier, and G. Fettweis, "Bit Error Rate Performance of Generalized Frequency Division Multiplexing," in IEEE Vehicular Technology Conference (VTC Fall'12), Quebec city, Canada, Sept. 2012, pp. 1–5.

[14] H.G. Feichtinger and T. Strohmer, Gabor Analysis and Algorithm - Theory and Applications, Birkhäuser, Boston-Basel-Berlin, 1998.

[15] P. Siohan, C. Siclet, and N. Lacaille, "Analysis and design of OFDM/OQAM systems based on filterbank theory," IEEE Transactions on Signal Processing, vol. 50, no. 5, May 2002, pp. 1170–1183.

[16] T. Ihalainen, T. H. Stitz, M. Rinne, and M. Renfors, "Channel Equalization in Filter Bank Based Multicarrier Modulation for Wireless Communications," EURASIP Journal on Advances in Signal Processing, vol. 2007, no. ID 49389, Aug. 2007, pp. 1–17.

[17] P. Tanguy H. Lin, P. Siohan and J-P. Javaudin, "An analysis of EIC for OFDM/OQAM systems," Journal of Communications (JCM), vol. 4, no. 1, Feb. 2008, pp. 52–60.

[18] G. Ndo, H. Lin, and P. Siohan, "FBMC/OQAM equalization: Exploiting the imaginary interference," in PIMRC'12, Sydney, Australia, Sept. 2012, pp. 2359 – 2364.

[19] F.J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," Proceedings of the IEEE, vol. 66, no. 1, Jan. 1978, pp. 51 – 83.

[20] 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 Tans. Audio and Electroacoust, vol. AU-15, Jun. 1967, pp. 17–20.

[21] T. Fusco, A. Petrella, and M. Tanda, "Data-Aided Symbol Timing and CFO Synchronization for Filter Bank Multicarrier Systems," IEEE Tans. on Wireless Communications, vol. 8, May 2009, pp. 2705–2715.

Communication Patterns: a Novel Modeling Approach for Software Defined Radio Systems

Andrea ENRICI, Ludovic APVRILLE and Renaud PACALET
Institut Mines-Telecom
Telecom ParisTech, CNRS/LTCI
Biot, France
Email: {andrea.enrici, ludovic.apvrille, renaud.pacalet}@telecom-paristech.fr

***Abstract*—Efficiently programming Software Defined Radio applications still remains a pending challenge. While most of the efforts are focused on the processing part of a design, communications are a great source of performance and portability issues that is often neglected. Within the frame of a Model Driven Engineering methodology for the design of dataflow processing applications, this paper proposes a novel approach to model complex communication interactions. This approach relies on communication patterns to capture communication protocols and standards at system-level, independently of computations. A case study for cognitive radio shows how communication patterns are efficiently used to generate cross-platform models that can be ported and refined to specific applications and architectures.**

***Keywords*—Software defined radio; hardware/software codesign; model-driven engineering**

I. Introduction

In the last decade, the demand for more flexible and reconfigurable solutions able to support multiple communication standards has lead to a shift from radios where the full range of capabilities was supplied by hardware elements to Software Defined Radios (SDRs) [1] where some or all of the physical layer functions are software defined [2]. This technological shift takes advantage of common Digital Signal Processing (DSP) algorithms, shared by modern air-services protocols and standards (e.g., Global Positioning System). A SDR system is composed of a *platform*, intended as the set of hardware elements (e.g., DSPs) and software layers (e.g., Operating System), on top of which a software *waveform* is executed. The latter is defined as the software application coordinating and configuring the platform in order to transform the information contained in the signals to be transmitted and received. Implementing a waveform in software adds greater flexibility to radio systems as their functionality can be changed by a simple software update without the need to change equipment. This added flexibility yields numerous advantages including increased service life time of equipments, cross-platform portability of the software, reduced costs and better ease in terms of system implementation, upgrade and maintenance. In order to adequately exploit the benefits of software waveforms, SDR platforms provide a way to implement computations and communications in a very generic and flexible way. However, the price to pay for this flexibility and genericness is an increase in the complexity of programming SDR platforms [3]. Currently, the latter are manually programmed by system experts in languages like C/C++. Nevertheless, such languages do not shield programmers from the system complexity, leading to higher development costs, longer time-to-market of new products, as well as reduced cross-platform portability of waveforms.

Thus, a current hot topic for the industrial and scientific community is to find an efficient way to automatically program such complex systems [4]. Currently, the most promising strategies rely on Model Driven Engineering (MDE) combining Domain Specific Modeling Languages (DSMLs) with transformation engines and generators to synthesize source code or alternative model representations [5]. Methodologies for the design of SDR systems employing the above strategies are discussed in Section II. These methodologies, however, either informally prioritize computations over communications or propose non-portable solutions that are specific for a given platform. A consequence of such approaches is that the limited expressive power of models restricts the systems that can be described and does not allow to efficiently exploit all the platform capabilities in terms of addressing and data transfers.

The work presented in this paper enriches DiplodocusDF [6], a methodology for design and code generation of heterogeneous processing applications of type dataflow, based on MDE and UML. The main contributions introduced here are (1) a novel feature to capture complex communication schemes, early at system-level, independently of computations and (2) a novel approach for modeling SDR systems in a portable way.

This paper is organized as follows: the related work is described in Section II followed by the context of our works, in Section III. Section IV describes our solution to model communications, its integration into DiplodocusDF and how the latter paves the way to a novel modeling approach. To show the benefits of our contributions, Section V applies them to a practical case study for cognitive radio. Conclusion along with the state and directions of our works are given in Section VI.

II. Related Work

SDR design is a hardware/software co-design topic that deals with most issues of design space exploration for embedded real-time systems. The complexity of SDR systems and their need to dynamically adapt to the environment (cognitive radio) make current methodologies for embedded real-time systems unsuitable to properly cope with all the requirements of SDR design (e.g., flexibility, heterogeneous processing, hardware abstraction). UML methodologies that reflect this statement, are those for the development of embedded applications that are based on the MARTE profile [7]. UML-MARTE in fact, lacks the notations to describe dataflow applications such as SDR in a pure abstract way and its

 ISBN: 978-1-61208-323-0

procedural approach leads to models with only one centralized controller, as mentioned in [6]. UML-MARTE has been successfully proposed for platform design, instead: the MOPCOM [8] project applies this profile to describe real time properties and to modeling in order to generate code for implementation and verification. On the other hand, closer to the aims of DiplodocusDF, the A3S project [9] provides a tool and a methodology to design, model and verify SDR systems in UML, targeting non-functional characteristics of both hardware and software components. A3S defines a UML profile, specific to SDR, but lacks the generation of validation and implementation code, as opposed to DiplodocusDF. Nevertheless, UML is not the only solution for abstract system modeling of SDR systems: other existing approaches are based on dataflow graph (DFG) representations. [10] is a MDE approach that proposes a lightweight programming model for describing SDR applications, by means of dataflow Models of Computation. The authors propose a minimally intrusive design flow that, however, needs preexisting tools, libraries and processes to breathe life into models. The paper [11] presents an interesting solution that, to our knowledge, shares the most with DiplodocusDF in terms of gathering within a unique framework several concepts and tools for heterogeneous design of SDR systems. The authors propose a methodology based on the Syndex tool [12], that allows executable code generation from high-level models through a series of graph transformations. The proposed approach is oriented to ultra-fast prototyping for heterogeneous platforms so as to deal with realistic implementations instead of simulation results. Applications are described as extended DFGs where nodes represent processing operations and edges represent data transfers. Similarly in the architecture graph, vertexes model hardware components and hyper-edges represent communication media. The key strength of the proposed methodology lays in its support for heterogeneous processing platforms (e.g., Field-Programmable Gate Arrays), the automatic code generation and the associated scheduling. With respect to communication modeling, [11] does not extend the way communications are handled in Syndex, so the critics related to graph-based approaches in Section III-C can be addressed to [11] also.

III. Context

In order to illustrate SDR platforms, this section provides an outline of the hardware architecture of Embb [13], the platform on which the implementation code generated by DiplodocusDF has been successfully executed. A more detailed description of DiplodocusDF follows.

A. The hardware platform Embb

The authors in [13], propose a new generic baseband architecture for SDR applications. An instance of such a platform is depicted in Fig. 1. It is composed of (1) DSP units, (2) a system interconnect and (3) a main CPU. The DSP units are in charge of executing the processing operations (e.g., FFT). They are equipped with a hardware accelerator as computational unit (Processing SubSystem, PSS), a DMA to transfer data, an internal memory mapped on the main processor memory and a microncontroller (μC) that allows to reduce interventions of the main CPU. The latter is linked to DSPs via the system interconnect that allows data-blocks and control information to be exchanged among units. The main processor executes the waveform control operations: it manages data-transfer operations, the computational units and the interface with the external environment (Fig. 1, yellow area).

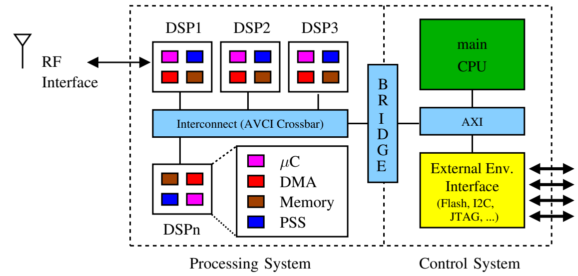


Fig. 1. The architecture of an Embb instance

Executing parallel applications on SDR platforms is not trivial because of both memory and computational resource pooling. This implies dense flows of data and control information being transferred among hardware units. Since these flows access shared resources (e.g., the bridge in Embb between the control and the processing systems), their impact over the system's performance cannot be neglected and urges for modeling techniques to properly take them into account when designing applications.

B. The DiplodocusDF methodology

DiplodocusDF [6] is a UML MDE methodology for the design of heterogeneous dataflow applications for real time embedded systems, in particular radio (such as Software Defined Radio), Fig. 2. It stems from DIPLODOCUS, [14], a UML Model Driven Engineering methodology for hw/sw partitioning of Systems on Chip at high abstraction level, currently implemented by the free software TTool [15]. The core strength of DIPLODOCUS is the automatic transformation of models for simulation and formal verification [16]. However, the DIPLODOCUS approach is too abstract to permit automatic code generation for SDR systems as models lack the necessary expressiveness to face the complexity of platforms and waveforms. DiplodocusDF is a first attempt to fill the aforementioned gaps; it enriches DIPLODOCUS with the following extensions:

- **A dataflow semantics:** an application is modeled as a dataflow graph, where nodes represent tasks (processing, routing, addressing operations) and edges are used to carry data-blocks and the related control parameters (e.g., r/w memory addresses).
- **A specialization of the architecture language:** heterogeneous platforms are represented as a network of computation nodes (e.g., DSPs), storage nodes (e.g., memories) and data-transfer nodes (e.g., bus) interconnected by communication edges.
- **An environment for automatic generation of executable code:** the description of a waveform mapped over a platform is translated in C-language code via an Intermediate Representation completed by the platform Application Programming Interface (API) and by a Run Time Environment for scheduling computations.

 ISBN: 978-1-61208-323-0

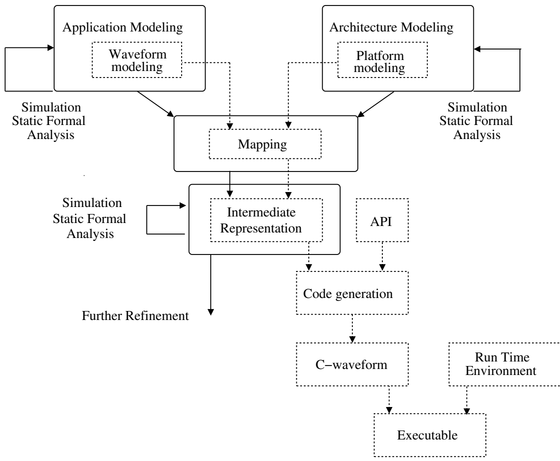


Fig. 2. The DIPLODOCUS (solid lines) and DiplodocusDF (dotted lines) methodologies

C. *Communication modeling in DiplodocusDF*

Since DiplodocusDF is a graph-based methodology, it models communication exchanges among processing operations by means of edges in the waveform DFG. At mapping stage, edges are projected over *buffers*, defined as memory regions of the platform storage elements, in a one-to-one fashion. Consequently, this forces the buffer of the producer operation to reside in the same memory as the one of the consumer operation. At mapping stage, the application graph is adapted to the platform addressing capabilities in order to take into account data transfers. This typically implies the injection of additional nodes in the application, e.g., to map multiple edges to the same buffer, to model data transfers that employ a DMA. This adaptation may also lead to alter the natural sequence of processing operations. As for the graph-based approaches of Section II, in DiplodocusDF application models, data transfers from a producer node to a consumer node can only be mapped on simple P2P paths in the platform. In other words, no complex data-transfer scheme with more than one DMA and with intermediate memories between the source and destination memories can be described. The source node and the destination node of a transfer must be able to directly access a storage element via either a bus or one single DMA. These characteristics restrict the design space in terms of waveforms and platforms that the methodology can describe.

IV. Communication Patterns

In addition to computations, the processing of radio signals includes communication exchanges between architecture nodes (e.g., DSP, bus). For instance, these communications represent exchanges for operations running on different units, or items (i.e., data, instructions) fetched from memory as part of one processing operation. These communications may be implemented by means of different standards and protocols (e.g., AMBA, PCI Express) within the same platform, thus making each transfer different in terms of performances and interactions among nodes. However, communications share common patterns independently of the mechanisms being used. Thus, our objective is not to model communication standards in details, but rather to model their underlying patterns at a high level of abstraction and adapt them to the transfer capabilities of a specific platform. This permits to describe the influence of communication interactions on the system's performance and to provide models with the expressive power to generate code for communications in a portable way. To reach this target, the initial concepts of *communication pattern* [17] are extended and the resulting contribution is integrated into an enriched DiplodocusDF methodology.

A. *Communication Patterns in a Nutshell*

A Communication Pattern (CP) describes a transfer between architecture actors. It is associated to a *communication flow*, between a source S and a destination D nodes in the application that are continuously connected by a set of edges. In other words, edges must build a path in the application graph that links S to D without interruptions. More formally, a communication pattern is defined as a tuple:

$$\mathcal{CP} = (\mathcal{N}_{arch}, \mathcal{E}_{arch}, \prec_{\mathcal{N}}, \mathcal{N}_{app}, \mathcal{E}_{app}, \mathcal{A}, \prec_{\mathcal{A}})$$

- $\mathcal{N}_{arch}$ is the set of architecture actors involved in the transfer;
- $\mathcal{E}_{arch}$ is the set of architecture edges connecting the architecture actors of $\mathcal{N}_{arch}$;
- $\prec_{\mathcal{N}} \subseteq \mathcal{N}_{arch} \times \mathcal{N}_{arch}$ is a total order relation among architecture actors;
- $\mathcal{N}_{app}$ represents the source S and destination D nodes from the application;
- $\mathcal{E}_{app}$ is the set of edges from the application graph that make up the communication flow from S to D;
- $\mathcal{A}$ is a set of actions performed by $\mathcal{N}_{arch}$ to accomplish the transfer;
- $\prec_{\mathcal{A}} \subseteq \mathcal{A} \times \mathcal{A}$ is a partial order relation among actions;

A CP is represented independently with respect to the waveform and platform models as an extended UML Sequence Diagram. Fig. 3 shows an abstract communication pattern for the communication flow of edge ed1 in Fig. 6. In Fig. 3, a generic data transfer is requested by a Controller actor and executed by a Transfer actor between a source and a destination storage actors. The actors (N_{arch}) in Fig. 3, are *architecture supernodes* representing generic architecture elements that are purely functional. Supernodes can be classified in three

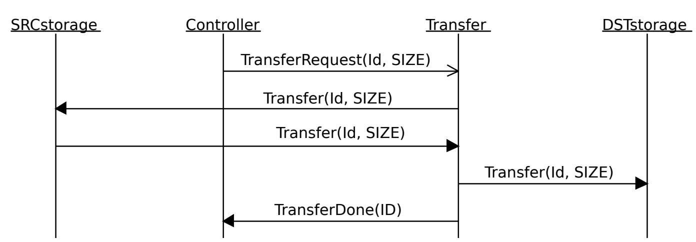


Fig. 3. An abstract communication pattern

classes: *storage* (e.g., memory) for item storing, *controller* (e.g., CPU, DSP) for control and configuration, *transfer* (e.g., bus, bridge) for routing and dispatching. Fig. 3 also shows some of the actions A among the architecture actors. An action is associated to an actor and provided with a type representing its functionality and a set of parameters to specify the latter. Actions are classified in three types: *Message Actions*, *Execution Specification Actions* and *Conditional Actions*.

Message Actions represent an abstraction of the interactions used by communication protocols and standards. A message action can either be blocking or non-blocking: the former blocks the sender until completion while the latter allows the sender to switch to another task (e.g., *TransferRequest()* in Fig. 3).

Execution Specification Actions model the processing time of an ongoing communication. The semantics of these actions depend on the underlying architecture node and can be observed after mapping CPs over the architecture. Two execution specification operators are defined: Exec, which represents the cost of actively executing a communication operation (e.g., the number of cycles taken by a bus to execute a transfer) and Delay, which is a generic time interval used when a node is not actively involved in a communication operation, e.g., to model the time taken to complete a DMA transfer from the CPU perspective.

Conditional Actions add conditional behavior to communication patterns to capture, for instance, precedence constraints in complex transfers. Conditional actions are of two types: the loop construct (Fig. 5, 10) and the if-else construct, respectively for iterative and conditional communication exchanges. A conditional action includes other actions and may be nested according to the scenario being described.

B. The Proposed Methodology

Following the above description of what a communication pattern is and looks like, it is now illustrated how describing communications with CPs results in a novel modeling approach. The added value of communication patterns is the separation of concerns between communications and computations that allows to elegantly capture the functionality of an application description. Since both communications and computations involve different sets of architecture elements, this separation of concerns translates into a separate mapping and refinement on the architecture graph. Fig. 4 illustrates such a novel approach that intends to replace the classic Y-Chart [18] scheme adopted by DiplodocusDF in the modeling phase of Fig. 2. Instead of projecting the application on the architecture in one single stage, followed by an adaptation stage, computations in the waveform DFG are mapped first and communications (CPs) are mapped next. In our approach, computations are atomic operations whose execution cannot be split over several processing nodes. Thus, mapping of computations is done in a single step by associating the application nodes onto the architecture processing nodes (level L2, Fig. 4). Conversely, as complex communication paths are expressed by multiple kinds of elements, a separate mapping step is required. In Fig. 4, the box corresponding to level L3 regroups this mapping in a single stage.

The top-most level of the methodology, L0 in Fig. 4, is represented by three sets of repositories containing the building bricks of the application, architecture and communication patterns. The application repository contains nodes for signal processing operations and edges for inter-node communication that are used to build the waveform DFG. For the architecture graphs and the communication pattern models, two separate repositories are provided, as supernodes in CPs have a higher

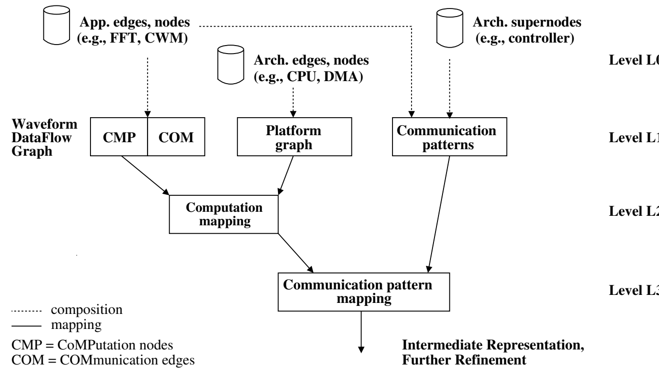


Fig. 4. The proposed modeling approach

level of abstraction with respect to nodes in the platform. Thus, at level L1, models consist in a waveform DFG (Fig. 6), a platform graph (Fig. 8) and stand-alone communication patterns (Fig. 3 and 9). At this level of abstraction, the purpose of such high-level models for communications (CPs) and computations (waveform DFG) is to express the application in a functional and portable way, that transcends from the actual platform implementation since the methodology aims at designing portable SDR applications.

Once level-L1 models are available, at level L2 computations are mapped over the architecture processing units according to their computational power (Fig. 8). Due to the communication capabilities of processing units, this mapping step constraints the mapping of communication patterns at level L3. Here, the abstract supernodes are mapped over the platform nodes: storage supernodes are associated to memories, controller supernodes to CPUs and transfer nodes to a network of DMAs, buses and bridges (Fig. 5, 10). It is important to state that mapping of storage and controller supernodes takes place in a one-to-one fashion. In fact, none of the two types of supernodes can be split over multiple platform units. Allowing such a one-to-many mapping would in fact imply additional transfers, within the set of mapped units, that were not described in the initial communication pattern. On the other hand, transfer supernodes nodes are associated to a network of transfer units in a one-to-many fashion to permit modeling of complex transfer paths requiring intermediate elements in between the source and the destination storage.

Fig. 5 shows the abstract communication pattern of Fig. 3,

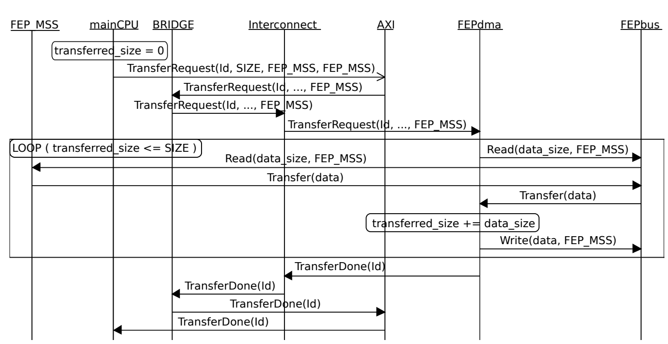


Fig. 5. The Communication Pattern of Fig. 3, after mapping level L3

mapped (level L3) on the architecture of Fig. 8. At level L2, SRC and FFT computations have been mapped to mainCPU and FEP (a DSP unit), respectively. Fig. 5 describes how data between SRC and FFT is moved within memory FEP_MSS (where SRCstorage and DSTstorage supernodes have been mapped) via the DMA FEPdma. mainCPU requests the DMA transfer via a network of transfer nodes made up of AXI bus, BRIDGE and Interconnect. Such a transfer is then executed by iteratively reading data from FEP_MSS, storing them into the intermediate FIFO of FEPdma and writing to FEP_MSS. The latter actions are governed by a loop conditional action that iterates according to the value of transferred_size. A notification is eventually sent back from FEPdma to mainCPU via Interconnect, BRIDGE and AXI.

V. Case Study: Welch Periodogram Detector

In this section, an implementation of the Welch Periodogram Detector (WPD) algorithm is taken from [6] as a scenario for a case study. WPD is a energy detection algorithm used for sensing the spectrum and detecting when a given frequency band can be opportunistically used. Fig. 6 shows the dataflow graph of level L1 for WPD, where only edges and processing operations are modeled, together with a FORK node to broadcast data. Here, a source (SRC) produces the input vectors whose frequency representation (FFT) is processed by the component-wise square of modulus (CWM). Next, two CWM output vectors are component-wise added (CWA) and the elements of the resulting vector are summed (SUM) and collected (SINK). On the other hand, Fig. 7 illustrates how the DFG of Fig. 6 is adapted, in DiplodocusDF, to Embb by injecting routing and addressing nodes (e.g., OVLP, DMA). Fig. 8 shows the architecture graph of level L1 for the portion of Embb (Fig. 1) relevant to the case study. The latter figure also displays mapping of computation operations (level L2) and the mapping of storage supernodes (level L3) for the CPs relevant to the case study. The Front End Processor (FEP) DSP [13], offers all the computational power required to process the WPD waveform. It is modeled as an interconnected subsystem with a CPU (FEP) for executing computations, a memory (FEP_MSS) for storing data and a DMA (FEPdma) for transfers to/from the memory via an internal bus (FEPbus). The remaining elements in Fig. 8 model the interconnect (Interconnect, TAVCI, BRIDGE) and the control part (mainCPU, mainBus, mainMemory, mainDMA) of Embb.

With respect to [6], the original case study is re-visited by adding the following requirement: let us suppose that sensing the spectrum with WPD is part of a larger scenario where the frequency representation of the input signal must be stored apart for later processing. Modeling such a requirement in DiplodocusDF or in one of the graph-based methodologies of Section II, would require a waveform re-design, e.g., to add a node collecting the output of FFT plus the related addressing and routing operations to perform the transfer. Instead, communication patterns can capture this exigency handily by adding one simple edge, ed8 in Fig. 6, to the waveform DFG of level L1. The communication pattern of level L1 for the communication flow corresponding to edges ed1 and ed8 is showed in Fig. 9. Here, data are transferred between Storage1 and Storage2 (edge ed1), then copied from the latter to Storage3 (edge ed8). Computations at level L2 are mapped to FEP, while at level L3 the transfer path can now be described with much greater flexibility: the designer can chose to copy data to FEP_MSS, mainMemory or any combination of the two. Similarly, the communications can be executed with or without DMA, via FEPdma, mainDMA, Interconnect, FEPbus or any combination of these elements. In this case study, it has been chosen to copy the FFT output to mainMemory by means of mainDMA.

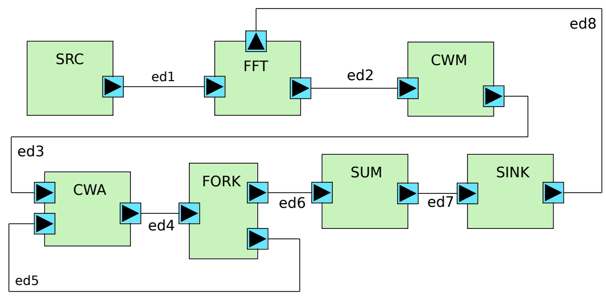


Fig. 6. WPD waveform DataFlow Graph of level L1

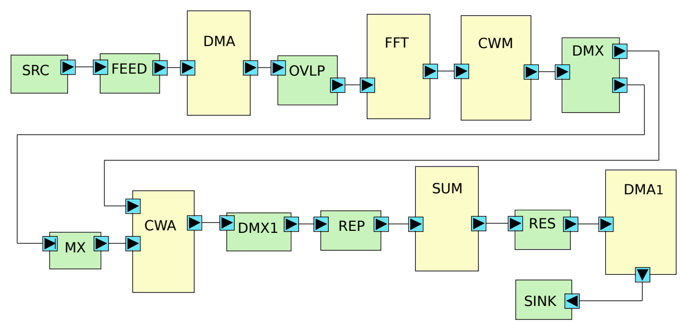


Fig. 7. WPD waveform DataFlow Graph of level L1 adapted to Embb

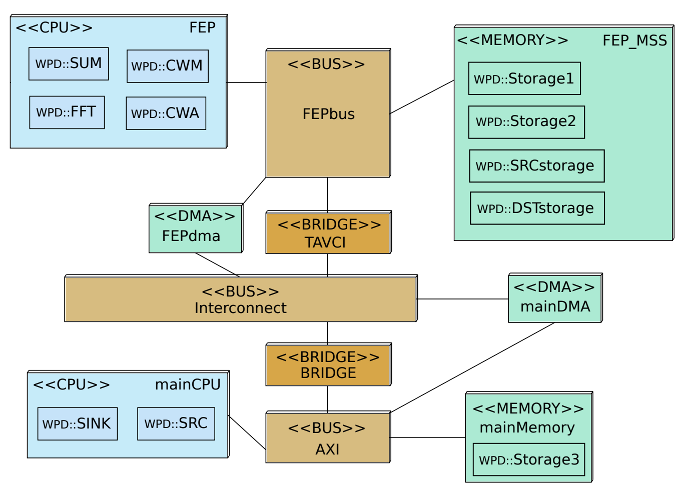


Fig. 8. The graph of the instance of Embb used for the WPD case study

For the sake of simplicity, Fig. 10 only illustrates the CP of level L3 for edge ed8. The complete CP of level L3 for both ed1 and ed8 can be composed by merging those in Fig. 5 and Fig. 10. In Fig. 10, the scenario is similar to that of Fig. 5, but more actors are involved and data are copied instead of being transferred. So, mainCPU programs a CopyRequest

 ISBN: 978-1-61208-323-0

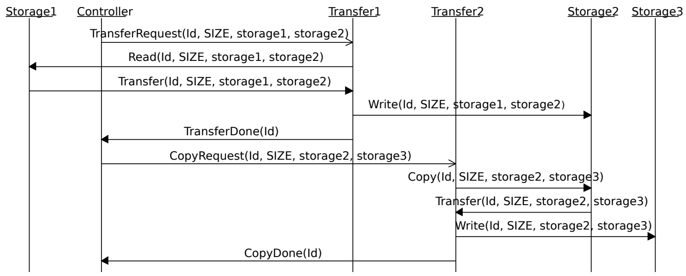


Fig. 9. Communication Pattern of level L1 for edges ed1, ed8

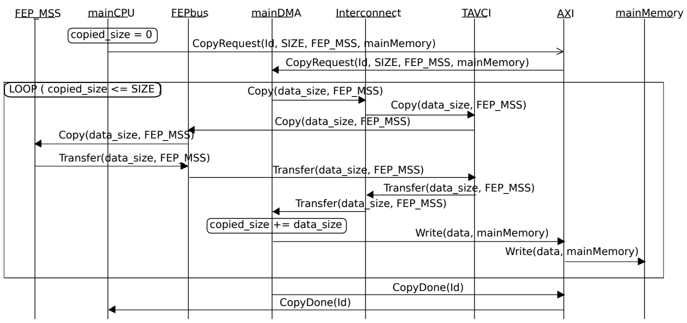


Fig. 10. Communication Pattern of level L3 for edge ed8

to mainDMA via the bus AXI. mainDMA then reads data from the source memory FEP_MSS via the transport network made up of Interconnect, TAVCI and FEPbus. The read data are stored in a FIFO in mainDMA and then written to the destination mainMemory via the bus AXI. The latter actions are iteratively executed until all data have been transferred according to the loop conditional action and an acknowledgment is sent to mainCPU by FEPdma via AXI bus.

VI. CONCLUSION AND FUTURE WORK

This paper described Communication Patterns, a novel feature for modeling the behavior of complex communication schemes at system-level. While CPs are presented here for SDR systems, they also represent a solution for other dataflow processing domains, e.g., image processing. As illustrated by the case study, our contributions provide the expressive power to describe complex multi-point transfer schemes that cannot be captured by traditional graph-based approaches that normally call for a re-design of the application. Moreover, CPs make application models portable by eliminating the need to adapt the latter to the addressing capabilities of a specific platform, thus leading to finer and faster designs. This paves the way to a novel modeling approach where waveform and platform graphs are disjoint and information contained in the models is separately mapped on the architecture in a extended Y-Chart fashion. The main gain of Communication Patterns results in the developer having complete control over communications, independently with respect to the rest of the application and portably with respect to different platforms. Our current works are dedicated to integrate CPs and the proposed methodology in TTool. At this stage, the advantage of using CPs is given in terms of modeling. However, once completely integrated into TTool, CPs will allow the user to separately investigate and extract information about the performance of data/control transfers, either via simulation or via code generation, without changing the rest of the system's models. Our future works will provide a more complete description of mapping and refinement rules for CPs. For instance, level L3 will be extended into intermediate mappings where each class of supernodes will be addressed separately, in order to target buffer addressing and indexing. These aspects are currently handled in DiplodocusDF waveforms by decorating edges with addressing parameters (e.g., read/write memory offsets) and instantiating dedicated nodes (e.g., OVLP in Fig. 7) that define how data are stored in memories. Our aim is to embed these functionalities in communication patterns and processing operations in order to achieve portable waveforms made up of pure dataflow representations like the one in Fig. 6.

REFERENCES

[1] J. Mitola III, *Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio*, Ph.D dissertation, Royal Institute of Technology (KTH), May 2000.

[2] The SDR forum, *http://www.wirelessinnovation.org*, [retrieved: December, 2013].

[3] B. Trask, *Two Thirds of SDR is SD*, in SDR WInnComm, 2010, pp. 197-198.

[4] O. Anjum et al., *State of the art baseband DSP platforms for Software Defined Radio: A survey*, in EURASIP Journal on Wireless Communications and Networking, vol. 2011, no. 1, 2011.

[5] D. C. Schmidt, *Model-Driven Engineering*, in IEEE Computer, vol. 39, no. 2, 2006, pp. 25-31.

[6] J. M. Gonzalez Pina, *Application Modeling and Software Architectures for the Software Defined Radio*, Ph.D dissertation, Telecom ParisTech, May 2013.

[7] The UML profile for MARTE, *http://www.omgmarte.org*, [retrieved: December, 2013]

[8] S. Lecomte, S. Guillouard, C. Moy, P. Leray and P. Soulard, *A co-design methodology based on model driven architecture for real time embedded systems*, in Mathematical and Computer Modelling, vol. 53, no. 34, 2011, pp. 471-484.

[9] S. Rouxel et al., *UML Framework for PIM and PSM Verification of SDR Systems*, in SDR Forum Technical Conference, 2005.

[10] C.C. Shen, W. Plishker, H.H. Wu and S.S. Bhattacharyya, *A Lightweight Dataflow Approach for Design and Implementation of SDR Systems*, in SDR-WInComm, 2010.

[11] C. Moy and M. Raulet, *High-Level Design for Ultra-Fast Software Defined Radio Prototyping on Multi-Processors Heterogeneous Platforms*, in Advances in Electronics and Telecommunications, vol. 1, no. 1, 2010, pp. 67-85.

[12] The AAA Methodology and Syndex, *http://www.syndex.org/*, [retrieved: December, 2013].

[13] N. -ul. -I. Muhammad, R. Rasheed, R. Pacalet, R. Knopp and K. Khalfallah, *Flexible Baseband Architectures for Future Wireless Systems*, in EUROMICRO DSD, 2008, pp. 39-46.

[14] L. Apvrille, W. Muhammad, R. Ameur-Boulifa, S. Coudert and R. Pacalet, *A UML-based Environment for System Design Space Exploration*, in IEEE ICECS, 2006, pp. 1272-1275.

[15] TTool, *http://ttool.telecom-paristech.fr/*, [retrieved: December, 2013].

[16] D. Knorreck, L. Apvrille and R. Pacalet, *Formal system-level design space exploration*, in Concurrency and Computation: Practice and Experience, vol. 25, no. 2, 2013, pp. 250-264.

[17] C. Jaber, *High-Level SoC Modeling and Performance Estimation Applied to Multi-Core Implementation of LTE eNodeB Physical Layer*, Ph.D dissertation, Telecom ParisTech, September 2011.

[18] B. Kienhuis, E.F. Deprettere, P. Van der Wolf and K.A. Vissers, *A Methodology to Design Programmable Embedded Systems - The Y-Chart Approach*, in IEEE SAMOS, 2002, pp. 18-37.