ARAGORN: How to Manage Cognitive Radios

P. Mähönen, M. Petrova, RWTH Aachen University, Wireless Networks, Aachen, Germany

Summary / Abstract

We summarize in this paper some of the goals and early results from ARAGORN project, which is implementing a key technological parts of cognitive radio prototype. We start by shortly discussing about different cognitive radio definitions. We specifically introduce a concept of Cognitive Resource Manager (CRM) that has been introduced by RWTH Aachen University as an architectural framework towards implementing flexible cognitive radios. We further describe the ongoing work that focused on defining necessary interfaces to manage complexity of flexible cognitive radio platforms. We also discuss about implementation challenges in our OFDM-based transceiver architecture.

1 Introduction

The cognitive radio technology has received a tremendous amount of interest during the last few years. The original cognitive radio (CR) approach was introduced by Mitola in [1], where the main idea was to make the radios smart and able to learn, decide and act based on the environmental perceptions. In a certain sense the use of term has inflated so that its meaning is not any more very clear. There are fundamentally two different types of cognitive radios in the literature. The first type could be described as spectrum agile and opportunistic radio. A very promising research has been done in the field of dynamic spectrum access (DSA) to build systems that will allow cooperative behaviour between the primary frequency users and opportunistic secondary frequency users [2]. The cognitive radios of second type are based on the Mitola's original, much more challenging, vision on combining the machine learning and model based cognitive radios. Several research groups have extended the cognitive radio paradigm in the context of wireless networks by introducing the concept of Cognitive Wireless Networks (CWN) [3][4][5][6].

The work in the Department of Wireless Networks at the RWTH Aachen University is mostly focusing on the second kind of cognitive radio, i.e. building a context sensitive, flexible and cross-layer optimization capable CRs. However, an important part of this work is also to consider a possible support for DSA technologies. The overall cognitive radio research program in Aachen is quite large. One of the corner stone projects for implementation and ISM-band operating CRs is a European Union funded project called ARAGON (Adaptive Reconfigurable Access and Generic interfaces for Optimisation in Radio Networks). The project was started officially in January 2008, although some preparatory work was done by the participants with their own funding before the start of the European Union funding. The project will run 30 months and its partners are RWTH Aachen University (coordinator), STMicroelectronics, Toshiba, Microsoft, University College London, Consorzio Ferrara Ricerche, Italy and University Ss. Cyril and Methodius, Skopje. The main technical aim of the project is to combine a strong team both from industry and academia comprising competence at all OSI-layers and implementation technologies, in order to build a prototype implementation of cognitive radio that is able to operate in ISM-bands and also supports connection management towards some licensed band technologies.

The rest of this paper is organized as follows: Section 2 gives an overview of the resource management in the future cognitive radio systems and introduces our cognitive resource management framework. Furthermore in this section, we discuss the architecture of several vital building blocks of the framework and highlight some real implementation aims of the project In Section 3 address some aspects of our work in cross-layer optimization using Genetic Algorithms (GAs). Finally, in Section 4 we conclude the paper.

2 Cognitive Resource Management

Providing various technical solutions for making the wireless embedded system self-configurable, context aware and optimized with respect to, e.g., throughput, delay, spectrum efficiency, energy, QoS and scalability has become a research trend in the last few years. Following the Mitola's groundbreaking work, we are aiming at to apply the cognition principles at systems level in order to enable cross-layer optimization and adaptive resource allocation. By cognition we mean ability of the system to use context information and machine learning techniques in the process of adaptation, auto-configuration and optimization. The design is done being aware of the practical complexity and difficulty on taking in account widely different characteristic time-scales of different protocol layers as pointed out in [7]. We have chosen to build a full optimization software framework for cognitive radios. The framework is called Cognitive Resource Manager (CRM), and it is providing the basic primitives, such as support for scheduling, conflict resolution and control loop establishment, that are needed to implement an efficient resource manager for cognitive radios. This approach is partially related to the concept of classical Radio Resource Managers (RRM) but the scope of our CRM is broader and most importantly the system is able to learn from its previous experience and to adapt flexibly based on context information.

2.1 Cognitive Resource Manager Framework

The conceptual architecture of the CRM framework is shown in Figure 1. The CRM is a central entity that can perform cross-layer optimization and efficient resource management, based on the information it gets from the application layer, the underlying data link and networking layers and from the operating system. The adaptation is done using adequate advanced reasoning methods. For example, CRM can optimally manage spectrum, flexibly adapt link parameters or allow the best possible resources for the applications running on top. The cognitive component in the system comes from the machine learning and optimization algorithms which will provide the capability to learn and reason. The key idea behind CRM is to achieve a close coordination between cross-layer optimization functionalities and machine-learning-based algorithms. The important feature of CRM is its modularity. It is not based on monolithic design, but rather looks like a specialized real-time operating system, which supports a building of cross-layer optimization based control loops. Some of the design principles of CRM are based on lessons learned from building complex robotics control systems and designing compact operating system.

In the optimization process CRM follows the approach based on using utility functions to describe the goals for optimizers. The design philosophy is in part inspired by He *et al.* contribution [13]. Instead of trying to design a highly optimized single purpose flexible radios system, we are aiming to provide system that can be modified flexible to support different situations. The relative tradeoff between the fairness and efficiency of resource allocation based on the QoS requirements of each user is handled based on estimating utility functions on user, link and network levels.

Introducing the learning capability and intelligence into the optimization process is a step forward in building auto-configurable and high performance wireless networks. One of the challenges is to provide adequate interfaces to exchange information between different entities within a single communication device and between them. The CRM obtains information from the lower layers through so called Unified Link-Layer API (ULLA) [10], from network and transport layer through the Generic Network Interface (GENI) and from applications through the Common Applications Requirements Interface (CAPRI). This information is used in the CRM by machine learning algorithms to make systems situation aware and to build a world model. Using all three generic interfaces CRM can also feed back information to the four major protocol layers and update the settings when necessary.

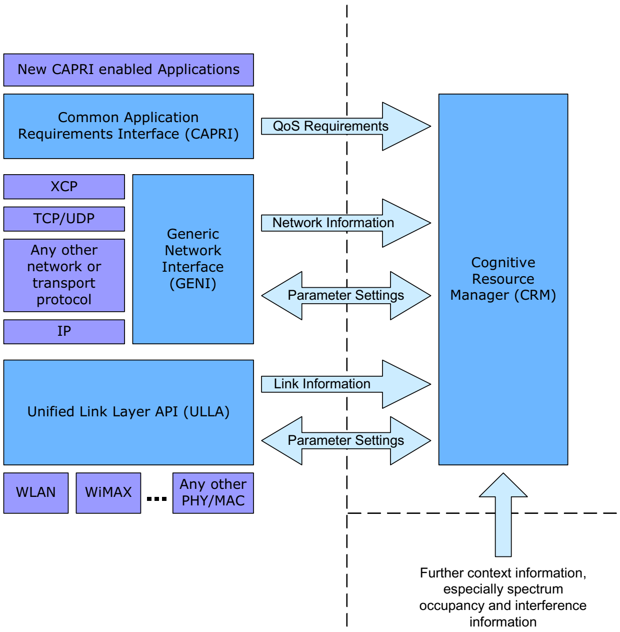

Figure 1 System architecture of a CRM enabled node

The CRM concept is, as far as we are aware, quite novel. There is, however, a similar approach introduced by Rieser [8] at Virginia Tech., known as Cognitive Engine (CE). There are some clear differences between the approaches; ours being currently more higher-OSI-layers and cooperative-networking based, although lately the frameworks have been started to converge towards a common vision. As mentioned above the CRM architecture is highly modular and it is based on flexible and extendible interfaces. One of the unique factors of CRM is that the optimization framework is strongly inspired by decomposition principle [18]. The underlying idea of CRM is that the individual optimization routines are actually seen as libraries or toolboxes. Coordination is enabled by the CRM which acts almost like *an operating system for cognitive radio.* The CRM is specifically responsible for scheduling and composition of different control loops. The optimization problems can be thus decomposed freely and efficiently.

When it comes to the actual implementation of the whole framework, we are currently considering using *subsumption architecture* combined with real-time operating system core as the starting point of CRM design [11]. The subsumption architecture is used to abstract the optimization goals into different layers. We strongly believe that the design of the cross-layered optimizer fits naturally with the subsumption based architecture. The protocol layers can be directly mapped to different layers of the subsumption architecture, with each layer solving a goal by abstracting the parameters of the lower level layers. For example, the source-rate control of TCP in the transport layer controls the injected traffic rate by abstracting the routing, data-link and physical layers. Furthermore the information flow between the layers is abstracted to remain at a suitable level so that each layer has the right amount of information to solve its optimization problem without burdening it with multitude of parameters which makes the optimization problem too complex. This is especially crucial in cognitive radio context since we have a large number of parameters that can be tuned. Hence, subsumption architecture aids in formulation of simplified and tractable optimization problems. Last but not least subsumption architecture leads to a modular design of the optimization layers. Each optimization problem can be easily plugged within the existing framework by identifying the appropriate layers and the input desired from the adjacent layers.

2.2 Interfaces

Generic and extensible interfaces are the key elements in the CRM framework. We are in the process of defining unified programmable interfaces for accessing information from different layers of the protocol stack. This is very important in order to provide easy flow of information and enable flexible configuration of different entities. Part of this work is done together with industry. Generally we feel that the current state of the art, where many of the interfaces are proprietary and mostly hidden from programmers can not enable development of smart radios and applications. However, the design of open interfaces is not easy due to commercial reasons, and also a number of research problems are related to the fact. One issue is how to make interfaces flexible and extendable. Another challenge is that a good interfaces should support different event based notifications and abstraction levels. Many of these challenges are studied in the ARAGORN project.

Flexible and feature rich interfaces enhance the scalability of the system and makes it "future proof" by providing powerful mechanism to introduce new technologies. In this section we will shortly discuss about the requirements of having CAPRI, GENI and ULLA in our framework. We are currently working on the CAPRI and GENI specifications whereas ULLA was originally designed in the earlier GOLLUM project [10]. The design of ULLA has also strongly affected on our views on how to build extendable APIs in general.

Different types of applications have different QoS requirements and expectations. For some applications speed and latency are crucial (e.g. video streaming or gaming), some applications can work only within a certain range (e.g. sensors), some depend on the type of device on which they are deployed (e.g. downsized version of an application on a PDA as compared to the laptop) and some need to be prioritized (e.g. emergency). By knowing these different QoS requirements the CRM can configure resources optimally. The main function of CAPRI is to communicate this information to the CRM. One very specific feature, and an innovative part, of CAPRI is that it allows utility (function) based description of requirements. Instead of using SLA (Service Level Agreement) type of static and parameters based architecture we allow applications to describe requirements as utility functions which describe value of "satisfaction" that the application derives from different situations. These functions are parsed by CRM through a simple interpreter. Moreover, CAPRI will allow that functions, or requirements, can also include some more lower-level constraints, e.g. radio interface or terminal equipment specific parameters. Thus in principle CAPRI allows applications to describe their utility as a set of optimization problem goals and constraints. This gives a possibility to define user or application centric constraints, requirements, utilities and recommended hints towards radio equipment. The utility based approach is a powerful concept that has been recently introduced also in the Internet research community [12][13][14].

In a similar fashion like CAPRI, the Generic Network Interface (GENI) will be able to gather information about the services, protocols, topology, etc. from the network and transport layers and expose this information to the CRM. GENI will allow CRM to configure the network and transport layer parameters accordingly. Furthermore the network layer contains information about the statistics of the network; these statistics can be used by the CRM to optimize the network in cooperative fashion.

For intelligent use of radio resources, knowledge on the available communication options and radio resources as well as resource usage patterns are very useful. The CRM provides central depository for information exchange. At the physical and link layer, ULLA has been identified to be a suitable candidate as generic link layer interface for ARAGORN.

ULLA is an interface we originally developed with our research partners to retrieve data link layer and physical layer information independent of the underlying technology (WLAN, Bluetooth, GSM, ZigBee, 2G/3G, etc.). Applications using ULLA do not have to be aware of the communication technology underneath since the interface offers universal monitoring and controlling interference for all technologies. The platform independence also gives us the freedom to easily extend the API for the new upcoming standards. The interface is not a minimum common set between technologies. There are different parameters available for different technologies, but the API is using always the same query command structure towards any technology. Moreover, in the case of common parameters, say finding out the bit-error rate of the link, the API query is always a similar looking. We anticipate that ULLA might be extended in this project to accommodate special requirements of the ARAGORN project not yet covered by the current design. Detailed description and the performance analysis of the API can be found in [15]. The core parts of ULLA reference implementation were published as an open source code through SourceForge.

The ULLA architecture is shown in Figure 2. It is composed of three main parts: the applications using ULLA which are also known as *link users*, the *ULLA core* and the *link providers,* which are abstractions of the network interface card in the device.

The ULLA core is the main part of the API. Its main task is to process the requests and the updates from the link users and link providers respectively. The interaction between the link providers and the ULLA core and the link users and the ULLA core is done through two independent interfaces, namely *link user interface* and *link provider interface*. Through the link user interface the applications can issue *commands* to the link providers, e.g., to scan the available links or request specific information (attribute values) from the link providers using *queries*. A ULLA Query Language, which can be seen as a subset of the SQL, was developed to specify the queries. For example, an application can request for all links that have latency smaller than 150 ms with the query: *SELECT linkId FROM ullaLink WHERE txLatency* < 150. Furthermore applications can subscribe for *notifications* if they want to be informed regularly about the updated values of the attributes or when the values reach certain threshold. The link provider interface tasks correspond to the ones of the link user interface. It is used for asking for updates in attribute values, for issuing commands, and for registering and deregistering links and link providers.

The current status with the interfaces is that ULLA has a relatively stable design and there is an existing code base for the future development. The most intensive research is focused towards CAPRI. The reason for this is that it is a strategically important part of interface structure of our cognitive radio framework, but also due to fact that we need to find new approaches to enable such interface. The traditional fixed API or even query engine type of approach adopted with ULLA is far from ideal for our purposes in the case of expressing utility function based objectives.

2.3 Implementation Focus

ARAGORN is a project which is strongly committed to build prototype implementation(s) of the key components of CR. One of our main objectives is to develop the first prototype of CRM and test those in small scale network. This work requires also definition and implementation work of key-interfaces. In the implementation domain the project is mainly limited to consider only different ISM-bands as its operating frequency bands. However, the implementation work at the RWTH Aachen University is also considering a wider flexible radio context.

The networking concepts will be prototyped with USRP-boards (also known as gnuRadio platform) [16]. However, there is a parallel and complementary work in progress in which we use much more powerful MIMO/OFDM-platform that allows the development of more flexible and advanced PHY-layer concepts. The chosen platform for OFDM-research is so called WARP-board from the Rice University [17].

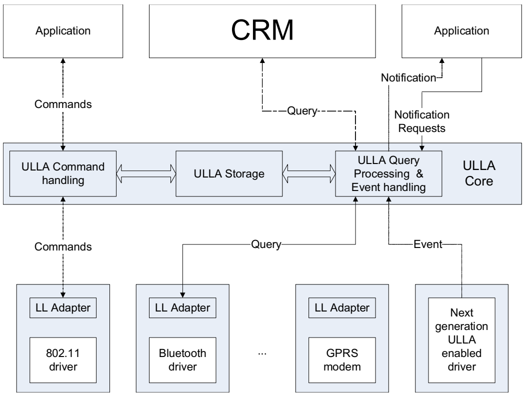

Figure 2 ULLA architecture

The current aim of the ARAGORN project is to demonstrate the first early prototype, which should include interfaces and higher layer functionalities, with IEEE802.11 devices and USRP-platform in 2009. Furthermore, the prototyping work towards OFDM-platform and mode advanced PHY-layer and MIMO-concepts by RWTH should be available for the first evaluation also in 2009.

3 Optimization with Genetic Algorithms: A Case Study

In this section we shortly describe one of the optimization toolboxes that are planned to be a part of final CRM framework. The description is given as a sort of case study, because a part of the system has been already implemented. The method we have chosen to introduce is Genetic Algorithm based optimization, since this has received a lot of attention. Many CR research groups have seen it as a promising heuristic optimization method.

Let us assume that we have the possibility to easily access and change parameters of wireless system in the different layers of the protocol stack. In principle, this gives us the opportunity to optimize the performance of the radio system in terms of end-to-end throughput, delay, power consumption, etc. by finding the right parameter settings. However, performing an optimization to achieve a certain quality of service is not a straightforward process due to different interdependences. In the simplest case the optimization in the wireless communication systems is oriented towards a single goal, for example maximization of the overall throughput or minimization of the bit error rate of single link. As the number of application running on the wireless systems and the variety of the services offered to the end user are constantly increasing, the optimization of the system performance gets another dimension. Very often there can be several objectives from different applications, and these should be optimized concurrently and as fairly as possible. Moreover the objectives are usually functions of not just one but a set of parameters. Further complication is generated by that fact that the final optimization functions tend to be often non-linear. The fact that several of these objectives might conflict with each other adds an additional complexity to the problem. For example, if we try, at the same time, to minimize the latency and the bit error rate (BER) a conflict will occur because the packet size affects both objectives in an opposite way. In order to come to an acceptable solution one should make a compromise and put more weight on one or on the optimization objective [20]. Theoretically, in case of constraint multi-objective optimization the optimal solutions lie on the so called Pareto optimal front. Further improvements are not possible due to the dependencies between the multiple objectives.

Solving a cross-layer optimization problem in such a setup is difficult. Most often analytical solutions are not possible. Also many traditional numerical methods might not be fast and flexible enough, or those are not able to converge to reasonable solutions. In some cases the global optimization problems might even be NP-hard. We have addressed this issue by applying evolution inspired machine learning methods such as Genetic Algorithms (GAs) to solve CR optimization problems [19]. Rieser [8] was one of the first researches applying GAs for cognitive radios. In our case study the goal was to minimize the power consumption, maximize the throughput, minimize the bit/packet error rate and minimize the transmission delay in wireless Orthogonal Frequency Division Multiplexing (OFDM) system using Carrier Sense Multiple Access with Collision Avoidance protocol (CSMA-CA) *at the same time*. It should be noted that we are making aiming at to medium-level time-scales, i.e., we do not try to optimize radio parameters against fading effect, but setup the longer time-scale target values for control loops. We aggregate the four objectives using a weighted sum approach to provide a single fitness function for the GA. This method has been also successfully implemented by Newman *et al.* [21]. Our approach maximizes the sum of the positively normalized, weighted, single objective scores of the parameter set solution. We consider parameters from physical, link and network layer in our cross-layer optimization problem, parameters such as modulation order and transmission power per subcarrier, minimum contention window, packet size, just to name few are considered in this framework. For technical details, the reader is referred to [22].

The obtained results show that GA-based optimization methods can beneficial in the context of cognitive wireless communication systems. Our GA implementation achieves comparable performance to the traditional methods such as bit loading with moderate complexity [23]. Furthermore the GA is scalable and can easily handle a large set of parameters. The convergence time is also very low and the algorithm can reach the optimal Pareto front quite fast. In this paper we do not want to advocate GAs as a magical optimization technology, but simply want to stress that it is a one heuristic optimization method which is worth of consideration, and definitely useful to have in the toolbox library of CRM.

Although GAs are promising on-line optimization tools we consider these mostly as off-line background computational tools, in the CRM context. GA modules can successfully work in background to explore *optimization possibilities*. The results can be used as an input to our world-models and the actual *dynamic* adaptation can be handled, e.g., through decision trees and fuzzy logic controllers. This is due to the fact that latter models provide computationally much faster adaptation in practical algorithms.

In order to better understand the potential of GAs in the future cognitive systems more extensive studies with large-scale networks and real hardware testbed implementations are required. In that domain, inside ARAGORN, we are currently experimenting not only with simulations, but also by using gnuRadio- and WARP-based testbeds.

4 Conclusions

The ARAGORN project itself is an ongoing research project. Hence, one should not make too far reaching conclusions from it. However, based on the early results and lessons learned from our other earlier and on-going research projects, we believe that cognitive radios can provide true benefits in the domain of context sensitive cross-layer optimization. We believe that the cognitive radio principle goes far beyond "simple" dynamic spectrum allocation. In fact, the smart cross-layer optimization might be actually the real main driver towards CRs in the future.

The CR based systems and applications are becoming reality due to emerging SDR based radio platforms. However, one has to face that fact that the system design challenges are not triv-ial ones. Designing and implementing just interfaces that allow the flexible exchange of data and reconfiguration of different parts of radios is a difficult task. Therein one has to also strike a balance between need to provide both enough openness and protection towards some proprietary technologies. The interface design needs to also go beyond the traditional static API-design. Finally, one should take in account that abstraction level needs to carefully chosen. Some other earlier approaches have succumbed under complexity and too heavy implementation, e.g. some of the CORBA based radio architectures are examples on this.

The initial results from ARAGORN-project and also by other research groups are promising. In our opinion the use of modern machine learning algorithms combined with decomposition based optimization can lead to significant improvements in the performance of the wireless systems.

5 Acknowledgements

The authors would like to thank their colleagues in ARAGORN project for fruitful discussions and collaboration. We thank the partial funding from European Union Framework Programme through ARAGON-project, and DFG and RWTH Aachen through UMIC-excellence cluster.

6 Literature

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

[2] Proc. of the 2005 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks. IEEE, November 2005

[3] Thomas, R.; Friend, D.; DaSilva, L.; MacKenzie, A. B.: Cognitive Networks: Adaptation and Learning to Achieve End-to-end Performance Objectives. IEEE Communications Magazine. Vol.44, No.12, Dec. 2006, pp. 51 – 57

[4] Nolan, K.; Doyle, L.: Teamwork and Collaboration in Cognitive Wireless Networks. IEEE Wireless Communications, Aug. 2007, pp. 22 – 27

[5] Mähönen, P.: Cognitive Trends in Making: future of networks. Proc. of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Sep 2004, pp. 1449 – 1454

[6] Petrova, M.; Mähönen, P.: Cognitive Resource Manager: Across-layer architecture for implementing Cognitive Radio Networks. Cognitive Wireless Networks (eds. Fittzek F. and Katz M.), Springer, 2007

[7] Kawadia, V.; Kumar, P. R.: A cautionary perspective on cross-layer design. IEEE Communication Magazine. No. 2, pp.3-11, Feb. 2005

[8] Rieser, C.: Biologically Inspired Cognitive Radio Engine Model Utilizing Distributed Genetic Algorithms for Secure and Robust Wireless Communications and Networking. Ph.D. Thesis, Virginia Tech, 2004

[9] Riihijärvi, J.; Wellens, M.; Mähönen, P.: Link-Layer Abstractions for Utility-Based Optimization in Cognitive Wireless Networks. Proc. of Crowncom 2006. Myconos: Greece, June 2006

[10] Sooriyabandara, M.; Farnham, T.; Efthymiou, C.; Wellens, M.; Riihijärvi, J.; Mähönen, P.; Gefflaut, A.; Galache, A.; Melpignano, D.; Rooijen, A.: Unified Link Layer API: A generic and open API to manage wireless media access. Computer Communications. Vol. 31, Iss. 5, 25 March 2008, pp. 962-979

[11] Brooks, R. A.: A robust layered control system for a mobile robot. IEEE Journal of Robotics and Automation. RA-2, April 1986, 14-23

[12] Shenker, S.: Fundamental design issues for the future internet. IEEE JSAC. Vol. 13, No. 7, Sept. 1995, pp. 1176-88

[13] He, J.; Rexford, J.; Chiang, M.: Don't optimize existing protocols, design optimizable protocols. ACM SIGCOMM Comp. Comm. Review. Vol. 38, Iss. 3., 2007, pp. 53-58

[14] Duan, X.; Niu, Z.; Zheng, J.: A dynamic utility-based radio resource management scheme for mobile multimedia DS-CDMA systems. Proc. of Globecom 2002. Vol. 1, Nov. 2002, pp. 804-808

[15] Open source release of the Unified Link-Layer API, http://ulla.sourceforge.net/

[16] Gnu Radio, http://www.gnu.org/software/gnuradio/

[17] WARP, http://warp.rice.edu/

[18] Mung Chiang; Low, S.H.; Calderbank, A.R.; Doyle, J.C.: Layering as Optimization Decomposition: A Mathematical Theory of Network Architectures. Proceedings of the IEEE. Vol.95, No.1, pp.255-312, Jan. 2007

[19] Whitley, D.: A Genetic Algorithm Tutorial. Statistics and Computing. No 4, pp.65–85, 1994

[20] Fonseca, C.M.; Fleming, P.J.: Multi-objective optimization and multiple constraint handling with evolutionary algorithms- Part I. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans. Vol.28, No.1, pp.26-37, January 1998

[21] Newman, T.; Barker, B.; Wyglinski, A.; Agah, A., Evans, J.; Minden, G.: Cognitive Engine Implementation for Wireless Multicarrier Transceivers. Wiley Wireless Communications and Mobile Computing Edition. Vol.7, Iss.. 9, pp. 1129 – 1142, May 2007

[22] de Baynast, A.; Mähönen, P.; Petrova, M.: ARQ-based cross-layer optimization for wireless multicarrier transmission on cognitive radio networks. Computer Networks.Vol 52, No.4, pp. 778-794, March. 2008

[23] Fischer, R.; Huber, J.: A new loading algorithm for discrete multitone transmission. In: Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM). Vol. 1, pp. 724–728, 1996