A Path Forwards for Cognitive Radio Research

(Invited Paper)

Dr. John M. Chapin
Vanu, Inc.
Cambridge, Massachusetts, USA

Dr. Linda Doyle
CTVR, Trinity College
Dublin, Ireland

Abstract—The Cognitive Radio research field has seen great enthusiasm and rapid growth in the number of active researchers, papers published, conference sessions and journal special issues. This paper presents the author's views on research topics and methodologies that will sustain momentum in the field over the next five years, particularly focusing on ways in which reasoning and learning may be effectively exploited and evaluated.

Keywords—radio communications systems; cognitive radio; dynamic spectrum access

I. INTRODUCTION

Cognitive Radio research began in earnest with Joseph Mitola's thesis in 2000 [1] and a US FCC policy statement of interest in 2002 [2]. Academic work accelerated with multiple conference sessions and workshops through 2004, leading to dedicated conferences beginning in 2005, including IEEE DYSPAN and CROWNCOM.

The initial pulse of interest by government and military leaders appears to be reaching its natural end. The US FCC has terminated one of its most aggressive cognitive radio policy initiatives, related to the use of interference temperature in spectrum management [3]. The DARPA XG program, which has funded the most advanced prototyping work in the US, will likely end in 2008. The first phase of the major European End-to-End Reconfigurability project (E2R) will finish in 2007.

From one point of view, the cognitive radio field is vigorously transitioning to commercialization. This is driven by the opportunity to exploit secondary access to TV whitespace for broadband wireless networks, which appears likely to become legal in the US from February 2009 and in other countries within a few years later.

However, the form of cognitive radio that is being commercialized for TV whitespace networks is somewhat limited in scope. Much of the potential benefit envisioned in academic research to date will not be delivered in these networks. In particular, the use of reasoning and learning techniques is conspicuously absent from current TV whitespace system proposals.

It appears to be an appropriate time for cognitive radio researchers to consider how best to sustain academic research momentum, and funding agency interest, in the more sophisticated aspects of the technology that are not being commercialized. This paper attempts to answer these questions.

II. TYPES OF COGNITIVE RADIOS

Cognitive Radio (CR) is generally defined as a type of radio in which communication systems are aware of their environment and internal state and can make decisions about their radio operating behavior based on that information and predefined objectives [4].

For the purposes of this paper we differentiate three levels of CR systems.

A *basic CR* is a radio system that senses and adapts to its environment, but does not necessarily use explicit reasoning or learning techniques. At the most primitive level, a 900 MHz ISM band cordless phone counts as a basic CR. At the start of each call it measures a number of available channels and selects the one with the lowest noise or interference level. Most current radios use some level of basic CR techniques in their operation. A more sophisticated form of basic CR is a dynamic spectrum access network, where nodes exchange spectrum measurements and run a distributed algorithm to decide which channels are safe for secondary use.

A *reasoning CR* is a system that adds deductive inference to improve on a basic CR. For example, the DARPA XG radio automatically determines legal limits on secondary spectrum operation through deduction from rule sets and information about current conditions [5].

A *learning CR* is a system that updates its decision-making based on the results of prior actions [4]. Standard AI techniques that may be used in learning cognitive radios include case-based learning and knowledge-based learning [6].

III. BARRIERS TO REASONING AND LEARNING CR

There is no question that basic CR techniques are highly beneficial for current and future radio networks. While much remains to be done both in technical and policy domains regarding dynamic spectrum access networks, the use of basic CR techniques for performance optimization within statically allocated spectrum is universal, sophisticated, and the subject of significant ongoing research and commercial investment.

However, despite enthusiasm from the research community in recent years, it is not so clear whether reasoning CR or learning CR will provide end-user benefits. The cost/benefit tradeoff of these techniques is unknown. Costs include implementation complexity, runtime overheads, and potentially

slow response to quickly changing conditions. As for the benefits, no "killer app" has yet been identified. (The next section gives some ideas for where the killer app may be lurking.)

While most of the potential costs are straightforward and widely known, there is one cost of reasoning and learning CR that we feel is significant and is underappreciated by the research community. That is the issue of the potentially unpredictable behavior of reasoning and learning CR systems.

Predictable behavior is highly prized in radio systems. The reasons are straightforward. Users need to be able to predict the level of service they will achieve with a communications network, to determine whether to trust their application to it and to decide how much they are willing to pay for it. Users also find unpredictable systems frustrating to use. Designers need a basis for choosing among design options and must assure that the system will meet its performance commitments. Regulators need to protect users against unsafe radiation levels and to protect other radio systems against harmful interference.

As a result, reasoning and learning techniques that reduce predictability will face significant barriers to adoption. In our view this is a fundamental constraint on the CR field. The potential for unpredictability should be considered at the early stages of any reasoning or learning CR research project. Methods to mitigate or bound the unpredictability should be developed as part of the project. A useful approach, perhaps not achievable in all situations, is to baseline a non-reasoning and non-learning design and provide a guarantee that the CR design never performs worse than the baseline.

One way to mitigate unpredictability is to keep the number of critical state variables small and to assure that all of them are externally visible. This enables the various stakeholders to form effective models that predict the system's behavior in different circumstances. As a simple example, consider the signal level "bars" shown on the display of almost all cell phones. Knowing the basestation signal level helps the user predict how a phone call will behave. This reduces frustration and supports external adaptations such as walking towards a window until system performance is acceptable. Similarly, making the inputs to the reasoner of a reasoning CR externally visible and comprehensible makes the device or system more predictable and hence more useful.

For learning CRs, the level of predictability can vary widely. At the predictable end of the range are systems that use learning solely to optimize pre-specified behaviors, for example by observing application communication patterns to improve MAC timer settings. At the unpredictable extreme are systems that dynamically generate new behaviors never specified by the designer. The challenge for research on learning CRs is to develop systems that behave as close as possible to the predictable end of this range, yet exploit learning to deliver improvements over the best non-learning design.

IV. APPLICATIONS OF REASONING AND LEARNING IN RADIO SYSTEMS

Today's state-of-the-art networks use basic CR techniques extensively. Although not called cognitive radios, most current networks sense and adapt extensively at the link level to optimize performance for varying channels, load types, and application requirements. Performance is known to be close to the achievable Shannon bound at the link level for carefully designed networks. As a result, the CR research field needs to look to other system aspects or to systems different in some fundamental way from current well-studied networks, to find opportunities for significant end-user benefits.

This section presents an admittedly personal view of areas where reasoning and learning in CR systems can be reasonably expected to deliver significant end-user benefits, without making system behavior unpredictable in ways that harm utility or safety.

A. Automate manual user/operator tasks

Reasoning and learning may be helpful to implement automatic responses to communications performance problems that today require manual response by the user or operator. Rather than seeking exotic problems, the highest payoff results are those that address common problems encountered every day by a large number of users, for example:

- Optimizing end-user QOS as a mobile device moves through predictable spatial coverage variations.

- Minimizing end-user cost when multiple access networks or plans are available offering different guarantees and tradeoffs.

- Automatically adapting system configuration to maintain maximum performance or coverage when faults occur.

B. Tolerate interference better

Reasoning and learning techniques can also provide significant performance improvements if used in characterizing and learning the behavior of other (interfering) systems and adapting at all layers to work better in that context. Current systems do this implicitly when adapting their PHY constellations but their adaptations are limited.

Current systems have largely been designed against an AWGN model. Most human-caused interference has significant higher level structure (e.g. cyclostationarity at the PHY layer, periodic transmissions at the MAC layer) that can be detected by a cognitive radio and used to minimize disruption caused by the interference.

Current systems have also largely been designed against known scenarios, for example a specified set of emitters and waveforms that are the system's neighbors in frequency or geography. When the neighbor set changes after deployment, the system's interference tolerance capability degrades. A CR that can adapt its interference tolerance over time to the changing environment will perform much better.

C. Address the novel problems of novel networks

Another fertile opportunity for high research payoffs is to focus on the performance of networks that are different in some fundamental way from today's well-studied cases (stationary wireless LANs, cellular mobile with fixed infrastructure). For example there are large unsolved problems in power minimization and outage rate minimization for MANETs; perhaps reasoning or learning can be a key part of the solution. Similarly the hot topic of disruption-tolerant networks may offer opportunities.

Dynamic spectrum access networks face particularly large variations in link-level pairwise bandwidth and connectivity due to varying spectrum availability. Delivering acceptable performance may require learning and reasoning: at the link level, to predict changes in the current spectrum hole and quickly discover alternate spectrum holes, or at the network level, to predict developing bottlenecks or derive effective responses.

D. Reduce or eliminate manual configuration

Reasoning and learning CR techniques can be exploited to implement a system that achieves the same performance level as a well-tuned traditional radio or network, but with less manual configuration effort required. A cognitive radio can do this through deducing the optimal configuration for its operating environment or though learning the user's usage patterns and needs.

Reduced configuration effort can have high benefits. For example the military must deploy wireless networks rapidly, in a fluid operational context. Automatic configuration mechanisms that assure network performance without human intervention significantly increase operational flexibility. In many commercial settings, networks can be too complex to optimize manually. An even more extreme case is future mesh networks that embody the vision of low-cost viral deployment in unlicensed or secondary access spectrum. These networks must be fully automatically configured to support their fundamental business model.

However, there is a significant challenge in this area. Often when engineers try to build a system that learns and adapts, they end up creating one that is harder to use and significantly more annoying than simpler systems. "Just do what I tell you to do!" is a common operator complaint. Research focused on automatic configuration and configuration improvement through learning should compare its results to a good scripting system, database based control layer, or graphical user interface.

E. Ensure Fair Coexistence

If dynamic spectrum access succeeds and becomes widely deployed, we can expect that multiple secondary networks will contend for spectrum access. In this context, the question arises: how should the systems be designed so they share spectrum resources appropriately? The desired sharing may be an equal division of spectrum resources among the systems, or an unequal one that maximizes some figure of merit. Achieving the desired sharing is called fair coexistence whether the division of resources is equal or not.

It is not clear whether current dynamic spectrum access system designs will achieve fair coexistence. For example, some TV whitespace network designs propose the use of sensitive cyclostationary detectors tuned for digital TV signals to determine whether a channel is occupied or available. What happens when that channel is occupied by another secondary user, one whose signal doesn't look like a digital TV signal? Multiple secondary users contending for access would have a high probability of jamming each other.

Fair coexistence is challenging because dynamic spectrum access networks cannot make assumptions about the characteristics of the occupying signal. A pure energy detector is known to have low sensitivity and a high vulnerability to hidden node problems. There are great research opportunities here for reasoning and learning. For example, the multiple nodes in a network should be able to cooperate to learn the characteristics of a new occupying signal that must be avoided, gradually improving the quality/fairness of coexistence while reducing the cost of detection.

F. Handle faulty/malicious nodes in a cooperative network

Cooperation among nodes in a CR network is a widely studied technique. For example, some researchers see it as a fundamental requirement to solve the hidden node problem in spectrum sensing [7].

In a real world deployment, not all the nodes in a CR network can be trusted to send valid data to their cooperating peers. This may occur because of a fault such as a defective sensor or software bug, or because an intruder takes over or spoofs a node and injects bad data into the CR control plane. These cases must be handled for a cooperative CR network to be reliable. For example, in 1973 an ARPANET router suffered a memory hardware error that caused it to claim a zero hop count path to all other nodes on the network. As a result of cooperative shortest-hop routing algorithms, all traffic was sent to that router and the network ceased to function [8]. The routing protocols had to be modified to reject obviously corrupted routing data inputs.

Ideally the CR network should be insensitive to arbitrary misbehavior of a small number of nodes, and its behavior should gracefully degrade as the number of problem nodes grows. This is formally known as Byzantine Fault Tolerance. While some practical implementations have been developed, the ideal response is difficult to implement and normally has high runtime costs [9]. Most real-world distributed systems offer some protection but don't handle fully Byzantine misbehavior by nodes in the network.

If reasoning can be used to effectively distinguish most corrupted messages from most valid ones, or learning the current situation and/or application goals can help reject malicious inputs, these CR techniques will have high value for end users. This value is not directly a performance improvement, but it can reduce the performance overhead required by other methods of tolerating bad cooperators, such as Byzantine agreement algorithms. Furthermore, using reasoning or learning for this will reduce the risk level of cooperative CR networks, and thereby make them useful for a wider range of applications.

V. COLLABORATING FOR RESEARCH PROGRESS

The previous section identified specific application areas of interest to the authors where reasoning or learning CR is reasonably likely to provide concrete end-user benefits compared to basic CR. We now move from application areas to recommendations for how researchers in the CR field can collaborate to support effective and rapid progress on applications like those described.

All the reasoning and learning application opportunities described in the previous section, as well as many others, can be unified into a single abstract topic. We call this the "universal question" of reasoning and learning CR research:

Can suitable local and network wide environmental observations be used in appropriately designed reasoning and learning processes to produce executable unilateral and multilateral (re)actions that bring significant user benefits?

Expressing the universal question this way is useful because it highlights the critical importance of the *research framework* in making effective progress on cognitive radios. We cannot examine the reasoning and learning processes, and the actions produced by them, in isolation. The framework provides the rest of what is necessary to investigate the question:

- A simulator or test platform to execute the CR device or network, for example [10]–[13].

- An environmental model or test setup, to provide observational data.

- A traffic model or test workload, together with spatial distribution and motion of the nodes.

- A clear definition of what metrics will be used to evaluate user benefits.

It is important for the vitality of the cognitive radio research field that, to the extent possible, research frameworks should be shared across research projects working on related topics. The significant benefits of sharing research frameworks include:

- Research results are comparable across projects, permitting synthesis and evaluation.

- Research ideas are more easily generalized and reused, reducing duplication of effort.

- Users can more easily determine which research results are relevant to their needs, improving the potential for technology transfer.

- Progress in the field is more quantifiable and visible, improving the prospects for continued funding support.

- The cost and time of developing research framework components, or adapting components from other research fields for the specific needs of CR research, can be amortized.

We believe it would be valuable for the CR field to hold a workshop in the near future specifically focused on developing and publishing common research frameworks. It is clear that there should be multiple frameworks rather than a single one.
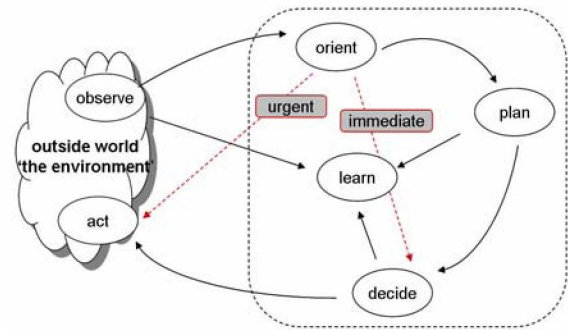

Figure 1: The cognitive cycle in a single cognitive radio

Ideally, the result of the workshop would be a relatively small number of well-documented frameworks spanning the types of research done by the bulk of researchers working in the field. More likely will be an amorphous set of partial frameworks, some of which will be used and evolved in future research projects. This would be a successful outcome in our view. The goal of the workshop should be that the CR field develops a habit of publishing, reusing, and improving research framework components over time.

The following sections highlight one critical area for framework discussion and publication: how the environment should be modeled when performing a simulation based evaluation of a CR algorithm, device or network. Our discussion is not intended as a definitive answer to this challenging question, which in any case cannot be answered outside the context of the specific research issue being investigated. Instead, this analysis is meant to stimulate the kind of discussion and reflection that would lead up to the recommended workshop.

Other topics not covered here that are critical for the workshop include: how should "significant benefits" be defined for different users and applications, and how should the environment be emulated in experiments that use real-world execution platforms rather than simulation.

A. The Need for Appropriate Environmental Models

CR is intrinsically an open system: it is affected by and affects its surrounding environment. (After all, the point is to be aware of the environment and adapt to it somehow!) As a result, the way the environment is modeled is absolutely critical in performing CR research.

To provide a useful perspective on this point, we can review the oft-used fundamental diagram of the cognitive radio cycle [1]. Figure 1 depicts the cognitive cycle for an individual radio node with its various stages. Figure 2 generalizes this to a network of nodes, which make cooperative multilateral decisions in addition to local unilateral ones.

Figure 1 shows that observations made on the environment are the starting point in the cycle. All else flows from this point – the reasoning, the decisions and the learning follow from the ability to observe, understand and make sense of the
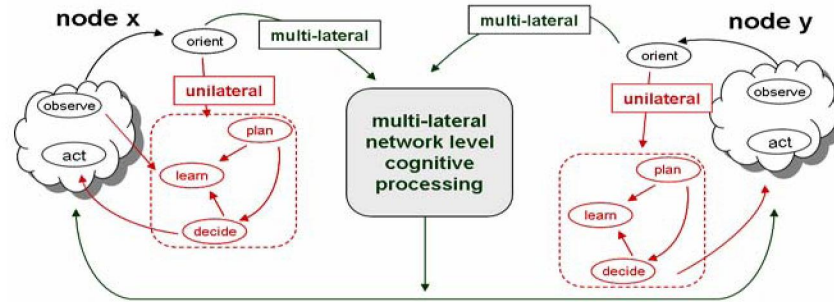
Figure 2: The double cognitive cycle in a cognitive network

environment. The environment or outside world is also the end of the cycle as all decisions are brought to bear in this space through making the changes or adaptations or reconfigurations that are needed.

In Figure 2 we again see how critical the modeling of this environment/outside world is for evaluating the behavior of a cooperating network. The observations made by the nodes, which should depend on their individual spatial-temporal contexts, and how the environmental model incorporates the results of actions by the network, largely determines the execution path of the network. Environmental effects can also affect network decision making itself, such as by delaying or reordering inputs to the decision algorithms, or preventing some nodes from receiving directives; these effects are fundamental to the behavior of large-scale systems.

Environmental models that are too simple, leaving out important effects, lead to incorrect evaluations of the technique or system being studied. Models that are too complex, containing many parameters, lead to evaluations that cannot be replicated, generalized, or compared to other results. Therefore there is no one ideal environmental model. Rather, different models are appropriate for use when investigating different research questions.

B. The Environmental Model Space

We reviewed the simulation based papers published in DYSPAN 2007, as a representative sample of recent research, to develop a broad idea of the various options for environmental models. We developed a 6-axis categorization of models for CR research.

(1) Propagation model – Propagation effects appropriate to consider for various experiments include not just a channel model but also antenna response, fading and shadowing due to motion of nodes through an obstructed environment, and even diffraction effects (e.g. for work where the hidden node problem is relevant).

(2) Interference model – Interference can be modeled at many different levels of realism: simple reuse distances, co-channel interference only, adjacent channel interference, intermodulation effects and non-linearities in transmitters and receivers. See for example [14] for a detailed discussion of various subtle interference affects that may have a significant impact on TV whitespace dynamic spectrum access networks. The sources of interference such as primary spectrum users may be treated as static or dynamic, and if dynamic, modeled with low-level behavior only or also with higher layer structure (e.g. packet formats) and temporal responses to interference (e.g. TCP backoff).

(3) Spatial model – The node count, distribution, mobility, and underlying topography are critical for evaluating research ideas. See for example [15] for a spatial statistical model grounded in measurements. In the real world, traffic and mobility patterns can be linked: users tend to use communications systems more heavily when physically in some places than in others, and the type of data exchanged changes based on velocity and other parameters of the spatial model. These correlations may be value to model in more sophisticated research frameworks.

(4) Node model – Large-scale real-world systems will face intruders, malicious nodes, and faulty nodes. These can create havoc for distributed algorithms that are not designed properly. Another issue is per node clock skew and drift. Clock issues can significantly affect distributed data fusion algorithms that assume observations of the environment by cooperating nodes are synchronized.

(5) Cost model – End-user benefits often depend on cost savings, so an appropriate resource cost model is essential. Potential items to model include: memory, processor and battery power used by the software, transmit power used by the radio, bandwidth and other communication resources used by a particular data flow, and end-to-end latency for completing operations. (There is a separate challenging issue of how end-user benefit is determined; it may be a function of cost savings, QOS improvements, risk reduction, and other factors.)

(6) Observation model – Observations gathered by a CR will normally be an approximation to the external reality, due for example to sensor noise, positioning errors, and sampling effects. CR systems need to be resilient to these effects, so if they are not modeled in the research there can be unknown limitations to the claimed results.

Given the model, the simulation environment may execute that model using a fixed approach where the environment behaves the same in every run, irrespective of behavior of the system being studied; a deterministic approach, where the environment reacts to the behavior of the system in a predetermined fashion; or a random approach such as a Monte Carlo simulation where the environment reacts probabilistically.

C. State of the Art in Environmental Models

Given the categorization described in the previous section, we went over the simulation based papers published in DYSPAN 2007 again to take stock of the state of the art in the CR field.

(1) Very few of the existing analyses use anything more than simple propagation models—circular disks around nodes and two ray propagation models—so the improvements gained may not bear out when the system is expanded to take more realistic situations into account.

(2) Very few analyses take complex interference of any sort into account. A lot of interference modeling is based on reuse distance, which is optimistic.

(3) Random node distribution is by far the most common spatial model, even though this does not correspond well to the real world.

(4) – (6) Issues raised by node, cost, and observational models are generally not considered. However, in terms of space in a paper, it may not always be possible to discuss all the details of the simulations used to derive the reported results.

With a few notable exceptions (e.g. [16]), the models used are not calibrated against reality. There are a number of measurements papers now published which could be used to justify values in the models, for example [17]–[19]. (One valuable area for research would be comparisons between models and reality, as well as measurements projects done to inform model construction.)

Sensitivity of the reported results to variations in environmental model parameters is not usually evaluated.

VI. CONCLUSION

We have highlighted concerns and opportunities for research on reasoning and learning cognitive radio systems. A primary concern is unpredictability of these systems. Opportunities include work to improve performance of novel networks, improve interference tolerance, reduce configuration cost and risk.

To make progress on these opportunities, we have recommended that the cognitive radio research community begin sharing research framework components, including simulators, environmental models, traffic models, and agreed upon ways of evaluating end user benefit. We have proposed a workshop be held to kick off collaboration on these topics.

REFERENCES

[1] J. Mitola, *Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio*. PhD Dissertation, KTH, May 2000.

[2] United States Federal Communications Commission, "Spectrum Policy Task Force Report," 2002. Online: http://www.fcc.gov/sptf/reports.html.

[3] United States Federal Communications Commission, Order in ET Docket No. 03-237, May 2, 2007. Online: http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-07-78A1.pdf.

[4] IEEE P1900.1, "Standard Definitions and Concepts for Spectrum Management and Advanced Radio System Technologies," draft version 0.30, April 5, 2007.

[5] G. Denker, D. Elenius, R. Senanayake, M.O. Stehr and D. Wilkins, "A Policy Engine For Spectrum Sharing," *IEEE DYSPAN 2007*.

[6] J. Gaeddert, K. Kim, R. Menon, L. Morales, Y. Zhao, K.K. Bae, and J. H. Reed, "Applying Artificial Intelligence to the Development of a Cognitive Radio Engine," Technical Report, Mobile and Portable Radio Research Group (MPRG), Virginia Tech, June 30, 2006.

[7] A. Sahai, N. Hoven, S.M. Mishra, and R. Tandra, "Fundamental Tradeoffs in Robust Spectrum Sensing for Opportunistic Frequency Reuse," Technical Report, Dept. of Electrical Engineering and Computer Science, UC Berkeley, March 2006. Online: http://www.eecs.berkeley.edu/~sahai/Papers/CognitiveTechReport06.pdf

[8] W.J. Goralski, *Juniper and Cisco Routing: Policy and Protocols for Multivendor IP Networks*. Wiley, September 2002, p. 10.

[9] M. Castro and B. Liskov, "Practical Byzantine Fault Tolerance," *Proceedings of the Third Symposium on Operating Systems Design and Implementation*. New Orleans, USA, February 1999.

[10] X. Zeng, R. Bagrodia, and M. Gerla, "Glomosim: A Library for Parallel Simulation of Large-Scale Wireless Networks," *12th Workshop on Parallel and Distributed Simulations*, 1998.

[11] Scalable Network Technologies, Inc., "Qualnet Network Simulator." Online: http://www.scalable-networks.com.

[12] S. McCanne and S. Floyd, Network Simulator Version 2. Online: http://www.isi.edu/nsnam/ns/

[13] D.X. Wei and P. Cao, "NS-2 TCP-linux: An NS-2 TCP Implementation with Congestion Control Algorithms from linux," in *ACM Workshop on NS-2 (ValueTools 2006)*, 2006.

[14] V. Petty, R. Rajbanshi, D. Datla, F. Weidling, D. DePardo, P. Kolodzy, M. Marcus, A. Wyglinski, J. Evans, G. Minden and J. Roberts, "Feasibility of Dynamic Spectrum Access in Underutilized Television Bands," *IEEE DYSPAN 2007*.

[15] P. Mähönen, M. Petrova and J. Riihijärvi, "Applications of Topology Information for Cognitive Radios and Networks," *IEEE DYSPAN 2007*.

[16] S. Jones, E. Jung, X. Liu, N. Merheb and I-J. Wang, "Characterization of Spectrum Activities in the U.S. Public Safety Band for Opportunistic Spectrum Access," *IEEE DYSPAN 2007*.

[17] M. McHenry, "NSF Spectrum Occupancy Measurements Project Summary," Technical Report, Shared Spectrum Company, Vienna, VA, August 2005. Online: http://www.sharedspectrum.com/?section=measurements.

[18] W. Turney, M. Karam, L. Malek and G. Buchwald, "VHF/ UHF Building Penetration Characteristics When Using Low Antenna Heights," *IEEE DYSPAN 2007*.

[19] O. Holland, P. Cordier, M. Muck, L. Mazet, C. Klöck and T. Renk, "Spectrum Power Measurements in 2G and 3G Cellular Phone Bands during the 2006 Football World Cup in Germany," *IEEE DYSPAN 2007*.