Power Consumption Optimization - A Cognitive Radio Approach

An He

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

Jeffrey H. Reed

William H. Tranter

R. Michael Buehrer

A. Lynn Abbott

Raghu Pasupathy

Feburary 25, 2011

Blacksburg, Virginia

Keywords: Power consumption optimization, Mobile and wireless communication, Cognitive Radio

Power Consumption Optimization - A Cognitive Radio Approach

An He

(ABSTRACT)

Power consumption is one of the most important aspects in mobile and wireless communications. Existing research has shown significant power reduction through limited radio reconfiguration based on the channel conditions, especially for short range sensor network applications.

A cognitive radio (CR) is an intelligent wireless communication system which is able to determine the most favorable operating parameters (cognition) based on the radio environment and its own capabilities and characteristics (awareness) and reconfigure the radio accordingly (reconfigurability).

This work leverages the advances in cognitive radio technology to dynamically implement favorable trade-offs in radio parameters to achieve more efficient use of radio resource (e.g., minimizing power consumption) on the required Quality of Service (QoS) of an application and channel. A CR-based approach enables us not only to adjust modulation, coding, and radiated power as in a conventional radio, but also to learn and to control component characteristics (e.g., the power amplifier (PA) efficiency characteristic) to minimize power consumption. Significant power savings using this approach are shown in this work for single input single output (SISO) systems and multiple input multiple output (MIMO) systems.

This work has a broad potential impact on the research of improving power efficiency of communication systems. It establishes a cognitive radio based methodology for system power consumption optimization. It emphasizes the difference between radiated power (power radiated from the transmit antenna) and the consumed power (power drawn from the power source, such as a battery). It provides a way to connect communication (which usually cares about radiated power, received signal to noise ratio, etc.) to hardware (which focuses on speed, efficiency, power consumption, etc.) and software (which emphasizes complexity, speed, etc.). This design methodology enhances the capability to jointly optimize communication, hardware, and software. In addition, this CR-based framework can be adapted for general radio resource management with various radio operation optimization targets, such as spectrum utilization.

Dedication

To my parents.

Acknowledgments

First of all, I would like to express my sincere appreciation to my advisors, Dr. Jeffrey H. Reed and Dr. William H. Tranter, for their guidance throughout my study at Virginia Tech. They have shared with me knowledge and vision about the field. They have given me freedom on research and trusted me in my capability and judgment. They have trained me to be an independent researcher. I could not complete my work without their help.

I have been fortunate to have Dr. A. Lynn Abbott, Dr. R. Michael Buehrer, and Dr. Raghu Pasupathy on my dissertation committee. The combination of their unique background and experience has often offered a new angle to the research and ways to better illustrate and sometimes defend the work. I also appreciate the consistent help from Dr. Claudio da Silva who served on my qualifying exam committee.

I would like to thank Dr. Srikathyayani Srikanteswara, Dr. Masoud Sajadieh, and Dr. Marian Verhelst from Intel and Dr. Kyung Kyoon Bae from Liberty University for offering great suggestions and feedback on the work.

Acknowledgment goes to Intel Corp., ETRI, and Wireless @ Virginia Tech affiliates for sponsoring the research.

My study at Virginia Tech would not be so enjoyable without my colleagues and friends at Wireless @ Virginia Tech, including Ihsan Akbar, Ashwin Amanna, Chris Anderson, Nitin Arora , Philip Balister, Mike Benonis, Soumava Bera, Kaigui Bian, Nikhil Challa, Venkata Gautham Chavali, Qinqin Chen, Xuetao Chen, Zhimin Chen, Cory Cooke, Thomas Cooper, Dinesh Datla, Hunter DeJarnette, Harpreet Singh Dhillon, Carl Dietrich, Zhenhua Feng, Manik Gadhiok, Joseph D. Gaeddert, Feng Ge, Carlos R. Aguayo Gonzalez, Xu Guo, Chuan Han, Mahmud Harun, S. M. Shajedul Hasan, William Headley, Benjamin Hilburn, Shawn Hymel, Eyosias Yoseph Imana, Ryan Irwin, Daniel J. Jakubisin, Joeng-o Jeong, Tao Jia, Bin Le, Jeong Heon

Lee, Chang Li, Xiaoxing Li, Qian Liu, Amy Malady, Rekha Menon, Maruf Mohammad, Sabares Moola, James Neel, Tim Newman, Ramesh Chembil Palat, Christopher Phelps, Matt Price, Sahana Raghunandan, Barathram Ramkumar, Shereef Sayed, Javier Schloemann, Yongsheng Shi, Jason Snyder, Michael Steiner, Benton Thompson, Thomas Tsou, Abid Ullah, Haris Volos, Ying Wang, Bei Xie, Yue Yan, and Youping Zhao. I would also like to express my gratitude to the MPRG staff, James Dunson, Jenny Frank, Nancy Goad, Cyndy Graham, Randall Nealy, Hilda Reynolds, and Shelby Smith for making MPRG a home for students.

In addition, it is great to have made many new friends out of the lab at Blacksburg, Xiaojun Chang, Chen Chen, Shijing Cheng, Bo Gao, Ming Gao, Yanpeng Hou, Yaoshun Jia, Canming Jiang, Shouwen Lai, Min Li, Lingrui Liao, Jianqiu Liu, Xiaomo Liu, Yifei Ma, Guanhong Pei, Yumin Qi, Yi Shi, Jiangming Teng, Chao Yang, Shuo Yao, Alan Wang, Ke Wang, Ruijia Wang, Ting Wang, Hao Wu, Haoran Wu, Wan Wu, Huadong Xia, Liguang Xie, Qing Yang, Yihong Yang, Xinhao Ye, Pengtao Yue, Bin Zhang, Jingyao Zhang, Xiaoyu Zhang, Xu Zhang, Yexin Zheng, among many others.

Finally, I would like to thank my family who have supported me throughout my life, especially my parents, who always encourage me and trust me.

Attribution

- Chapter 3: A. He, S. Srikanteswara, K. K. Bae, J. H. Reed, and W. H. Tranter, Energy consumption minimization for mobile and wireless devices - a cognitive approach, IEEE Transactions on Consumer Electronics, vol. 56, no. 3, pp. 1814-1821, Aug. 2010.
  - S. Srikanteswara is with Intel Corporation, the sponsor of the research this dissertation was derived from. Srikanteswara attended the monthly project meeting and offered suggestions and comments from an industry perspective and made the outcome of this research applicable to industry practice.
  - K. K. Bae was with the Wireless @ Virginia Tech at Virginia Polytechnic Institute and State University while the project started. Bae attended the monthly project meeting and offered suggestions and comments especially on multiple input multiple output systems and cognitive engine design.
  - J. H. Reed is with Virginia Polytechnic Institute and State University and is the co-chair of the dissertation committee. Reed offered consistent guidance throughout the research and dissertation preparation.
  - W. H. Tranter is with Virginia Polytechnic Institute and State University and is the co-chair of the dissertation committee. Tranter offered consistent guidance throughout the research and dissertation preparation.
- Chapter 4: A. He, S. Srikanteswara, K. K. Bae, T. R. Newman, J. H. Reed, and W. H. Tranter, M. Sajadieh, and M. Verhelst, Power consumption minimization for MIMO systems - a cognitive radio approach, IEEE Journal on Selected Areas in Communications, vol. 29, no. 2, pp. 469-479, Feb. 2011.

- S. Srikanteswara is with Intel Corporation, the sponsor of the research this dissertation is derived from. Srikanteswara attended the monthly project meeting and offered suggestions and comments from an industry perspective and made the outcome of this research applicable to industry practice.
- K. K. Bae was with the Wireless @ Virginia Tech at Virginia Polytechnic Institute and State University while the project started. Bae attended the monthly project meeting and offered suggestions and comments especially on multiple input multiple output systems and cognitive engine design.
- T. R. Newman was with Virginia Polytechnic Institute and State University while the manuscript was prepared. Newman offered suggestions and comments on the preparation of the manuscript.
- J. H. Reed is with Virginia Polytechnic Institute and State University and is the co-chair of the dissertation committee. Reed offered consistent guidance throughout the research and dissertation preparation.
- W. H. Tranter is with Virginia Polytechnic Institute and State University and is the co-chair of the dissertation committee. Tranter offered consistent guidance throughout the research and dissertation preparation.
- M. Sajadieh is with Intel Corporation, the sponsor of the research this dissertation is derived from. Sajadieh attended the monthly project meeting and offered suggestions and comments from an industry perspective and made the outcome of this research applicable to industry practice.
- M. Verhelst is with Intel Corporation, the sponsor of the research this dissertation is derived from. Verhelst attended the monthly project meeting and offered suggestions and comments from an industry perspective and made the outcome of this research applicable to industry practice.

- Chapter 5: A. He, K. K. Bae, T. R. Newman, J. Gaeddert, K. Kim, R. Menon, L. Morales, J. Neel, Y. Zhao, J. H. Reed, and W. H. Tranter, A survey of artificial intelligence for cognitive radios, IEEE Transactions on Vehicular Technology, vol. 59, no. 4, pp. 1578-1592, May 2010.
  - K. K. Bae was with the Wireless @ Virginia Tech at Virginia Polytechnic Institute and State University during the project. Bae contributed to the subject of case study in the manuscript.

- T. R. Newman was with Virginia Polytechnic Institute and State University while the manuscript was prepared. Newman offered suggestions and comments on the preparation of the manuscript.
- J. Gaeddert is with Virginia Polytechnic Institute and State University. Gaeddert contributed to the subject of artificial neural network in the manuscript.
- K. Kim was with Virginia Polytechnic Institute and State University while the manuscript was prepared. Kim contributed to the subject of hidden Markov model in the manuscript.
- R. Menon was with Virginia Polytechnic Institute and State University while the manuscript was prepared. Menon contributed to the subject of metaheuristic algorithms in the manuscript.
- L. Morales was with Virginia Polytechnic Institute and State University while the manuscript was prepared. Morales contributed to the subject of case-based system in the manuscript.
- J. Neel was with Virginia Polytechnic Institute and State University while the manuscript was prepared. Neel contributed to the subject of game theory in the manuscript.
- Y. Zhao was with Virginia Polytechnic Institute and State University while the manuscript was prepared. Zhao contributed to the subject of rule-based system in the manuscript.
- J. H. Reed is with Virginia Polytechnic Institute and State University. Reed oversaw the project that the research was derived from and the preparation of the manuscript.
- W. H. Tranter is with Virginia Polytechnic Institute and State University. Tranter oversaw the preparation of the manuscript.

• Chapter 6: A. He, S. Srikanteswara, K. K. Bae, J. H. Reed, and W. H. Tranter, Designing a cognitive engine framework for radio power consumption optimization, to be submitted.

  - S. Srikanteswara is with Intel Corporation, the sponsor of the research this dissertation is derived from. Srikanteswara attended the monthly project meeting and offered suggestions and comments from an industry perspective and made the outcome of this research applicable to industry practice.
  - K. K. Bae was with the Wireless @ Virginia Tech at Virginia Polytechnic Institute and State University while the project started. Bae attended the monthly project meeting and offered suggestions and comments especially on multiple input multiple output systems and cognitive engine design.

- J. H. Reed is with Virginia Polytechnic Institute and State University and is the co-chair of the dissertation committee. Reed offered consistent guidance throughout the research and dissertation preparation.
- W. H. Tranter is with Virginia Polytechnic Institute and State University and is the co-chair of the dissertation committee. Tranter offered consistent guidance throughout the research and dissertation preparation.

- Appendix A: A. He, J. Gaeddert, K. K. Bae, T. R. Newman, J. H. Reed, L. Morales, and C. Park, Development of a case-based reasoning cognitive engine for IEEE 802.22 WRAN applications, ACM Mobile Computing and Communications Review, vol. 13, no. 2, pp. 37-48, Apr. 2009.
  - J. Gaeddert is with Virginia Polytechnic Institute and State University. Gaeddert contributed to the subject of performance metrics, simulation environment setting, and genetic algorithm and hill climbing search algorithm design in the manuscript.
  - K. K. Bae was with the Wireless @ Virginia Tech at Virginia Polytechnic Institute and State University during the project. Bae oversaw the project the research was derived from and the preparation of the manuscript.
  - T. R. Newman was with Virginia Polytechnic Institute and State University while the manuscript was prepared. Newman offered suggestions and comments on the preparation of the manuscript.
  - J. H. Reed is with Virginia Polytechnic Institute and State University. Reed oversaw the project the research was derived from and the preparation of the manuscript.
  - L. Morales was with Virginia Polytechnic Institute and State University while the manuscript was prepared. Morales offered suggestions and comments on the preparation of the manuscript.
  - C. Park is with Electronics and Telecommunications Research Institute, the sponsor of the research this dissertation was derived from. Park offered suggestions and comments on the preparation of the manuscript.

Contents

List of Figures

List of Tables

List of Abbreviations

ADC	Analog to Digital Converter
AI	Artificial Intelligence
AM	Adaptive Modulation
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
CSI	Channel State Information
CSIR	Channel State Information at Receiver
CSIT	Channel State Information at Transmitter
CT	Cognitive Transmission
DAC	Digital to Analog Converter
DSA	Dynamic Spectrum Access
FEC	Forward Error Correction
LNA	Low Noise Amplifier
PAR	Peak to Average Ratio
RFE	Radio Front End
SNR	Signal to Noise Ratio

Chapter 1

Introduction

1.1 Motivation

Power consumption[1] is one of the most important aspects in mobile and wireless communications. Recent research has shown significant power reduction through limited radio reconfiguration based on the channel conditions, especially for short range sensor network applications. At the system level [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19], various system power consumption models have been proposed for power optimization by adapting modulation, coding, antenna configuration, and radiated power to minimize power consumption. Although a lot of the work has been focused on the radiated power [1, 2, 3, 4, 5, 6, 7], some researchers [8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18] have noticed the difference between the radiated power (power radiated from the transmit antenna) and the consumed power (power drawn from the power source, such as a battery). On the other hand, the research at the circuit level has been focused on improving the component (e.g., power amplifier (PA)) efficiency. For example, in [20, 21, 22], dynamic voltage biasing and dynamic current biasing techniques are used to adjust the operating point of the PA according to the PA output signal envelope so that the PA power consumption can be reduced when the required output power level is low and the PA efficiency at low output power levels can be improved.

A cognitive radio (CR) [23, 24] is an intelligent wireless communication system which is able to determine the most favorable operating parameters (cognition) based on the radio environment and its own capabilities and

[1]It is important to emphasize that in general power consumption and energy consumption are two different but closely related concepts. In this work, we use them interchangeably in the general discussion. We will clearly distinguish them as needed in the specific context.

characteristics (awareness) and reconfigure the radio accordingly (reconfigurability). Awareness, learning, and decision making reside within the cognitive engine (CE). CR has been widely researched for dynamic spectrum access to improve spectrum utilization [25, 26, 27, 28]. However, it has a much wider area of application.

This work seeks to leverage advances in cognitive radio technology to dynamically implement favorable trade-offs in radio parameters to minimize power consumption for the required Quality of Service (QoS) for a particular application and channel. By doing so, more efficient use of radio resource can be achieved. A CR enables us not only to adjust modulation, coding, and radiated power, but also to learn and to advise the adaptation of component characteristics (e.g., the PA efficiency characteristic[2]) to minimize power consumption.

1.2 Contributions

The contributions of this work consist of the following elements:

1. Power Consumption Optimization Framework Using Cognitive Radio:
   This work proposes a methodology of using a CR framework for power consumption optimization[3]. The framework enables learning of the radio component characteristics (e.g., the PA efficiency) which is necessary for power consumption optimization. Based on the learned component characteristics, radio parameters and component characteristics can be jointly adapted to minimize power consumption given the channel and meet the QoS requirement of an application. As the PA usually dominates the power consumption in medium and long range wireless applications, this work focuses on the impact of the PA characteristics on the power consumption. Other components can be integrated into the framework as needed. This framework can also be applied to optimize radio operation to achieve additional goals. The capability of controlling radio parameter adaptation is especially useful for future radios, where complex digitally controlled analog components are used. For example, a wideband Motorola RF transceiver IC [29] has thousands of possible parameter combinations to control the programmable internal filter banks, synthesizer, amplifier gain, etc. It is almost impossible to test all

[2]Note that this dissertation focuses on Class A, Class B, and some practical PAs. These types of PAs are widely used in current wireless communication systems as the linearity of a PA is one of the most important design concerns.

[3]This cognitive process itself consumes power. Therefore, additional work shall be conducted to understand the power consumption of this process and to draw a conclusion on overall power consumption. This work focuses on understanding the achievable power savings first.

the combinations manually. A CR based parameter adaptation framework can be very useful in this case by trying different combinations, learning corresponding performance, and deciding appropriate configuration based on learned characteristics.

2. Power Savings for Single Input Single Output Communication Systems using the CR Framework:
   This work investigates the relationship between the radiated power, the PA power consumption, and the PA efficiency characteristic in single carrier single input single output (SISO) communication systems. A unified PA efficiency model characterizing theoretical Class A, Class B, and practical PAs is adopted and enables the analysis of the impact of different radio configurations and channel conditions on power efficiency. With PA efficiency knowledge, the CR framework is able to minimize the power consumption by adapting radio parameters, such as, modulation, coding rate and coding gain, and radiated power. In addition, radio parameters and radio component characteristics (i.e., PA efficiency characteristic) are jointly adapted to obtain further system power consumption reduction for given channel and QoS requirement.

3. Power Savings for Multiple Input Multiple Output Communication Systems using the CR Framework:
   This work further leverages the results from information theory and the capabilities of a CR (e.g., the awareness of the component capabilities and characteristics) and develops a theoretical framework to minimize power consumption for multiple input multiple output (MIMO) communication systems. The power consumption minimization problems under a sum rate constraint for MIMO systems under various scenarios (e.g., channel state information availability at the transmitter and the receiver and antenna correlation at the transmitter) are mathematically formulated. Several heuristic numerical algorithms are developed to solve the constraint optimization problems and evaluated by simulation.

4. Cognitive Engine Development for Power Consumption Optimization:
   A cognitive engine is used to facilitate learning and decision making in the CR framework. A CR can not only learn the channel conditions as in conventional radios, but is also aware of the radio (component) capabilities and characteristics. The CE can be implemented with different techniques and different levels of complexity according to the requirement of the application. As a proof of concept, this work presents and simulates a case-based reasoning cognitive engine reference design for the application of power consumption minimization.

1.3 Organization of the Dissertation

This dissertation consists of journal papers on each topics delineated above. To assist the readers, before each chapter a transition cover page is prepared summarizing the work discussed in this chapter and the connection to other chapters.

To be specific, the dissertation is organized as follows. Chapter 2 reviews the finding on power consumption breakdown in a battery powered mobile and wireless communication device, the existing work on power reduction using radio reconfiguration. Chapter 3 (Energy Consumption Minimization for Mobile and Wireless Devices - A Cognitive Approach) presents the CR framework for power consumption optimization and power savings for SISO systems. Chapter 4 (Power Consumption Minimization for MIMO Systems - A Cognitive Radio Approach) shows the power savings results for MIMO systems. Chapter 5 (A Survey of Artificial Intelligence for Cognitive Radios) reviews some existing cognitive engine design work and artificial intelligence techniques employed in the CE implementation. Chapter 6 (Designing a Cognitive Engine Framework for Radio Power Consumption Optimization) discusses the CE designed for power consumption optimization. Chapter 7 concludes the dissertation.

Here is a list of chapters and corresponding detailed citations of the journal publications.

- Chapter 3: A. He, S. Srikanteswara, K. K. Bae, J. H. Reed, and W. H. Tranter, Energy consumption minimization for mobile and wireless devices - a cognitive approach, IEEE Transactions on Consumer Electronics, vol. 56, no. 3, pp. 1814-1821, Aug. 2010.

- Chapter 4: A. He, S. Srikanteswara, K. K. Bae, T. R. Newman, J. H. Reed, W. H. Tranter, M. Sajadieh, and M. Verhelst, Power consumption minimization for MIMO systems - a cognitive radio approach, IEEE Journal on Selected Areas in Communications, vol. 29, no. 2, Feb. 2011.

- Chapter 5: A. He, K. K. Bae, T. R. Newman, J. Gaeddert, K. Kim, R. Menon, L. Morales, J. Neel, Y. Zhao, J. H. Reed, and W. H. Tranter, A survey of artificial intelligence for cognitive radios, IEEE Transactions on Vehicular Technology, vol. 59, no. 4, pp. 1578-1592, May 2010.

- Chapter 6: A. He, S. Srikanteswara, K. K. Bae, J. H. Reed, and W. H. Tranter, Designing a Cognitive Engine Framework for Radio Power Consumption Optimization, to be submitted.

Chapter 2

Background

This chapter starts with findings on power consumption breakdown in battery powered mobile and wireless communication devices and shows the importance of understanding the transmitter power consumption, especially, the power consumption related to the power amplifier, which is the theme of this dissertation. Then, it highlights the most relevant existing work on radio power reduction using radio reconfiguration in the literature and touches upon the concept of cognitive radio which is a convenient tool for the power consumption optimization that this work deals with. Finally, this chapter discusses the gaps and limitations of the existing work and highlights the reasons and importance of the research carried out in this work to advance the state of the art.

2.1 Power Consumption Breakdown in Mobile and Wireless Communication Devices

It is a very challenging task to accurately quantify the power consumption of each component/module in mobile wireless communication devices. And the exact breakdown is platform and application dependent. Usually, this is done by the manufacturers and kept to themselves. There are very limited results in the publicly available literature on the detail power consumption of real communication systems. This section summarizes these results.

In [30], the power consumption breakdown of a 3G phone running video streaming is provided in a pie chart.

Although no specific numbers are provided in [30], it is estimated from the pie chart that approximately 30% of the total power is consumed by the transmitter including the PA, 18% by the receiver, 22% by the baseband and the mixed signal processing including the analog to digital converter (ADC) and the digital to analog converter (DAC), and 30% by the user interface hardware including display, speakerphone, etc.

[31], referring to [30], provides some power consumption numbers for a multimedia cell phone operated in the 384kb/s video streaming mode. The power consumption of the RF receiver and the cellular modem is about 1200mW, the user interface (the audio, the display, the keyboard backlight, etc) about 1000mW, the application processors and memories about 600mW, and the mass memories about 200mW. These numbers approximately match the pie chart in [30].

In addition to the direct results on power consumption breakdown in [30, 31], the power consumption breakdown can be estimated using simulation. For example, given the models and numbers in [18], simple calculation shows that at a distance of 100 meters the radio front end (RFE) module, mainly the PA, accounts for more than 75% of the system power consumption, which includes the RFE power consumption, the baseband signal processing power consumption, the MIMO detector power consumption, and the FEC power consumption. At further distance, the RFE takes an even larger portion of the system power consumption since the needed radiated power increases more than linearly with the distance (usually, $P_{rad} \propto d^{3\sim4}$) while the power consumption of other parts stays almost constant. One possible reason this result is different from that in [30, 31] is that [18] excludes power consumption of user interface, receiver, and complex video processing in baseband signal processing.

Comparing the result calculated from [18] and that in [30, 31], we observe that the exact power consumption breakdown depends on many factors, such as, applications, system architecture, hardware choices, and software implementation. However, in general, for medium and long range applications, the transmitter module consumes about half the total power of the mobile and wireless devices. The baseband signal processing including the ADC and the DAC consumes significant amount of power, especially for the multimedia applications, such as video streaming, where date throughput and data processing rate is high. The user interface module consumes comparable amount of power with the baseband signal processing module.

The PA is usually a major power consuming component in the RFE module due to the high radiated power requirement of the wireless device for medium and long range coverage and the low power efficiency of the PA for required linearity. For example, Skyworks provides many PA products for various wireless applications [32], including 2G and 3G cellular applications, WiFi and other ISM band applications. Most of these PAs provide maximum output power around 20 - 30 dBm with efficiency much less than 50%. Even if the PA efficiency is 50%, the power consumption of the PA is about 23 - 33 dBm for these wireless applications. As we will see in the later chapters, the PA efficiency usually peaks at the maximum output power and decreases at lower output power levels. The dynamic range of the radio radiated power can be tens of dB and the average output power can be 10 dB lower than the maximum output power [33]. The average PA efficiency is then much lower than its maximum efficiency so that the PA power consumption is much more significant.

In medium and long range wireless communications, the RFE module usually dominates the system power power consumption, of which the PA takes a significant portion. This work focuses on the impact of the PA on the system power consumption of mobile and wireless devices. Other significant power consuming components, such as, ADC, DAC, mixer, digital baseband signal processing unit, and filters, can be integrated into the work as needed. Note that most of the existing work on system power consumption minimization [9, 10, 13, 14, 15, 16, 17, 18] deals with short range or sensor network applications where power consumption of these components becomes comparable to that of the RFE module.

2.2 Existing Work on Power Reduction Using Radio Reconfiguration

The research on power reduction through radio reconfiguration is closely related to the research on transmission rate improvement using adaptive modulation. Many ideas come from the adaptive modulation work. This section first briefly reviews the idea of adaptive modulation for transmission rate improvement. Then, it summarizes the work on power (radiated power and consumed power) optimization through radio reconfiguration.

2.2.1 Adaptive Modulation

Adaptive modulation (AM) was first proposed in 1960's [34]. However, the research on AM did not thrive until 1990's when enabling technologies, such as, reconfigurable hardware and accurate channel estimation techniques, became available and the demand for spectrally efficient communications increased. The AM system changes its transmission parameters to meet certain QoS requirements (e.g., bit error rate (BER)) based on the channel state information. In this way, the AM system can achieve higher transmission rate under good channel conditions and improve average transmission rate compared to the fixed transmission system which is usually designed for worst case scenarios. The transmission parameters adapted include radiated power [34], symbol rate [35], error correction coding [36], modulation [37], or the combination of any of these parameters [38, 39, 40, 41, 42, 43, 44, 45, 46, 47]. In addition, AM can be integrated with MIMO and OFDM techniques to achieve even higher transmission rate [48, 49, 50, 51].

2.2.2 Power Minimization through Radio Reconfiguration

There are other important aspects in wireless communications in addition to the transmission rate. After tremendous improvement in transmission rate, the power minimization problem, a problem closely related to the transmission rate maximization problem, has emerged as an important research topic in wireless communications, especially, in battery powered mobile and wireless communications. Many concepts in the transmission rate maximization using AM are adapted for the power minimization. Recent research has shown significant power reduction through limited radio reconfiguration based on the channel condition in various wireless communication applications. The research can be generally categorized into two groups: radiated power minimization problem and consumed power minimization problem. In this dissertation, the radiated power refers to the output power from the transmit antennas and the consumed power refers to the total power consumed by the radio (i.e., the power drawn from the power supply of the radio). Since the purpose of this research is to reduce the consumed power of mobile and wireless communication devices, this section focuses on consumed power minimization with a brief review of radiated power minimization.

Radiated Power Minimization

The required radiated power has to compensate for the radio propagation loss between the transmitter and the receiver so that the received signal to noise ratio (SNR) is sufficient for the receiver to demodulate and decode the transmitter signal at a target BER. Transmission parameters can be adapted to minimize the required radiated power given the target transmission rate and the channel condition. The selected work on radiated power minimization is summarized in Table 2.1.

Table 2.1: Summary of existing work on radiated power minimization

Ref.	**Parameters**	**Constraints**	**Target**	**Year**
[1]	Trellis coded modulation, power distribution	Constant transmission rate, constant BER	Overall radiate power in OFDM systems	1999
[2]	Radiated power, transmission rate	Signal to interference ratio, transmission rate, overall radiated power	Overall radiated power in Code Division Multiple Access systems	2002
[3]	Modulation, antenna diversity scheme	SNR, overall network throughput, average user throughput	Overall uplink radiated power in multiple access networks	2004
[4]	Modulation, power distribution	Transmission rate, BER	Radiated power in MIMO-OFDM systems with Channel State Information at Transmitter (CSIT) and Receiver (CSIR)	2004
[5]	Modulation, radiated power	Transmission rate, BER	Radiated power in IEEE 802.11b links	2005
[6]	Modulation, radiate power	Transmission rate, BER	Radiated power in MIMO systems with imperfect channel state information	2006
[7]	Modulation, coding	Average transmission rate, BER	Time Division Multiple Access systems with quantized CSIT.	2008

System Power Minimization

Radiate power is of particular interest to capacity and interference analysis. However, consumed power is also important. The consumed power includes power consumed by all components of the radio. Past systems tended to rely on fixed hardware and software and thus had fixed power consumption. This might be part of the reason why the radiated power minimization problem has been investigated much more thoroughly than the consumed power minimization problem has. However, today's systems can be quite flexible, offering a variety of hardware and software configurations that can be changed on-the-fly. Key to enabling cognitive radio hardware is the availability of good power models. In reality, it is very challenging to accurately model the power consumption of a radio. First, modern radios usually allow several operation modes. Second, within each operation mode, the number of tunable radio parameters and the number of choices for each radio parameter can be large. Third, for a specific combination of the set of radio parameters, the implementation related issues, such as, hardware component variation, and software runtime variation, make an accurate power consumption model prohibitively complex to allow mathematical analysis. Therefore, simplification and approximation is inevitable in almost all power consumption models developed. System power consumption model is key, and thus in this section, we review such models as well as the system power minimization approaches.

Various system power consumption models discussed in the literature are summarized in Table 2.2. In general, the system power consumption consists of transmit power consumption and circuitry power consumption. The transmit power consumption usually characterizes power consumption related to PA at the transmitter. The circuitry power consumption usually characterizes power consumption related to other electronic circuits, such as, ADC, DAC, low noise amplifier (LNA), mixer, frequency synthesizer, and filters.

Table 2.2: Summary of existing system power models

Ref.	**Transmit power consumption**, P_{tr}	**Circuitry power consumption**, P_{cr}	**Year**
[8]	$P_{tr} = P_{rad}$ (P_{rad}, radiated power)	Linear function of the symbol rate	2001
[9]	$P_{tr} = \alpha P_{rad} + \beta$ (α and β, two constants characterizing PA efficiency)	Viterbi decoding: sum of switching and leakage power; Receiver electronics: constant; Radio startup power: constant	2002
[10]	$P_{tr} = P_{rad}/\eta$ (η, PA drain efficiency)	LNA: constant; Frequency synthesizer: constant; Mixer: constant; Filters: constant	2003

Continued on next page

Table2.2 – Continued from previous page

Ref.	**Transmit power consumption, P_{tr}**	**Circuitry power consumption, P_{cr}**	**Year**
[11]	$P_{tr} = c_t P_{rad}$ (c_t, a constant characterizing power used for modulation and amplification)	Source coder: linearly proportional to the number of operations	2003
[12]	Model in [10]	Model in [8]	2003
[13]	$P_{tr} = P_{rad}\frac{\xi}{\eta}$ (ξ, signal peak to average ratio (PAR))	All component models in [10]; ADC: linearly proportional to sampling frequency; DAC: sum of static and dynamic power consumption; Viterbi decoder: linearly proportional to the bit rate	2005
[14]	Adapt model in [13] for MIMO	Adapt models in [13] for MIMO	2004
[15]	Model in [13]	Receive power: constant; Sleeping mode power: constant	2007
[16]	Model in [13]	Models in [13]	2008
[17]	Model in [13]	Models in [13]	2009
[18]	Adapt model in [13] for MIMO	All component models in [13]; Baseband signal processing: linearly proportional to bandwidth; Forward error correction (FEC) decoder: linearly proportional to the processing rate; MIMO decoder: linearly proportional to the processing rate	2010

Using the system power models listed in Table 2.2, various system power minimization approaches have been proposed. The research on system power consumption minimization is summarized in Table 2.3.

Table 2.3: Summary of existing work on system power consumption minimization

Ref.	**Parameters**	**Constraints**	**Application**	**Year**
[8]	Modulation, transmission rate, radiate power	Constant BER, delay	Wireless packet scheduling systems	2001
[9]	Coding rate, constraint length, radiated power, processor supply voltage	BER, latency, range	Wireless micro sensor networks	2002
[10]	Modulation, radiated power	BER, transmission rate, delay, peak power	Short range applications	2003
[11]	Source coding, channel coding, radiated power	Constant end-to-end distortion	Wireless multimedia applications	2003
[12]	Modulation, radiated power, radio on/off	BER, average transmission rate	Wireless links	2003
[13]	Modulation, coding, radiate power	BER, delay, transmission rate	Short range applications	2005
[14]	Modulation, Alamouti based MIMO schemes	BER, Delay, transmission rate	MIMO and virtual MIMO sensor networks	2004
[15]	Modulation, subchannel allocation	BER	Cognitive sensor networks	2007
[16]	Modulation, radiate power, packet length, retransmission, overhead	BER	Sensor networks	2008
[17]	Modulation, radiated power	Average BER	Short range single-hop and multi-hop applications	2009
[18]	Modulation, coding, number of antennas, MIMO detection algorithm, bandwidth, radiated power	Packet error rate, transmission rate	Short range MIMO-OFDM applications	2010

2.3 Cognitive Radio

The term "Cognitive Radio" was first coined by Joseph Mitola [23]. Although there is no universal definition on cognitive radio, several definitions, such as those in [23, 24, 52, 53], show the most important common characteristics of a CR. A CR is an intelligent wireless communication system which is able to determine the most favorable operating parameters (cognition) based on the radio environment and its own capability (awareness) and reconfigure the radio accordingly (reconfigurability). By doing this, more efficient use of radio resource can be achieved. Awareness and cognition (learning, and decision making) reside in the cognitive engine. Here, a CE is defined as an "intelligent" agent that manages the cognition tasks in a CR, where intelligence denotes behavior that is consistent with a specified goal [53].

To date, the driver and the most common application of CR is spectrum management, to be specific, dynamic spectrum access (DSA). DSA is motivated by the lack of unassigned spectrum for new wireless applications and the observation of the inefficient use of most of the spectrum assigned under static allocation [54]. DSA aims at refarming the unused or inefficiently used spectrum, such as that within the TV spectrum band. Some examples include the dynamic frequency selection in IEEE 802.11 [55] and IEEE 802.16 [56], and CR for TV white space in IEEE 802.22 [57]. Other potential applications of CR include link reliability improvement [58], automated interoperability for public safety systems [59], cognitive networking [60], spectrum markets [61], femtocells [62], self-organizing networks, cooperative relaying and networks [63], smart grid communications [64], and vehicular networks [65]. CR promises to dramatically improve spectrum access [54], capacity [66], and link performance [67] while also more closely connecting the behavior of the network to the needs of the user [68].

2.4 Opportunities in Advancing the State of the Art

From the review of the power reduction research shown in Tables 2.1-2.3, it is obvious that the research community has started to realize the difference between the radiated power and the consumed power. A closer look at the transmission power consumption models adopted in the literature (see Table 2.2) shows a clear path in understanding the relationship between PA output power and power consumption. This difference was much ignored first [8] where the PA radiated power was taken as consumed power. Then, several very similar linear and signal characteristics independent models were developed [9, 10, 11, 12] where the PA efficiency was modeled as a constant value less than 1 to characterize the imperfect power conversion from its DC power supply to its output. Later, this model was refined by incorporating signal PAR characteristic, first shown in [13, 14], then adopted by other researchers in the community [15, 16, 17, 18].

At this point, this is no reason for ignoring this difference in power consumption related research. However, we find out this model only suitable for theoretical Class A PA with an underline assumption that the maximum power of the signal is the same as the maximum PA output power in its linear region. Therefore, in this work, we develop a PA model which can characterize theoretical Class A and Class B PAs and even some realistic PAs in the whole linear region. It is obvious that due to a change of model some results would not hold and we investigate some typical scenarios while much more research opportunities remain.

In addition to the analysis of the potential power savings, this works adopts a CR as a tool to optimize system power consumption of mobile and wireless communication devices. A CR can not only learn the channel conditions and adapt its modulation, coding and radiated power as in conventional radios, but is also aware of the radio (component) capabilities and characteristics (e.g., PA efficiency characteristics). The knowledge of radio capabilities and characteristics and application helps to further optimize system power consumption. The power consumption optimization process considered in this work is rather complex and involves many different functional entities and layers, such as, radio platform supervision (e.g., PA efficiency characteristic extraction), environment measurement, radio configuration decision making, policy and constraint validation and communication with higher layer entities. A cognitive radio turns out to be a convenience framework to capture all these essential functions to the power savings.

Chapter 3

Power Optimization for SISO Systems Using Cognitive Radio

This chapter[1] shows how cognitive radio can help to optimize radio power consumption for SISO systems. It starts the discussion on the power optimization work with a relatively simple scenario, SISO systems. On this subject, we first bring up the idea of using cognitive radio as a tool to better optimize radio power consumption and a CR framework which enables such optimization. Then, the theoretical foundation and simulation results for power savings are presented for SISO systems with theoretical Class A and Class B and some practical PAs using the CR framework. A unified PA efficiency model characterizing theoretical Class A and Class B and some practical PAs is adopted in this work and facilitates the formulation and analysis of the power consumption optimization problem.

Note that the efficiency for theoretical Class A and Class B and some practical PAs can be modeled as:

$$\bar{\eta} = \frac{P_t}{P} = \eta\left[P_t\right] = \left(\frac{P_t}{P_{max}}\right)^{\alpha} \cdot \eta_{max}, \tag{3.1}$$

where P_t is the PA output power, P is the PA power consumption, P_{max} is the maximal PA output power,

η_{max} is the maximum PA efficiency which is always less than 1, and α characterizes the PA efficiency characteristics and is $0 < \alpha \leq 1$ for theoretical Class A and Class B and some practical PAs considered in this work. Hence, any radio configuration minimizing the radiated power, P_t, minimizes the consumed power, P. This conclusion generally holds for any SISO system whose PA efficiency characteristic can be modeled as (3.1) with $0 < \alpha \leq 1$.

However, this conclusion does not necessarily hold for the energy case as energy is a product of power and time (the faster it transmits, the more power it requires while the quicker it finishes). Therefore, this chapter studies energy consumption minimization per a fixed amount of data to be transmitted under a constant BER constraint for systems with theoretical Class A and Class B and some practical PAs.

In addition to the results included in this chapter, the average energy savings (battery lifetime improvement) can be roughly calculated given the probability density function of the distance between the radios. For example, assuming a uniform distribution for the distance in Figure 3.3, average energy savings are about 48% for $\alpha = 1$, which almost doubles the battery lifetime.

(The following is the published journal paper on this subject.)

Energy Consumption Minimization for Mobile and Wireless Devices - A Cognitive Approach

3.1 Abstract

Energy consumption for mobile and wireless communication devices, such as cell phones, has long been an important aspect for both designers and customers. This paper shows how a cognitive radio (CR) framework can help to reduce system energy consumption of a mobile and wireless communication device based on the application quality of service requirement, the channel condition, and the radio capabilities and characteristics. The CR framework enables not only adaptation of modulation, coding rate, coding gain, and radiated power as conventional adaptive modulation (AM) scheme, but also joint adjustment of radio component characteristics (e.g., power amplifier (PA) characteristics) to achieve high energy efficiency. A unified PA efficiency model characterizing theoretical Class A, Class B, and practical PAs is adopted and enables the analysis of the impact of different radio configurations and channel conditions on energy efficiency. Significant energy savings (up to 90%) using the proposed CR framework for systems with theoretical PAs

and with a realistic PA can be achieved compared with the conventional AM approach in simulation. This framework can also be used to manage other radio resources.

3.2 Introduction

As the mobile and wireless communication devices become more and more powerful, system energy consumption becomes one of the important aspects to be considered by design engineers and end users. As one of the methods reducing energy consumption in a radio device while meeting operational requirements, it has been shown that radio reconfiguration can improve energy efficiency. At the system level, for example, in [6, 8, 9, 13, 18, 69], according to the application quality of service (QoS) requirement along with the channel condition, modulation, coding, antenna configuration, and radiated power are adapted to minimize energy consumption, especially for short range sensor network applications. At the circuit level, for example, in [20, 21, 22], power amplifier (PA) biasing voltage and/or current are adjusted according to the PA output signal envelope to improve the PA efficiency at low output power levels.

However, the system level approach and the circuit level approach are investigated mostly independently. In [70], we have proposed a cognitive radio (CR) framework enabling synergies between the conventional system level adaptation and the conventional circuit level adaptation to further improve radio resource usage (e.g., energy consumption) based on not only the application QoS requirement and the channel condition as in conventional radios, but also the knowledge of radio (component) capabilities and characteristics rendered by CR. The proposed CR framework has achieved significant (up to 95%) energy savings for a mobile and wireless communication device using a theoretical Class A PA with two biasing levels [70].

Given the example of significant energy savings using CR framework shown in [70], this paper investigates the theoretical foundation of the CR-based energy optimization framework and greatly extends the previous work in the following ways. First, a unified PA efficiency model applicable to the theoretical Class A and Class B PAs is proposed. This unified model facilitates convenient and tractable analysis of the impact of PA efficiency characteristics and channel conditions on energy efficiency. An energy efficiency metric is also defined based on this unified model. Second, in addition to the theoretical Class A and Class B PAs, the efficiency characteristic of a realistic PA is also considered for the investigation of energy savings. It has been shown that the proposed unified efficiency model fits the realistic efficiency characteristics very well. Hence, this unified model helps to extend mathematical analysis to systems even with certain practical PAs and enable fast calculation of power/energy consumption avoiding a long lookup table needed for efficiency characteristics with high resolution. For all cases considered in this paper, the energy optimization CR framework achieves significant energy savings compared with the conventional adaptive modulation approach.

In practice, a cognitive radio can enable the additional energy saving over adaptive modulation. It can provide necessary information on radio environment and radio component capabilities and characteristics, determine favorable configuration for the QoS requirement and accumulate knowledge on the interaction between radio environment and radio component characteristics which is usually hard to model analytically. There are many interesting potential gains CR can provide as compared with conventional radio, the main objective of this paper is the potential gains that CR can offer in terms of energy consumption optimization. The analysis on the enabling capabilities (awareness and cognition) and the tradeoff between power savings and complexity will be presented in a future paper.

This paper is organized as follows. Section 3.3 reviews the CR-based optimization framework. Section 3.4 presents the system power and energy consumption models used in this paper and analyzes the energy efficiency of different radio configuration given different PA efficiency models. Section 3.5 evaluates the potential energy savings of the proposed scheme through simulations for different PA efficiency characteristics, including practical PA efficiency characteristics, under various channel conditions. Section 3.6 concludes the paper.

3.3 Energy Optimization Framework Using CR

The energy optimization CR framework proposed in our earlier work is shown in Figure 3.1 [70]. In this framework, the solid lines and boxes are components in a conventional radio and the dashed lines and the cognitive engine (CE) block are new components enabling CR capabilities.

Conventionally, adaptive modulation (AM), dating back to 1960's [34], is used to adapt the radio configuration (e.g., modulation, coding, and radiated power) to channel changes for certain communication goal (e.g., maximizing rate or minimizing radiated power) with little consideration of radio component characteristics in the optimization process. On the other hand, in a CR, the CE learns the radio (component) capabilities and characteristics and controls the radio configuration through the bidirectional connections between the CE and various component blocks in Figure 3.1. The CR can change the radio configuration to optimize the use of radio resource based on the application QoS requirement, the channel condition, and the knowledge of

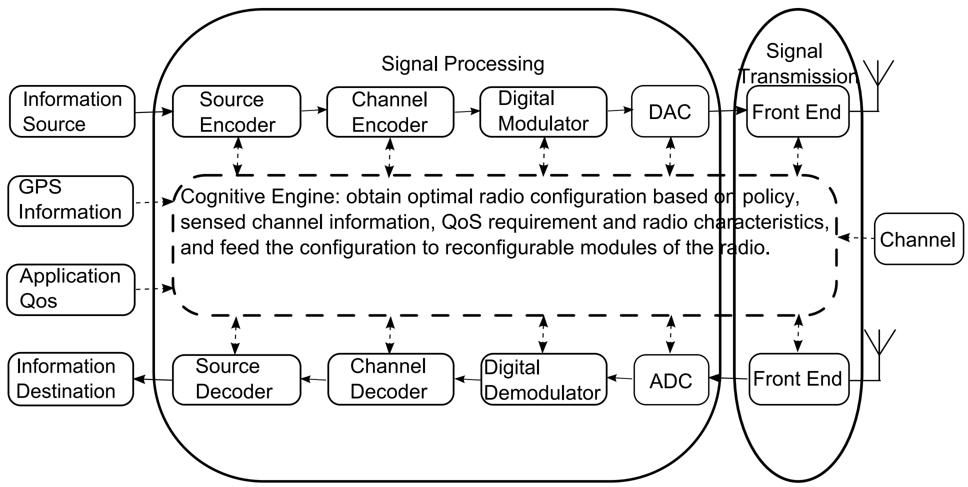

Figure 3.1: CR energy optimization framework.

component characteristics and capabilities. This paper names the CE-enabled dynamic transmission scheme as cognitive transmission (CT). In general, the CT knows more information about the radio capabilities and characteristics and makes better use of the information than the conventional AM does. Two levels of CT enabled by the capabilities of CR are investigated in this paper: conventional AM with component knowledge (e.g., knowledge of the efficiency characteristic of the PA in the system) and conventional AM with component adaptation (e.g., the adaptation of the PA biasing leading to favorable PA operation mode).

Another important aspect of CT is the acquirement of component characteristics. This information can be obtained in several ways. First, the product data sheet usually provides a good starting point for this knowledge. Second, field measurement and observation can refine the knowledge for specific radio and environment. Of course, this would need certain hardware support, such as, small sensors, in the radio. This support, however, would be necessary for future radio if further performance improvement is desired. Similar sensors have already been deployed on the motherboard of a personal computer for real time monitoring and diagnosing purposes. Similar technology would be expected to appear on radios in the near future. This paper assumes such support and component characteristics (i.e., PA efficiency characteristics) are available to the CE. The learning process of CE will be left as a future study for further investigation in more details.

3.4 System Power and Energy Consumption Models

For mobile and wireless devices used in medium and long range mobile and wireless communications, such as cellular phones, the PA usually dominates the system power and energy consumption. Therefore, among various important parameters of a communication system, this paper focuses on the impact of the PA on the radio energy consumption. Other significant energy consuming components, such as, analog to digital converters (ADC), digital to analog converters (DAC), mixers, digital baseband signal processing units, various filters, and user interface components, can be incorporated in the CR framework for different applications.

3.4.1 Radiated Power and Radiated Energy Models

The required average radiated power for a given QoS requirement (i.e., bit error rate (BER)) can be expressed as [71]:

$$P_t\left[d, M, R_c, G_c, R_s\right] = \frac{\left(\frac{E_b[M,Rc,Gc]}{N_0}\right)_{req.} \cdot R_b\left[Rs, M\right] \cdot N_0 \cdot PL\left[d\right] \cdot Lm}{G_t \cdot G_r}, \tag{3.2}$$

where $\left(\frac{E_b[M,Rc,Gc]}{N_0}\right)_{req.}$ is the required signal to noise ratio (SNR) per bit for a given QoS requirement depending on modulation scheme (M), coding rate (R_c), and coding gain (G_c) used. R_b is the bit rate given by $R_b = R_s \cdot M$ (R_s symbol rate, M modulation order or the number of bits per symbol). N_0 is the noise power spectral density. G_t and G_r are the transmitting antenna gain and the receiving antenna gain, respectively. $PL\left[d\right]$ is the path loss at distance d with a path loss exponent n, and Lm is the link margin. G_t, G_r and Lm depend on the radio and system setup and are assumed fixed and known in the following discussion. We also assume R_s is fixed and known, which suggests the bandwidth in the system is fixed.

The total radiated energy can then be expressed as:

$$E_t\left[d, M, R_c, G_c, R_s\right] = P_t\left[d, M, R_c, G_c, R_s\right] \cdot T, \tag{3.3}$$

where T is the total transmission time for a task in seconds.

3.4.2 System Power and Energy Consumption Models

In this paper, PA power consumption is the main focus in discussing system power consumption since the PA usually dominates the power consumption of a radio in medium and long range wireless applications. The system power and energy consumption can then be expressed as follows.

$$P\left[d, M, R_c, G_c, R_s, \eta\right] = \frac{P_t\left[d, M, R_c, G_c, R_s\right]}{\eta}, \tag{3.4}$$

and

$$\begin{aligned} E\left[d, M, R_c, G_c, R_s, \eta\right] &= P\left[d, M, R_c, G_c, R_s, \eta\right] \cdot T \\ &= \frac{P_t\left[d, M, R_c, G_c, R_s\right]}{\eta} \cdot T \\ &= \frac{E_t\left[d, M, R_c, G_c, R_s\right]}{\bar{\eta}}, \end{aligned} \tag{3.5}$$

where $\bar{\eta}$ is the average PA efficiency (over the underline waveform going through it), which is modeled using an engineering approximation as [70, 72] in this paper

$$\bar{\eta} = \eta\left[P_t\right], \tag{3.6}$$

where $\eta\left[P_t\right]$ is the instantaneous PA efficiency at the average output level P_t. In general, a more accurate PA average efficiency model can be developed if the distribution of the power of the underlying signal is available [72].

Using the theoretical instantaneous PA efficiency expressions in [73], the average PA efficiency for Class A and Class B PAs can be collectively written as

$$\bar{\eta} = \eta\left[P_t\right] = \left(\frac{P_t}{P_{max}}\right)^{\alpha} \cdot \eta_{max}, \tag{3.7}$$

where P_{max} is the maximum PA output power, η_{max} is the maximum PA efficiency, and α is the efficiency exponent depending on the type of the PA. To be specific,

$$\begin{cases} \eta_{max,A} = 0.25 & \alpha_A = 1, \quad \text{Class A;} \\ \eta_{max,B} = 0.785 & \alpha_B = 0.5, \quad \text{Class B.} \end{cases} \tag{3.8}$$

The theoretical efficiency characteristics of Class AB PA generally falls in between Class A and Class B PAs. In order to make the problem analytically tractable, this paper assumes that the theoretical efficiency of a Class AB PA can also be represented as (3.7) where $\eta_{max,A} < \eta_{max,AB} < \eta_{max,B}$ and $\alpha_B < \alpha_{AB} < \alpha_A$. Note that this unified PA efficiency model, (3.7), also fits well for some realistic PA efficiency characteristics as shown in Section IV-B3). In other words, this unified PA efficiency model enables convenient and tractable mathematical analysis on system energy efficiency of different radio configuration (e.g., the following modulation energy efficiency analysis) during product development phase.

The system energy consumption in (3.5) can be expressed by using (3.2), (3.3), and (3.7), as

$$\begin{aligned} & E\left[d, M, R_c, G_c, R_s, \eta\right] \\ = & \tfrac{P_t\left[d, M, R_c, G_c, R_s\right]}{\bar{\eta}} \cdot T \\ = & \tfrac{D}{M \cdot R_s} \cdot \tfrac{P_{max}^{\alpha}}{\eta_{max}} \cdot P_t\left[d, M, R_c, G_c, R_s\right]^{1-\alpha}, \end{aligned} \tag{3.9}$$

where $D = T \cdot R_b = T \cdot M \cdot R_s$ is the amount of data to be transmitted.

The energy efficiency of different radio configuration can then be compared based on (3.2) and (3.9). The relative energy efficiency metric is defined as the ratio of the energy consumption of two radio configurations and it provides insight on how one scheme compares with another in terms of energy consumption given the same application requirement.

$$\begin{aligned} m &= \frac{E_1\left[d, M, R_c, G_c, R_s, \eta\right]}{E_2\left[d, M, R_c, G_c, R_s, \eta\right]} \\ &= \left[\frac{\left(E_b\left[M_1, R_{c,1}, G_{c,1}\right]/N_0\right)_{req,1}}{\left(E_b\left[M_2, R_{c,2}, G_{c,2}\right]/N_0\right)_{req,2}}\right]^{1-\alpha} \cdot \left(\frac{M_2}{M_1}\right)^{\alpha}. \end{aligned} \tag{3.10}$$

Note that this definition assumes that different configuration uses the same PA, antenna, etc. Hence, the factors based on them are canceled out. This is the case for most radios where reconfiguration is usually limited to modulation, coding, and output power. We call this "soft" reconfiguration. However, this metric can be easily extended to include reconfigurable PA, antenna, etc., where the fundamental characteristics of these component can be changed as needed. In that case, we call it "hard" reconfiguration. This paper discusses the soft reconfiguration in details since it characterizes the reconfiguration capability of most radios so far. We do discuss the extra benefit that can be achieved for radios with certain hard reconfiguration capability, such as changing PA biasing level to change its efficiency characteristics.

Using (3.10), in terms of energy efficiency of different radio configuration, we have

$$\begin{cases} m < 1, & \text{configuration 1 is more efficient;} \\ m = 1, & \text{configuration 1 and 2 are equally efficient;} \\ m > 1, & \text{configuration 2 is more efficient.} \end{cases} \tag{3.11}$$

It is clear from (3.10) and (3.11) that the relative energy efficiency of each radio configuration varies, depending on the PA efficiency characteristics, in our model, α in (3.7). This is also confirmed in simulation results. In practice, the energy efficiency of each radio configuration depends on the efficiency characteristics of the actual PA used in the radio.

Then, for a delay-insensitive application, the system energy consumption minimization problem with a QoS constraint can be mathematically defined as

$$E_{min}\left[d\right] = \min_{\{M,R_c,G_c,\bar{\eta}\}} \left\{E\left[d, M, R_c, G_c, R_s, \bar{\eta}\right]\right\}|_{QoS}. \tag{3.12}$$

In contrast, the transmit energy minimization problem with the same constraint usually associated with conventional AM is formulated as

$$E_{t,min}\left[d\right] = \min_{\{M,R_c,G_c\}} \left\{Et\left[d, M, R_c, G_c, R_s\right]\right\}|_{QoS}. \tag{3.13}$$

Due to the nonlinear relationship between the radiated power and the consumed power of the PA in (3.6) and (3.7), the radio configuration minimizing radiated energy obtained in conventional AM does not necessarily minimize consumed energy. Note that although this paper focuses on a delay-insensitive application, the optimization problem defined in (3.12) and the following discussion and simulation can be extended for a delay-sensitive application by adding a data rate constraint. In addition, the settling time constraints need to be considered to in delay-sensitive applications to meet the delay constraint. These factors will be considered in our future work.

Table 3.1: Simulation parameters

Parameter	**Value or Range**
Symbol rate, R_s	constant 10^6
QoS (BER) requirement, P_b	constant 10^{-6}
Distance, d	200 - 1000 m
Path loss exponent, n	3
Modulation, M (number of bits per symbol)	QPSK, $M = 2$; 16QAM, $M = 4$; 64QAM, $M = 6$
Coding rate, R_c	1 (no coding); $\frac{3}{4}$; $\frac{2}{3}$; $\frac{1}{2}$
Coding gain, G_c	0 - 8 dB
PA efficiency, η	low biasing, η_l (maximum output power $P_{max,l} =$ 20 dBm); high biasing, η_h (maximum output power $P_{max,h} = 32$ dBm)

3.5 Simulation Results

3.5.1 Performance Metric and Simulation Environment

The performance metric (energy savings) is defined as

$$S_E = \frac{E_{AM} - E_{CT}}{E_{AM}} \cdot 100\%, \tag{3.14}$$

where E_{CT} is the radio energy consumption when operating in CT mode, and E_{AM} the radio energy consumption when operated in AM mode.

The performance of the proposed energy minimization CR framework is simulated in the radio channel with power law path loss and additive white Gaussian noise (AWGN). Convolutional code with variable coding rate and coding gain is used. The PA characteristics (e.g., efficiency) are assumed known to the CE. Table 3.1 summaries the simulation parameters which have been used in [70] earlier.

The energy minimization problem in (3.12) has been solved with exhaustive search in this paper since the search space is limited. More computationally efficient algorithms (e.g., making use of experience) can be developed for practical applications. The discussion of such algorithms is out of the scope of this paper.

3.5.2 Simulation Results

The energy savings using the CR framework for single channel communication system with Class A PA under a relatively static channel environment (fixed path loss exponent) has been investigated in detail in [70]. This paper focuses on the impact of component characteristics (i.e., different PA efficiency characteristics) and environment characteristics (i.e., different path loss exponents) on energy savings. It also investigates the energy savings using the CR framework for a system with a realistic linear PA[2] [74], which has been used in many different communication devices.

We partition our CT simulation into two levels. The first level CT extends the conventional AM to minimize energy consumption by incorporating the knowledge of PA efficiency characteristics in the optimization. In this level of CT, the PA is biased at a fixed level and its maximum output power is sufficient to cover the target distance range. The transmission parameters (modulation scheme, coding rate, coding gain, and radiated power) are adapted to minimize energy consumption for various distances with the knowledge of the PA efficiency characteristics. The second level CT extends the first level CT by allowing adaptation of PA efficiency characteristics. The second level CT jointly adapts the PA efficiency characteristics and the transmission parameters used in level one CT. We call this kind of PA a gear shifting PA [70]. It is assumed that the PA gear shifting is achieved by adapting the PA biasing levels as proposed in [20, 21, 22] in order to reduce PA power consumption at low output power levels. Two biasing levels (low and high) are adopted in the gear shifting PA, resulting in two different PA efficiency characteristics (e.g., see Figure 3.2 for those based on a theoretical Class A PA efficiency characteristics. The fixed biasing level of the PA adopted in the level one CT is the same as the high biasing level of the gear shifting PA used in level two CT. In other words, the PA efficiency characteristic used in level one CT corresponds to the curve for high biasing level in Figure 3.2, whereas for level two CT, both curves in Figure 3.2 can be used as needed.

Impact of PA Efficiency Characteristics

As shown in [70], the proposed CR framework and CT scheme can achieve significant energy savings for a communication system with Class A PA and level two CT can achieve even further savings compared with level one CT. This section focuses on the impact of different PA efficiency characteristics on the energy savings.

As discussed earlier, the energy efficiency of a transmission scheme depends on the efficiency characteristics

[2]The Texas Instruments (TI) TRF4002.

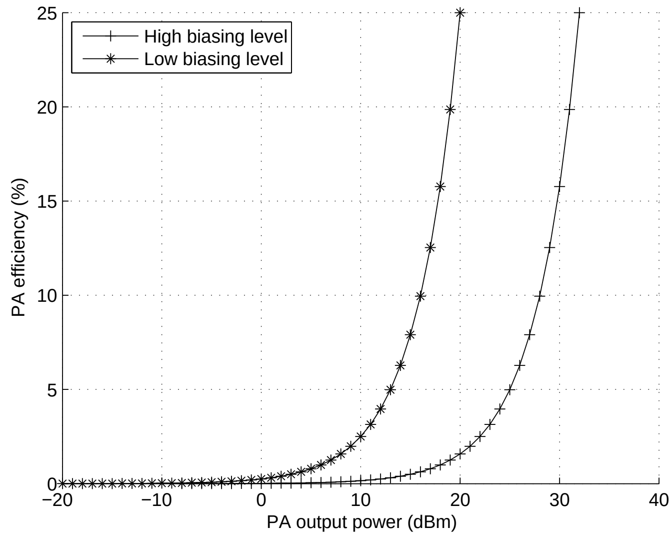

Figure 3.2: Gear shifting PA efficiency characteristics with two biasing levels based on theoretical Class A PA.

of the PA used in the radio. The critical values for the efficiency component, α, for the modulation schemes investigated in this paper can be obtained as follows. For example, at $BER = 10^{-6}$, the required SNR values for uncoded QPSK, 16QAM, and 64QAM are 10.5 dB, 14.5 dB, and 19 dB, respectively. In order for 16QAM to be more energy efficient than QPSK, using (3.10) we have

$$m = \left(\frac{10^{10.5/10}}{10^{14.5/10}}\right)^{1-\alpha} \cdot \left(\frac{4}{2}\right)^{\alpha} > 1. \tag{3.15}$$

Solving (3.15), we have $\alpha > 0.57$. Similarly, we can obtain critical values for other scenarios. Hence, we have

$$\begin{cases} 0.5 \le \alpha \le 0.57 & QPSK; \\ 0.57 \le \alpha \le 0.72 & 16QAM; \\ 0.72 \le \alpha \le 1 & 64QAM. \end{cases} \tag{3.16}$$

This result[3] is an important practical extension of the energy efficiency concept in most digital communications textbooks, such as [71]. It shows that the energy efficiency of a particular modulation scheme depends on the PA efficiency characteristics of the radio where the modulation is used. Note that (3.16) holds when-

[3] To better understand this result, an extreme case might be helpful. Consider a Class A PA case where $alpha = 1$. From (3.7), the consumed power of a Class A PA is the same no matter what the radiated power is. In this case, although a higher order modulation, such as 64QAM, requires a higher radiated energy per bit, it takes shorter time to transmit the data and therefore consumes a smaller amount of energy compared with a lower order modulation, such as QPSK.

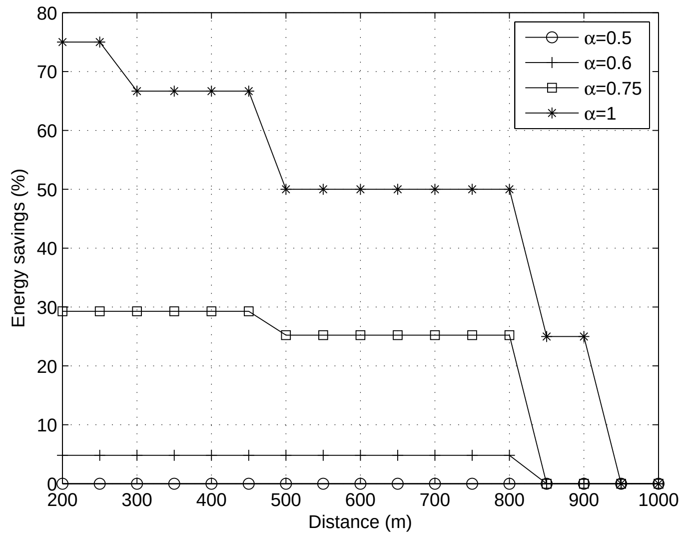

Figure 3.3: Energy savings comparison for different PA efficiency characteristics using level one CT.

ever the maximum output power of the PA can support that modulation at certain distance. This analysis can be readily extended to other modulation schemes and coded cases. Furthermore, implementation loss in practical systems does not affect this metric as in most systems the implementation loss is usually the same across different modulation and its impact is then canceled out.

To understand the impact of different PA efficiency characteristics on energy savings, the energy savings using level one CT for systems with four different PA efficiency characteristics ($\alpha = 0.5$, $\alpha = 0.6$, $\alpha = 0.75$, and $\alpha = 1$) are shown in Figure 3.3. The corresponding modulation choices for these efficiency characteristics are shown in Figure 3.4. It is clear from the simulation results that energy savings and modulation choices are consistent with that predicated by (3.16). Whenever a modulation order higher than QPSK can be chosen, some amount of energy can be saved. This is because conventional AM tends to choose QPSK in order to minimize radiated energy without considering the imperfect energy conversion in PA. Hence, to minimize energy consumption due to both radiation and PA operation, higher order modulation can be used to take advantage of more favorable PA operation point to save energy. In Figure 3.4, the modulation order decreases at further distances. This is because the required radiated power of the favorable modulation at further distances exceeds the maximal PA output power and only lower order modulation can be supported at those distances.

Further energy savings can be achieved using level two CT for all cases discussed. For example, the energy savings with two levels of CT for $\alpha = 0.75$ and $\alpha = 1$ are shown in Figure 3.5. Clearly from Figure 3.5,

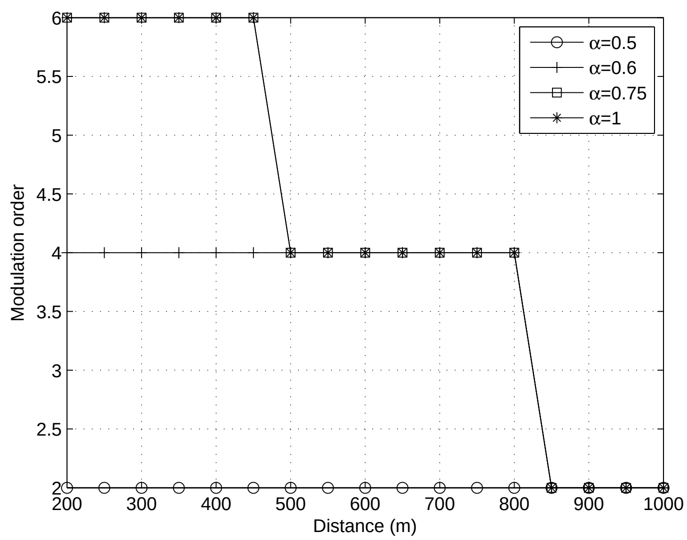

Figure 3.4: Modulation comparison for different PA efficiency characteristics using level one CT.

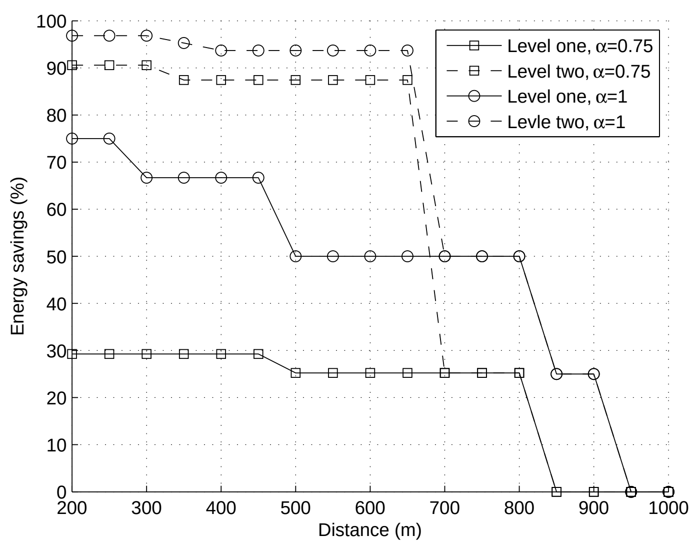

Figure 3.5: Energy savings using two levels of CT for different PA efficiency characteristics.

additional energy savings are available using level two CT than corresponding level one CT for short distances, where the radiated power requirement is lower. In addition, higher energy savings can be achieved with larger value of α which corresponds to worse PA efficiency performance.

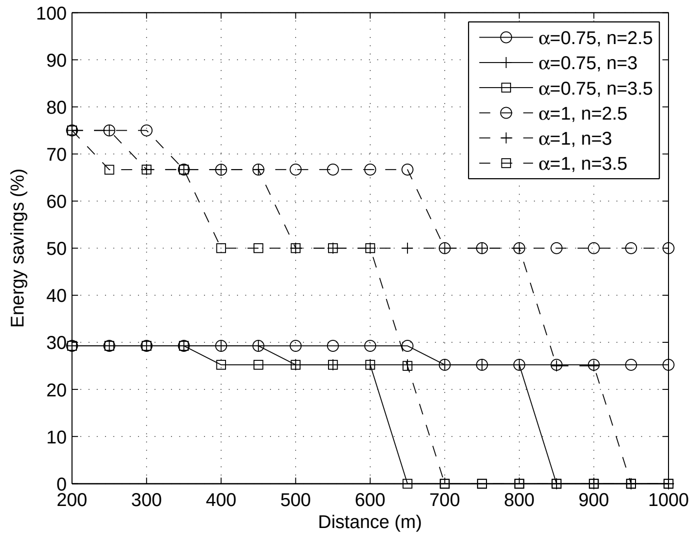

Figure 3.6: Energy savings using level one CT for different α under various channel path loss exponents.

Impact of Channel Conditions

The above section shows that significant energy savings can be achieved using the proposed CR framework and CT scheme for a communication system with different PA efficiency characteristics under a channel with a fixed path loss exponent. This section investigates the impact of different path loss exponents on the energy savings.

The energy savings for systems using level one CT for $\alpha = 0.75$ and $\alpha = 1$ under some typical channel path loss exponents ($n = 2.5$, $n = 3$, and $n = 3.5$) are shown in Figure 3.6. Observe that higher savings are achieved at the same distance under favorable channel conditions (smaller path loss exponent) for both PAs since better channel condition allows more choices in terms of possible radio configuration at the same distance. Again, additional savings can be achieved using level two CT for all channel conditions. These results are omitted here to avoid repetition.

System with a Realistic PA

After the investigation of the impact of different PA efficiency characteristics and channel conditions on the performance of the proposed energy optimization CR framework, this section considers the energy savings using the two levels of CT for a system with a realistic PA under a channel with a path loss exponent $n = 3$.

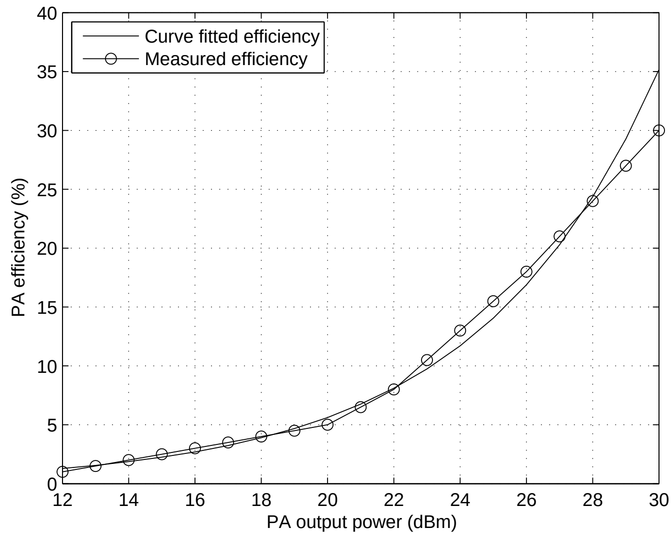

Figure 3.7: The realistic PA efficiency characteristics and the curve fitted efficiency model.

The efficiency characteristics of the realistic PA shown in Figure 3.7 are based on the product data sheet [74]. Also shown in this figure is the curve fitted model of the realistic PA efficiency to the PA efficiency model in (3.7) under the minimum mean square error (MMSE) criterion. The curve fitted efficiency model ($\eta_{max} = 0.35$ and $\alpha = 0.8$) is reasonably close to the realistic efficiency characteristics. This suggests the applicability of the proposed PA efficiency model in (3.7) for some practical PAs with unspecified type. This unified model enables mathematically tractable analysis of a practical system and fast calculation of power/energy consumption avoiding a long lookup table needed for efficiency characteristics with high resolution.

The energy savings are shown in Figure 3.8. The associated gear shifting PA efficiency characteristics are shown in Figure 3.9. Both levels of CT achieve great energy savings compared with AM and level two CT achieves even further energy savings than level one CT. The performance of the two levels of CT converges at a distance of 700 m where the required radiated power exceeds the allowable radiated power of the PA at the low biasing level and both schemes actually use the same PA (the PA with the high biasing level and hence higher maximum output power).

It is also verified by the simulation that the change of PA biasing happens at exact the same distance where the performance of the two levels of CT converges. These results are compatible with that obtained using theoretical PA efficiency characteristics in [70] and in the above sections. In addition, since the curve fitted PA efficiency for the practical PA has a value $\alpha = 0.8$, its energy savings performance is expected to fall

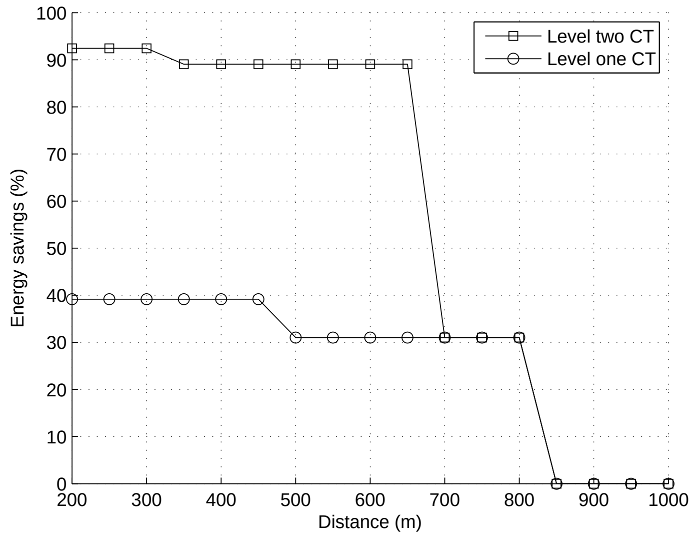

Figure 3.8: Energy savings for level one and level two CT for the realistic PA.

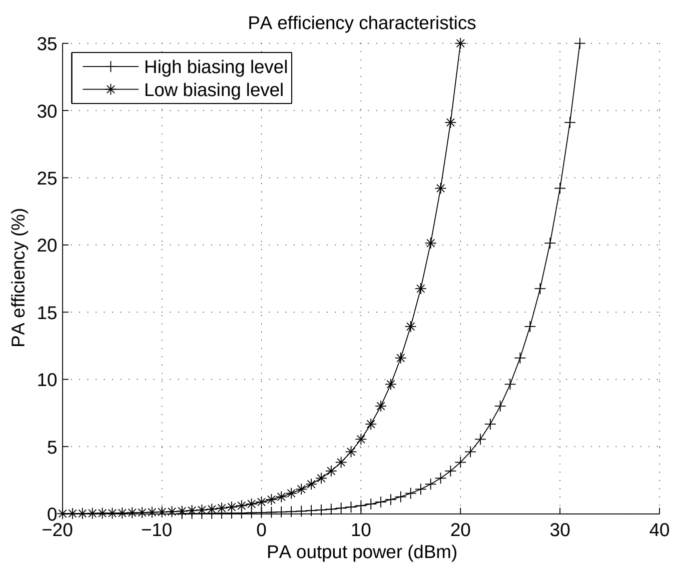

Figure 3.9: Gear shifting PA efficiency characteristics with two biasing levels based on the efficiency characteristics of the realistic PA.

between that for $\alpha = 0.75$ and $\alpha = 1$ in Figure 3.5. This is confirmed in Figure 3.8.

3.6 Conclusion

This paper provides the theoretical foundation and simulation results for energy savings using an energy optimization cognitive radio framework. A unified mathematical PA efficiency model applicable to theoretical Class A, Class B, and Class AB, and practical PAs is proposed in this paper. This model enables convenient and tractable mathematical analysis and optimization for system energy consumption during development phase. The impact on system energy efficiency of different PA efficiency characteristics (including realistic PA efficiency characteristics) and various channel conditions has been investigated. The results show that the proposed cognitive radio framework can utilize radio resources more efficiently than the conventional adaptive modulation approach by matching radio capabilities and characteristics with channel conditions and application QoS requirements. All these benefits stem from incorporating the knowledge of the radio capabilities and characteristics in the optimization process. The cognitive radio framework enables such utilization. This framework can be applied to optimize the use of radio resources for other purposes such as spectral efficiency improvement as well.

Chapter 4

Power Optimization for MIMO Systems Using Cognitive Radio

After the investigation on SISO systems, this chapter[1] moves on to a more complicated topic, MIMO systems, and shows how cognitive radio can help to optimize radio power consumption for MIMO systems. Leveraging results from information theory and capabilities of a CR (e.g., the awareness of the component capabilities and characteristics), a theoretical framework on the system power consumption minimization for MIMO systems is developed. This chapter first mathematically formulates the system power consumption minimization problem under a sum rate constraint for MIMO systems. The impact of channel correlation and partial channel state information at the transmitter is considered. Numerical algorithms are then developed to solve the constrained optimization problems. The simulation results on power savings and relative performance of the numerical algorithms are discussed. This optimization in this work only uses spacial multiplexing. Other MIMO techniques, such as beamforming might be included in the optimization formulation and new results might be obtained accordingly.

The result in this work is based on simulation of a confidence interval of 95%. This confidence interval determines the number of simulation runs to determine the average performance metrics, such as, power savings and number of active antennas.

Note that the simulation results in this chapter indicate a high possibility that one of the numerical algorithms, to be specific, the branch optimization algorithm, might actually be optimal as its performance is always hardly distinguishable from that of the exhaustive search based quasi-optimal algorithm for all the scenarios investigated in this work. Unfortunately, I was not able to prove it. This might be meaningful future work as it might give hints in solving many related problems in MIMO as well as MIMO-OFDM systems.

(The following is the accepted journal paper on this subject.)

Power Consumption Minimization for MIMO Systems - A Cognitive Radio Approach

4.1 Abstract

This paper shows how cognitive radio (CR) can help to optimize system power consumption of multiple input multiple output (MIMO) communication systems. Leveraging results from information theory and capabilities of a CR (e.g., the awareness of the component capabilities and characteristics), a theoretical framework is developed to minimize the system power consumption of MIMO systems while still considering radiated power. This paper mathematically formulates the system power consumption minimization problem under a sum rate constraint for MIMO systems. The impact of channel correlation and partial channel state information at the transmitter is considered. Numerical algorithms are developed to solve the constrained optimization problem. The simulation results show that significant power savings (e.g., up to 75% for a 4×4 MIMO system with Class A power amplifiers) can be achieved compared to conventional power allocation schemes. The results also show that the more computationally efficient suboptimal heuristic algorithms can achieve power savings comparable to the exhaustive search algorithm.

4.2 Introduction

This paper shows how cognitive radio (CR) can be used to optimize system power consumption of multiple input multiple output (MIMO) communication systems by dynamically reconfiguring the radio based on the required Quality of Service (QoS), the channel condition, and the knowledge of the platform (component) capabilities and characteristics.

Recent research has shown significant power reduction through radio reconfiguration based on channel conditions and QoS requirements, mostly in short range and sensor network applications. In [9, 14, 13, 15, 70, 75, 76], various system power consumption models have been proposed and used for power optimization by adapting modulation, coding, antenna configuration, and radiated power to channel conditions and QoS requirements.

The dual MIMO capacity problems, [77, 78], rate maximization (maximizing rate under a total power constraint) and power minimization (minimizing power under a sum rate constraint), have been widely investigated. Various power and bit allocation algorithms have been proposed to solve both problems. However, most investigations focused on the received power or the radiated power. For example, in [3], the modulation scheme and the antenna diversity scheme are adapted to minimize overall uplink radiated power in a multiple access network under average user throughput and overall network throughput constraints. In [6], the modulation scheme is adapted to minimize radiated power in MIMO systems with imperfect channel state information (CSI) under bit error rate (BER) and transmission rate constraints. The system power consumption of MIMO systems, on the other hand, has only received limited attention. In [14], the Alamouti code based MIMO scheme and modulation are jointly adapted to minimize system power consumption under throughput and delay constraints for sensor networks. In [75], the number of antennas, modulation, coding, MIMO detector, bandwidth, and radiated power are jointly adapted to minimize system power consumption under packet error rate and throughput constraints for short range applications. However, the fundamental relationship between rate and system power consumption has not been fully investigated.

With the advance of CR technologies, some capabilities of a CR can be adopted to optimize radio system power consumption further. A CR [23, 24] is an intelligent wireless communication system which is able to determine the most favorable operating parameters for the application QoS requirement (cognition) based on the knowledge of the radio environment and its capability and characteristics (awareness) and reconfigure the radio accordingly (reconfigurability). By doing this, radio resources can be used more efficiently. In [15], in addition to modulation adaptation, best available subchannel is dynamically detected and chosen for transmission to minimize power consumption under a bit error rate (BER) constraint in a CR sensor network setting. Moreover, a CR can not only learn channel conditions as in a conventional radio, but also be aware of platform (component) capabilities and characteristics. The knowledge of component capabilities and characteristics helps system power consumption optimization. A power optimization framework using CR for given QoS requirements based on channel conditions and radio capabilities and characteristics has been proposed in [70, 76]. Significant power savings are demonstrated using this framework.

This paper extends our previous work in [70, 76] to MIMO systems. The research leverages information theory results and cognitive radio capabilities (e.g., learning capability in obtaining radio capabilities and characteristics) for the development of a theoretical framework for system power consumption minimization of MIMO systems. Instead of focusing on specific MIMO techniques, such as those investigated in [14, 75], this paper mathematically formulates a generic system power consumption minimization problem with a sum rate constraint for MIMO systems and develops numerical solutions. This paper intends to provide some heuristic algorithms and show the potential benefit of this new approach in system power consumption minimization. A more rigorous approach to this problem will be studied in a more general cognitive radio based radio resource management framework. The work focuses on medium and long range applications where the power amplifier (PA) usually dominates the system power consumption. Other power consuming components, such as, digital signal processor, mixer, and low noise amplifier, can be integrated into the framework as future work.

The power saving approach proposed in this paper can be enabled by cognitive radio. The awareness of a cognitive radio can provide necessary information on radio component capabilities and characteristics as well as radio environment. The reasoning and learning enables the radio to determine favorable configuration for the QoS requirement and accumulate knowledge on the interaction between radio environment and radio component characteristics which is usually hard to model analytically. The focus of this paper is to emphasize the potential gains of this cognitive radio approach. The detailed analysis on the enabling capabilities (awareness and cognition) and the tradeoff between power saving and complexity will be presented in another paper.

This paper is organized as follows. Section 4.3 discusses the system model. Section 4.4 formulates the system power consumption minimization problem with a sum rate constraint. Section 4.5 discusses numerical algorithms to solve the constraint optimization problem. Section 4.6 evaluates the potential power reduction of the proposed framework by simulation. Section 4.7 concludes the paper.

4.3 System Model

A MIMO system with t transmit antennas and r receive antennas can be modeled as [79]

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{4.1}$$

where $\mathbf{x} \in \mathbb{C}^{t\times 1}$ is the transmitted signal vector, $\mathbf{y} \in \mathbb{C}^{r\times 1}$ is the received signal vector, $\mathbf{H} \in \mathbb{C}^{r\times t}$ is the channel matrix, and $\mathbf{n} \in \mathbb{C}^{r\times 1}$ is the noise vector, which is modeled as a zero mean complex Gaussian vector with independent and equal variance real and imaginary parts and covariance $\mathrm{E}\left\{\mathbf{n}\mathbf{n}^{\dagger}\right\} = \mathbf{I}_{\mathrm{r}}\sigma_{\mathrm{n}}^2$ where σ_n^2 is the noise variance and $\mathbf{I}_r$ is a $r \times r$ identity matrix. $\mathbf{H}$ is assumed to be independent of $\mathbf{x}$ and $\mathbf{n}$.

The radiated power from the transmit branches is

$$\tilde{\mathbf{p}} = \mathrm{diag}\left(\mathbf{Q}\right), \tag{4.2}$$

where $\mathbf{Q}$ is the covariance matrix of the transmitted signal vector $\mathbf{x}$, $\mathbf{Q} = \mathrm{E}\left\{\mathbf{x}\mathbf{x}^{\dagger}\right\}$ and $\mathrm{diag}(\cdot)$ is a vector consisting of the diagonal elements of the underlying matrix.

Therefore, the total radiated power becomes

$$\tilde{P}\left(\mathbf{Q}\right) = \sum_{n=1}^{t} \tilde{p}_n, \tag{4.3}$$

where $\tilde{p}_n$ is the nth element of $\tilde{\mathbf{p}}$ and the radiated power from transmit branch n.

The power consumption from branch n is modeled as

$$\hat{p}_n = \hat{p}_{PA,n} = \tilde{p}_n / \bar{\eta}_n, \tag{4.4}$$

where $\hat{p}_n$ is the nth element of the power consumption vector from transmit branches $\hat{\mathbf{p}}$, $\hat{p}_{PA,n}$ is the power consumption of the PA from transmit branch n , $\bar{\eta}_n$ is the average efficiency of the PA from transmit branch n. Note that only the power consumed by the PA is considered in (4.4) as this paper focuses on medium and long range applications where the PA at the transmitter usually dominates the system power consumption. Other power consuming components can be integrated into (4.4) as future work.

Thus, the system power consumption is

$$\hat{P}\left(\mathbf{Q}\right) = \sum_{n=1}^{t} \hat{p}_n = \sum_{n=1}^{t} \tilde{p}_n / \bar{\eta}_n. \tag{4.5}$$

An engineering approximation model is adopted for average PA efficiency [70, 76, 72].

$$\bar{\eta} = \eta\left(\tilde{P}\right), \tag{4.6}$$

where $\eta\left(\tilde{P}\right)$ is the instantaneous PA efficiency at the average output level $\tilde{P}$. A more accurate average efficiency model can be developed if the distribution of the power of the underlying signal is known [72]. With the instantaneous PA efficiency equations in [73], the average PA efficiency for Class A and Class B PAs can be collectively written as

$$\bar{\eta} = \eta\left(\tilde{P}\right) = \left(\tilde{P}/P_{max}\right)^{\alpha} \cdot \eta_{max}, \tag{4.7}$$

where P_{max} is the maximal PA output power, η_{max} is the maximal PA efficiency, and α is the efficiency exponent depending on the type of the PA. To be specific,

$$\begin{cases} \eta_{max} = \eta_{max,A} = 0.5 & \alpha = \alpha_A = 1, \quad \text{Class A;} \\ \eta_{max} = \eta_{max,B} = 0.785 & \alpha = \alpha_B = 0.5, \quad \text{Class B.} \end{cases} \tag{4.8}$$

Note that in a practical system, a cognitive radio would enable learning of the efficiency relationship (e.g., (4.7) and (4.8)).

Therefore, the system power consumption from branch n in (4.4) can then be expressed as

$$\hat{P}_n = \frac{P_{max}^{\alpha}}{\eta_{max}} \cdot \tilde{p}_n^{1-\alpha} = \begin{cases} \frac{P_{max}}{\eta_{max,A}}, & \text{Class A;} \\ \frac{P_{max}^{1/2}}{\eta_{max,B}} \cdot \tilde{p}_n^{1/2}, & \text{Class B} \end{cases}. \tag{4.9}$$

It is interesting to note that, according to (4.9), the power consumption of a Class A PA is independent of its output power level. In other words, the power consumption of a Class A PA is constant over its whole range of output power level. Note that in this work we assume that the PAs in all branches are the same, which is common in practice. However, this assumption can be removed by replacing P_{max}, η_{max}, and α with their corresponding values for each branch if the PA in each branch is different for certain application requirement.

4.4 System Power Consumption Optimization for MIMO Systems

The learning capability of a CR helps minimize system power consumption. A CR learns not only channel conditions as in a conventional radio, but also platform (component) capabilities and characteristics needed for system power consumption optimization. This section formulates the system power minimization problem with the knowledge on PA efficiency.

Based on the system model in Section 4.3, the system power consumption minimization problem with a sum rate constraint can be formulated mathematically. This paper investigates the problem under two common scenarios whose corresponding capacity results and capacity achieving power allocation algorithms are available in the information theory literature [79, 80] so that performance comparison of the proposed framework and the conventional MIMO system is possible. These scenarios are:

- Perfect channel state information at the receiver and the transmitter (CSIR-CSIT).
- Perfect channel state information at the receiver and channel distribution information at the transmitter (CSIR-CDIT).

In this work, the CSI refers to the channel matrix, $\mathbf{H}$. Therefore, under CSIR or CSIT, $\mathbf{H}$ is known at the receiver or the transmitter, respectively. Under CDIT, the statistical distribution of $\mathbf{H}$ is known at the transmitter.

For the CSIR-CSIT case, the system power consumption minimization with a sum rate constraint is formulated as

$$\hat{P} = \mathrm{E}_{\mathbf{H}}\left[\min_{\mathbf{Q}:C(\mathbf{Q})=C_{tgt}} \hat{P}\left(\mathbf{Q}\right)\right], \tag{4.10}$$

where $\mathrm{E}_{\mathbf{H}}\left[\cdot\right]$ is the expectation operation with respect to the channel matrix $\mathbf{H}$, C_{tgt} is the target rate, and the achieved rate for the input covariance matrix $\mathbf{Q}$ is given by [79, 80]

$$C\left(\mathbf{Q}\right) = \log\left|\mathbf{I}_r + \mathbf{H}\mathbf{Q}\mathbf{H}^{\dagger}\right|. \tag{4.11}$$

In this case, since every channel realization $\mathbf{H}$ is known at the transmitter, the transmitter can adapt the transmit signal covariance matrix $\mathbf{Q}$ accordingly and obtain one optimal transmit signal covariance matrix

$\mathbf{Q}$ for each channel realization $\mathbf{H}$ under the target rate constraint C_{tgt}. The system power consumption is averaged over all the channel realizations.

For the CSIR-CDIT case, the system power consumption minimization with a rate constraint can be formulated as

$$\hat{P} = \min_{\mathbf{Q}:\mathrm{E}_{\mathbf{H}}[C(\mathbf{Q})]=C_{tgt}} \hat{P}\left(\mathbf{Q}\right), \tag{4.12}$$

where $C\left(\mathbf{Q}\right)$ is given by (4.11). In this case, since the transmitter only knows the ensemble characteristics of the channel realization $\mathbf{H}$, it has to determine the transmit signal covariance matrix $\mathbf{Q}$ according to the distribution of the channel realization $\mathbf{H}$, resulting in only one transmit signal covariance matrix $\mathbf{Q}$ for all channel realizations abiding by the same channel distribution under the target rate constraint C_{tgt}.

Since the PA efficiency characteristics $\bar{\eta}$, depending on the type of PA, can be nonlinear as shown in (4.7) and (4.8), the analytical solutions to the constrained minimization problems defined in (4.10) and (4.12) can only be obtained for certain class of PA. Therefore, numerical algorithms are developed to solve the constrained optimization problems.

In contrast to the constrained system power consumption minimization problems defined in (4.10) and (4.12), the problems that have been widely investigated in information theory are defined as [77, 78]

$$\tilde{P} = \mathrm{E}_{\mathbf{H}}\left[\min_{\mathbf{Q}:C(\mathbf{Q})=C_{tgt}} \tilde{P}\left(\mathbf{Q}\right)\right], \tag{4.13}$$

and

$$\tilde{P} = \min_{\mathbf{Q}:\mathrm{E}_{\mathbf{H}}[C(\mathbf{Q})]=C_{tgt}} \tilde{P}\left(\mathbf{Q}\right), \tag{4.14}$$

for the CSIR-CSIT case and the CSIR-CDIT case, respectively. These constrained minimization problems focus on radiated power rather than system power consumption. Various power allocation algorithms have been developed to solve these problems. However, the solutions to (4.13) and (4.14) may not be the same as those to (4.10) and (4.12) due to the nonlinear relationship between the radiated power and the consumed power. This paper solves the constrained system power consumption minimization problems under several channel conditions and uses the solutions to the conventional constrained radiated power minimization problems as comparison baseline to demonstrate the benefit achieved with the knowledge of the radio component characteristics using the proposed framework.

4.5 Numerical Algorithms

This section discusses numerical algorithms to solve the constrained system power consumption minimization problem defined in (4.10) and (4.12) in Section 4.4. It is assumed that any branch of the MIMO system can be set to inactive independently as needed to reduce power consumption. Deactivating a transmit branch results in a change in **H**, i.e., the column of **H** corresponding to the inactive transmit branch is removed from **H**. In addition, for the same number of active transmit branches, different branch combination results in different **H**.

Conventionally, the water filling algorithm [77] is used to solve the problem defined in (4.13) [79, 80]. For the problem defined in (4.14), equal power allocation provides the optimal solution [79, 80]. However, due to the nonlinear relationship between the radiated power and the consumed power, the algorithms solving problems (4.13) and (4.14) are not necessarily the optimal solutions to the constrained system power consumption minimization problems defined in (4.10) and (4.12). For this reason, this paper proposes several algorithms to solve the constrained system power consumption minimization problems defined in (4.10) and (4.12). The optimal solutions to the constrained radiated power minimization problems defined in (4.13) and (4.14) are used as the comparison baseline.

4.5.1 Quasi-optimal Algorithm - Exhaustive Search

One way to find a quasi-optimal solution to the constrained minimization problem is exhaustive search. This approach tests all antenna configuration and power allocation combinations for a given power allocation step size to find a solution that is negligibly different than the true optimal solution. The pseudo code describing the exhaustive search process is in Table 4.1.

For a MIMO system with t transmit branches, the total number of power allocation combinations which has to be searched by the algorithm, is

$$\begin{aligned} N_{es} &= \sum_{n=1}^{t} \begin{pmatrix} t \\ n \end{pmatrix} \left(\left\lceil \frac{P_{max} - P_{min}}{P_{step}} \right\rceil + 1 \right)^{n} \\ &= \left(\left\lceil \frac{P_{max} - P_{min}}{P_{step}} \right\rceil + 2 \right)^{t} - 1, \end{aligned} \tag{4.15}$$

where P_{max}, P_{min}, and P_{step} are the maximal output power, the minimal output power and the step size of

Table 4.1: Quasi-optimal Algorithm - Exhaustive Search

- **Initialization**
  1. Create and store N_{es} power allocation combinations in $\mathbf{Q}_c$.
  2. Power allocation combination index: $i = 0$.
- **Iteration**
  1. $i = i + 1$.
  2. If $i \leq N_{es}$
     (a) Pick power allocation $\mathbf{Q} = \mathbf{Q}_c(i)$.
     (b) Calculate and store corresponding achieved rate, $C(\mathbf{Q})$.
     (c) Calculate and store corresponding radiated power, $\tilde{\mathbf{p}}$.
     (d) Calculate and store corresponding consumed power, $\hat{\mathbf{p}}$.
     (e) Calculate and store corresponding system power consumption, $\hat{P}(\mathbf{Q}) = \sum_{n=1}^{t} \hat{p}_n$.
     (f) Go to step 1.
- **Decision**
  1. If any power allocation achieves target rate, C_{tgt}
     (a) Return power allocation and antenna configuration achieving target rate with minimal power consumption.

     else

     (a) Return power allocation and antenna configuration achieving maximal rate with minimal power consumption.

the power allocation, respectively, and $\lceil \cdot \rceil$ is the ceiling function. This number is related to the computational complexity and the memory footprint of the algorithm. For a large power range, a small step size, or a large number of branches, the execution time and memory footprint of this algorithm can be prohibitive.

4.5.2 Suboptimal Algorithm 1 - Branch Adaptation

In order to reduce the computational burden and the memory requirement, a suboptimal algorithm is proposed. The water-filling algorithm and equal power allocation are used to allocate power for each branch combination for problems defined in (4.10) and (4.12), respectively. Hence, the time and memory consuming exhaustive search is avoided. This algorithm differs from the conventional water-filling algorithm and the equal power allocation algorithm in a way that the branches in the conventional water-filling algorithm and the equal power allocation algorithm are fixed while the choice of branches in this proposed algorithm is adapted to minimize system power consumption. Therefore, we call this algorithm branch adaptation. The pseudo code describing the branch adaptation algorithm is in Table 4.2.

The total number of power allocation combinations to be searched by the algorithm is

$$N_{ba} = \sum_{n=1}^{t} \begin{pmatrix} t \\ n \end{pmatrix} = 2^t - 1. \tag{4.16}$$

The branch adaptation algorithm results in much less computational burden and smaller memory requirement as compared to the exhaustive search algorithm.

4.5.3 Suboptimal Algorithm 2 - Branch Minimization

The suboptimal algorithm 1 can be further simplified based on the observation in the MIMO systems employing Class A PAs: the more transmit branches the MIMO system uses, the more power the system consumes. The power consumption of the Class A PA is the same no matter what the output power is. Therefore, it makes sense to use the minimal number of branches (PAs) as the system power consumption is linearly proportional to the number of active branches as long as the target rate can be achieved. In suboptimal algorithm 2, the number of active branches increases only when the current number of active branches cannot satisfy the rate requirement. We call this algorithm branch minimization. As in branch adaptation algorithm, the water-filling algorithm and the equal power allocation algorithm are used to

Table 4.2: Suboptimal Algorithm 1 - Branch Adaptation

- **Initialization**
  1. Number of active transmit antennas: $n_t = 0$.
  2. Power allocation: $\mathbf{Q} = \mathbf{0}$.
- **Iteration**
  1. $n_t = n_t + 1$.
  2. If $n_t \leq t$
     (a) Calculate and store power allocation, $\mathbf{Q}$, (water filling algorithm for problem (4.10) and equal power allocation for problem (4.12)), for target rate, C_{tgt}.
     (b) Calculate and store corresponding radiated power, $\tilde{\mathbf{p}}$.
     (c) Calculate and store corresponding consumed power, $\hat{\mathbf{p}}$.
     (d) Calculate and store corresponding system power consumption, $\hat{P}(\mathbf{Q}) = \sum_{n=1}^{t} \hat{p}_n$.
     (e) Go to step 1.
- **Decision**
  1. If any power allocation achieves target rate, C_{tgt}
     (a) Return power allocation and antenna configuration achieving target rate with minimal power consumption.

     else

     (a) Return power allocation and antenna configuration achieving maximal rate with minimal power consumption.

Table 4.3: Suboptimal Algorithm 2 - Branch Minimization

- **Initialization**
  1. Number of active transmit antennas: $n_t = 0$.
  2. Power allocation: $\mathbf{Q} = \mathbf{0}$.
- **Iteration**
  1. $n_t = n_t + 1$.
  2. If $n_t \leq t$
     (a) Calculate and store power allocation, $\mathbf{Q}$, (water filling algorithm for problem (4.10) and equal power allocation for problem (4.12)), for target rate, C_{tgt}.
     (b) Calculate and store corresponding radiated power, $\tilde{\mathbf{p}}$.
     (c) Calculate and store corresponding consumed power, $\hat{\mathbf{p}}$.
     (d) Calculate and store corresponding system power consumption, $\hat{P}\left(\mathbf{Q}\right) = \sum_{n=1}^{t} \hat{p}_n$.
     (e) If any power allocation achieves target rate, C_{tgt}
        i. Return power allocation and antenna configuration achieving target rate with minimal power consumption

        else if $n_t \geq t$

        i. Return power allocation and antenna configuration achieving maximal rate with minimal power consumption

        else

        i. Go to step 1.

allocate power for each branch combination for problems defined in (4.10) and (4.12), respectively. The pseudo code describing the branch minimization algorithm is in Table 4.3.

In the worst case, the total number of power allocation combinations for branch minimization algorithm is the same as that for the branch adaptation algorithm. The actual average number of combinations the branch minimization algorithm has to evaluate, depending on the distribution of the channel state and the target data rate, is lower than that of the branch adaptation algorithm.

Note that the branch minimization algorithm provides better performance for systems with Class A PAs than for systems with Class B PAs since the power consumption of systems with Class B PAs depends not only on the number of active branches but also the actual radiated power from each branches. This difference is observed in the later simulation results. However, the branch minimization algorithm is simpler than the exhaustive search algorithm and the branch adaptation algorithm. It may be useful for systems with Class B PAs when fast adaptation is desired.

4.6 Performance Evaluation and Simulation Results

This section evaluates the performance of the proposed optimization framework and algorithms by simulation. We assume a 4×4 MIMO system with identical branches ($r = t = 4$). Maximal output power for each branch is 1 W ($P_{max} = 1$ W) and an inactive branch radiates zero power ($P_{min} = 0$ W) and consumes zero power. In the exhaustive search algorithm, a step size $P_{step} = 10^{-3}$ W is used. The channel is memoryless flat Rayleigh faded. For the CSIR-CSIT case, both uncorrelated and correlated fading are considered. For the CSIR-CDIT case, the fading is assumed to be uncorrelated. For the uncorrelated case, $\mathbf{H}$ is assumed to be a complex Gaussian random matrix with independent and identically distributed (i.i.d.) entries, each entry having independent real and imaginary parts with zero mean and equal variance $\frac{1}{2}$ [79]. Hence, the amplitude of the element of the channel matrix follows a Rayleigh distribution. Correlated channel matrix is generated using uncorrelated channel matrix and covariance matrices at transmitter and receiver as shown in [81]. The maximum average signal to noise ratio (SNR) per channel[2] under maximum radiated power is assumed to be 10 ($\bar{\gamma}_{max} = 10$) to achieve the desired rate range. In other words, the radiated power from each branch can be scaled (from 0 to P_{max}) so that the received SNR satisfies the requirement for the given target rate. Monte Carlo simulation consisting of 10^5 trials is performed and one channel matrix realization is drawn for each trial.

The system power consumption reduction is defined as

$$P_{saving} = \frac{P_{con} - P_{cog}}{P_{con}} \cdot 100\%, \tag{4.17}$$

where P_{con} is the system power consumption of a MIMO system with a conventional power allocation scheme (i.e., the water-filling algorithm for the CSIR-CSIT case and the equal power allocation for the CSIR-CDIT case), and P_{cog} the system power consumption of a MIMO system with the proposed CR framework. The MIMO systems employing Class A PAs and Class B PAs are simulated for each case.

4.6.1 CSIR-CSIT Case

The proposed framework is evaluated for the CSIR-CSIT case under the uncorrelated and correlated Rayleigh fading environment, respectively. The comparison baseline is a MIMO system with the water-filling algorithm

[2] $\bar{\gamma}_{max}$ is defined as $\bar{\gamma}_{max} = \frac{P_{max}}{\sigma_n^2 \cdot r \cdot t} \mathrm{E}\left[\|\mathbf{H}\|_F^2\right]$, where $\|\mathbf{H}\|_F$ is the Frobenius norm of $\mathbf{H}$. In this work, we assume $\sigma_n^2 = 0.1$, so that $\bar{\gamma}_{max} = 10$.

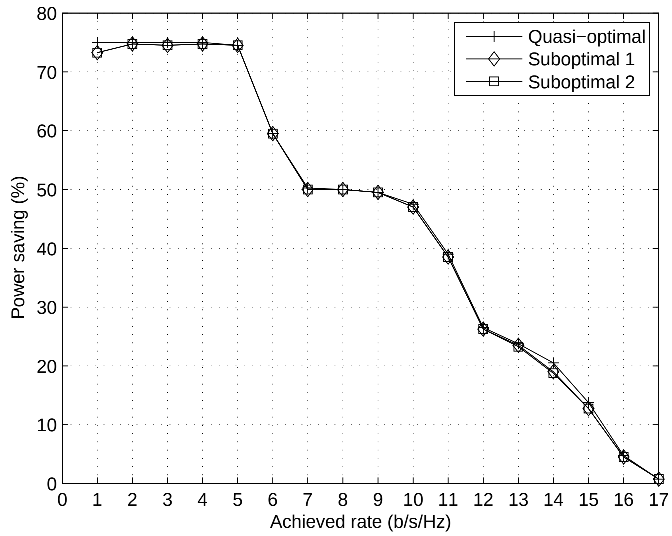

Figure 4.1: Power savings for MIMO systems with Class A PAs for CSIR-CSIT case.

which is the optimal power allocation scheme for the CSIR-CSIT case in terms of radiated power minimization [79, 80].

CSIR-CSIT with Uncorrelated Fading

The performance of the proposed framework for MIMO systems with Class A and Class B PAs under uncorrelated Rayleigh fading is discussed in this section.

The power savings and the corresponding antenna configuration of the proposed framework for MIMO systems with Class A PAs under different rate constraints are shown in Figures 4.1 and 4.2. Up to 75% power savings can be achieved depending on the rate requirements. The power savings decrease as the target rate increases. The proposed framework and algorithms tend to use as few transmit branches as possible. As mentioned in Section 4.3, the power consumption of Class A PA is the same for all output power levels. Therefore, the more transmit branches the MIMO system uses, the more power it consumes. At lower target rate, the proposed framework uses fewer transmit branches. As the target rate increases, more transmit branches are activated (see Figure 4.2). In other words, the power consumption of the proposed framework increases as the target rate. On the contrary, the system power consumption of the conventional approach is the same over all target rates since it always uses all transmit branches no matter what the target rate is. Therefore, the power savings decrease as the target rate increases. The decrease in power savings also

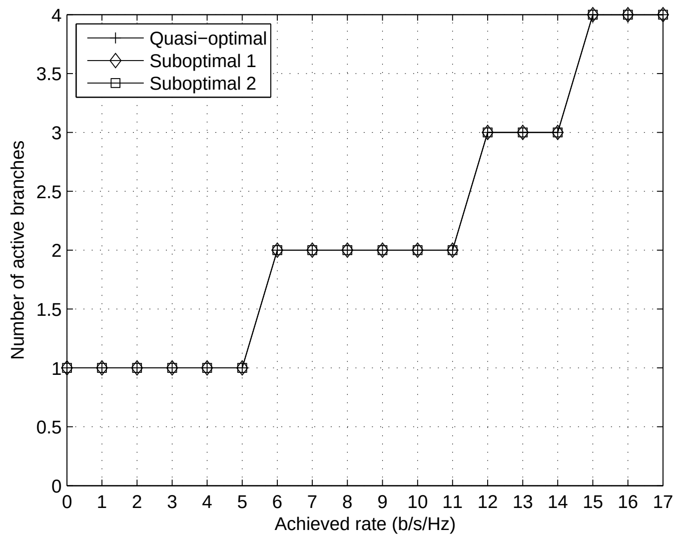

Figure 4.2: Antenna configuration for MIMO systems with Class A PAs for CSIR-CSIT case.

indicates that the power efficiency of conventional MIMO systems is better at higher rate region. Note that the power savings results are average power savings and obtained using Monte Carlo simulation. Due to the statistical nature of the channel, it is rare that all channels simultaneously fall in deep fade. Therefore, when the transmit antennas can be selected based on the knowledge of the channel matrix, it is highly likely that some good channels can be found for a given target rate. In other words, the number of antennas is expected to be similar even for different channel realizations given the same target rate.

The power savings and the average number of active antennas results in Figures 4.1 and 4.2 are almost identical for the exhaustive search, the branch adaptation, and the branch minimization algorithms. This is because the system power consumption and power savings for the Class A PA case are determined by the number of active antennas. The exhaustive search, the branch adaptation, and the branch minimization algorithms arrive at the same number of active antennas for the same target rate. The results suggest the branch minimization algorithm provides the best tradeoff between power savings and algorithm complexity for the Class A PA case.

The power savings and the corresponding antenna configuration results of the proposed framework for MIMO systems with Class B PAs are shown in Figures 4.3 and 4.4. Up to 30% of power savings can be achieved depending on the rate requirements. Compared with the Class A PA case, lower power savings are observed for the Class B PA case. This is because Class B PAs have higher efficiency than Class A PAs at all output power levels. Hence, the power consumption penalty for using more branches in conventional MIMO systems

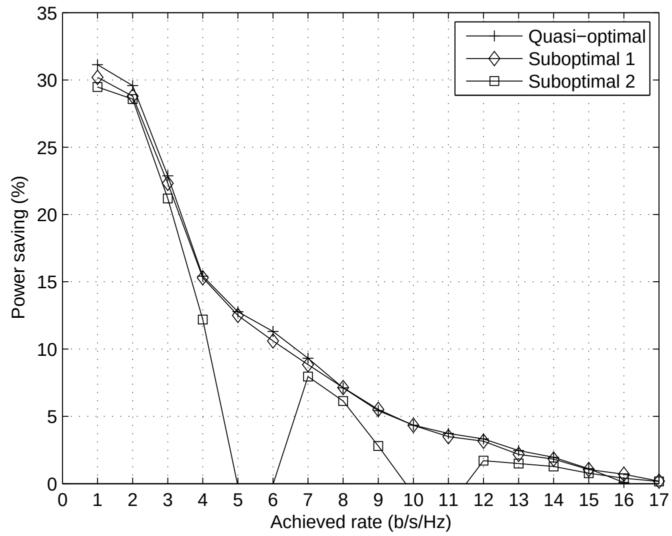

Figure 4.3: Power savings for MIMO systems with Class B PAs for CSIR-CSIT case.

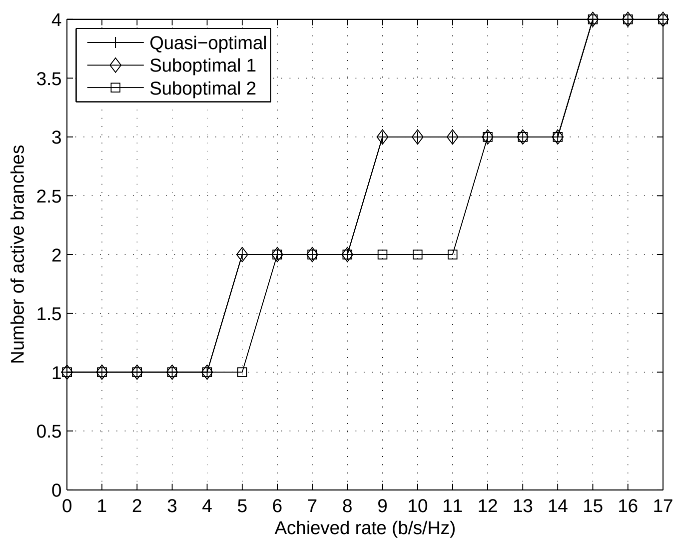

Figure 4.4: Antenna configuration for MIMO systems with Class B PAs for CSIR-CSIT case.

is less severe for the Class B PA case.

Note that the power savings and the average number of active antennas results in Figures 4.3 and 4.4 are almost identical for the exhaustive search and the branch adaptation algorithms. However, the performance gap between the exhaustive search and the branch minimization algorithms is larger than that in the Class A PA case. This observation is consistent with earlier discussion on the branch minimization algorithm that its performance may be inferior to the branch adaptation algorithm for systems with Class B PAs. For example, at target rate of 5 b/s/Hz, there is a large gap in power savings between the exhaustive search algorithm (and the branch adaptation algorithm) and the branch minimization algorithm as shown in Figure 4.3. From Figure 4.4, the exhaustive search algorithm (and the branch adaptation algorithm) uses 2 branches while the branch minimization algorithm uses only 1 branch at that rate. This suggests at certain rates, for MIMO systems with Class B PAs, using more transmit branches can reduce system power consumption. Overall, the results suggest the branch adaptation algorithm provides the best tradeoff between power savings and algorithm complexity for the Class B PA case.

CSIR-CSIT with Correlated Fading

This paper also investigates the performance of the propose framework and algorithms for the CSIR-CSIT case with correlated Rayleigh fading.

Channel measurement results show that the channel covariance matrix can be modeled as [82, 83]

$$\mathbf{R}_{\mathbf{H}} = \mathbf{R}_{\mathbf{H}}^{T_x} \otimes \mathbf{R}_{\mathbf{H}}^{R_x}, \tag{4.18}$$

where $\mathbf{R}_{\mathbf{H}}$, $\mathbf{R}_{\mathbf{H}}^{T_x}$ and $\mathbf{R}_{\mathbf{H}}^{R_x}$ are the channel covariance matrix, the covariance matrix at the transmitter, and the covariance matrix at the receiver, respectively, and $\otimes$ stands for Kronecker product operation.

The underlying correlated channel can then modeled as [81]

$$\mathbf{H} = \left(\mathbf{R}_{\mathbf{H}}^{R_x}\right)^{1/2} \mathbf{H}_w \left[\left(\mathbf{R}_{\mathbf{H}}^{T_x}\right)^{1/2}\right]^{\dagger}, \tag{4.19}$$

where $\mathbf{H}_w$ is a stochastic channel matrix with i.i.d. elements, and $(\cdot)^{1/2}$ and $(\cdot)^{\dagger}$ denote the matrix square root operation and matrix conjugate transpose operation, respectively. In this paper, the elements of $\mathbf{H}_w$ are assumed to be i.i.d. complex Gaussian random variables having independent real and imaginary parts with zero mean and variance $\frac{1}{2}$.

The covariance matrices at the transmitter and the receiver are modeled by the exponential correlation model

[84] for equally spaced antennas

$$\mathbf{R}_{\mathbf{H}}^{T_x} = \begin{bmatrix} 1 & \rho_T & \dots & \rho_T^{t-1} \\ \rho_T & 1 & \dots & \rho_T^{t-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_T^{t-1} & \rho_T^{t-2} & \dots & 1 \end{bmatrix} \tag{4.20}$$

and

$$\mathbf{R}_{\mathbf{H}}^{R_x} = \begin{bmatrix} 1 & \rho_R & \dots & \rho_R^{r-1} \\ \rho_R & 1 & \dots & \rho_R^{r-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_R^{r-1} & \rho_R^{r-2} & \dots & 1 \end{bmatrix}, \tag{4.21}$$

respectively, where ρ_T and ρ_R are covariance between adjacent antennas at the transmitter and the receiver, respectively.

As this work focuses on PAs at the transmitter, this section only shows the simulation results for systems with channel correlation at the transmitter. Similar results can be obtained for systems with correlation at the receiver and at both transmitter and receiver.

The power savings and the corresponding antenna configuration results obtained with the proposed framework for MIMO systems with Class A PAs are shown in Figures 4.5 and 4.6 for different channel correlation at the transmitter. In these figures, the results are plotted within their corresponding achievable rate region (i.e., 17 b/s/Hz for $\rho_T = 0$, 14 b/s/Hz for $\rho_T = 0.2$, and 8 b/s/Hz for $\rho_T = 0.8$). The proposed framework and algorithms achieve significant power reduction. In general, the capacity decreases as the antenna correlation increases [80]. In other words, more antennas are usually needed to achieved the same target rate as correlation increases. Therefore, the power savings decrease. This relationship between the power savings and the number of active antennas is clearly demonstrated in Figures 4.5 and 4.6 for MIMO systems with Class A PAs.

The power savings and the corresponding antenna configuration results obtained with the proposed framework for MIMO systems with Class B PAs are shown in Figures 4.7 and 4.8 for different channel correlation at the transmitter. As in the Class A PA case, the results are plotted within their corresponding achievable rate region. The proposed framework and algorithms achieve significant power reduction. Different from what observed for the Class A PA case, the power savings for a channel with higher correlation could be

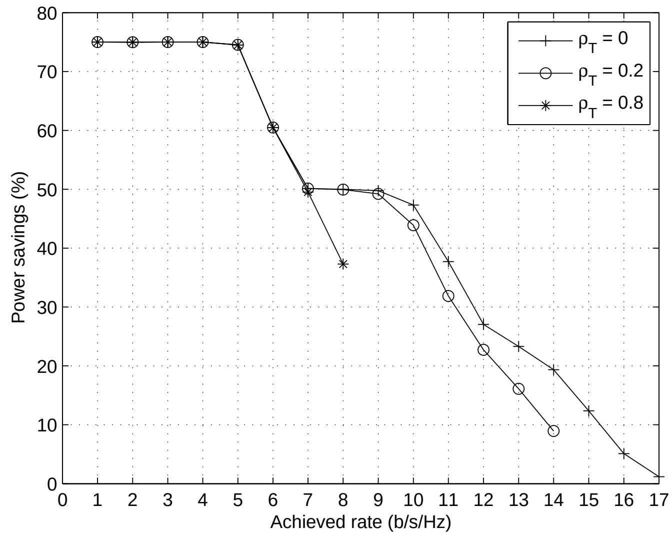

Figure 4.5: Power savings for MIMO systems with Class A PAs with channel correlation at the transmitter for CSIR-CSIT case.

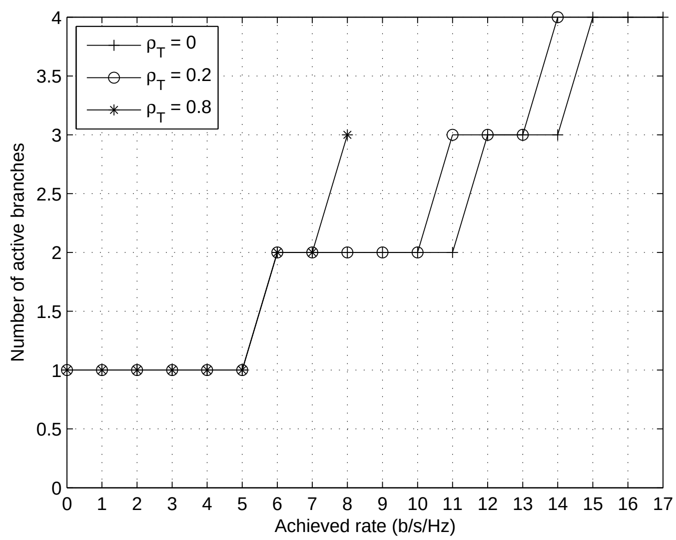

Figure 4.6: Antenna configuration for MIMO systems with Class A PAs with channel correlation at the transmitter for CSIR-CSIT case.

higher than that with lower correlation in some target rate region. This is largely due to the nonlinear relationship between the radiated power and the power consumption for the Class B PA case instead of constant power consumption for the Class A PA case. This nonlinear relationship results in possible power savings with more antennas within certain target rate region. This observation is also confirmed by the simulation

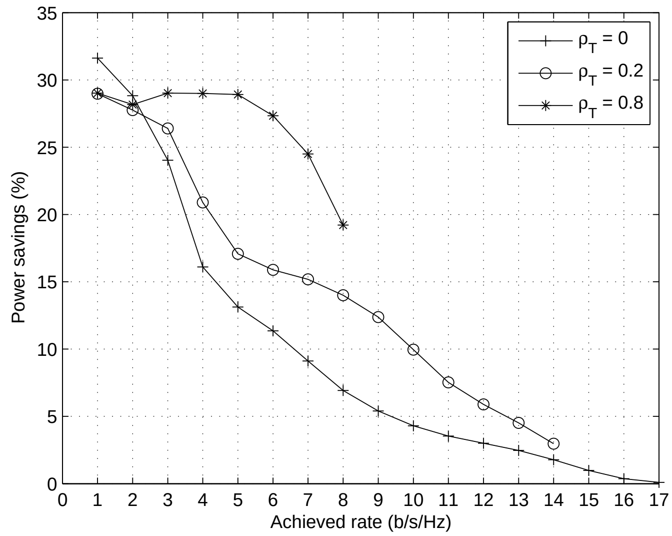

Figure 4.7: Power savings for MIMO systems with Class B PAs with channel correlation at transmitter for CSIR-CSIT case.

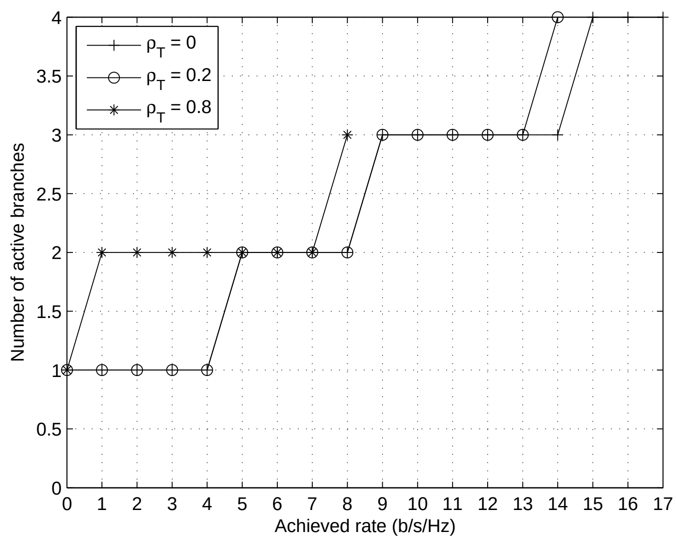

Figure 4.8: Antenna configuration for MIMO systems with Class B PAs with channel correlation at transmitter for CSIR-CSIT case.

results in Figures 4.3 and 4.4 for uncorrelated MIMO systems with Class B PAs. Comparing the power savings and the number of active antennas results between the exhaustive search algorithm (quasi-optimal) and branch minimization algorithm (suboptimal 2) in Figures 4.3 and 4.4, it is clear that at some target rates, using more antennas can achieve further power savings.

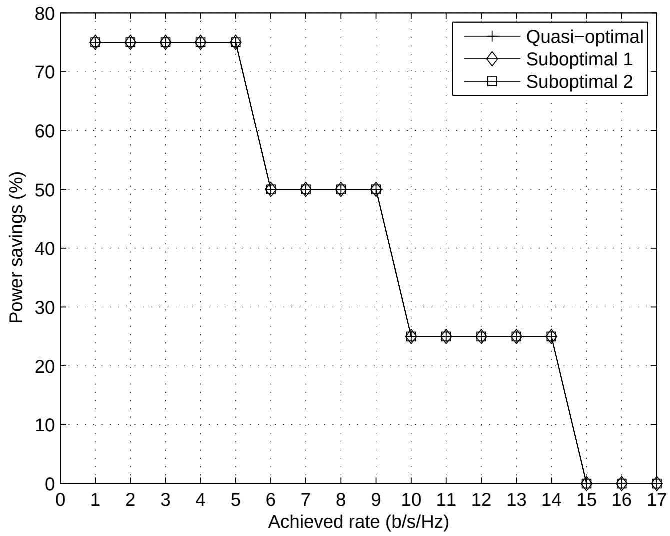

Figure 4.9: Power savings for MIMO systems with Class A PAs for CSIR-CDIT case.

Based on the results observed for the Class A PA case and the Class B PA case, it is necessary for the framework to dynamically decide the antenna configuration for different channel conditions and PA characteristics.

4.6.2 CSIR-CDIT Case

The proposed framework is evaluated for the CSIR-CDIT case with uncorrelated Rayleigh fading. The comparison baseline is a conventional MIMO system with the equal power allocation algorithm which is optimal for the CSIR-CDIT case in terms of radiated power minimization [79, 80]. MIMO systems with Class A PAs and Class B PAs are investigated.

The power savings and the corresponding antenna configuration results of the proposed framework for MIMO systems with Class A PAs are shown in Figures 4.9 and 4.10. Up to 75% of power savings can be achieved depending on the rate requirements. Similar trend in power savings and the number of active antennas is observed in the CSIR-CDIT case as in the CSIR-CSIT case with Class A PAs.

As in the CSIR-CSIT case, the exhaustive search, the branch adaptation, and the branch minimization algorithms achieve identical results in power savings and the number of active antennas for the CSIR-CDIT case as shown in Figures 4.9 and 4.10. The results suggest the branch minimization algorithm provides the best tradeoff between power savings and algorithm complexity for the Class A PA case.

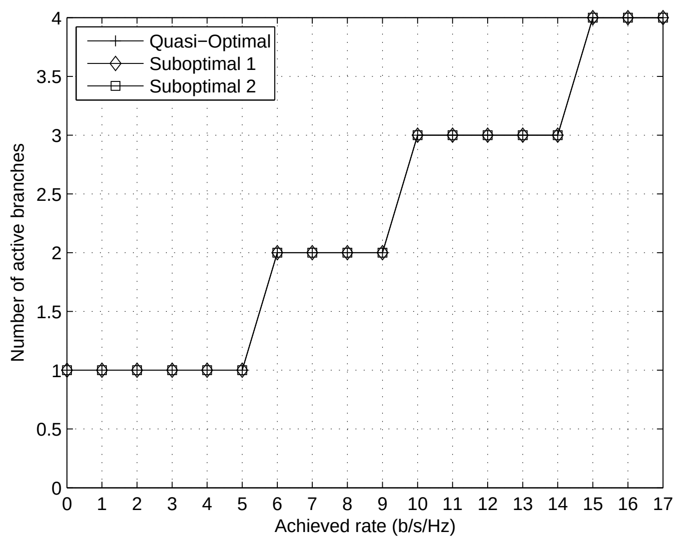

Figure 4.10: Antenna configuration for MIMO systems with Class A PAs for CSIR-CDIT case.

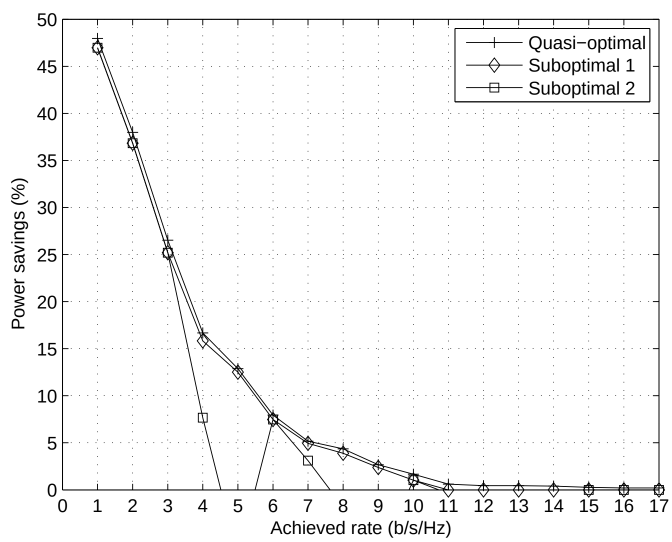

Figure 4.11: Power savings for MIMO systems with Class B PAs for CSIR-CDIT case.

The power savings and the corresponding antenna configuration results of the proposed framework for MIMO systems with Class B PAs are shown in Figures 4.11 and 4.12. Up to 50% of power savings can be achieved depending on the rate requirements. Similar trend in power savings and the number of active antennas is observed for the CSIR-CDIT case as for the CSIR-CSIT case with Class B PAs under uncorrelated fading. The power savings for the CSIR-CDIT case is higher than that for the CSIR-CSIT case at low rate region.

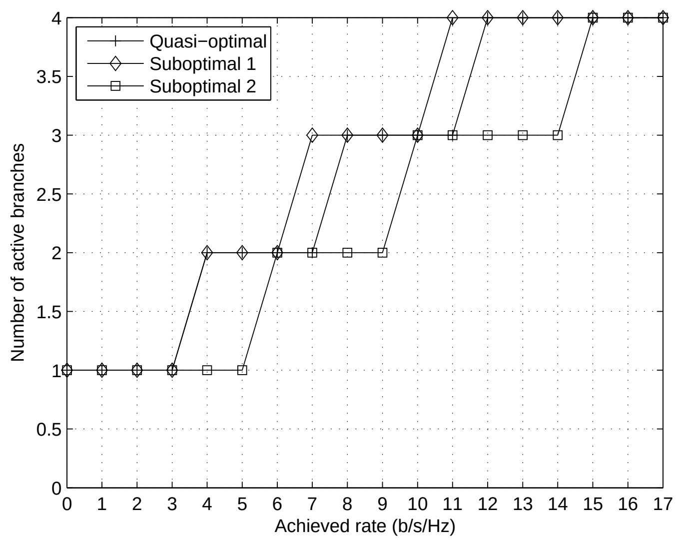

Figure 4.12: Antenna configuration for MIMO systems with Class B PAs for CSIR-CDIT case.

As the CSIR-CSIT case has more accurate information on the channel, the transmitter (both conventional and cognitive) can be more power efficient than that for the CSIR-CDIT case. Hence, the room for further power savings is less for the CSIR-CSIT case.

As the CSIR-CSIT case, the exhaustive search and the branch adaptation algorithms achieve almost identical results in power savings and the number of active antennas as shown in Figures 4.11 and 4.12 for the CSIR-CDIT case. The performance gap between the exhaustive search and the branch minimization algorithms is larger than that for the Class A PA case for the same reason as in the CSIR-CSIT case. The results suggest the branch adaptation algorithm provides the best tradeoff between power savings and algorithm complexity for the Class B PA case.

4.7 Conclusion

This paper discusses the theoretical framework of system power consumption optimization for MIMO systems using cognitive radio. The system power consumption minimization problem with a rate constraint for MIMO systems is formulated and numerical algorithms to the constrained minimization problem are developed. The power savings achieved by the proposed framework and power allocation algorithms in comparison with the conventional power allocation scheme is evaluated by simulation for 4×4 MIMO systems with Class A PAs and with Class B PAs, respectively. The impact of availability of channel state information at the transmitter and the receiver and the correlation between the transmit antennas are investigated. The simulation results show that significant system power savings (e.g., up to 75% for MIMO systems with Class A PAs and 30% for MIMO systems with Class B PAs when the channel state information is known at both the transmitter and the receiver, and 75% for MIMO systems with Class A PAs and 50% for MIMO systems with Class B PAs when the channel state information is known at the receiver and the channel distribution information is known at the transmitter) can be achieved using the proposed framework. In addition, the power savings for the Class A PA case and for the Class B case set the boundaries for the popular Class AB operation. This framework can be applied to other MIMO structures other than the 4×4 MIMO example, which has been proposed in recent wireless standards. For example, 4×4 MIMO is optional for LTE and IEEE 802.11n and mandatory for LTE-A. For other MIMO structures, such as, 2×4 or 4×2, the maximum number of parallel virtual channels and the maximum achievable rate may be different under the same channel characteristics assumption. However, the trend on average power savings is similar.

The framework and results obtained in this paper can also be used to improve power efficiency of existing MIMO systems. The knowledge of PA efficiency characteristics and its impact on system design are essential for the power savings. For example, a simple yet effective modification of a conventional system, which is allowed to activate/deactivate branches, is to minimize the number of active antennas. This is the suboptimal 2 algorithm. It achieves good performance for systems with Class A PAs and some power savings for systems with Class B PAs. Note that additional devices, added complexity, and associated power consumption are critical aspects to be considered when modifying the existing systems. Based on our experience on implementing a cognitive engine for wireless regional area networks (WRAN) devices [85], we think the balance between power savings and additional cost can be found. This balance depends much on specific system requirement and requires extensive studies. A detailed cost study would have to be done carefully and elaborately. In additional to cost, another important decision on how fast the antenna configuration is adapted has to be made in modifying existing systems or designing new systems. To maximize power savings, we need to shut down the unused branches (at least PAs). Due to the inertia of analog circuits (take some time to power up and power down), this cannot be done instantaneously. It, e.g., requires a few micro-seconds (depends heavily on the design). There is a balance between how fast we can adapt and how much power savings we can achieve. In this paper, we assume the adaptation of antenna configuration can keep up with channel changes. However, in real systems, it may not be possible to adapt antenna configuration (shutting down / powering up branches, etc) as fast as channel changes. It may not be desirable to change antenna configuration during transmission of a data packet, either. On the order of packet rate, it is possible to adapt antenna configuration. Therefore, the proposed approach needs to be carefully revised for real system implementation.

In addition to power consumption optimization detailed in this paper, the cognitive radio based approach may be useful to optimize radio operation for other goals. Unique to a CR system are its awareness, learning, and decision making capabilities, which enable some popular applications, such as dynamic spectrum access (DSA) [25] to improve radio spectrum efficiency. In this paper, we focus on the benefit CR can bring in terms of power savings, which is one aspect of many other optimizations CR is expected to handle. Furthermore, the new universal mobile telecommunications system (UMTS) proposals advocate self-configuring and self-organizing wireless networks [86] which is able to automatically manage wireless networks during operation. A CR based approach may be beneficial to this new automatic management task through online learning and monitoring of network operation, integrating learned knowledge about network operation in network optimization, and dynamically reconfiguring network to improve network efficiency. In summary, a cognitive radio based radio resource management framework may be applied to many applications and provide unique benefits beyond a conventional radio resource management framework.

Chapter 5

Review on Artificial Intelligence Techniques for Cognitive Engine Design

After the investigation on power savings for SISO and MIMO systems, this chapter[1] deals with the enabling function in this work, the cognitive engine. It reviews several widely used artificial intelligence (AI) techniques: artificial neural networks, meta-heuristic algorithms, hidden Markov models, rule-based systems, ontology-based systems, and case-based systems. Factors that influence the choice of AI techniques, such as responsiveness, complexity, security, robustness, and stability are discussed. These factors are illustrated in the discussion of two cognitive engine design cases to provide readers with a more concrete understanding.

(The following is the published journal paper on this subject.)

A Survey of Artificial Intelligence for Cognitive Radios

5.1 Abstract

Cognitive radio (CR) is an enabling technology for numerous new capabilities such as dynamic spectrum access, spectrum markets, and self-organizing networks. To realize this diverse set of applications, CR researchers leverage a variety of Artificial Intelligence (AI) techniques. To help researchers better understand the practical implications of AI to their CR designs, this paper reviews several CR implementations that used the following AI techniques: artificial neural networks, meta-heuristic algorithms, hidden Markov models, rule-based systems, ontology-based systems, and case-based systems. Factors that influence the choice of AI techniques, such as responsiveness, complexity, security, robustness, and stability are discussed. To provide readers with a more concrete understanding, these factors are illustrated in an extended discussion of two CR designs.

5.2 Introduction

With roots arguably tracing back to early cellular systems and personal communications systems, cognitive radio (CR) research has greatly expanded since the concept was first formalized in late 1990's [23]. The assumption of CR or CR-like capabilities formed the basis for the FCC's tentative approval to unlicensed devices to operate as secondary users in the TV spectrum (the so-called "white space") [87], the OFCOM's plans for the digital dividend [88], and the plans for future US military networks [89]. CR has been proposed for a wide range of applications such as automated interoperability for public safety systems [59], cognitive networking [60], spectrum markets [61], femtocells [62], self-organizing networks, cooperative relaying and networks [63], smart grid communications [64] and vehicular networks [65]. CR promises to dramatically improve spectrum access [54], capacity [66], and link performance [67] while also more closely tying the behavior of the network to the needs of the user [23].

While all of these applications are recognized as in the domain of "cognitive radio," there is much disagreement on exactly what is, and is not, a CR with seemingly everyone having their own definition (e.g., see [23, 24, 53, 57, 90, 91]). Nonetheless, the following attributes are commonly expected from a CR and form a baseline set of assumptions for the remainder of this paper.

1. Observation - collect information about the operating environment, capability, and characteristics of the radio;

2. Reconfiguration - change the operation parameters of the radio;

3. Cognition - understand the environment and capability of the radio (awareness), make informed decisions on actions (reasoning), and learn the impact of these actions on the performance of the radio as well as the performance of the network in which the radio is embedded (learning).

These three attributes can be embodied in a common element known as a cognitive engine (CE). This paper defines a CE as an "intelligent" agent that manages the cognition tasks in a CR, where intelligence denotes behavior that is consistent with a specified goal [53]. The CE can be implemented as an independent entity interacting with the radio transceiver (e.g., reconfigurable radio transceiver implemented with software defined radio (SDR)), or as a collection of interacting entities with each entity fulfilling a specific role. Given the input from its environment or user (observation), the CE analyzes and classifies the situation and determines the appropriate response to the stimulus (cognition), and carries out the decision (reconfiguration). As an example, this response can be adapting radio parameters, such as, channel coding scheme, modulation scheme and operation frequency, given user requirements, current environment conditions, and previous experiences at the CE.

How to implement the various aspects of a CE is an active area of research in which considerable attention is given to each of these three attributes. This paper focuses on the practical issues associated with realizing cognition in a radio, which we view as an application of artificial intelligence (AI) to the radio domain, while only considering the aspects of observation and reconfiguration as needed for clarity. Specifically, this paper surveys the state-of-the-art in the use of AI in CR to ascertain available choices for implementing a practical CR and the relative merits of various proposed techniques in differing applications. The paper takes the following structure: Section 5.3 reviews AI techniques proposed for use in a CE and presents examples of their applications. Section 5.4 discusses practical issues with CR such as convergence time, implementation complexity, stability and how to handle multiple coexisting AIs, perhaps within a single CE. Section 5.5 reviews two CEs developed using different AI techniques. Finally, Section 5.6 concludes with some observations on the different AI techniques.

5.3 Artificial Intelligence Techniques for Cognitive Radio

Awareness, reasoning, and learning are the basic components of a CR discussed in this paper. Despite various definitions of these terms, in this paper, awareness refers to the process of extracting the information regarding environment and radio itself for a specific purpose. Reasoning is defined as the process of finding an appropriate action in response to a particular situation with a system target (e.g., maximum operating lifetime, maximum robustness, lowest-cost communications) based on user application quality of service (QoS) requirement (e.g., latency, bit error rate (BER)) and willingness to share resources and collaborate with other devices in the network. Learning is defined as the process of accumulating knowledge based on the observed impact upon applying the action. Typically, these processes complement each other to improve the operation of the CR process as a whole. In other words, awareness, reasoning, and learning in a CR interact and influence each other. Among these three, awareness is the starting point of a CR process and is the foundation of learning and reasoning. Learning can improve reasoning by enriching the knowledge or experience used in reasoning process. Powerful reasoning can improve the efficiency of learning by providing good examples for learning in return. This section presents some AI techniques that have been proposed throughout literature as possible candidates for CR. They are presented in the order of historical development. The discussion on general AI and its applications is out of the scope of this paper, on which abundant information may be found at [92], the website provided by the Association for the Advancement of Artificial Intelligence.

5.3.1 ANN

The first artificial neural was presented by the neurophysiologist Warren McCulloch and the logician Walter Pits in 1943 for the study of human brain. The idea of artificial neural network (ANN) then be applied to computational models. Modeled on a nerve plexus, an ANN is nothing more than a set of nonlinear functions with adjustable parameters to give a desired output [93]. Different types of ANNs are separated by their network configurations and training methods, allowing for a multitude of applications. However, they all are comprised of neurons interconnected to form a network. Each artificial neuron usually produces a single output value by accumulating inputs from other neurons. While there are many types of ANNs available in the literature, only those most common and applicable to CR are presented here.

MLPN

Multi-layer linear perceptron networks (MLPNs) are comprised of layers of neurons, each a linear combination of the previous layer's outputs. Generally, the weights for the linear combination are chosen randomly before training and can be updated using several methods, such as back propagation (BP) [93], genetic algorithm, or combinations of methods. The performance of such training algorithms is dependent upon the size of the network and its application. Hybrid training methods can be used to extract the best features from each, such as pre-training a network with a genetic algorithm, and then refining the output using BP.

NPN

By introducing nonlinearity into the network through perhaps squaring the inputs, or cross-multiplying two inputs, the network can be customized to fit the sample set. Although multilayer nonlinear perceptron networks (NPNs) can provide highly flexible and dynamic results, their network configuration must often reflect the data which they represent. Furthermore, BP for training the weights in the neurons may be slow to converge, requiring significant processing time to achieve precise results [93].

RBFN

Similar to NPN, radial basis function networks (RBFNs) have a built-in distance criterion with respect to a center (a radial nonlinear function) in its hidden layer. This transformation has the advantage of preventing the network from settling into local minimals, a common problem with perceptron networks. The function itself is usually Gaussian, but Euclidean distance and others have also been used [93]. Training is often implemented with the gradient descent method.

Application of ANN to CR

Because of their ability to dynamically adapt and be trained at any time, ANNs are able to “learn” patterns, features, and attributes of the system they describe. The term “learn” refers to the fact that the neurons are stored in computer memory, the outputs of which can be adjusted systematically to yield a new result for a new situation, and remember the results. The attributes can be highly nonlinear, complex, and numerous, yet ANNs can be constructed by only a few examples, thus reducing the complexity of the solution. For this reason they have long been used to describe functions, processes, or classes that are otherwise difficult to formulate analytically. Therefore, ANN not only can be used to classify or recognize received stimuli, but also to assist in solution adaptation process.

The ANN has been adopted in spectrum sensing for CR [94, 95, 96]. In [94], Fehske et al. develop an ANN-based signal classifier utilizing the extracted cyclostationary signal features. The combination of cyclostationary analysis and ANN provides efficient and reliable signal classification and reduces the online processing time by performing significant amount of computation offline. In [95], Cattoni et al. use the ANN to classify different IEEE 802.11 signals (the complementary code keying (CCK) signal and the orthogonal frequency division multiplexing (OFDM) signal) based on the frequency features. In [96], Zhu et al. evaluate an ANN-based spectrum sensing algorithm for wireless mesh networks. The simulation results show that the ANN-based algorithm achieves better performance in accuracy and speed than the Bayesian based algorithm.

The ANN has also been used for radio parameter adaptation in CR [58, 97, 98]. In [58], Reed et al. develop a CR testbed using Tektronix test equipment as RF hardware and a PC running Open Source SCA Implementation (OSSIE [99, 100]) for different waveforms. The ANN determines radio parameters for given channel states with three optimization goals, including meeting BER, maximizing throughput, and minimizing transmit power. In [97], Hasegawa et al. propose an ANN-based distributed optimization algorithm for large scale cognitive wireless clouds which consists of many heterogeneous terminals and networks.

Baldo et al. propose to use the ANN to characterize the realtime achievable communication performance in CR [101]. Since the characterization is based on runtime measurement, it provides certain learning capability that can be exploited by CE. The simulation results demonstrate good modeling accuracy and flexibility in various applications and scenarios.

In addition, the ANN has been used for pattern classification in a pattern based transmission for CR [102, 103] where the transmission bit string is mapped to a signal pattern at the transmitter and the received pattern is classified and mapped back to a bit string at the receiver using the ANN.

5.3.2 MA

Explicit relations between parameters of a CR and desired performance metrics are usually not available. Therefore, search algorithms based on mathematical relations cannot be applied to find optimal parameters with respect to performance metrics. Instead, metaheuristic algorithms (MA) [104] can be applied to computationally hard problems to search through the solution space while learning and establishing the requisite relationships. Although the term “metaheuristic” was probably first mentioned in 1986 [105], it can be traced back to earlier work on stochastic optimization methods in 1950’s [106]. Several selected metaheuristic algorithms are presented here and their relative merits are summarized in Table 5.1.

Table 5.1: Characteristics of metaheuristic techniques

Decision Process	**Key Benefits**	**Drawbacks**
Classical Techniques	Provides globally optimal solutions for class of convex optimization problems; Convergence properties are well-analyzed.	Could yield sub-optimal (non-desirable) solutions for ill-behaved functions; Branch-and-bound, clustering and multi-start techniques that enhance performance need access to global information in addition to being computationally intensive.
Genetic Algorithms	Well-investigated for wireless applications.	Convergence has not been fully investigated; Efficiency depends on proper parameter selection.
Simulated Annealing	Asymptotically converges to globally optimal solution with probability 1; Easy to implement.	Convergence rate can be slow; Only converges to a global optimal as time goes to infinity for a finite search space.
Tabu Search	Easy to implement.	Efficiency depends on proper parameter selection.
Ant Colony Optimization	Can easily adapt to changes in real-time.	Inferior to simulated annealing for local searches.

Evolutionary/Genetic Algorithms/GAs

GAs [107], which are a particular class of evolutionary algorithms, draw their inspiration from genetic evolution and natural selection of species in nature.

The definitions of chromosomes and fitness functions are fundamental to the description of a GA. Chromosomes are abstract representations of candidate solutions. A fitness function, closely correlated with the objective of the algorithm or optimization process, quantifies the desirability of the solution. Candidate solutions are evaluated on the basis of the values they generate for the fitness function (called fitness levels), which characterize the performance of candidate solutions. An ideal fitness function should lend itself to fast computation, since it takes several evaluations to produce a single generation, and several generations to produce a useful result. A GA maintains a population of candidate solutions for a given problem. The fitness of the population is evaluated and multiple individuals (based on fitness levels) are selected to form a new population by "reproduction" (combination of candidate solutions) and "mutation" (incorporation of some new solution trait to the current solution). This new population then becomes the current population for the next iteration (or generation). In this process, "unfit" elements eventually die out and are replaced by solution offspring with increased fitness levels.

SA

Simulated annealing (SA) [108] is a simple approach for global optimization in a large search space. SA is motivated by the annealing process in metallurgy, a process involving controlled slow cooling of a heated melt to reduce or remove defects and achieve perfect crystallization of the material.

At each step, the SA algorithm considers some neighbors of the current state s, and probabilistically decides either to move the system to state s' or stay in state s. The probabilities are chosen so that the system ultimately tends to move to states of lower energy. Typically this step is repeated until the system reaches a state which is good enough for the application, or until a given computation budget has been exhausted. The local search space size is usually a function of the current energy level or sometimes time from start. This way, the algorithm initially wanders in a broad area of the search space containing good solutions, ignoring small features of the energy function and as it moves toward lower energy regions, the search space becomes narrower and narrower.

TS

Tabu search (TS) [109] enhances the performance of a search method by using a memory structure.

The basic elements of TS are memory structures called Tabu list. The list ensures that a recent move is not repeated or reversed. TS uses memory in different ways to guide the search procedure and the role of memory can change as the search proceeds. When the search procedure is in some region with more acceptable solutions, it might be more advantageous to intensify or focus the search. Such intensification can be carried out by prioritizing solutions which have common features with the current solution. This can be done with the introduction of an additional term in the objective function which penalizes solutions far from the present one. At other times, it might be useful to spread the exploration space of the search algorithm and this can be done by diversification. As before, diversification can be achieved by introducing an additional term in the objective function which penalizes solutions that are close to the present solution. Dynamic weights are attached to the intensification and diversification terms such that intensification and diversification phases alternate during the search.

ACO

Ant colony optimization (ACO) [110] is inspired by the behavior of ants in finding shortest paths from their colonies to food sources.

Ants initially wander randomly and upon finding food return to their colony while laying down pheromone trails. Upon finding a pheromone trail, other ants follow this trail rather than randomly wandering. Thus a successful pheromone trail (a trail that leads to food) is continually reinforced. Also, the pheromone trail starts to evaporate over time. Hence the pheromone trail is more attractive if the time taken to travel the path is shorter as this gives the pheromone lesser time to evaporate. The ants are thus successful in finding the shortest path to a food source. ACO mimics this ant behavior with "simulated ants" walking around a graph representing the problem to solve and finding locally productive areas. ACO algorithms search in parallel over several constructive computational threads based on local problem data and a dynamic memory structure containing information on the quality of previously obtained result. These algorithms thus combine apriori information about the structure of a promising solution with a posteriori information about the structure of previously obtained good solutions.

Application of MA to CR

The metaheuristic techniques presented here can not only be used for reasoning or finding optimal solution with objective/utility function, but also be used for learning with the aid of training examples when the relationship between parameters and a desired performance measure is not well understood. The objective of learning is to identify a hypothesis or a rule set from the search space that maximizes the fit of the training examples to the target concept or in other words, to identify a hypothesis set or a set of rules that is consistent with the training examples. Although the characteristics of each search algorithm are different as can be seen in Table 5.1, a common challenge in the application of metaheuristic techniques is the formulation of extensive examples for target scenarios.

Among the various metaheuristic algorithms, the GA has been widely adopted to solve multiobjective optimization problem and dynamically configure the CR in response to the changing wireless environment [67, 111, 112, 113, 114, 75, 115, 116]. In [67], Rondeau et al. apply GA to adapt the radio parameters of a SDR to the changing radio environment. In their implementation, the fitness function dynamically links and weights different objects according to the link conditions and user application requirements. In [111], Newman et al. design a GA-based CE to control the radio parameters for single and multicarrier transceivers. This paper derives a set of fitness functions to guide the search direction of GA and investigate the trade-off between the convergence time and the size of the GA search space. In [112], Hauris uses the GA to adapt the parameters of CRs on autonomous vehicles. These autonomous vehicles form a geographically varying dynamic wireless network for communication and information sharing among the vehicles and the base station. In [113], Park et al. validate the applicability of GA-based radio parameter adaptation for the cdma2000 forward link in a realistic scenario with Rician fading. In [114], Thilakawardana et al. investigate a GA-based cell-by-cell dynamic spectrum allocation scheme achieving better spectral efficiency than the fixed spectrum allocation scheme. A new solution encoding technique is proposed to reduce the GA convergence time. In [75], Kim et al. implement a software testbed for CR with the spectrum sensing capability and a GA-based CE to optimize radio parameters for dynamic spectrum access (DSA).

5.3.3 HMM

Hidden Markov Model (HMM) was first introduced in late 1960's. It is a convenient and mathematically tractable statistical model to describe and analyze the dynamic behavior of a complex random phenomenon [117] which can be modeled as a Markov process with observable and unobservable states. The HMM generates sequences of observation symbols by making transitions from state to state, one symbol per transition. However, the states are hidden and only the output is observable. In general, a real world process can be expressed as a random process producing a sequence of observation symbols or patterns with hidden parameters generating the observables. The symbols or patterns may be discrete or continuous depending on the specific processes.

An HHM can be completely specified in a compact form, $\lambda = (\mathbf{A}, \mathbf{B}, \pi(1))$, where $\mathbf{A}$ is state transition probability matrix of dimension $N \times N$, $\mathbf{B}$ the observation symbol probability matrix of dimension $K \times N$, $\pi(1)$ the initial state probability vector of dimension $N \times 1$, N the number of states, and K the number of distinct observation symbols per state.

Three basic problems for an HMM

There are three key problems associated with HMM in real world applications [117]:

1. Evaluation or recognition problem: Given the parameters of the model, λ, compute the probability of a particular observation sequence. The forward-backward algorithm solves this problem.

2. Decoding problem: Given the parameters of the model, λ, and the observation sequence, find the sequence of hidden states that best explains the observation sequence. The Viterbi algorithm solves this problem.

3. Training or learning problem: Given an observation sequence, find the most likely set of state transition and observation symbol probabilities. In other words, this estimates an HMM, λ, using observation sequence. This problem is the subset of expectation maximization (EM). The Baum-Welch Algorithm (BWA) solves this problem.

Application of HMM to CR

An HMM can be built for a specific system to explain and characterize the occurrence of the observed symbols or patterns. This model can then be used to identify the sequences of observations with the same pattern by choosing the model that would most likely produce the observed sequences. Therefore, HMM can be used as an observation process of CE to recognize or classify received stimuli and can achieve awareness. In addition, since it can reproduce the training sequences, it can be used for prediction. Furthermore, learning can be accomplished by creating new models.

HMMs have been applied to CR research. Rondeau et al. propose to model the wireless channel online using HMM for CR [118]. The HMM is trained using the GA with data from a broadband channel sounder in a line of sight additive white Gaussian noise channel.

HMMs have also been used for spectrum sensing in CR [119, 120]. In [119], Kim at al. propose to use the HMM to process signal cyclostationary features for primary signal detection in CR. The HMM-based spectrum sensing approach can detect and classify signals at low SNRs with only limited information on signal bandwidth. In [120], Ghosh et al. validate the existence of a Markov chain model for wireless channel utilization with real-time measured data in the paging band and formulate the spectrum sensing problem using an HMM.

In addition, HMMs have been used for spectrum occupancy prediction [121, 122]. In [121], Akbar et al. develop an HMM-based DSA algorithm where the HMM models and predicts the spectrum occupancy of the licensed radio bands for CR networks. The paper shows that the HMM-based DSA algorithm can achieve significant signal to interference ratio improvement compared to the traditional carrier sense multiple access (CSMA) based approach.

5.3.4 RBS

In a rule-based system (RBS), rules are extracted from a specific application area (automatically or manually) and used in decision making for that domain. It is a natural way of encoding a human expert's knowledge in a narrow area into an automated system. The idea of RBS was used in the development of DENDRAL, one of the oldest expert system developed in 1964. A typical RBS consists of the following fundamental elements [123]:

- Rule base: Contains a list of permanent rules.
- Inference engine (IE): Infers information or takes action based upon the input and the rule base.

Rules are usually expressed in the following form [123]:

IF *conditions* THEN *actions*.

The input is tested against the conditions, and the actions are taken if the conditions are satisfied. In general, the order in which the rules are executed is not critical to the final result as long as all rules are executed [123]. This is different from most procedural programs where the order is important.

Operation of Inference Engine

Generally speaking, there are two broad sets of inference engines: forward chaining and backward chaining [123]. In a forward-chaining IE, new data trigger rules whose conditions are met and conclusions are drawn based on these rules. Conclusions can further trigger other rules to generate new conclusions. This process continues until no more rules can be triggered. In other words, a forward-chaining IE is data-driven. On the contrary, a backward-chaining IE is goal-driven. In a backward-chaining IE, a particular goal is set to be proved and rules whose conclusions match the goal are identified. The conditions of those rules are then set as goals for new rules. This process continues until no new conditions are identified.

Application of RBS to CR

The most significant advantage of a RBS is its simplicity. A radio can deduce actions for an input quickly using a RBS. Given the proper domain knowledge, the rule base can be built automatically or manually.

During the rule derivation process, relevant and irrelevant features can be handled differently for a rule. However, this process is usually tedious for a complex domain. The accuracy of the RBS system depends on the completeness and accuracy of the underlying rule base. If the domain is not perfectly understood, the RBS system might return inappropriate responses. This is often the case for some CR applications, such as, dynamic spectrum access (DSA). One possible solution to this issue is to assign certainty values to rules [123] and make use of statistical tools, such as Bayesian analysis. Although this approach leads to a less rigorous conclusion, it could provide a good guess for situations that are not perfectly understood. Another choice can be to combine the RBS with the case-based system and make use of experience to compensate for the inaccuracy in the RBS.

The rule-based reasoning cognitive engine (RBR-CE) has been designed for CR [124, 125, 126]. In [124, 125], Reed et al. design and evaluate a RBR-CE for IEEE 802.22 WRAN applications. The developed RBR-CE can achieve similar performance to the CE based on the genetic algorithm with a lower computational complexity. In [126], Clancy et al. develop a RBR-CE using predicate calculus, focusing on a generic CR architecture enabling learning in the reasoning process. The learning results are expressed in predicate calculus and therefore, can be used in the CE. The operation of the CE is demonstrated for capacity maximization and DSA problems.

Fundamental to any rule-based reasoning system is the generation of the rule database. In [127], Weingart et al. develop a systematic method to derive the rule database through automatic experiments over a vast parameter space using statistical analysis of variance and design of experiments. Important parameters across the protocol layers and relationships between parameters and performance metrics are identified and modeled through the experiments. Based on the extracted rules, optimal configuration can be drawn for specific channel condition and application requirement.

5.3.5 OBS

Although “ontology” was not deliberately defined in computer science until early 1990s by Tomas Gruber [128], it has been used in AI community since 1980s. An ontology is an “explicit”, “shared”, “formal” representation of a set of concepts within a domain and relationships among these concepts [129]. As a formal representation, the ontology becomes machine-understandable. In addition, the ontology needs to be accepted or shared in a community to make it useful. In ontology-based systems (OBSs), the ontology is used to reason about the attributes of the domain of interest.

An ontology usually includes the following basic components [129]:

- Classes: a set of objects in the modeled domain.
- Instances: individuals belonging to classes in the modeled domain.
- Attributes: properties of objects.
- Relations: links between various entities.

Ontology Languages

An ontology is expressed using an ontology language to facilitate machine processing. A number of ontology languages have been developed for various ontologies. Here, we focus on web-based ontology languages due to the immense impact of World Wide Web. There are three major web-based ontology languages:

- XML Topic Maps (XTM) [130]: This is an ISO standard that allows relations among any number of entities.
- Resource Description Framework (RDF) [131]: This is a World Wide Web Consortium standard that only allows relations between two entities.
- Web Ontology Language (OWL) [132]: THis is a World Wide Web Consortium standard that only allows relations between two entities. It is an extension to RDF.

Note that RDF and OWL are fundamental technologies for semantic web, a web of data (computer manipulatable) instead of a web of documents (human readable) [133].

Application of OBS to CR

Using ontology languages, an OBS is able to deduce facts logically. A radio equipped with OBS can understand the capability and characteristics of itself and other radios using logic deduction. This understanding as well as understanding of the environment helps the radio to deduce optimal operating parameters. Note that the logic derivation in the inference engine might be impractically long [132]. For example, the time requirement of OWL-Full, the most sophisticated level of OWL, is undecidable, which means some derivation cannot be completed within a finite amount of time.

Ontology language has been investigated for CR development [134, 135]. An important example is DARPA's neXt Generation (XG) policy language framework [134]. This framework is based on an extensible ontological framework to facilitate future extension of spectrum rules and related concepts. It allows both regulatory and system policies to be expressed and enforced. A good tutorial on the properties of a formal language for CR is provided in [135]. Kokar et. al. start the investigation with the OWL and extend the capability of the OWL as the areas that the OWL cannot support are identified. This paper concludes that a good language for CR should be able to express rules, introduce new functions, and specify radio behaviors.

In [136], ontology-based reasoning (OBR) is used to achieve self-awareness and interoperability among SDR nodes. With an example on radios negotiating equalizer training sequence structure, the authors show self-awareness can improve the interoperability of SDRs. OWL is used in the OBR.

In addition to self-awareness and interoperability, OBR is also used to apply spectrum policy to DSA [137, 138]. A prototype framework is built for CR using OBR [137, 138]. Ontologies and rules are combined to achieve a knowledge-driven differential-response capability, which, as defined by the authors, is the capability of reasoning about a failure in an attempt and identifying alternative actions to satisfy the goal using knowledge of radio technology, policy, goals, and other contextual information. In other words, the CE learns from its failures and avoids the same failures in the future.

5.3.6 CBS

The case-based system (CBS) can trace its root to the work of Roger Schank on the model's of dynamic memory in early 1980s. It is an AI area that focuses on using previous similar experiences, or cases, to guide the problem solving process and to obtain a solution [139]. In a CBS, a solution to the new problem is created by selecting the cases that are most relevant to the problem, narrowing down the selected cases to a single case, and adapting this case to fit the current scenario. Adapting the parameters of the case can be viewed as an optimization problem. The purpose of the initial similar case retrieval is to allow the optimization process to begin at a point closer to the goal in a search space by finding similar cases. This reduces the time and processing needed to optimize the parameters the system is looking for. Upon new solutions obtained from case adaptation, the case database is updated with the new cases.

The characteristics of CBS [139] include the capability to solve problems within partially understood domains, the capability to provide unique explanation, and the close resemblance to actual human reasoning process. One of the key issues of CBS is that the performance relies on previous cases. If previous cases have been solved incorrectly, it is possible for mistakes to propagate onto new cases. In addition, for a complex domain where the system requires a large case database to represent its characteristics, populating and searching such a case database can be time consuming and sometimes difficult. In this case, integration with other techniques, such as with a rule-based system, may be necessary to improve the performance and reduce the time to build and search the case database.

Modules in a CBS

In general, a CBS may contain the following functional modules [139]:

- Case representation and indexing module: format the input information such that it can be understood by other modules of the system.
- Case selection and retrieval module: search the case database and obtain cases that satisfy the request under certain criteria.
- Case evaluation and adaptation module: evaluate the performance of the retrieved case for the new problem based on some criteria and modify the case if its performance is not satisfactory.
- Case database population and maintenance module: populate initial case database, insert new cases, update existing cases, and remove redundant cases in the case database as necessary.

Application of CBS to CR

Given the currently observed environment and radio objectives, a CR can use a CBS to determine an acceptable solution (action) for current environment based on the existing case in a case database. As a case database may not include all possible situations a CR may encounter in operation, a CR needs to learn new cases when it encounters new situations, generate new actions for the new situations, and update the case database with the new cases. These are the general tasks of the case-based system or case-based reasoning (CBR) in a CR.

CBR has been investigated for CE design recently [140, 85, 141]. In [140, 85], Reed et al. design a case-based reasoning cognitive engine (CBR-CE) to obtain radio parameters for IEEE 802.22 WRAN applications. The performance of the CBR-CE is evaluated under various radio scenarios and compared to several multi objective optimization only algorithms, including the hill climbing search and a genetic algorithm. The simulation results show that the CBR-CE can achieve comparable performance with less complexity after appropriate training/learning. The learning process of the CBR is also simulated and discussed. In [141], Khedr et al. design a CE using CBR and fuzzy logic to determine the channel type (flat fading vs. frequency selective fading, fast fading vs. slow fading) for WiMAX systems.

In [142], Le et al. propose a CE architecture and the use of CBR in a CE. The functionalities of the building blocks in the cognition cycle and the CE are discussed, including environmental awareness, case-based learning, multi-objective optimization, and hardware-portable interface. The implementation of the building blocks is also suggested.

5.4 Practical Issues and Interactions of Cognitive Engines

5.4.1 Implementation of a Single CE

Artificial neural networks are mathematical emulations of biological neural networks, used primarily for non-linear pattern matching and statistical modeling. They are able to describe complex relationships between multi-dimensional data sets, and are trainable online to contend with fluctuations in trends. ANNs are excellent candidates for classification. However there is little theory to link the particular application with the required network size, nor the specific type of network. Furthermore, over-training a network to recognize only the initial data set is a possible problem.

Metaheuristic or search algorithms are very efficient when the rules to be learned are in the form of searching for a set of parameters that optimize a given performance metric. Furthermore, the effectiveness of the algorithms can itself be improved by using the algorithms in conjunction with learning mechanisms such as prior knowledge-based learning (as is done in Soar [143] and Prodigy [144]) and inductive learning where new search-rules are formulated on the basis of training examples and/or patterns observed in previous iterations of the search. The biggest challenge for Metaheuristic search algorithms is the formulation of the hypothesis space. By definition, these techniques only try to find the best hypotheses from the search space and cannot create new hypotheses beyond the search space. Hence the formulation of a comprehensive search space, that includes all possible hypotheses or relationships between contributing factors, is extremely critical to the performance of these algorithms.

An HMM based approach can analytically model a complicated stochastic process using the observation sequence. Both classification and predication can be achieved using an HMM. However, the development of an HMM requires a good training sequence and the training process can be computationally complex. Other AI techniques, such as GA, to improve the model training efficiency [118]. In addition, RBS and CBS can help the HMM to determine the required observation duration.

The most significant advantage of a RBS is that it is simple to understand and that it can tackle unforeseen scenarios. RBS also has the ability to include only relevant features while formulating a rule. However, the rule derivation process could be tedious, and it requires perfect domain knowledge which may not always be available. In practice, RBS can be combined with CBS and OBS to better deal with unfamiliar domain. Without characterization and abstraction of radio scenarios, the number of rules in the knowledge base could be prohibitively large for implementation. Tradeoff has to be made between the levels of abstraction and the details of the solutions for adaptation [125].

An OBS provides a good approach to enable the radio to understand the characteristics, capabilities, and constraints of itself and others in the radio environment. In addition, this understanding can be further used in its logical deduction. One challenge for OBS lies in the development of ontology. This process usually requires perfect domain knowledge and extensive processing. In addition, the logic derivation process using the ontology can sometimes be time consuming [132]. In order to improve the efficiency and robustness of OBS, CBS and RBS can be incorporated to gather effective experience and to reduce the work load of the logic deduction.

Similar to human learning and reasoning process, a CBS can develop as it operates. This characteristic is especially appealing to CR design as a CR needs to be able to operate in an unfamiliar environment. On the other hand, appropriate training of the CBS before deployment can reduce the processing time and improve the robustness of the system [85]. In addition, prior knowledge about the expected deployment environment can help to reduce the size of the case database, to accelerate the case retrieval time, and more importantly, to reduce the impact of irrelevant patterns.

These preceding tradeoffs are summarized in Table 5.2.

Table 5.2: Comparison of different AI techniques

Alg.	**Strengths**	**Limitations**	**Options**
ANN	Ability to describe a multitude of functions Conceptually easy to scalable Excellent for classification Can identify new patterns	Training may be slow depending on network size Possible over training No theory to link application with required network	Can use other learning techniques in the training phase (i.e., GA) Can be combined with RBS
MA	Excellent for parameter optimization and learning involving relationship between parameter values Can use other learning techniques in the training phase (i.e., GA)	Formulation of rule space is difficult when learning or optimization is not restricted to parameter values	Can be used in conjunction with RBS Learning can also be used in the search process
HMM	Can model complicated statistical processes Good for classification Easily scalable Can predict based on experiences	Requires good training sequence Computationally complex	Based on previous knowledge, CBS and RBS can help HMM determine the observation duration for a specific application and overcome issues with new situations
RBS	Simple implementation Ability to tackle unforeseen situations Ability to include only relevant features while formulating a rule	Tedious rule derivation process Requires perfect domain knowledge which is not always available	Can be combined with CBS and OBS to better deal with unfamiliar domain

Continued on next page

Table5.2 – Continued from previous page

Alg.	Strengths	Limitations	Options
OBS	Ability to logically deduce Ability to understand the capabilities and characteristics of its own and others	Requires perfect domain knowledge to develop ontology Low efficiency for sophisticated ontology and ontology language	Can be combined with CBS and RBS to improve efficiency and robustness
CBS	Close to human reasoning Can work in a chaotic situation with lots of variables Allows fast acquisition of knowledge Allows learning in the absence of domain knowledge	Relies solely on previous case Requires large case memory Might include irrelevant patterns	Can be combined with RBS and OBS to yield a more robust problem solving system that does not rely solely on experience

5.4.2 Coexistence of Multiple CEs

As highlighted by the movement toward cognitive femtocells, centralized control of hundreds and thousands of clusters adapting in real-time quickly becomes infeasible. So we can expect that there will be numerous CEs each controlling a subset of all adaptive radios. Further simplifications could be made if instead of implementing a single CE process that manages all aspects of a radio's (or subnet's) behavior, several quasi-independent processes could be used to manage different aspects of the system's behavior, e.g., different processes for spectrum utilization, transmit power, group membership, and routing. Likewise independent processes could be used to make different observations (e.g., location, presence of a primary user, link states elsewhere in the network) and to guide radio behavior (e.g., ascertaining policy and user objectives) and to differentiate radios in the network. Thus rather than realizing a CE as a singular AI, computational considerations dictate the deployment of multiple AIs each controlling their own domain, but influenced by the choices of the other AIs in the radio or network.

So, whether we consider each device in a wireless network to be controlled by a singular autonomous CE, or we consider the adaptations of each device to be jointly determined by several CEs distributed across the protocol stack, or we consider multiple devices controlled by a single CE in a cognitive network [60], it is clear that the deployment of CR will introduce countless situations where the performance and thus the choices of each CE is influenced by the choices of the multitude of other CEs in the environment. These interactions can easily spawn an infinite sequence of adaptations that never converge, can yield an unstable network whose behavior radically changes with small changes in the environment, or can produce a network with decidedly suboptimal performance (e.g., a tragedy of the commons [145]). Examples of this unstable behavior arising from even simple cognitive processes are illustrated in [146, 66]. The analysis and design of networks with numerous interactive processes is further complicated by market realities where platform and process implementations will vary by vendor. Thus it is insufficient to designs AI process for a CE that only considers the behavior of the AI in isolation.

While these interactions makes CR networks much more complex to analyze and design, the interaction of multiple networked AIs has a natural analog - the interaction of multiple (somewhat) networked natural intelligences we call human society - that can be drawn upon to guide the practical design of coexisting AIs. Extending informally back to the dawn of human history and more formally since the 18th (Waldgreave), 19th (Cournot), 20th (Zermelo) centuries depending on the desired level of formalism, game theory provides a collection of models and analytical techniques that permit the prediction of likely outcomes of the interactions of intelligent decision makers. Generally, the basic model of these interactions - a game - consists of at least a set of players, which correspond to the intelligent decision makers, sets of available adaptations for each player, and a set of preferences over the possible outcomes from these interactions (usually expressed as *utility functions* - analogous to the objective or cost functions discussed in Section 5). The analogy between the AIs in CR networks and the natural intelligences in human society was so clear that almost immediately after the term cognitive radio was publicly coined, papers were written suggesting the use of game theory in the design and analysis of the interactions of cognitive radios [147]. By leveraging analysis techniques from game theory, CE researchers are able to predict likely operating states for interacting CR AIs as the points from which no CE would choose to unilaterally adapt away from lest its own performance suffers (points called Nash equilibria) and characterize convergence and stability properties of complex networks of CEs [146].

Further, by drawing on the relatively rich body of literature game theory has developed for humans, CR researchers have identified numerous game models that allow for the researchers to shape the way their AIs will interact in the field. This includes:

- Supermodular games as applied to distributed power control [148]
- Supermodular games for beamforming and power control in MIMO systems [149]
- Repeated games with punishment to encourage the forwarding of packets [150]
- Auction theory for power control and resource allocation [151]
- Potential games as applied to distributed spectrum management [66]
- Potential games as applied to routing algorithms [152]
- Potential games for topology control algorithms [153]
- Stochastic games with strategic learning applied to spectrum management [154]

In general these game models are adopted because they yield the desired behavior when CEs interact in the desired domain, but also dictate what properties the AIs must adopt to ensure the networked deployment achieves the desired results. For instance, cognitive processes designed assuming interactions that can be modeled as a supermodular game (loosely, a game with strategic complementaries between players) require the AIs to exhaust its search space so that each adaptation is individually"optimal" to ensure convergent behavior. While this may not be practically achievable in the general case, it has proven quite doable for various power control applications [148]. Likewise, when the interactions can be modeled as a potential game (a game where there exists an emergent function that increases monotonically with every unilateral selfish adaptation by an AI), there are virtually no further restrictions on the design of cognitive processes - only informed selfish unsynchronized behavior is required. However, ensuring the AIs satisfy the requisite conditions for a potential game is not as readily done as for other game models [146].

5.5 Case Study

This section presents cognitive engine development efforts by researchers at the Wireless @ Virginia Tech as a case study for applying AI techniques to CR. Two design examples are presented, one using an ANN and the other using CBR. The first example uses ordinary lab testing equipment to build fast cognitive radio prototype. It also proves that in general an AI technique (e.g., an ANN) can be chosen to accomplish complicated parameter optimization in the CR for given channel state and application requirement. The

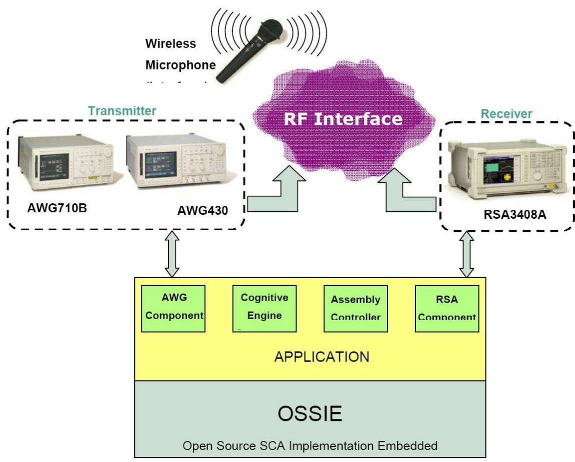

Figure 5.1: Block diagram of the CoRTekS.

second example builds upon the observation from the first one and develops a refined CE framework and process flow based on CBR. The CBR-based framework can better facilitate the interaction among awareness, reasoning, and learning in the CR.

5.5.1 ANN-Based CE Development

As a proof of concept, a cognitive radio testbed, Cognitive Radio Tektronix System (CoRTekS), has been developed by researchers at the Wireless @ Virginia Tech [58]. The block diagram of this testbed is shown in Figure 5.1. It consists of a transmitter (Tektronix arbitrary waveform generators (AWGs), a mixer, filters, an amplifier, and an omni-directional antenna), a receiver (Tektronix real-time spectrum analyzer (RSA) and an omni-directional antenna), and a personal computer (on which software defined radio and cognitive engine run).

The CE controls the transmitter and receiver by determining the best set of radio parameters for the given channel state and application requirement. An image is transmitted over the air where potential interference exists. As shown in Figure 5.2, the image is successfully received with the current channel condition and radio parameters.

As the first attempt at CE implementation, a classic AI technique, i.e., ANN, has been chosen to drive the CE. In the CE, an MLPN ANN assists in deciding the parameters to be used in the next block of transmissions. The CE uses the ANN to determine how closely a set of parameters meet a set of goals, and then chooses the set of parameters which maximizes the utility functions. The radio can change modulation

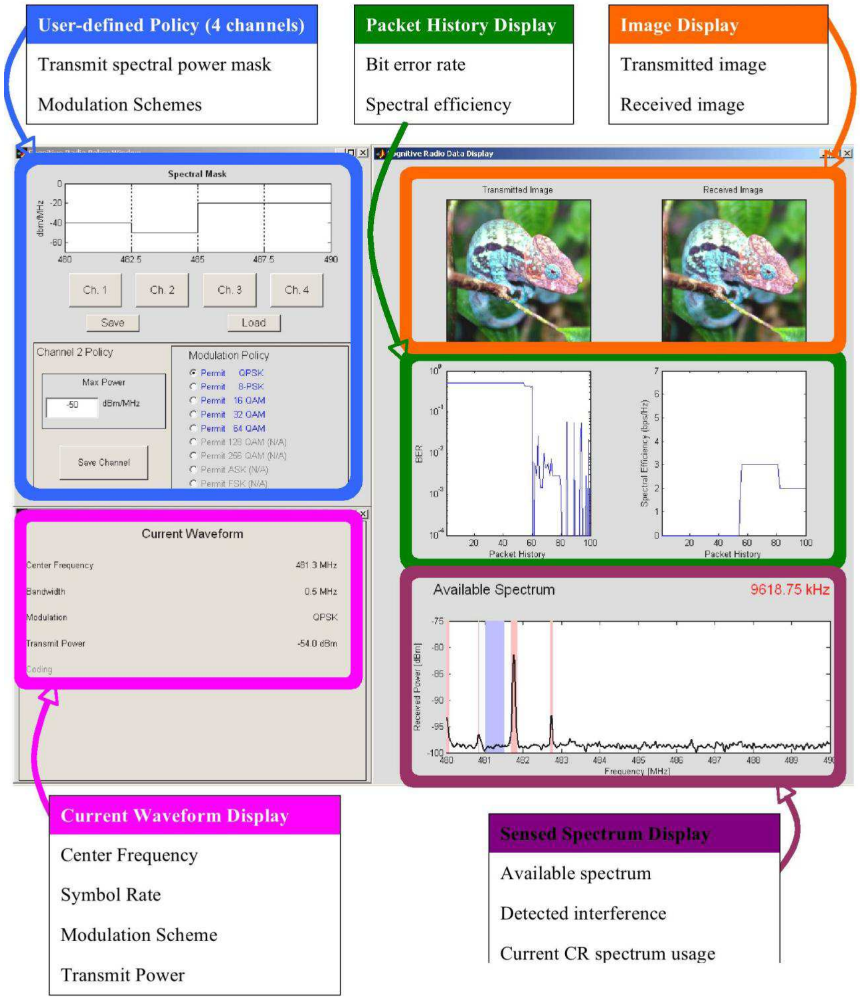

Figure 5.2: Screen shot of the cognitive radio testbed.

type, transmission power, and frequency to minimize interference, avoid the channels primary users use, and optimize three main goals: meeting QoS (BER), maximizing throughput, and minimizing transmit power. Periodically, the CE re-trains the ANN based on decisions it has made and results observed. Gradually, the ANN learns a better parameter set for a given situation by drawing relationships between radio parameters and system performance.

In order to protect primary users while the CR operates, spectrum sensing and policy are enforced in the testbed. Spectral information shown in Figure 5.2 indicates available spectrum, detected interference, and current spectrum usage. On the other hand, active policies, as demonstrated in Figure 5.2, include the transmit spectral power mask and allowable modulation schemes. Each solution adopted by the radio, e.g.,

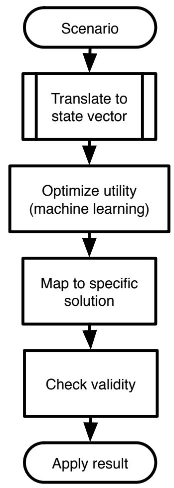

Figure 5.3: Architecture for a CE.

the waveform (radio configuration, such as, operating frequency, bandwidth, modulation scheme, channel coding scheme, and transmit power) used in the screen shot, complies with both policies.

5.5.2 CBR-Based CE Development

After the validation of the feasibility of applying AI techniques in CE design, further efforts were taken to investigate the architecture of the CE and the performance of different AI techniques. Due to the inherent structure and flexibility of a CBR system, various AI techniques can be employed in different modules of a CBR system to tailor the CE to a specific application. A CBR based CE has been developed for IEEE 802.22 WRAN applications. This work has been reported in [85] in detail. Here, we review the architecture of the CE and the observation based on this development.

An architecture of the CE is shown in Figure 5.3 [85]. The CE consists of three major processes:

1. Orientation: maps the current scenario (e.g., primary user detected) to a state vector that the optimizer (e.g., CBR) can understand.

2. Reasoning and learning: develops and optimizes the solution (e.g., a case in CBR case database) to the current problem under the policy.

3. Solution mapping and validity checking: maps the solution to a specific action and validates it against the policy and regulation before it is applied to the radio.

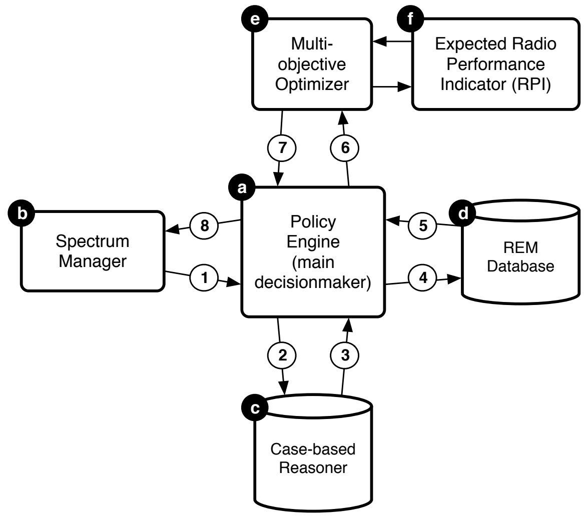

Figure 5.4: Framework of the CBR-CE for the WRAN application.

Furthermore, a CBR-based framework has been proposed for the IEEE 802.22 WRAN applications as in Figure 5.4 [85]. The arrows and the number markups in Figure 5.4 indicate a general processing flow of the CE. The CBR-CE is designed in modules with well-defined interfaces so that each module can pass on necessary information properly. This modular approach makes easier to replace any functional block with an equivalent processing element. This approach also makes the framework useful for testing and evaluating different algorithms.

The functionalities of the modules in this proposed framework are as follows.

- Spectrum Manager (SM): Monitors the radio environment, interfaces with the physical radio hardware, and allocates resources according to the solution received.

- Policy Engine (PE): Interprets policies, including standards, regulations, and customer specifications, to guide the operation of the CBR. It checks the legitimacy of a candidate solution before returning this solution to the SM.

- Case-Based Reasoner (CBR): Provides candidate solutions based on the request from the PE module.

- Radio Environment Map (REM) Database: Stores scenario-specific parameters about the system, such as geographic features, network and service availability, spectrum information, radio location and activities, policies, and experiences [125].

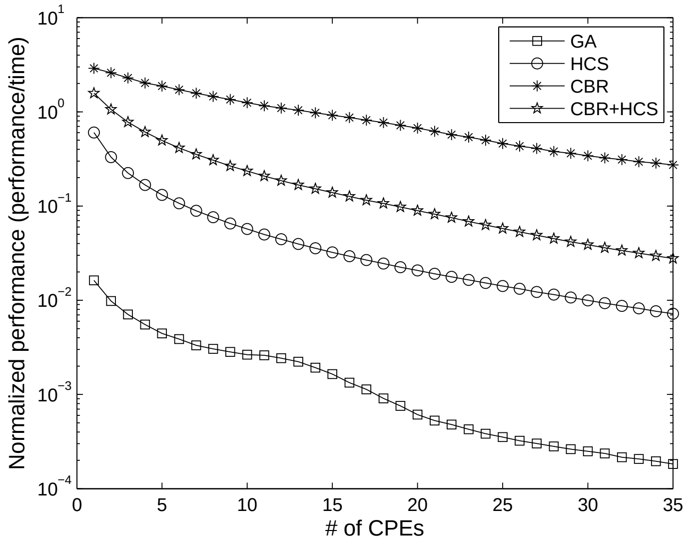

Figure 5.5: Performance comparison of different cognitive engines.

- Multi Objective Optimizer (MOO): Adapts parameters at the PHY and MAC layers based on the solution provided by the CBR. The algorithms investigated include hill climbing search (HCS), the GA, and a combination of HCS and GA. Note that other algorithms can be added to the MOO. The algorithms can also be combined to achieve a better balance between performance and the execution time.

- Expected Radio Performance Indicator (RPI): Evaluates the anticipated performance for a given solution before applying it in the real environment. The result is used in the MOO to determine whether QoS requirements are satisfied and whether further optimization is necessary.

Using the framework in Figure 5.4, we emulate the operation of a CR base station (BS) and customer premise equipment (CPE). A GA based CE (GA-CE), an HCS-based CE (HCS-CE), a CBR based CE (CBR-CE), and a CE based on the combination of CBR and HCS (CBR+HCS-CE) are used to find optimal radio configurations for specific channel conditions and application QoS requirements. In the simulation, a BS allocates resources to multiple CPEs with different QoS requirement. The complexity of the problem depends on the number of CPEs. So do the execution time and the performance. The simulation result we obtained (see Figure 5.5) indicate that a tradeoff between CE performance and execution time needs to be explored for specific applications. In this figure, the performance of each CE is normalized by the its corresponding execution time for different number of CPEs. For the IEEE 802.22 WRAN applications where the response time of the CE is critical, a CBR-CE with or without MOO (CBR+HCS) provides a sound balance between the achieved performance and the execution time. The detailed discussion and more results can be found in [85].

5.5.3 Observations from the Case Study

The examples studied above show that although various AI techniques can be used for a particular application, a CE must be carefully designed observing the tradeoff between performance and complexity determined by the application requirement.

Another important observation is that training and learning is necessary for the cognitive engine to achieve acceptable performance. In the ANN based CE, it takes a significant amount of cases (simulated or measured) to train the ANN. Similarly, after appropriate training, the CBR-CE performs well for new scenarios. In addition, the designed CBR-CE is capable of performing real time learning.

5.6 Conclusions

AI techniques lie at the heart of CR and understanding the tradeoffs in the selection and design of AI processes is critical to a successful CR design. This article reviewed several artificial intelligence techniques - ANNs, metahueristic algorithms, hidden Markov models, rule-based systems, ontology-based systems, and case-based systems - that have been proposed for providing the cognition capability in a CE. While we saw where AI techniques have been applied to numerous CR applications [87, 88, 89, 59, 60, 61, 62, 63, 64, 65, 54, 66, 67], many implementations remain rudimentary, perhaps due to the interdisciplinary nature of the field and perhaps because products are just beginning to appear.

We saw where the appropriateness of AI techniques varied by application and implementation. The decision in choosing one or some AI techniques over other techniques in CE design needs to be made based on the application requirement, considering the tradeoffs among response time, processing complexity, training sample availability, robustness, etc. In addition, the learning capability of the AI technique needs to be considered and exploited in designing a CE as learning is critical to the performance of autonomously deployed CRs.

Additionally, we noted that the design of AI processes should consider how the processes perform when deployed in the context of other AI processes and not just performance in isolation. While game theory provides an approach to make this problem analytically tractable, any analysis must begin with knowledge of expected operation of other AIs that may be encountered which is problematic assuming different implementations from different vendors and if the AI processes self-direct their evolution. Thus developing an inter-CE (and thus inter-AI) etiquette that is robust to variances in AI implementations will be important to the success of large-scale deployments of CRs. Ultimately, the implications of the etiquette will need to flow down all the way to the choice of performance metrics that guide the AI processes' decisions [155].

Our experience suggests that for a robust cognitive engine, different AI techniques should be used to complement their relative strengths and weaknesses. A CE needs to be tailored for the desired application by identifying and appropriately combining different reasoning/learning AI algorithms. Attractive combinations we have found include combining CBS, OBS, and RBS or combining HMMs with GAs or RBS. Given that the best combination of AI techniques varies by application, CE designs should accommodate mechanisms to change AI combinations as applications change. If use of a single AI technique is required, we feel that the modularity of CBR comes closest to achieving the cognitive capabilities and application flexibility while still maintaining suitability for real-time systems.

In addition to the choice of AI techniques, the CE training process is crucial to performance. Training could be hastened by using cooperative training techniques, but to limit potential security vulnerabilities, these sources should be authenticated and externally learned behaviors evaluated against self-generated field measurements.

Last but not least, although secondary spectrum access enabled by dynamic spectrum access has received the most interest in the CR community and arguably does not require AI, we believe the greatest payoff from CR will come from the ability to support self-organizing networks that can continuously improve the management of heterogeneous network elements and radio resources far beyond what the CR's designers could conceive. To achieve this goal, the application of AI techniques to CR will need to be further refined and extended to a meta-cognitive process. To learn how to design radios that learn is a journey of exploration that has just begun.

Chapter 6

Cognitive Engine Design for Radio Power Consumption Optimization

This chapter[1] discusses the design of a cognitive engine for radio power consumption optimization using a case-based structure. In Chapters 3 and 4, we have shown the theoretical framework for power consumption reduction and the potential power savings when radio component capabilities and characteristics are considered in the radio adaptation process in addition to application quality of service requirements and channel condition. A cognitive radio is a good conceptual method of achieving the potential savings as it is able to observe and learn radio environment and radio component characteristics, make decision on favorable radio configuration to achieve the required application quality of service based on the knowledge on the environment and the radio, and reconfigure the radio based on the decision. In Chapter 5, we have reviewed several artificial intelligence techniques that can be used in a cognitive radio to facilitate this cognition process and based on our experience a case-based framework which can integrate different techniques for different subtasks works well for a cognitive radio. This chapter presents the work on designing such a cognitive engine for power consumption optimization.

(The following is the manuscript under preparation on this subject.)

Designing a Cognitive Engine Framework for Radio Power Consumption Opti-

[1] This chapter is based on reproduction of a manuscript to be submitted to a journal. The complete citation of the work is as follows: A. He, S. Srikanteswara, K. K. Bae, J. H. Reed, and W. H. Tranter, "Designing a Cognitive Engine Framework for Radio Power Consumption Optimization," to be submitted.

mization

6.1 Abstract

In our previous work, we have demonstrated significant power savings beyond conventional power optimization can be achieved using a cognitive approach. This paper presents the design of a cognitive engine framework which enables the savings. A cognitive engine (CE) is the critical component performing learning, reasoning, and decision making in a cognitive radio (CR). CR and CE have been widely researched for dynamic spectrum access to improve spectrum efficiency. This paper discusses using CR for radio power consumption minimization. This application can also serve as an example of a cognitive approach to radio resource management. In this paper, a CE framework is designed and its dynamics in optimizing power consumption are studied by simulation. In addition, this CE can also facilitate learning of radio hardware characteristics and monitoring of radio hardware operation, which is not only essential to achieving the power savings, but also important for maintaining and trouble shooting future complex radio systems.

6.2 Introduction

Our previous work [156, 157] has demonstrated significant potential power savings beyond conventional power optimization approaches when radio capabilities and characteristics are incorporated in the radio operation optimization process. This paper focuses on one of the enabling technologies, a cognitive engine (CE) framework that facilitates learning of radio environment and radio capabilities and characteristics, reasoning and decision making on favorable radio configuration, and control of radio reconfiguration. This paper discusses the design philosophy of the CE framework and the general processing flow. It also presents some simulation results using the developed CE framework.

CE design has been presented in several papers in the literature, mainly focusing on the application of dynamic spectrum access (DSA). For example, in [67, 85, 114, 125, 126, 141], CE has been designed to determine optimum radio configuration for DSA applications. In [134, 137], the enforcement of spectrum policy for DSA in CE has been discussed. In [136], the awareness of the radio software has been introduced using an example where radios negotiate the equalizer training sequence. However, none of the above CE development has addressed learning of radio hardware capabilities and characteristics and incorporating this knowledge in the radio resource management.

The CE in this paper is tailored toward the radio resource management. In particular, this work studies the dynamics of the CE for power consumption minimization in response to changing application requirements and radio environment. It also shows a feasible way of learning radio capabilities and characteristics and monitoring radio operation which has not been addressed in previous CE design. This learning can extract the relationship of a particular component in responding to a certain signal waveform. For example, it can learn the PA efficiency characteristics for different input waveforms. It would be tedious to characterize for all possible waveform during the design phase. In addition, a CR based approach can learn the characteristics even for new waveforms that the radio is not originally designed for. The learning can extend to capture the interaction between different modules and that between an individual component and the overall system. Furthermore, the capability of controlling radio parameter adaptation is especially useful for future radios, where complex digitally controlled analog components are used. For example, a wideband Motorola RF transceiver IC [29] has thousands of possible parameter combinations to control the programmable internal filter banks, synthesizer, amplifier gain, etc for various communication needs. It is almost impossible to test all the combinations manually. A CR based parameter adaptation framework can be very useful in this case by trying different combinations, characterizing corresponding performance, and deciding appropriate configuration based on learned characteristics.

This paper is organized as follows. Section 6.3 presents the CE framework and the processing flow. Section 6.4 discusses the simulation settings and analyzes the results. Section 6.5 concludes the paper.

6.3 Design of a Cognitive Engine Framework

A generic CE for the radio resource management given application quality of service (QoS) requirement, radio environment knowledge, and radio platform capabilities and characteristics needs to

- communicate with applications,
- monitor radio environment,
- obtain radio platform characteristics and control radio platform functionalities, and
- determine favorable radio configuration satisfying QoS requirement based on channel and radio characteristics.

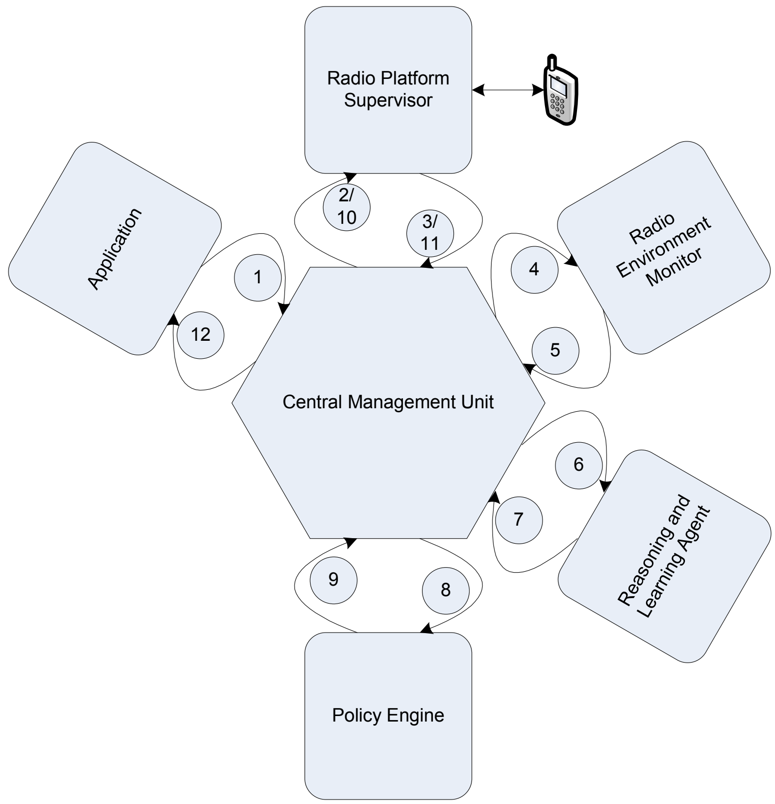

Figure 6.1: Cognitive engine framework and flowchart for power consumption optimization.

In addition to the above functional requirement, the structure of learning and decision making needs to be considered in the framework.

6.3.1 CE Framework and Processing Flow

A CE framework and processing flow for power consumption optimization (or general radio resource management) fulfilling the above requirements is shown in Figure 6.1. The functionalities of the components in the CE framework are as follows.

- Central management unit: This module handles all the control processes in the CE and directly communicates with other functional blocks. All inter-block communication has to go through this unit. It collects radio platform and radio environment information from radio platform supervisor and radio environment monitor, respectively, and sends the information to the reasoning and learning agent along with the application QoS requirement received from the application module. This design enables a relatively simple interface and a flexible processing flow where certain functional blocks can be easily bypassed or replaced to achieve better performance tradeoff during the radio operation. During an application, when the central management unit receives updates from other modules (e.g., channel gain matrix update from the radio environment monitor) it formulates an appropriate inquiry with the update to the reasoning and learning agent for possible new radio configuration.

- Application: This module is the interface between applications and the central management unit of the CE. It converts an abstract application request (e.g., email, voice, and video) to one of the predefined radio QoS requirements (e.g., a combination of data rate, delay, and bit error rate (BER)). It also notifies the application of the CE results. These notifications can include but are not limited to whether the requested application can be executed under current environment, what the expected performance is, and what alternative options are if the original requirements cannot be satisfied. For example, under certain condition (e.g, bad radio link or low battery), the radio can support audio but video. If a user request a video phone call, instead of declining such a request, the CE can suggests via this module an alternative, a regular phone call such that part of the user request can be accommodated under the current condition. This function is not currently supported but will be a nice feature if a cognitive radio resource management approach with radio hardware learning capability is adopted.

- Radio platform supervisor: This block supervises the radio platform. It is not only responsible of reconfiguring different radio components upon the order of the central management unit as in a software defined radio, but also in charge of learning the capabilities and characteristics of various radio components. This additional learning capability provides valuable information that the reasoning and decision making module can take advantage of and is the foundation for further optimizing the operation of conventional radios. In addition, this learning capability can be used for radio monitoring and trouble shooting. When the observation varies greatly from what has been remembered, the radio platform supervisor can perform diagnostic scanning and pass the scanning result to the central management unit for further analysis. For example, this approach can help understand the radio aging issue or even the human factor in radio operation (e.g., hand effect on antenna radiation). This monitoring and trouble shooting capability is beyond the focus of this paper. In this work, the radio platform supervisor learns the PA efficiency characteristics.

- Radio environment monitor: This module monitors the dynamic radio environment. In addition, it can store certain statistical channel information depending on the system requirements, such as, memory size, responding time, and granularity. The statistical channel information may include channel availability which can be used in determining the best channel for a specific type application. For example, for a voice application, a channel with longer available time is preferred to minimize channel switching which may degrade application QoS. On the other hand, for a data downloading application, more frequent channel switching is acceptable as the user can tolerate certain data rate variation.

- Reasoning and learning agent: This is the "brain" of the framework. It determines appropriate radio configuration that achieves the goal (e.g., minimizing radio power consumption in this work) based on the application QoS requirement, the channel condition, and radio characteristics. The reasoning process can naturally incorporate learning from experience, both direct and indirect experience. Here the direct experience refers to what the module has experienced as the radio operates, whereas the indirect experience refers to what has been gathered from either simulation or other radios and stored for the module to use. The indirect experience can help the radio to greatly reduce the time to learn and adapt during its initial operation in a new environment. On the other hand, the direct experience is what the radio experiences itself and can provide most relevant information for decision making. Similar concepts of direct and indirect experience apply to the learning of radio component capabilities and characteristics as well where indirect experience can be obtained from data sheets and direct experience can be obtained from measurement during radio operation. Direct experience is usually more pertinent to a specific radio. In this module, combination of several reasoning and decision making techniques can be used. This is discussed in anther section in more details.

- Policy engine: The policy engine deals with policies and constraints, such as the spectrum mask for a specific frequency band. It validates a radio configuration determined by the reasoning and learning agent based on policies and constraints. Specific design constraints can also be incorporated in this block. Some policy consideration can be enforced in the reasoning and learning process. This increases the possibility that the returned solution from the reasoning and learning agent complies with policy. However, it also increases the complexity of the reasoning and learning process in finding a possible solution. This balance needs to be carefully studied for specific design situation. For example, in this work, we impose constraints on maximum output power from each branch and fixed bandwidth in the reasoning process. This module validates whether the possible solution complies with the spectrum mask for a specific channel. Other constraints can be incorporated as needed.

A typical CE processing flow is also shown in the CE framework in Figure 6.1.

① Service request: The application module converts an abstract application request received into a QoS requirement that the radio can understand and passes that to the central management unit.

②/③ Radio platform information retrieval: The central management unit retrieves radio capabilities and characteristics from the radio platform supervisor. This information is used by the reasoning and learning agent.

④/⑤ Radio environment information retrieval: The central management unit retrieves current channel condition and possibly historical spectrum information from the radio environment monitor. This information is used by the reasoning and learning agent to quantify channel quality.

⑥/⑦ The reasoning and learning agent determines favorable radio configuration based on the QoS requirement, the radio platform information, and the radio environment information received from the central management unit. Multiple possible radio configurations can be returned to the central management unit depending on the system specification. In the case that the requested QoS requirement may not be satisfied on the current platform or under the current radio condition, the reasoning and learning agent may suggest an alternative option with corresponding QoS.

⑧/⑨ Radio configuration validation: The policy engine validates the radio configurations forwarded from the central management unit against policies and constraints. Although the spectrum policy is most common, other policy and constraint can be incorporated. As we mention earlier, some policy information can be included and considered in the reasoning process, which may lead to added complexity in reasoning. This paper assumes that no explicit policy is considered in the reasoning process.

⑩/⑪ Radio platform reconfiguration: After the policy engine validates the possible configuration, the central management unit can then request the radio platform supervisor to reconfigure the radio using the new configuration. After the radio is properly configured and ready for the task, the radio platform supervisor notifies the central management unit with a success.

⑫ Service launching: When the central management unit receives a successful reconfiguration message from the radio platform supervisor, it notifies the application module to send the message to allow the requested application to be launched. When the requested service cannot be supported for some reason, the central management unit will also send a message with either reasons or alternative options.

In the above process, some steps may occur iteratively. For example, the radio platform information retrieval may happen only once when a service request is received. However, the CE may need to circle through radio environment information retrieval, radio configuration determination and validation, and radio platform reconfiguration multiple times during the duration of an application when the radio environment changes. It is also necessary to perform radio platform information retrieval again if some abnormal phenomena are observed.

6.3.2 Learning and Reasoning

Learning and reasoning is a unique and important element of this cognitive radio approach to radio power optimization. The learning process resides in multiple modules in the framework. First, the framework learns radio component characteristics and capabilities. As shown in [156, 157], the knowledge of radio component characteristics is essential to achieving the power savings beyond conventional mechanisms. The radio platform supervisor is responsible for this learning process. This work focuses on the impact of the PA efficiency characteristics on the radio power efficiency since the PA is usually one of the most significant power consuming components in a radio in medium to long range applications, such as cellular services. The basic PA efficiency characteristics can usually be obtained either from component data sheet or from measurement at the design stage. However, it is very tedious if is not impossible to capture the run time characteristics, including specific waveform effect, temperature effect, aging effect, process variation, and so on. The radio platform supervisor needs to learn them. The learning of hardware characteristics can be enabled by some new measurement IC products. For example, a small form factor and low power consumption device, the ADM1191 - digital power monitor by Analog Device [158], can be integrated in the radio to measure power consumption. Note that this learning process only needs to be performed once in a relatively long period of time (in days, weeks, or even months) compared with application duration depending on how stable the component characteristics are and how accurate the characteristics need to be. Therefore, the overhead incured due to such learning can be minimized. In addition, a similar approach can be used to monitor the operation of other radio components and the results can be used for radio malfunction troubleshooting.

Another important place where learning plays a vital role is in the reasoning and learning agent. In this module, learning and reasoning intertwine with each other similar to some other CE designs [85, 114, 126]. In a survey of artificial intelligence (AI) techniques for CR [159], we have found out that careful selection of a combination of several AI techniques is critical for a specific application of CR. In addition, the survey and our practice show that the case-based reasoning (CBR) approach provides a nice structure for a CE to incorporate learning and reasoning for applications such as radio resource allocation [85]. Using the CBR based structure, different AI techniques can be integrated in the CE to implement different functions. Therefore, in this work, the CBR is also adopted in the reasoning and learning agent. Since this work focuses on the CR framework and its operation and the detailed CBR based CE designed has been discussed in [85] for a similar radio resource management application, we highlight the essence of the reasoning and learning process.

Case based reasoning [139] is a reasoning process based on experiences (cases). This process resembles certain human reasoning process where a decision is made based on similar experience instead of rigorous deduction. Hence, CBR possesses many features similar to its human counterpart. For example, it can arrive at a decision quickly as step by step derivation is avoided. In addition, it can come up with a solution to a problem where there is limited knowledge about the problem or missing steps in existing analytical solution finding procedure by trying and modifying a known solution to a similar problem. These features make CBR particularly appealing to radio resource management since the knowledge (model) about the radio environment is usually ambiguous and it is important to make decision quickly in response to changing environment to explore the benefit that dynamic radio resource management can bring. Therefore, CBR is adopted here for radio power consumption minimization.

The functional components of CBR naturally include [139]:

- Case representation and indexing: formats the experience so that it can be later used by the reasoner and retrieved accordingly;
- Case selection and retrieval: finds the case in the case database that matches the request;
- Case evaluation and adaptation: evaluates the retrieved case using a performance model or by applying the solution and observing the outcome;
- Case learning and case database maintenance: retains useful experience by creating a new case or adapting an existing case and manages the case database to maintain a balance between case granularity and retrieval efficiency.

This work adopts the same architecture as in [85] and implements it for the purpose of power consumption optimization.

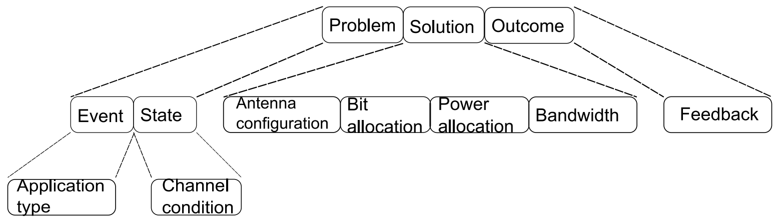

Figure 6.2: Case structure.

One concept that is worth additional attention is the definition of a case as it needs to be tailored for each specific application of CBR. A case, a fundamental element in CBR, is a piece of contextualized experience that "teaches a lesson fundamental to achieving the goals of the reasoner [139]." As shown in Figure 6.2, the content of a case includes three components: problem description which describes the most distinguishing features of a problem, solution which explains how the problem was solved, and outcome records the result of applying the solution. The design of a case is important to the performance of the CBR. A case needs to capture necessary information for future use. The representation of information may lead to vastly different performance in terms of memory requirement, retrieving speed, and accuracy. This aspect is studied in more details in the simulation results.

6.4 Simulation Environment and Results

A CE framework is designed and evaluated by simulation to study the dynamics of the framework in response to the changes in application and radio environment. This section presents the simulation environment and some results. Since some results on a CBR based CE for a similar radio resource allocation application have been discussed in [85], this section focus on the essence of the reasoning and learning process, such as, the learning of the radio hardware component characteristics, the case design, and the online and offline learning.

6.4.1 Simulation Environment

The simulation considers a single cell radio environment with a base station (BS) and a mobile station (MS). The current simulation focuses on a single link between the BS and the MS. Multiuser/multilink scenario will be studied in future work. We assume that the MS is moving at a constant speed within the cell range. This motion is mathematically modeled as a two dimensional (2D) random walk. At a constant observation rate, the location of the MS and the radio channel between the MS and the BS are estimated. It is assumed

Table 6.1: Simulation Settings

Parameter (unit)	**Value or Range**
Number of cells	1
Cell radius (m)	1000
Path loss exponent	3
Number of BSs	1
Number of antennas on BS	4
Maximum output power from each BS antenna (W)	1
Number of MSs	1
Number of antennas on MS	4
Maximum output power from each MS antenna (W)	1
Channel central frequency (MHz)	900
MS speed (km/h)	3.6
Type of application, spectral efficiency (b/s/Hz): Example application (exponential duration (min): mean)	Low, 2: voice (3); Medium, 5: Internet music (4); High, 8: Internet video (7)
Application arrival model (/min)	Poisson process with arrival rate $\frac{1}{20}$

that the radio can be reconfigured at the same rate if it is needed. Both BS and MS are equipped with multiple antennas and the channel gain on each antenna experiences uncorrelated Rayleigh fading along with path loss depending on the distance between the BS and the MS. Other channel models can also be adopted. The change of radio channel during the period of an application may lead to reconfiguration of the radio given the optimization target. This is determined by the reasoning and learning agent. The system supports three categories of applications based on the spectral efficiency assuming a fixed bandwidth for all applications, namely, low rate (2b/s/Hz), medium rate (5b/s/Hz), and high rate (8b/s/Hz). The application request arrives according to a Poisson process, and different types of application requests appear with equal probability. The detailed system settings are summarized in Table 6.1 for easy reference.

6.4.2 Learning of Radio Component Characteristics

As one of the important features of the CR framework, the radio platform supervisor is able to learn radio hardware characteristics, e.g., PA power efficiency characteristics in this case. Since the PA power efficiency characteristics are relatively stable, the learning can be performed only once in a long period of time and the results can be stored in memory for future retrieval. Every time the characteristics learning is performed, the memory is updated with the new results. This update will also be able to detect any significant changes in the characteristics and alert the central management unit of possible malfunction of the component. An important tradeoff in this process is between the accuracy and the memory space and speed. In this framework, the radio platform supervisor curve-fits the measurement data to the two-parameter PA efficiency

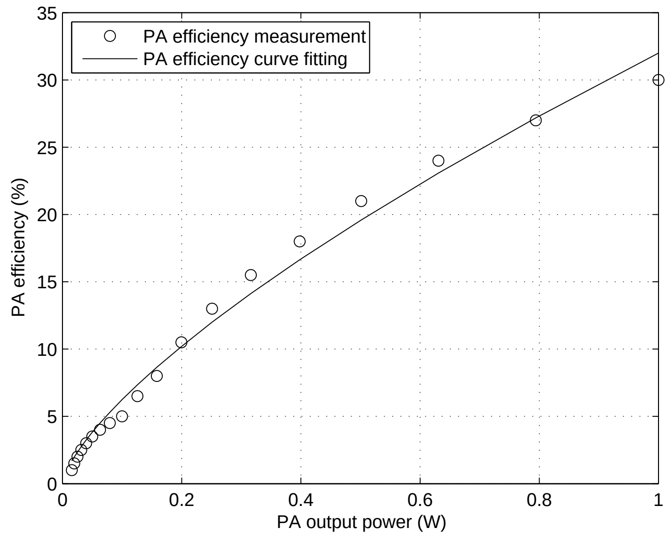

Figure 6.3: PA efficiency characteristics parameter extraction via MMSE curve fitting.

model in (6.1) first used in [76, 160]

$$\bar{\eta} = \eta\left[P_t\right] = \left(\frac{P_t}{P_{max}}\right)^{\alpha} \eta_{max}, \tag{6.1}$$

where P_t is the PA output power, P_{max} is the maximum PA output power, η_{max} is the PA efficiency at its maximum output power level, and α is the efficiency exponent depending on the particular PA used. This model based approach requires only two parameters and captures the efficiency characteristics of the theoretical Class A and Class B PAs and some real PA based on minimum mean square error (MMSE) criterion [156]. On the other hand, if a table lookup approach is adopted for efficiency knowledge, a large memory space is inevitable for required accuracy. The lookup table approach may also lead to relatively longer processing time compared with the model based approach.

In this work, a real PA (TRF4002 [74] by Texas Instruments which has been used in several wireless products) is assumed to be used in the system. A measurement IC (e.g., ADM1191 - digital power monitor [158] by Analog Devices with upto 400K Hz) is assumed to be integrated in the radio platform to monitor the PA power consumption. The "measured" [2] PA efficiency characteristics are shown as circles in Figure 6.3. Also shown in this figure is the MMSE curve-fitting result of the realistic PA efficiency characteristics to the analytical model in (6.1). The extracted parameters for this PA are $\eta_{max} = 0.32$ and $\alpha = 0.7$. As expected,

[2] This measurement is emulated by using the PA efficiency curve in the product data sheet of the realistic PA, TRF4002.

this set of parameters falls in between Class A and Class B PA parameters. The following results are all based on this extracted realistic PA efficiency model.

6.4.3 Case Design

The average signal to noise ratio (SNR) per channel in this work is defined as

$$\bar{\gamma} = \frac{P}{\sigma_n^2 \cdot r \cdot t} \mathrm{E}\left[\|\mathbf{H}\|_F^2\right], \tag{6.2}$$

where P is the radiated power from a transmit antenna, σ_n^2 is the noise power spectral density, r and t are the number of transmit and receive antennas, respectively, and $\|\mathbf{H}\|_F$ is the Frobenius norm of channel gain matrix $\mathbf{H}$.

The Frobenius norm can be expressed as [161]

$$\|\mathbf{H}\|_F^2 = \sum_{i=1}^{r}\sum_{j=1}^{t} |h_{i,j}|^2 = \mathrm{tr}\left(\mathbf{H}^\dagger \mathbf{H}\right) = \sum_{i=1}^{\min(r,t)} \lambda_i^2, \tag{6.3}$$

where $h_{i,j}$ is the element on the i-th row and j-th column of the channel gain matrix $\mathbf{H}$, $\mathrm{tr}(\cdot)$ is the trace function, and λ_i is the i-th singular value of channel gain matrix $\mathbf{H}$ and λ_i^2 is the i-th eigenvalue of $\mathbf{H}^\dagger\mathbf{H}$.

The element of the channel gain matrix, $h_{i,j}$, captures path loss and Rayleigh fading effects with two independent multiplicative components. The path loss at a distance d with a path loss exponent n, $PL(d,n)$, is defined as [162]

$$PL(d,n) = \left(\frac{d}{d_0}\right)^n \cdot PL(d_0,n), \tag{6.4}$$

where $PL(d_0,n)$ is the path loss at the reference distance d_0 with the path loss exponent n. The Rayleigh fading component [162], $g_{i,j}$, is modeled as a complex Gaussian random variable with zero mean and equal variance $\frac{1}{2}$ for its real and imaginary parts. Since path loss is about average power and Rayleigh fading is about amplitude, the channel amplitude gain $h_{i,j}$ can be expressed as

$$h_{i,j} = \sqrt{\frac{1}{PL(d,n)}} \cdot g_{i,j}. \tag{6.5}$$

Therefore, $h_{i,j}$ is a complex Gaussian random variable with zero mean and equal variance $\frac{\sigma^2}{2} = \frac{1}{PL(d,n)}$ for its real and imaginary parts.

Hence, we have

$$\mathrm{E}\left[\|\mathbf{H}\|_F^2\right] = r \cdot t \cdot \mathrm{E}\left[|h_{i,j}|^2\right] = \frac{r \cdot t}{PL(d,n)}, \tag{6.6}$$

and

$$\bar{\gamma} = \frac{P}{\sigma_n^2 \cdot r \cdot t}\mathrm{E}\left[\|\mathbf{H}\|_F^2\right] = P \cdot \frac{1}{\sigma_n^2 \cdot PL(d,n)}. \tag{6.7}$$

In the simulation, we use normalized power spectral density $\sigma_n^2 = 1$ and the maximum average SNR per channel under maximum output power at the cell edge is set to 3 dB to achieve desired application QoS.

The design of a case is critical to the performance of CBR. Since the channel information is important in determining antenna combination and rate allocation, it needs to be properly presented in a case. Naturally, we can either keep the channel gain matrix $\mathbf{H}$ or the corresponding singular values λ_i in a case. It is important to estimate the number of cases that the case database should hold. This parameter determines the memory space that needs to be reserved and also the retrieval time and the power consumption for retrieval in the CBR. To determine which form of information to store in a case, we investigate the probability density functions (PDFs) of the element of the channel gain matrix and the corresponding singular value, respectively.

The PDF of the element of the channel gain matrix (real or imaginary part) can be expressed using a mixture distribution [163] as follows:

$$f_H(h) = \int_S f_S(s) \cdot f_H(h;s) ds, \tag{6.8}$$

where $f_H(h;s)$ is the PDF of the real or imaginary part of $h_{i,j}$ parameterized by s and $f_H(h;s) = \frac{1}{\sqrt{2\pi \cdot s/2}} e^{\frac{-h^2}{2 \cdot s/2}}$, $f_S(s)$ is the PDF of s, the variance of $h_{i,j}$, and $s = \sigma^2 = \mathrm{var}\left[h_{i,j}\right] = \frac{1}{PL(d,n)}$. For each value of s in the set S, $f_H(h;s)$ is a PDF with respect to $h_{i,j}$. This mixture distribution reflects the fact that the variance of the element of the channel gain matrix is different when the MS is at different location (different distance between the MS and the BS).

Assume the location of the MS in the cell is uniformly distributed, i.e.,

$$f_D(d) = \frac{d^2}{R^2}, \tag{6.9}$$

where d is the distance between the MS and the BS and R is the cell radius. It is easy to show that the variance of the element of the channel gain matrix, s, follows

$$f_S(s) = \frac{d_0^2}{R^2} \cdot \frac{2}{n} \cdot (PL_0)^{-\frac{2}{n}} \cdot s^{-\frac{2}{n}-1}. \tag{6.10}$$

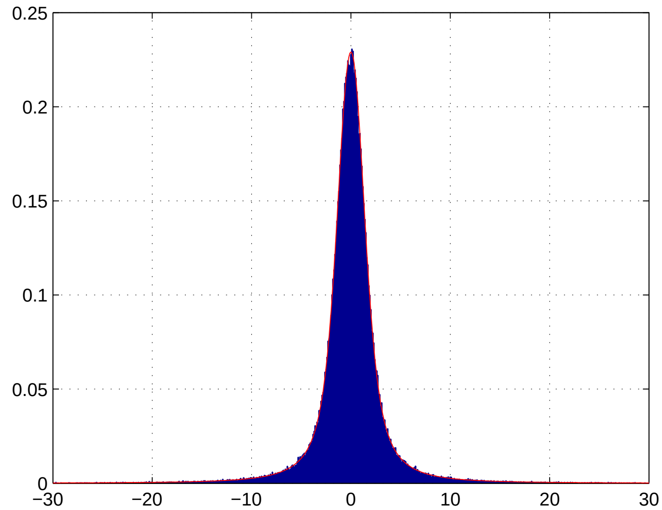

Figure 6.4: PDF of the element of the channel gain matrix.

The resulting PDF of the real or imaginary part of the element of the channel gain matrix is plotted in Figure 6.4. The estimated PDF based on histogram is also shown in this figure. The estimated PDF matches closely the theoretical expression of the PDF.

Similarly, we can obtain the PDF of the corresponding singular values of the channel gain matrix. The limiting density of the eigenvalues of a Wishart matrix ($\mathbf{W} = \frac{1}{t}\mathbf{X}\mathbf{X}^\dagger$) is given by [164]

$$f_X(x) = \frac{1}{2\pi c x}\sqrt{(b_+ - x)(x - b_-)}, \tag{6.11}$$

where $\mathbf{X}$ is of dimension $r \times t$ whose element $x_{i,j}$ is independent and identically distributed (i.i.d.) with zero mean and unit variance, $c = \frac{r}{t}$, and $b_\pm = (1 \pm \sqrt{c})^2$. With simple derivation, we can show that the limiting density of the eigenvalues of a matrix $\mathbf{V} = \mathbf{Y}\mathbf{Y}^\dagger$ where $\mathbf{Y}$ is of dimension $r \times t$ whose element $y_{i,j}$ is i.i.d. with zero mean and variance σ^2 can be expressed as

$$f_Y(y) = \frac{1}{2\pi c t \sigma^2 y}\sqrt{(t\sigma^2 b_+ - y)(y - t\sigma^2 b_-)}. \tag{6.12}$$

Furthermore, due to the relationship $z^2 = y$ between the singular value, z, of the matrix $\mathbf{Y}$ and the eigenvalue,

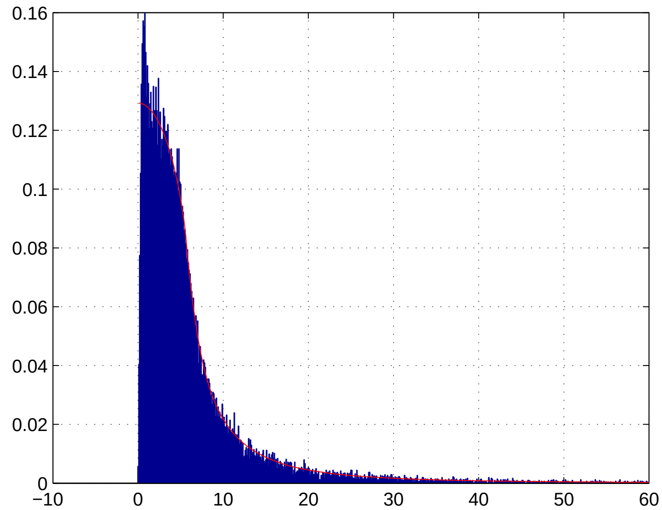

Figure 6.5: PDF of the singular values the channel gain matrix.

y, of the matrix $\mathbf{V}$, the limiting density of the singular values of the matrix $\mathbf{Y}$ can be expressed as

$$f_Z(z;s) = \frac{1}{\pi ctsz}\sqrt{\left(tsb_+ - z^2\right)\left(z^2 - tsb_-\right)}, \tag{6.13}$$

where $s = \sigma^2$.

Therefore, the PDF of the singular values of the channel gain matrix with random variance can be expressed using a mixture distribution as follows:

$$f_\lambda\left(z\right) = \int_s f_S(s) f_Z(z;s) ds. \tag{6.14}$$

This PDF is plotted in Figure 6.5 along with the estimated PDF based on histogram.

Using the obtained PDFs, we can determine the ranges of the channel gain and the singular value in which the channel gain and the singular value fall with a probability greater than or equal to 95%. This range is approximately $[-8, +8]$ and $[0.1, 22]$ for the channel gain and the singular value, respectively. These ranges can be used to estimate the required memory space for storing such information. For example, if a step size of 1 is used for both the channel gain and the singular value, there are 17 steps and 22 steps for the channel gain and the singular values, respectively. For a 4×4 system considered in this work, it may generate approximately 17^{16} and 23^4 cases, respectively, if we assume the channel gains and the singular values are independent. It is obvious that it requires much less memory space to store the singular values in a case than the channel gains in a case. Note that the actual number of cases in a case database can exceed these numbers since when the radio operates long enough, it experiences scenarios with small probability and corresponding cases are added to the case database. On the other hand, the singular values are required in determining the actual power and rate allocation. Therefore, it is a better choice to store singular values in a case.

6.4.4 Dynamics of Reasoning and Learning Agent

After studying the learning of radio hardware component characteristics and the case design, this section investigates the dynamics of the reasoning and learning agent in responding to various types of applications and channel conditions.

Using the PA efficiency characteristics learned by the radio platform supervisor, the CR framework determines appropriate radio configuration to minimize power consumption. This work focuses on the CSIR-CSIT scenario considered in Chapter 4 as an example, where the channel state information is known at the receiver and the transmitter in a multiple input and multiple output (MIMO) system. As in Chapter 4, the power savings metric is defined as

$$S = \frac{P_{con} - P_{cog}}{P_{con}} \cdot 100\%, \tag{6.15}$$

where P_{con} is the power consumption using a conventional power allocation approach, P_{cog} is the power consumption using the proposed CR based power allocation approach.

Offline Learning vs. Online Learning

For many artificial intelligence techniques, training is a critical step for its operation. The case based reasoning can actually work without any training and learns as it operates. However, the CBR can also learn (accumulate experience) in an offline fashion (the offline learning) where it is trained with limited typical scenarios and improve its learning process. To differentiate, we call the learning process during the CBR operation the online learning.

In the offline learning process, the CBR starts with an empty case database and is provided with purposely created training scenarios. When the CBR sees a training scenario, it first tries to find a similar case in its case database. If it cannot find one, it then uses the heuristic algorithm developed in Chapter 4 to determine

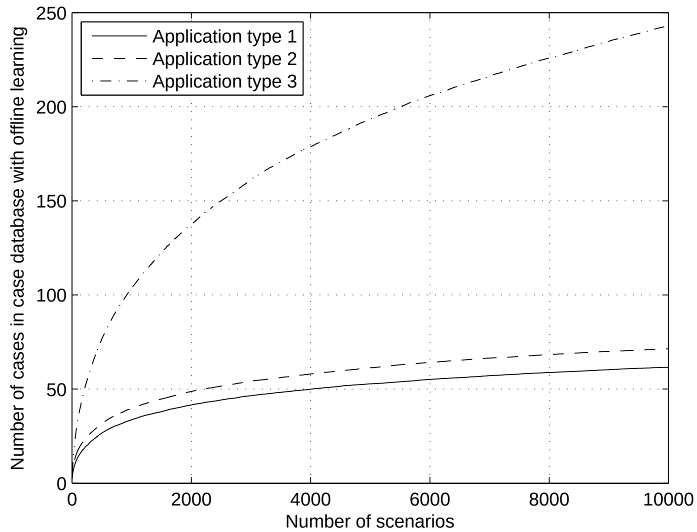

Figure 6.6: Offline learning for the CBR.

an appropriate solution. This solution is then validated and the validated solution is stored in the case database for future use. The online training process is similar to the offline training except that in the online training, the CBR learns with actual scenarios.

Figure 6.6 shows the average offline learning progress for the three types of applications investigated in this work. The results are averaged over a sufficient number of sample offline learning processes such that a 95% confidence interval is guaranteed. In each offline training process, 1×10^4 training scenarios are used and each scenario corresponds to one distinct location and each scenario has one channel observation. It is clear that the rate of case accumulation in the case database is fast when the number of cases in the case library is small. This rate slows down when the number of cases increases further. Finally, it tends to saturate for application type 1 and type 2 where the number of cases in the case database is presumably small. It is natural to observe that the number of cases in the case database depends on the application type. In other words, the case database contains more cases for the application with higher rate requirement. This is because under the same channel condition, the application with higher rate requirement may require more antennas at the transmitter which leads to a larger number of singular values. Hence, it has the opportunity to create cases using more active antennas. These scenarios are created using the channel model discussed above. After the offline learning, the CBR accumulates on average 61, 71, and 242 cases for the three types of applications, respectively. It is also possible to further study the characteristics of the accumulated cases (e.g., the statistical characteristics of the cases) for the different applications and create a set of typical cases

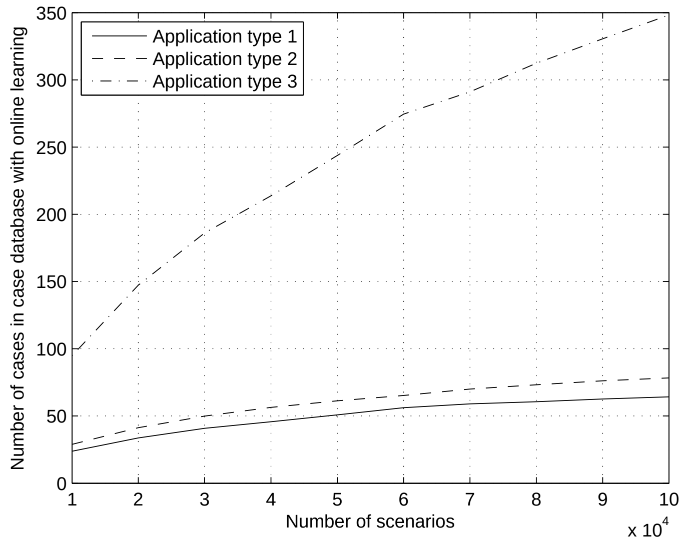

Figure 6.7: Online learning for the CBR.

for each application. These sets of typical cases can then be directly load into the the case database as the offline learning results.

In comparison, Figure 6.7 shows the average online learning progress for the three types of applications investigated. Each offline training process consists of 50 applications of a specific type and each application consists of many channel observations. Based on the average result, to obtain the similar number of cases in the case database, the online learning process requires 8×10^4, 8×10^4, and 5×10^4 channel observations, respectively. In compared with 1×10^4 channel observations used in the offline learning process, much more channel observations are required in the online learning process to accumulate a similar amount of cases in the case database. This is because in the offline learning process the training samples are created such that they are spread out in the target coverage area whereas in the online learning process the observations from each application session tend to lump around a specific location in the coverage area (i.e., the initial location of the radio determined at the start of each application session). Generally speaking, the offline learning process accumulate typical experience and the online learning process collects more detailed experience under the specific environment.

From the result on the comparison of the offline learning and the online learning, the offline learning can accumulate typical experience quickly. On the other hand, the benefit of using the online learning is that the radio learns as it operates and that it does not consume additional resource for learning. Therefore, in this framework, we combine both types of learning to exploit their benefits. In other words, first the radio

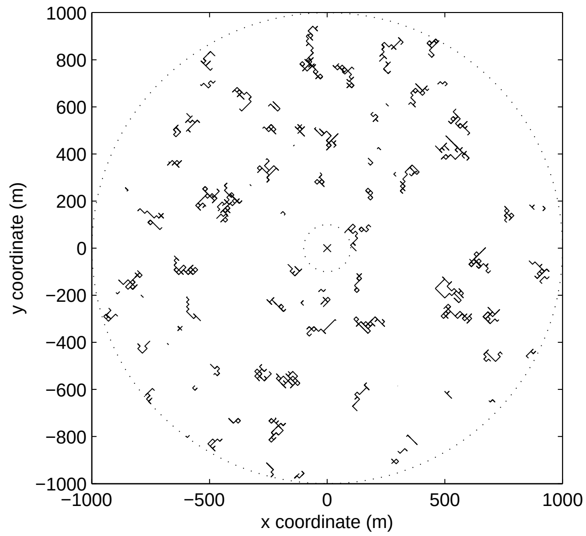

Figure 6.8: MS location for different applications.

performs the offline learning for the CBR to gain some general experience. Then, it use the trained case database in the following operation and performs online learning to gain specific experience.

Power Savings and Time Complexity

To study the dynamics of the CE framework, a temporal simulation is performed where 100 independent random application requests for each application type are generated and corresponding performance results are observed. Each application type is generated equally likely. At the beginning of each application, a random MS location uniformly distributed in the cell range is created and the MS moves at a constant speed according to a 2D random walk process during the application duration. Figure 6.8 shows a sample MS motion trajectory in the cell range. At a constant frequency, the MS location is noted as a point on one trajectory. At the mean time, the channel realization at that location is observed and used in determining a preferred radio configuration for the given application type. If the new configuration is different from the current configuration, the radio needs to be reconfigured.

It is important to study the power savings and the time complexity performance using the CBR approach compared with the exhaustive/heuristic algorithms. Figures 6.10 and 6.9 show the average power savings that can be achieved using a numerical algorithm and the relative power savings performance of the cognitive approach with respect to the numerical algorithm approach, respectively.

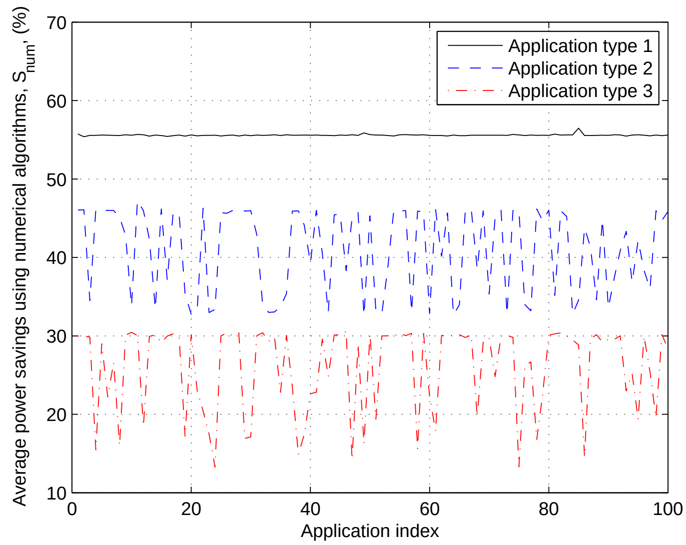

Figure 6.9: Average power savings using a numerical algorithm (branch optimization in Section 4.5.2) for three types of applications.

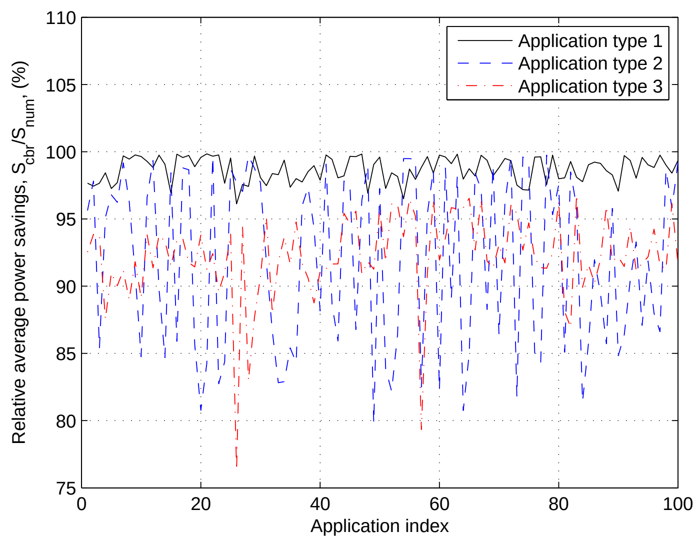

Figure 6.10: Relative power savings for three types of applications.

In addition, the relative time complexity of the cognitive approach with respect to the numerical algorithm approach is shown in Figure 6.11.

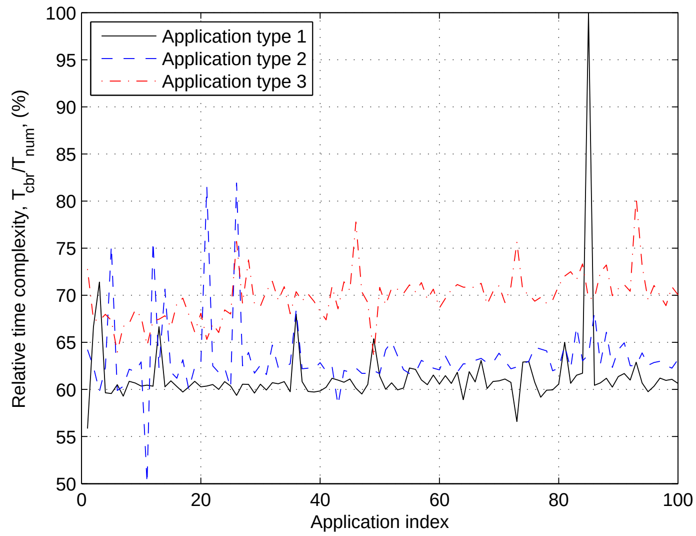

Figure 6.11: Relative time complexity for three types of applications.

Note that relative power savings and the relative time complexity are defined as:

$$S_{rel} = \frac{S_{cbr}}{S_{num}} \cdot 100\%, \tag{6.16}$$

and

$$T_{rel} = \frac{T_{cbr}}{T_{num}} \cdot 100\%, \tag{6.17}$$

respectively.

It can be seen that the CBR based approach can achieve on average about 90% power savings with less than 75% time complexity depending on the type of application. Note that this time reduction is based on one of the heuristic algorithms, which is already much faster than the pure exhaustive search algorithm. Here we use the heuristic algorithm as baseline since in practical implementation the time consuming exhaustive search approach would not be chosen as it cannot keep up with the time requirement. Observe that the relative time complexity for application type 1 is lower than that of application type 2 and type 3. This is because the case database for application type 1 is generally smaller than that of type 2 or type 3. The difference between the time complexity is partly due to the search of the case databases of different size. It is interesting to notice that the power savings are confined in a certain region for each applications even if the MS is wandering around the cell. For curiosity, Figure 6.12 shows the average power savings when the initial locations of the MS are either at the cell edge or at the cell center. This result confirms that the

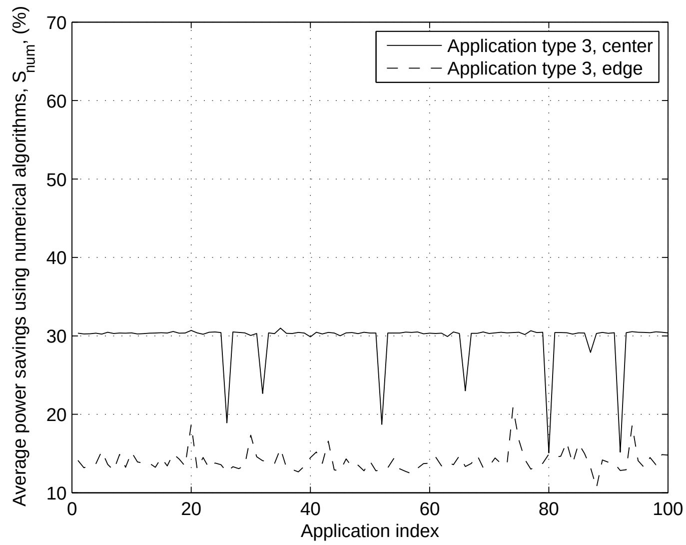

Figure 6.12: Average power savings using a numerical algorithm for type 3 application at either cell edge or cell center.

power savings in Figure 6.9 should fall in between the two extreme cases.

6.5 Conclusion

This paper discusses a cognitive radio approach to radio power consumption optimization and the design of a case based cognitive engine that realizes the additional power savings beyond conventional approaches for mobile and wireless communications. The proposed CE framework incorporates the learning of radio hardware component characteristics and the radio operation optimization for power reduction using the learned radio characteristics. The capability of learning hardware component characteristics has a wide potential not only for radio operation optimization but also for radio operation monitoring and problem trouble shooting and diagnosis. This learning capability can provide valuable information for applications such as the radio self-organizing networks in the Long Term Evolution (LTE). This paper also investigates the tradeoff between the offline learning approach and the online learning approach and designs a combined learning approach for the cognitive engine where the offline learning quickly equips the cognitive engine with typical cases for operation and the online learning then obtains specific experience for close match. In addition, the power savings and time complexity of the case based approach are compared with the numerical approach we studied in our previous work. The result shows that the case based approach can achieve almost all the power savings with lower time complexity by exploiting the experience recorded as a case in the case database. Hence, the case based approach provides a flexible way to balance the potential power savings and the time complexity given the actual radio operation requirement. For example, for a more time constraint environment, the case based CE can be adjusted such that faster adaptation is achieved with lower potential power savings.

Chapter 7

Conclusions

7.1 Research Summary

Power consumption has been a critical aspect in wireless and mobile communications. Previously, a tremendous amount of work has been devoted to this field from the device level to the system level. This work seeks to advance the understanding and the optimization of system power consumption taking advantage of the emerging cognitive radio technologies.

Specifically, this work first proposes a methodology of using a cognitive radio framework for power consumption optimization. The framework enables learning of the radio component characteristics (e.g., the PA efficiency) which is necessary for power consumption optimization. Using the learned component characteristics, radio parameters and component characteristics can be jointly adapted to minimize power consumption given the channel and meet the QoS requirement of an application. This framework can be extended to optimize radio operation to achieve additional goals. This work focuses on the impact of the PA characteristics on the power consumption since the PA usually dominates the power consumption in medium and long range wireless applications. Other components can be integrated into the framework as needed.

Then, this work goes on with the potential power savings using the CR approach for the SISO systems. The relationship between the radiated power, the PA power consumption, and the PA efficiency characteristic in SISO systems is investigated. A unified PA efficiency model characterizing theoretical Class A, Class B, and practical PAs is adopted and enables the analysis of the impact of different radio configurations and channel conditions on power efficiency. With PA efficiency knowledge, the CR framework is able to minimize the power consumption by adapting radio parameters, such as, modulation, coding rate and coding gain, and radiated power. In addition, further system power consumption reduction for given channel and QoS requirement can be achieved if radio parameters and radio component characteristics (i.e., PA efficiency characteristic) are jointly adapted.

The potential power savings using the CR approach for the MIMO systems is also investigated. Leveraging the results from information theory and the capabilities of a CR (e.g., the awareness of the component capabilities and characteristics), we develop a theoretical framework to minimize power consumption for MIMO systems. The power consumption minimization problems under a sum rate constraint for MIMO systems under various scenarios (e.g., different channel state information availability at the transmitter and the receiver and antenna correlation at the transmitter) are mathematically formulated. Several heuristic numerical algorithms are developed to solve the constraint optimization problems and the power savings are evaluated by simulation.

After understanding the potential power savings, we develop a cognitive engine to enable the savings. This cognitive engine facilitates learning and decision making in the CR framework. A CR can not only learn the channel conditions as in conventional radios, but is also aware of the radio (component) capabilities and characteristics. In this work, a case-based reasoning cognitive engine is designed and evaluated. A combined learning approach is adopted in the cognitive engine where the offline learning process quickly equip the CE with typical experience and the online learning accumulates specific experience while the radio operates. With the combined learning, the CE is able to achieve most of the power savings from the numerical algorithms with significantly lower time complexity.

7.2 Future Work

This work focuses on power consumption optimization for a wireless link and shows significant savings using the cognitive approach. SISO and MIMO systems are investigated. It is natural to extend the work to orthogonal frequency division multiplexing (OFDM) as MIMO and OFDM are two enabling technologies widely used in emerging communication standards.

In addition to the work on a single link, it is important to understand the network aspect where the individual links affect each other. It is observed that both SISO and MIMO systems can benefit in power efficiency by transmitting at high output power levels. This makes the research on potential interference and interference mitigation and management critical in a network where high output power level from each node is favorable for individual node. In a centralized network, scheduling might be used to alleviate the interference issue.

This work assumes linear PA and operation in linear region. If this assumption is relaxed by accepting certain controllable distortion by allowing the PA to run in nonlinear region, better efficiency might be achieved. However, the nonlinear operation introduces distortion which has to be carefully compensated. The tradeoff between allowable signal distortion and achievable power efficiency is an interesting topic to study.

All the above topics including this work focus on the gain, the power reduction. It is equally important to study the cost, the necessary additional computation burden to achieve the gain, e.g., the power consumption of the cognitive radio framework. Some interesting work has been developed on computation power consumption. It would be great if this kind of study can be integrated with the power reduction study. Several key topics in integration include computation power consumption models and computation complexity models for various algorithms and operations and their relationship.

7.3 Publications

This dissertation is a collection of selected journal publications on relevant research topics. This section summarizes all publications produced in the PhD study in a reversed chronological order.

7.3.1 Journal and Book Chapter Publication

1. A. He, S. Srikanteswara, K. K. Bae, J. H. Reed, and W. H. Tranter, Designing a Cognitive Engine Framework for Radio Power Consumption Optimization, to be submitted to a journal

2. A. He, A. Amanna, T. Tsou, X. Chen, D. Datla, T. R. Newman, S. M. Hasan, H. Volos, J. H. Reed, and T. Bose, Green Communications: A New Paradigm for Power Efficient Wireless Systems, submitted to Journal of Communications, Jan. 2011.

3. J. H. Reed, A. He, and K. K. Bae, Smart antennas, in Designing Software and Cognitive Radios, Wiley, forthcoming.

4. A. He, S. Srikanteswara, K. K. Bae, T. R. Newman, J. H. Reed, W. H. Tranter, M. Sajadieh, and

M. Verhelst, Power consumption minimization for MIMO systems - a cognitive radio approach, IEEE Journal on Selected Areas in Communications, vol. 29, no. 2, pp. 469-479, Feb. 2011.

5. A. He, S. Srikanteswara, K. K. Bae, J. H. Reed, and W. H. Tranter, Energy consumption minimization for mobile and wireless devices - a cognitive approach, IEEE Transactions on Consumer Electronics, vol. 56, no. 3, pp. 1814-1821, Aug. 2010.

6. A. He, K. K. Bae, T. R. Newman, J. Gaeddert, K. Kim, R. Menon, L. Morales, J. Neel, Y. Zhao, J. H. Reed, and W. H. Tranter, A survey of artificial intelligence for cognitive radios, IEEE Transactions on Vehicular Technology, vol. 59, no. 4, pp. 1578-1592, May 2010.

7. A. He, J. Gaeddert, K. Bae, T. Newman, J. H. Reed, L. Morales, and C. Park, Development of a case-based reasoning cognitive engine for IEEE 802.22WRAN applications, ACM Mobile Computing and Communications Review, vol. 13, no. 2, pp. 37-48, Apr. 2009

7.3.2 Conference Publication

1. A. He, S. Srikanteswara, K. Bae, T. R. Newman, J. H. Reed, W. H. Tranter, M. Sajadieh, and M. Verhelst, Power consumption minimization for MIMO systems using cognitive radio, in SDR'09, Washington D.C., Dec. 1-4, 2009.

2. A. He, S. Srikanteswara, K. Bae, T. R. Newman, J. H. Reed, W. H. Tranter, M. Sajadieh, and M. Verhelst, System power consumption minimization for multichannel communications using cognitive radio, in COMCAS 2009, Tel Aviv, Israel, Nov. 9-11, 2009.

3. A. He, S. Srikanteswara, J. H. Reed, X. Chen, W. H. Tranter, K. Bae, and M. Sajadieh, Minimizing energy consumption using cognitive radio, in IPCCC 2008, Austin TX, Dec. 7-9, 2008.

Appendix A

Development of a Case-Based Reasoning Cognitive Engine for IEEE 802.22 WRAN Applications

This chapter[1] discusses the design of a case-based reasoning cognitive engine for IEEE 802.22 WRAN applications in details.

(The following is the published journal paper on this subject.)

Development of a Case-Based Reasoning Cognitive Engine for IEEE 802.22 WRAN Applications

A.1 Abstract

On Nov. 4 2008, the Federal Communications Commission adopted rules for unlicensed use of television white spaces. The IEEE 802.22 Wireless Regional Area Networks (WRAN) standard is the first IEEE standard utilizing cognitive radio (CR) technology to exploit the television white space. A decision engine that is able to respond to the changes in the radio environment is necessary to efficiently exploit underutilized spectrum resources and avoid interfering with the licensed systems (e.g., TV services). This paper discusses the development of a case-based reasoning cognitive engine (CBR-CE) for the IEEE 802.22 WRAN applications. The performance of the CBR-CE is evaluated under various radio scenarios and compared to that of several multi objective search based algorithms, including the hill climbing search (HCS) and the genetic algorithm (GA). The simulation results show that the developed CBR-CE can achieve comparable utility with faster adaptation than the search based cognitive engines after appropriate training / learning. The learning process of the CBR is also simulated and discussed.

A.2 Introduction

The IEEE 802.22 Wireless Regional Area Networks (WRAN) standard is the first IEEE standard based on the cognitive radio (CR) technology to exploit the underutilized spectrum (frequency white space) allocated to the television (TV) broadcast services [57, 165]. Sharing spectrum resources can only be achieved by minimizing interference to such primary users (PU) as not to degrade their service quality. PUs for WRAN systems consist of TV stations, TV receivers, TV translators, TV boosters, and wireless microphones. The secondary user (SU), the unlicensed user of the licensed spectrum band, may include base stations (BS) and customer premise equipment (CPE) nodes. The IEEE 802.22 WRAN standard defines the operation of the SU in an unlicensed, fixed point-to-multipoint wireless communication manner. The operation of the SU in the TV band must abide by the IEEE 802.22 WRAN standard such that the SUs' operation causes minimal interference with the PUs' operation.

In correspondence with the IEEE's standardization process, the Federal Communications Commission (FCC) has finally adopted rules for unlicensed use of television white spaces on Nov. 4, 2008 [87]. The rules adopted allow various wireless devices, fixed and portable, to operate in the unused television spectrum with sufficient protection to the PUs, such as televisions and wireless microphones. In order to guarantee no harmful interference to PUs, the unlicensed devices have either to access a database of the registered PUs or to sense the radio environment.

In order to utilize the unused spectrum while protecting the PU service quality, the SU must be aware of and adapt to the PU activity as well as typical wireless channel impairments (e.g., multi-path fading). Much of this information can be obtained through the SU's sensing module and is used appropriately to choose available spectrum bands and radio configurations to accommodate specific application requests and quality of service (QoS) requirements. This process suggests the SU be cognitive of its environment and respond to such changes in accordance with the IEEE 802.22 standard and puts strict constraints on SU's performance and response time. In order to meet these requirements a cognitive engine (CE) that is able to manage a variety of scenarios is necessary.

This paper discusses the benefits of applying case-based reasoning (CBR), an artificial intelligence (AI) style approach, to CE design for IEEE 802.22 WRAN applications. In general, a major advantage of an AI approach is its flexibility. An AI approach can perform reasonably well under the circumstance where the system model is not well known or the objective functions may change. On the other hand, for a problem with specific requirements and a precise model, a heuristic approach might be more efficient. However, slight changes to the system model or the objective function might break it. It is desirable that a CR system is flexible to respond to various situations, even those that are not planned at design time. Therefore, an AI approach is preferable and provides a good tradeoff between performance and flexibility for the design of a generic system under the loosely defined environment. In this paper, we focus on one AI-style approach, the CBR.

CBR is an emulation of human learning, understanding, and problem solving process based on experience (cases) [139, 166]. In [124, 125], we have proposed a cognitive engine architecture that is able to incorporate both case- and knowledge-based reasoning and investigated the performance of a knowledge-based CE for WRAN applications. A case-based approach (case-based decision theory) is also proposed for CE design in [142]. Specifically, this paper extends the research in [124, 125] by developing a case-based reasoning cognitive engine (CBR-CE) for WRAN applications and evaluating its performance under a variety of scenarios. Both the training and evaluation processes of CBR-CE are discussed. The training and evaluation processes elaborate the learning and reasoning processes in the CBR-CE. As our results demonstrate, the CBR-CE provides a computationally efficient means for adapting to the dynamic wireless environment.

The remainder of this paper is organized as follows: Section A.3 discusses the architecture of a generic cognitive engine for the IEEE 802.22 WRAN applications along with recent research on CE design, the motivation for developing the CBR-CE, and performance metrics and utility functions; Section A.4 overviews the basic concepts of CBR; Section A.5 details the framework and the processing flowchart of the CBR-CE; Section A.6 presents the simulation results under a variety of scenarios and discusses the learning and reasoning process of the CBR-CE; Section A.7 concludes the paper and proposes some further work.

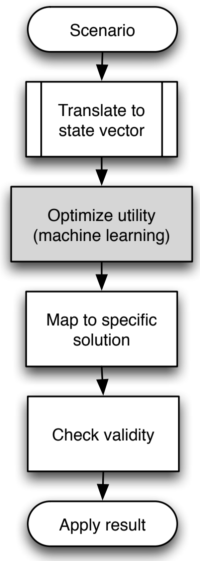

Figure A.1: Generic Architecture for the WRAN CE.

A.3 Design Methodology

A.3.1 Cognitive Engine Architecture

A generic process diagram of the WRAN CE is shown in Figure A.1. The CE consists of three major processes: orientation, reasoning, learning and decision-making, and solution mapping and validity check.

In the CE design for the WRAN applications, the "scenario" includes the following information:

- General spectrum condition: is affected by PU's activities and characterized by the spectrum availability. The spectrum availability specifies the TV channel status [57] (occupied, candidate, active, and null) due to the operation of the PU.
- Policy on spectrum (spectrum mask): specifies the maximum transmission power allowed in one or a series of adjacent TV channels [57].
- Specific radio environment between the CPE and the BS: specifies the radio channel condition such as path loss, fading and shadowing.
- Service request of CPE or BS: indicates the type of service and the associated QoS of CPE or BS.

The orientation process translates the scenario information into a service request vector and a policy vector. The service request vector includes the information on the specific radio environment and the service request. The policy vector includes the information on the general spectrum condition and other constraints. The reasoning, learning and decision-making process takes the service request vector as input and the policy vector as a constraint on solution and determines a possible solution. If the anticipated performance of this solution does not satisfy a certain criterion, the solution will be further adapted using adaptation mechanisms until the performance criterion is met or an adaptation time limit is reached. The solution mapping and validity check process maps the solution to a specific action in a vector of operational parameters and validates it against the policy and regulation before applied to the radio.

In this modular CE design all the processes in Figure A.1 can be implemented using various approaches as the application programming interfaces (API) between modules have been clearly defined. For example, a knowledge-based CE can be implemented [125]. In this paper, a CBR-CE is investigated.

A.3.2 Recent Research on Cognitive Engine Development

Various AI techniques have been investigated for CE design. The genetic algorithm (GA) has been applied to CE design recently. Rondeau et al. apply GA to control software defined radios [67]. The paper focuses on the adaptation mechanism of a CE that adjusts the radio parameters using GA in a changing radio environment. Newman et al. also implement a GA-based CE to control multicarrier transceivers [111]. Several fitness functions are defined to guide the GA search direction based on the numerical analysis on the relationships between environmental parameters and transmission parameters. The way to weigh different objectives and the trade-off between the convergence time of the GA and the size of the search space are also discussed.

Weingart et al. [167] apply Design of Experiments (DOE) technique to the CR systems to statistically identify the factor or combinations of factors (radio configuration and radio condition) that impact significantly the response (performance goals, such as bit error rate (BER), latency, and throughput) and generate a model for the system. The learned model is then used to generate a specific configuration for the CR according to the performance goals and environment condition. The learning of the model can be done either offline or in real-time. However, for large sets of factors and responses, it needs to be done offline at current computing capability. In contrast, the CBR approach, as we show in Section A.6.3, is fast and can run in real-time even though the parameter set is large.

Case- and knowledge-based reasoning has been proposed for CE design recently. Reed et al. [124, 125] propose a CE architecture that incorporates case- and knowledge-based reasoning and investigate the performance of a knowledge-based CE for WRAN applications. In addition, a radio environment map (REM) database is included to facilitate the reasoning and learning process. Le et al. also propose to use case-based reasoning in CE design [142]. The functionalities of the building blocks in the cognition cycle and the CE are discussed, including environmental awareness, case-based learning, multi-objective optimization, and hardware-portable interface.

A.3.3 Motivation of Case-Based Reasoning Cognitive Engine

As we have discussed in the previous section, interesting results have been demonstrated in designing CE using different AI techniques [125, 142, 67, 111]. However, as those researchers point out, there is much to be done. We have conducted a survey on the AI techniques to understand their applicability to CE design [168]. Six AI techniques are investigated, including case-based reasoning, knowledge-based reasoning, search-based reasoning, hidden Markov models, artificial neural networks, and cooperative reasoning. Among these techniques, case-based reasoning is suggested to be promising for the IEEE 802.22 WRAN applications in obtaining optimal radio configuration under uncertain radio environment and service requests [168].

CBR has been studied and applied in many fields [139, 166]. It is a reasoning process based on experiences (cases), which resembles human reasoning process to some extent. Generally speaking, there are two styles of the CBR depending on its purpose [139]:

- Problem solving CBR: suggests solutions and reminds possible problems in various tasks such as design, planning, and diagnosis.
- Interpretive CBR: justifies solutions in tasks such as situation classification, adversarial reasoning, and effect projection.

Among the advantages of using CBR, it can work with limited knowledge of the system and provide the solution quickly [139]. This makes CBR particularly appealing to the CE design for WRAN applications since the environment where a WRAN CE usually operates can hardly be predicted/understood precisely especially considering the existence of wireless microphones. Also, the WRAN CE needs to respond promptly to the changes in the radio environment and be flexible for various service requests [57, 165]. A problem solving case-based reasoning CE seems to be suitable for the IEEE 802.22 WRAN applications. It proposes solutions based on the cases (experiences) in the case database, revises the old solutions for the new situation, evaluates the performance before applying the solution in the real environment, collects and analyze the feedback information, and updates the case database based on the analysis.

In [124, 125, 142], case-based reasoning has been considered in the CE design. However, no tangible results (simulation or implementation) have been demonstrated. In this paper, we extend the work in [124, 125] to evaluate the applicability and feasibility of the CBR in CE development. The training and evaluation processes are discussed to elaborate the learning and reasoning processes in the CBR. In addition, we compare the performance and execution time tradeoff between CBR-CE and GA-based CE.

A.3.4 Performance Metrics and Utility Functions

Defining an appropriate utility function to evaluate the performance of a specific solution is crucial to the design of the CE. The result of this evaluation is used to qualify one solution over another, and is also used to compare the performance of CE implementations based on different AI techniques. This result can be used to select a specific implementation for certain application. In the CBR-CE design, we employ the utility function and utility metrics used in the knowledge-based CE [125]. The utility metrics considered include the BER, the data rate, and the transmission power.

The utility / performance of an individual CPE is defined as

$$u_{cpe} = f\left(\frac{1}{P_b}, \frac{1}{\dot{P}_b}\right)^2 f\left(R_b, \dot{R}_b,\right)^2 f\left(\frac{1}{P_t}, \frac{1}{\dot{P}_t}\right)^1 \tag{A.1}$$

where the utility function $f(\cdot)$ is defined as

$$f\left(x, \dot{x}; \eta, \sigma\right) = \frac{1}{2}\left\{\tanh\left[\log\left(\frac{x}{\dot{x}}\right) - \eta\right]\sigma + 1\right\}, \tag{A.2}$$

and $\dot{P}_b$, $\dot{R}_b$, and $\dot{P}_t$ are the target metrics for the BER, the data rate, and the transmission power (in linear scale) determined by the application requirements, respectively. The threshold and the spread parameters, η and σ, respectively, are chosen such that the utility is 0.95 when the metric (x) achieves the target ($\dot{x}$) and is 0.05 when the metric is one decade away from the target. As shown in Figure A.2, the utility function is monotonically increasing and is bounded by $0 < f < 1$. Furthermore, the utility exhibits diminishing returns for increasing the metric beyond its goal. The superscripts in (A.1) weigh the importance of achieving different goals. Since $0 < f < 1$, the larger the superscript of a particular goal, the higher the importance of

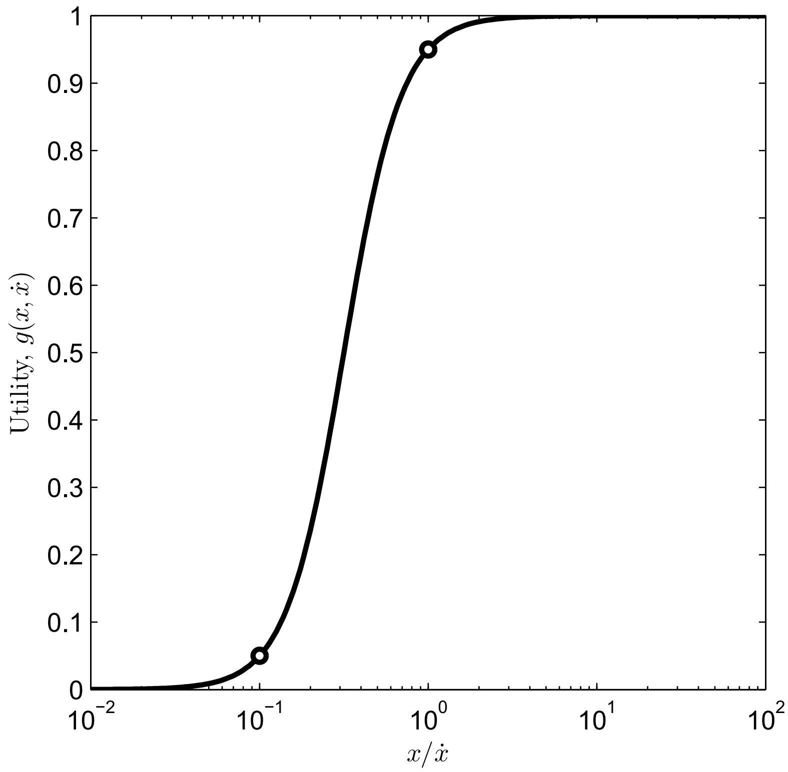

Figure A.2: Generic Utility Function.

that goal. There are other choices for the utility function. In (A.1), a multiplicative function instead of an additive function is chosen since it is more sensitive to a low value of any attribute. In (A.2), the slope of the curve is steeper when the attribute value is further away from the target value. Therefore, the combination of (A.1) and (A.2) produces significant difference between the utility value of a solution achieving all three goals and that of a solution achieving only two goals while at the same time restrains the utility value in a finite region of $(0, 1)$.

The global utility is then defined as:

$$u_{global} = \prod_{k} (u_k)^{\omega_k}, \tag{A.3}$$

where $\omega_k > 0$ is the weight assigned to the individual utility metric u_k.

A.4 Case-Based Reasoning Overview

The CBR is the core module of the CBR-CE. This section focuses on the design of the CBR module. Since a case is the fundamental element in a CBR system, its concept is discussed first. Then, the functionalities of different CBR components are explained in details.

A.4.1 Concept of Case

A case, the fundamental element in a CBR system, represents an experience. All other components in a CBR system take a case either as input or as output. A case is defined in [139] as follows.

> "A case is a contextualized piece of knowledge representing an experience that teaches a lesson fundamental to achieving the goals of the reasoner."

Note that although a case represents an experience, not every experience is a useful case. The essence of a case lies in its capability of "teaching a useful lesson" [139]. In other words, an experience that is worth remembering by the reasoner is what helps the reasoner to achieve a goal, warns the reasoner against a potential failure, or specifies an unexpected situation.

A.4.2 Components of CBR

Case-based reasoning refers to the reasoning process based on previous recorded experiences (cases). A case-based reasoner is an entity that performs case-based reasoning. In the discussion of this paper, CBR can refer to either case-based reasoning or case-based reasoner according to the context. In general, CBR consists of case representation and indexing, case selection and retrieval, case evaluation and adaptation, case learning and case database maintenance [139, 166]. In the following, we discuss these modules in details.

Case Representation and Indexing

Case representation formats the input information such that this information can be understood by other modules in the CBR. A case usually consists of two parts: the content and the indexes [139]. The content of a case records the experience or the lesson it teaches. It contains the following information.

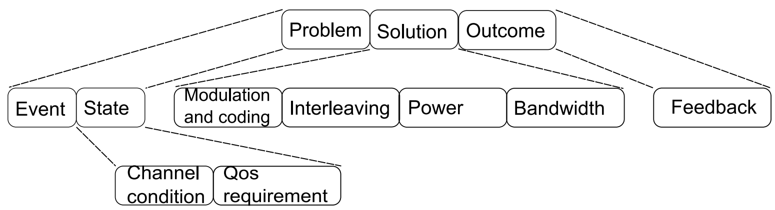

Figure A.3: Case structure.

- Problem description: describes the relevant experience. It specifies the detailed information about the problem including the radio environment and the service request with its QoS requirement.

- Solution: explains how the problem was solved in the past. It specifies a possible radio configuration for the problem specified in the problem description.

- Outcome: records the result of applying the solution. It specifies the feedback from the real environment (e.g., success or failure) after the solution was applied to the CPE.

The other important aspect of a case is its indexes, which specify the context where the content of a case is gained and where it is useful and describe the distinguishing features of a case.

To assist the readers, as an example, the case structure used in the CBR-CE design is shown in Figure A.3. The "Problem" field specifies the general event (e.g., new CPE service request, and PU detection), the channel condition (path loss), and the QoS requirement (BER). This field is used to form the index to a case. The "Solution" field includes information on modulation and coding scheme, interleaving, transmission power, and bandwidth use. These are the system parameters the CE can adjust. The "Outcome" field includes feedback information (achieved utility). It guides the case update in the case database. More details on the scenario settings and the range of attribute values are discussed in Section A.6.2.

Case Selection and Retrieval

The case selection and retrieval module searches the case database for cases that satisfy the request to a certain extent. In our CBR-CE design, for example, a policy guided retrieval process is developed. The policy vector obtained from the policy engine module contains a spectrum mask for a specific frequency band [57]. The cases are first retrieved from the case database according to their indexes. Then, the retrieved cases are checked against the policy vector. Those cases that comply with the policy vector will be collected for further processing. We call them valid cases. The retrieval criterion is the most important aspect in effective case retrieval [166]. General weighted sum based retrieval criteria do not work well in our particular environment due to the policy consideration in the WRAN system. For example, shorter Euclidean distance between two situation does not always guarantee a closer match of the two situations. This is why the policy guided retrieval is applied in the CBR-CE.

The case selection and retrieval module returns the best matched case or cases among the valid cases according to a selection criterion. The selection in our CBR-CE is based on the performance metric discussed in Section A.3.4. One or multiple valid cases with highest utility are returned. The choice of the selection criterion is an open question in many CBR applications. The utility metric we use is a choice with acceptable performance. Other metrics and utility functions such as those discussed in [111] can be adopted easily in this process in the modular design of the CBR. For example, it is easy to implement a case selection functionality with a metric on transmission power, such as selecting a case with the minimum transmission power.

Case Evaluation and Adaptation

The case evaluation module evaluates the performance of the retrieved case from the case selection and retrieval module either by using a performance model or by applying the solution and observing the outcome. The retrieved case is expected to be applicable to the new problem. However, the applicability or the performance is not guaranteed. The retrieved case is recorded as a solution to a previous problem and retrieved as a possible solution to the new problem due to the similarity between the old problem and the new one. If the performance of the retrieved case is not satisfactory, the case is modified by the case adaptation module. Note that the case evaluation and case adaptation can be recursive in order to obtain an appropriate solution. It increases the processing time. Therefore, it is employed only if it is needed.

Case Learning and Case Library / Database Maintenance

A case library is an ensemble of similar cases. Several case libraries can be organized in a hierarchical structure into a case database for efficient retrieval.

Experiences are remembered by the CBR system as cases in the case database. The CBR gains additional information, or learns, by solving new problems or receiving feedback. As the experience increases, more cases are accumulated in the case database. New experience may also be incorporated in the case database by updating the existing cases. In this way, the size of the case database will not increase linearly with the number of new problems and may finally converge. This phenomenon is observed in the simulation results in Section A.6.3. Although a larger case database does not always guarantee better performance for a specific problem, a CBR system can generally return a better solution with a larger case database than with a smaller one. This performance improvement reflects the learning capability of a CBR system. We discuss the learning process in details with the simulation results.

The case database maintenance module is critical to the learning process and reasoning process of the CBR-CE. When a solution to a new problem is formulated, this solution needs to be "remembered" by the CBR for future use. The outcome of applying this solution to the new problem should be "remembered", too. As more and more cases are recorded in the case database, the CBR gets more and more efficient in the sense that it can utilize cases in the case database to the new problem instead of finding a solution from scratch using computationally intensive optimization. Another important consideration in case database maintenance is the balance between the number of cases in the case database and the amount of time used in case retrieval. A larger case database covers a larger problem space. However, this usually also means a longer case retrieval time and larger memory space. Although memory size is becoming less of a concern in modern system design due to advancements in memory size, processing time is always a big concern in many wireless applications. For example, the IEEE 802.22 standard specifies stringent time requirements upon detection of the primary user (e.g., Channel Move Time (2 s) and Channel Closing Transmission Time (0.1 s)) [165] to protect the PU. Therefore, an appropriate case granularity needs to be defined and redundant cases removed from the case database to improve execution time.

A.5 Case-Based Reasoning Cognitive Engine Framework

The framework of the CBR-CE for the IEEE 802.22 WRAN applications is shown in Figure A.4. Also shown in Figure A.4 is a generic processing flowchart of the CE indicated by arrows with number markups.

Since the CBR-CE is designed in modules with well-defined interfaces, each functional block can be replaced with an equivalent processing element with ease. This modular design makes the framework handy for testing and evaluating different algorithms.

A.5.1 CE Module Functionalities

The functionalities of the modules of the cognitive engine framework are as follows.

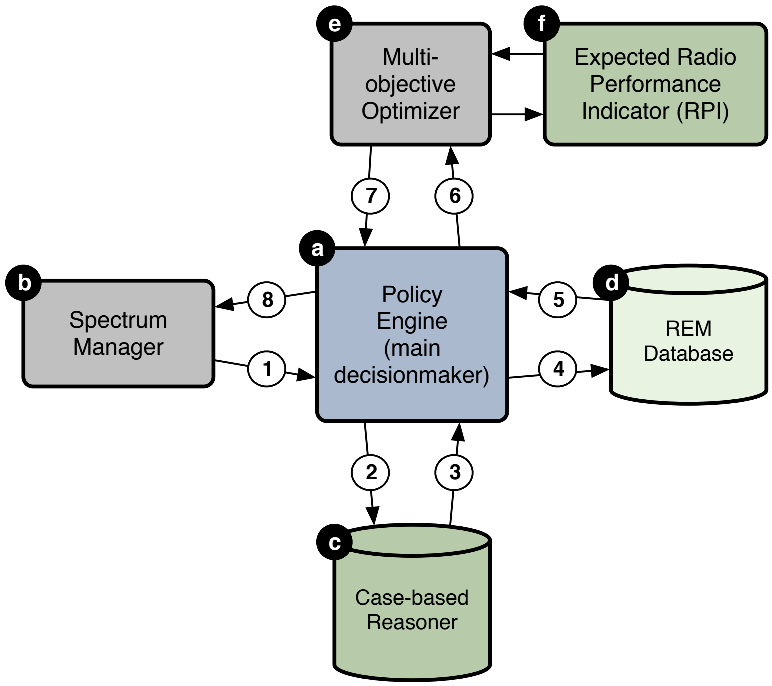

Figure A.4: Framework of the CBR-CE for the WRAN application.

Spectrum Manager (SM)

The SM monitors the radio environment and interfaces with the physical radio hardware. It abstracts the information from the spectrum sensing module for radio environment status and exchanges information with the radio environment map module. In addition, it allocates resources including channels and subcarriers according to the solution from the multi objective optimizer module.

Policy Engine (PE)

The PE module specifies the general policies, including the standard, the regulation, and the customized specification. This general policy information is used to guide the operation of the case-based reasoning module. The PE checks the legitimacy of the solutions returned by the case-based reasoning module and multi objective optimizer module.

Case-Based Reasoner (CBR)

The CBR module is the core module of the CBR-CE. It is responsible for providing candidate solutions based on the request from the PE module. As we have discussed in details in Section A.4, the building blocks of the CBR include case representation and indexing, case selection and retrieval, case evaluation and adaptation, and case learning and case library/database maintenance.

Radio Environment Map (REM) Database

The REM database is used to store scenario specific parameters about the system, such as geographic features, network and service availability, spectrum statistics, radio location and activities, policies, and experiences [169]. The FCC suggests a database to store the incumbent user service for the unlicensed user [87]. This database can be incorporated into the REM database.

Multi Objective Optimizer (MOO)

The MOO adapts the solution returned by the CBR to satisfy specific QoS requirement of the new problem. The algorithms investigated in this paper include:

- Hill climbing search (HCS): sequentially makes small changes to the solution and improves the solution a little bit each time.
- Genetic algorithm (GA) search: uses evolutionary search methods to find a better solution.
- Combination of HCS and GA: uses HCS and GA sequentially to find a better solution.

The returned solution from the CBR provides a good starting point for the MOO algorithms. Note that this case adaptation process (MOO module) can be bypassed if the returned solution provides satisfactory performance to reduce the execution time. This is observed in our simulation results and is the reason why an independent MOO module is designed in the CBR-CE framework in Figure A.4. In addition, other adaptation algorithms can be added to the MOO easily. The algorithms can also be combined to achieve a better balance between performance and execution time.

Expected Radio Performance Indicator (RPI)

The RPI module evaluates the anticipated performance for a given solution with the utility function before applying it in the real environment. The result is used in the MOO to determine whether certain QoS requirement is met and whether further optimization is necessary.

A.5.2 Functional Processing Flow

Enumerated below is the functional description of the processing flow of the WRAN CE described in Figure A.4. Each process is indicated by a functional example of the information passed through the interface:

1. Event triggered: Primary user is detected on channel #29; channel #29 is now occupied; CPE #51 working on this channel needs to be evacuated. SM passes a CPE evacuation request to PE. This request includes the information on the event (CPE evacuation) and radio environment information.

2. Query CBR for abstract solution: PE queries CBR for solution/action that should be applied to CPE #51 with specific information on the event and the CPE.

3. Return solution: CBR employs the policy guided retrieval process as discussed in Section A.4.2 to identify valid cases in the case database and returns a best matched case in the case structure defined in Figure A.3 containing specific radio configuration in the "solution" field.

4. Query REM for specific spectrum / channel: PE queries REM for best candidate channel.

5. Return solution: REM informs PE the best candidate channel, Channel #32.

6. Execute MOO: PE sends the solution with specific channel and radio configuration to MOO. MOO evaluates the expected performance of this solution using RPI. If the anticipated performance is better than a threshold, the optimization at MOO stops. Otherwise, MOO optimizes the solution and calculates the anticipated performance until the threshold is met or the optimization times out.

7. Return specific solution: MOO returns the optimized solution with anticipated performance to PE. PE checks the legitimacy of the solution. CPE #51 should use QPSK with rate 1/2 convolutional code in channel #32 with bandwidth 1MHz; it should give a utility of 0.92. The solution complies with the policy. Note if MOO returns an illegitimate solution, PE informs MOO to run again if possible. Otherwise, a failure is declared to SM.

8. Return specific solution to SM: PE returns final solution to SM.

9. Apply solution to the radio, measure actual performance, and update databases (not shown in Figure A.4): SM reconfigures the CPE to work on channel #32 with modulation, coding, and bandwidth specified in the returned solution from PE. Actual utility is measured and returned as feedback. This feedback can be used to update the case database through the case database maintenance as in Section A.4.2 and the REM database. Actual utility for CPE #51 in channel #32 is 0.83, and the case database and the REM database are updated.

Note that some processes, such as the MOO and the RPI, might run recursively in the processing flow in order to obtain an acceptable solution.

A.6 Performance Evaluation and Simulation Results

A.6.1 Testing Methodology

The testing and the evaluation of the cognitive engine for WRAN applications are challenging due to demanding and unpredictable environments. In this paper, we apply a scenario driven testing methodology proposed in [125] where the CE is tested in various scenarios depicted in a series of extensible markup language (XML) files. Since the GA based CE (GA-CE) provided a consistent performance close to that provided by the exhaustive search based CE and requires much less time to operate than the exhaustive search based CE [125], in this paper the GA-CE is used as the bench mark for performance comparison. The performance of the CBR-CE, the GA-CE, and the HCS based CE (HCS-CE) are investigated under various radio environments. Note that since the CBR is based on experience, a training process or an initial learning process is crucial to its performance. This training process is investigated through simulation.

A.6.2 Training and Evaluation Scenarios

In the training and the evaluation of the CBR-CE, the radio links between the BS and CPEs are randomly generated by Monte Carlo simulations. Here, the training process mainly focuses on the operation of the CBR in circumstance where little experience has been accumulated or few cases have been added to the case database. The evaluation process, on the other hand, investigates the operation of the CBR when sufficient

Table A.1: Simulation Settings

Parameter	**Value or Range**
Number of BSs	1
Cell radius	20 km
Number of new connections	1 - 35
Distribution of CPEs	Uniformly distributed
Type of service, $\dot{R}_b$ and $\dot{P}_b$	Voice: 10kbps, 10^{-2} Video: 100kbps, 10^{-3} Low speed data: 250kbps, 10^{-6} High speed data: 750kbps, 10^{-6}
Multiplexing/Duplex	OFDMA/TDD
Number of subcarriers	2048
Maximal number of active channels supported by the BS	1

Table A.2: Adjustable Parameters in the WRAN CE.

Parameter	Value or Range
Frequency channel	VHF / UHF (54 - 862 MHz)
CPE and BS transmission power	Up to 4 Watts EIRP, subject to EIRP profile defined by 802.22 Standard
Modulation scheme	QPSK, 16QAM, 64QAM
Channel coding (convolutional coding)	1/2, 2/3, 3/4
Number of uplink and downlink subcarriers	Variable 4 - 256

cases have been collected in the case database. In other words, the main difference between the training process and the evaluation process is the sufficiency of the cases in the case database. Note that although learning exists in both processes, it is more clearly observed in the training process than in the evaluation process as we see in the simulation results. Therefore, we use the training process to explain the learning process of the CBR.

The CBR-CE is developed in C++ and run on a personal computer with 3 GHz CPU clock rate and 2 GB RAM. The detailed simulation settings in a single cell system are in Table A.1.

The simulation evaluates the CE capability to perform dynamic spectrum access and optimize spectrum and power usage. Various operational parameters for both BS and CPE can be configured and optimized by the WRAN CE. These adjustable parameters are listed in Table A.2. Although the BS can only support one active channel at a time [57], it is assumed to be able to operate in the frequency range 54 - 862 MHz.

The training and the evaluation scenarios for the CBR-CE in the simulation are CPE service requests, i.e., N new CPE nodes request service to a BS in a cell simultaneously. This scenario is of particular interest for two reasons. First, it is a common scenario in WRAN applications. Second, in the scenario where PU is detected and some CPEs need to be evacuated, the CPE evacuation can be looked as CPE service request

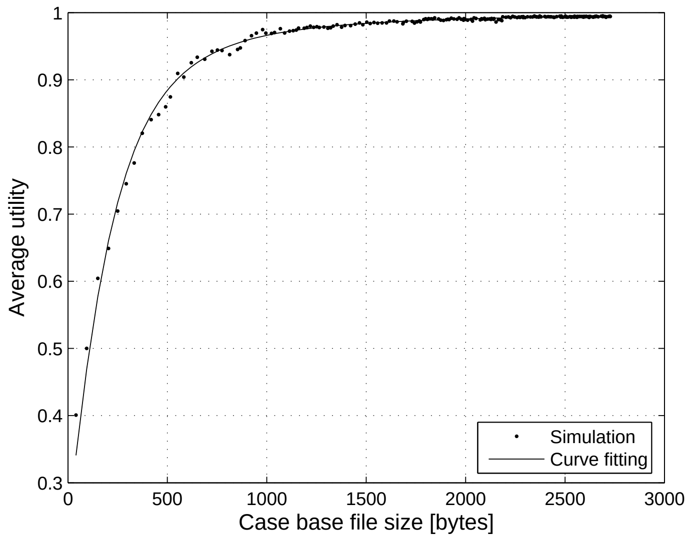

Figure A.5: CBR-CE utility vs. case database size.

with higher priority with regard to other CPE service request.

A.6.3 Training / Learning Results

Figure A.5 and Figure A.6 show the training / learning process of the CBR-CE and the impact of the case database size on the performance and the execution time of the CBR-CE. No restriction is enforced on the size of the case database in the simulation. The trailing / learning process of the CBR starts with an empty case database. As the CBR experiences new scenarios, good results are recorded in the case database. Therefore, the size of the case database increases as the CBR encounters more scenarios. The cases in the case database are continuously updated when better solutions are obtained, i.e., a case with higher utility. Two kinds of update are considered in the simulation.

- Add a new case: When the CBR encounters a new scenario or a scenario that is remotely similar to previous one, the CBR retrieves a related case and adapts the case to the new scenario. This solution is then added to the case database as a new case.
- Adapt an existing case: When the CBR encounters a reoccurred scenario or a scenario that is very similar to previous one, the CBR retrieves a closely matched case and optimizes this solution to obtained a better solution. The existing case in the case database is then replaced by the new

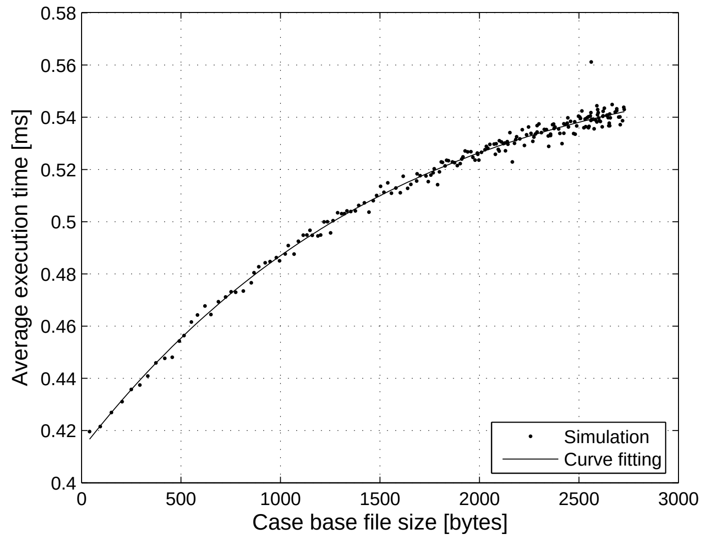

Figure A.6: CBR-CE execution time vs. case data size.

solution.

The increase in the size of the case database indicates the accumulation of the experience, or learning, of the CBR as it experiences more and more scenarios (as well as time elapses). The improvement in the performance of the CBR as the increase in the size of the case database shows that the CBR works better with more experience as in Figure A.5. Note that, in Figure A.5, the performance improves very fast when the size of the case database is relatively small and the performance is poor. This performance improvement tends to slow down as the size of the case database becomes larger and the performance becomes better. This phenomenon can be better understood with a comparison to the human learning process. For instance, imagine a process where we learn to play tennis or any other sport. We can usually learn to start playing tennis in a relatively short amount of time. We spend more time in practice then we improve our skills. However, it will take much longer, if not forever, for us to be ATP champions. The result in Figure A.5 indicates the overtraining issue with CBR. In other words, excess training only gets marginal return in terms of performance improvement (average utility in particular) in CBR. In our simulation, no performance degradation has been observed using CBR under excess training. This might be due to the capability of CBR to update its database using better solutions. Our design benefits from this observation such that a very limited amount of training and database size can provide satisfactory performance. The small training set and the small database result in fast execution result in Figure A.6.

Figure A.6 suggests that the execution time of the CBR increases almost linearly with the size of the case database. This is because the retrieval algorithm used in CBR needs to search through all the cases in the case database for the solution. As the number of cases increases, the searching time increases as well. Other algorithms [166] can be used in case retrieval to improve the time performance, but this straightforward approach provides a useful baseline. Note that the size of the case database depends on the complexity of the problem and the granularity of the case. In our CBR-CE, the performance is improved at the expense of an almost linear increase in execution time (as observed in Figure A.6) as the size of the CBR case database increases. Note that the absolute value of the execution time varies with the platform where this simulation is run. However, what is more interesting to note is the nearly linear time complexity. As we see in the performance comparison in the following section, this is a significant benefit of applying CBR-CE in time stringent applications such as the IEEE 802.22 applications compared to other AI techniques.

The saturation in performance observed in Figure A.5 and the linear increase in execution time observed in Figure A.6 indicate that a limited number of cases (about 150 cases in our simulation with a case database size of 3kB) is sufficient to provide a good balance between performance and execution time. This number will be different for different applications and needs to be determined through experience.

One common issue with CBR and other AI techniques that need training is that these AI systems might need to be retrained under certain conditions. In CBR system, this issue is incorporated into the learning process. As we see in the evaluation process where the CBR is initially trained, the case database is still updated when a better solution or a new solution is obtained. In other words, the training / learning process and the evaluation process of the CBR cannot be clearly separated from each other. The training / learning process is an evaluation process with little experience; and the evaluation process is simply a training / learning process with great experience. In the mean time, we can imagine that even the CBR is well trained for a certain environment, it might need to be retrained when it operates in a significantly different environment where all the experience becomes meaningless. Should that happen, the CBR would pick up new experience as it operates in the new environment.

A.6.4 Evaluation Results

Several CEs are developed for evaluation purposes including a variety of combinations of algorithms. The CEs tested in this paper are tabulated in Table A.3. As we see in the simulation results shown in Figure A.7 and Figure A.8, different CEs have different performance and execution time. The results are obtained

Table A.3: List of evaluated CEs.

Engine name	Description
GA	GA with 1024 generations
HCS	HCS with 4 iterations
CBR	CBR without further optimization
CBR+HCS	CBR with single iteration HCS

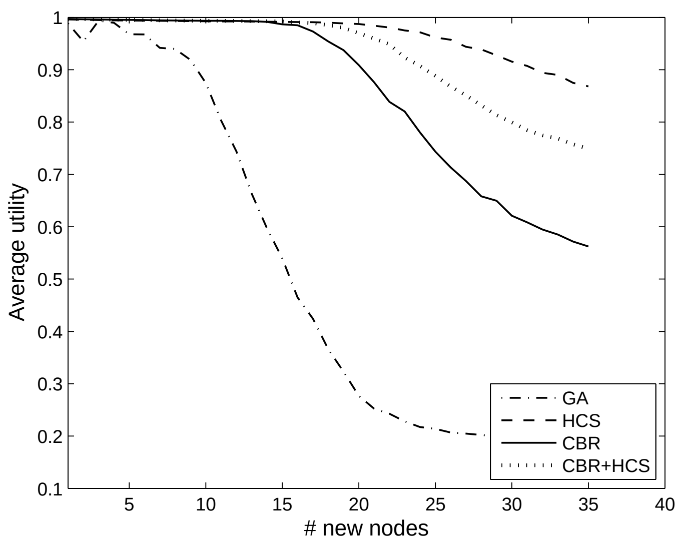

Figure A.7: Utility results for various CEs.

upon 1000 trials for each scenario (random scenarios are generated for each trial).

In Figure A.7, all engines except the GA-CE provide reliable solutions for $N \leq 15$, N the number of CPEs connected to the base station. The performance starts to degrade after $N \geq 15$ ($N \geq 10$ for GA-CE) in terms of average utility. In other words, the CEs can successfully handle complicated scenarios, such as simultaneous new service requests. It is interesting to note that the HCS-CE performs consistently well in the simulation. The HCS algorithm predictably adjusts each parameter and it can guarantee finding the local maximum of the solution space. The results above suggest that the solution space for this particular simulation either has few if any local maximum, or has equivalent local maximums, thus allowing the HCS to reliably find a good solution.

The execution time of various CEs is compared in Figure A.8. Observe that the CBR-CE executes much faster than other CEs as expected. Note that the slope of the CBR-CE execution time vs. problem complexity

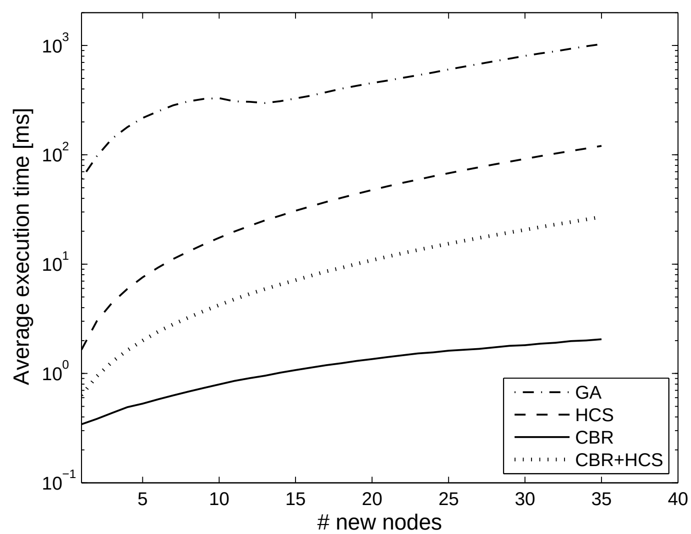

Figure A.8: Execution time results for various CEs.

(the number of nodes requests service simultaneously) is not as steep as other CEs. This renders the CBR-CE a desirable feature in solving complex problems where the response time is crucial (e.g., IEEE 802.22 WRAN applications) since its execution time increases slower than that of other CEs. Surprisingly, for most practical scenarios ($N \leq 10$) the solution from CBR-CE needs no further refinement from the MOO if an appropriately trained case database is used. This provides a good justification of separating the MOO module in the CBR-CE framework shown in Figure A.4 so that the MOO can be easily bypassed to reduce execution time further when the performance is acceptable. On the other hand, due to the nature of the GA and the chosen utility function, after a certain point only a few parameters significantly affect the total utility. Because the GA is a stochastic search algorithm, the probability that the crossover or mutation affects these particular parameters is low, especially for complex scenarios (e.g., $N \geq 10$). Therefore, a significant portion of the search time is wasted on regions of the search space that do not affect the utility in a positive manner.

As seen in Figure A.7 and Figure A.8, a tradeoff between CE performance in terms of the average utility and the average execution time needs to be determined. For the IEEE 802.22 WRAN applications where the response time of the CE is critical, a CBR-CE with or without further multi objective optimization provides a sound balance between the average utility and the average execution time.

A.7 Conclusions

In this paper, a case-based reasoning cognitive engine for IEEE 802.22 WRAN applications is presented. Simulation results show that the CBR-CE can achieve comparable utility performance and fast adaptation with appropriate training. Further, the size of the case database is small for a moderate problem, which leads to significantly faster processing. In addition, the case adaption process can be bypassed to further reduce execution time given a properly trained case database. The fast execution characteristic makes CBR-CE a good choice for time stringent applications such as IEEE 802.22 WRAN applications. The simulation results indicate that the CBR-CE can be used with little experience. This means that the CBR-CE can be applied to the system that needs to be deployed in unknown or unfamiliar environment and the system performance can improve gradually as it operates. This is useful for IEEE 802.22 WRAN applications since rural environment may not have detailed radio environment information. In addition, the architecture of the CBR-CE is generic and the CBR-CE can be easily adapted to other applications.

Although results on the proposed cognitive engine are promising, there are several open questions that we would like to address in our future work. For example,

- What are other possible indexing schemes and what is the performance of the CBR-CE with other indexing schemes?
- What are other possible selection criteria in CBR and what is the performance of the CBR-CE with other criteria?
- What is the appropriate case granularity for this application and how can it be determined for other applications?

The CBR-CE is designed in C++ in anticipation of its real-time implementation with OSSIE, an open source software defined radio framework developed in our group [99].

Bibliography

[1] S. K. Lai, R. S. Cheng, K. B. Letaief, and R. D. Murch. Adaptive trellis coded MQAM and power optimization for OFDM transmission. In *1999-Spring IEEE 49th Vehicular Technology Conference (VTC1999-Spring)*, pages 290–294, Tokyo, Japan, May 16-20 1999.

[2] X. Duan, Z. Niu, and J. Zheng. Downlink transmit power minimization in power-controlled multimedia CDMA systems. In *2002 IEEE 13th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2002)*, pages 1102–1106, Lisboa, Portugal, September 15-18 2002.

[3] Z. Han and K. J. R. Liu. Power minimization under throughput management over wireless networks with antenna diversity. *IEEE Trans. Wireless Commun.*, 3(10):2170–2181, November 2004.

[4] A. Pandharipande. Adaptive modulation for MIMO-OFDM systems. In *2004-Fall IEEE 60th Vehicular Technology Conference (VTC2004-Fall)*, pages 1266–1270, Los Angeles, CA, September 26-29 2004.

[5] T. Mo and C. W. Bostian. A throughput optimization and transmitter power saving algorithm for IEEE 802.11b links. In *2005 IEEE Wireless Communications and Networking Conference (WCNC 2005)*, pages 57–62, New Orleans, LA, March 13-17 2005.

[6] Z. Wang, C. He, and A. He. Robust AM-MIMO based on minimized transmission power. *IEEE Commun. Lett.*, 10(6):432–434, June 2006.

[7] X. Wang, A. G. Marques, and G. B. Giannakis. Power-efficient resource allocation and quantization for TDMA using adaptive transmission and limited-rate feedback. *IEEE Trans. Signal Processing*, 56(9):4470 – 4485, September 2008.

[8] C. Schurgers, O. Aberthorne, and M. B. Srivastava. Modulation scaling for energy aware communication systems. In *2001 International Symposium on Low Power Electronics and Design (ISLPED 2001)*, pages 96–99, Huntington Beach, CA, August 6-7 2001.

[9] R. Min and A. Chandrakasan. A framework for energy-scalable communication in high-density wireless networks. In *2002 International Symposium on Low Power Electronics and Design (ISLPED 2002)*, pages 36–41, Monterey, CA, August 12-14 2002.

[10] S. Cui, A. J. Goldsmith, and A. Bahai. Modulation optimization under energy constraints. In *2003 IEEE International Conference on Communications (ICC 2003)*, pages 2805–2811, Anchorage, AK, 11-15 May 2003.

[11] X. Lu, E. Erkip, Y. Wang, and D. Goodman. Power efficient multimedia communication over wireless channels. *IEEE J. Select. Areas Commun.*, 21(10):1738–1751, December 2003.

[12] C. Schurgers and M. B. Srivastava. Energy optimal scheduling under average throughput constraint. In *2003 IEEE International Conference on Communications (ICC 2003)*, pages 1648–1652, Anchorage, AK, 11-15 May 2003.

[13] S. Cui, A. J. Goldsmith, and A. Bahai. Energy-constrained modulation optimization. *IEEE Trans. Wireless Commun.*, 4(5):2349 – 2360, September 2005.

[14] S. Cui, A. J. Goldsmith, and A. Bahai. Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks. *IEEE J. Select. Areas Commun.*, 22(6):1089–1098, August 2004.

[15] S. Gao, L. Qian, D. R. Vaman, and Q. Qu. Energy efficient adaptive modulation in wireless cognitive radio sensor networks. In *2007 IEEE International Conference on Communications (ICC 2007)*, pages 3980–3986, Glasgow, Scotland, June 24-28 2007.

[16] T. Wang, W. Heinzelman, and A. Seyedi. Minimization of transceiver energy consumption in wireless sensor networks with AWGN channels. In *2008 46th Allerton conference on Communication, Control, and Computing*, Monticello, IL, September 23-25 2008.

[17] Q. Chen and M. C. Gursoy. Energy-efficient modulation design for reliable communication in wireless networks. In *2009 43rd Annual Conference on Information Sciences and Systems (CISS 2009)*, pages 811–816, Baltimore, MA, March 18-20 2009.

[18] H. S. Kim and B. Daneshrad. Energy-constrained link adaptation for MIMO OFDM wireless communication systems. *IEEE Trans. Wireless Commun.*, 9(9):2820–2832, September 2010.

[19] B. Niu, O. Simeone, O. Somekh, and A. M. Haimovich. Ergodic and outage sum-rate of fading broadcast channels with 1-bit feedback. *IEEE Trans. Veh. Technol.*, 59(3):1282–1293, March 2010.

[20] A. A. M. Saleh and D. C. Cox. Improving the power-added efficiency of FET amplifiers operating with varying-envelope signals. *IEEE Trans. Microwave Theory Tech.*, 31(1):51–56, January 1983.

[21] G. Hanington, P-F. Chen, P. M. Asbeck, and L. E. Larson. High-efficiency power amplifier using dynamic power-supply voltage for CDMA applications. *IEEE Trans. Microwave Theory Tech.*, 47(8):1471–1476, August 1999.

[22] J. Deng, P. S. Gudem, L. E. Larson, D. F. Kimball, and P. M. Asbeck. A SiGe PA with dual dynamic bias control and memoryless digital predistortion for WCDMA handset applications. *IEEE J. Solid-State Circuits*, 41(5):1210–1221, May 2006.

[23] J. Mitola and G. Q. Maguire. Cognitive radio: making software radios more personal. *IEEE Personal Commun. Mag.*, 6(4):13–18, March 1999.

[24] S. Haykin. Cognitive radio: brain-empowered wireless communications. *IEEE J. Select. Areas Commun.*, 23(2):201–220, February 2005.

[25] I. F. Akyildiz, W-Y. Lee, M. C. Vuran, and S. Mohanty. Next genereation/dynamic spectrum access/cognitive radio wireless networks: A survey. *Elsevier Computer Networks Journal*, 50(13):2127–2159, September 2006.

[26] Q. Zhao and B. M. Sadler. A survey of dynamic spectrum access. *IEEE Signal Processing Mag.*, 24(3):79–89, May 2007.

[27] L. Berlemann and S. Mangold. *Cognitive Radio and Dynamic Spectrum Access*. Wiley, New York, NY, 2009.

[28] E. Hossain, D. Niyato, and Z. Han. *Dynamic Spectrum Access and Management in Cognitive Radio Networks*. Cambridge University Press, New York, NY, 2009.

[29] G. Cafaro, T. Gradishar, J. Heck, S. Machan, G. Nagaraj, S. Olson, B. Stengel R. Salvi, and B. Ziemer. Cellular phones as embedded systems. In *2007 IEEE Radio Frequency Integrated Circuits Symposium (RFIC 2007)*, pages 189–192, Honolulu, HI, June 3-5 2007.

[30] Y. Neuvo. Cellular phones as embedded systems. In *2004 IEEE International Solid-State Circuits Conference (ISSCC 2004)*, pages 32–37, San Francisco, CA, February 15-19 2004.

[31] O. Silven and K. Jyrkka. Observation on power-efficiency trends in mobile communication devices. *EURASIP Journal on Embedded Systems*, 2007(1), January 2007.

[32] Skyworks Solutions Inc. Power amplifier for wireless applications. `www.skyworksinc.com`.

[33] J. F. Sevic. Statistical characterization of RF power amplifier efficiency for CDMA wireless communication systems. In *1997 Wireless Communications Conference*, pages 110–113, Boulder, CO, USA, August 11-13 1997.

[34] J. Hayes. Adaptive feedback communications. *IEEE Trans. Commun. Technol.*, 1(16):29–34, February 1968.

[35] J. Cavers. Variable-rate transmission for Rayleigh fading channels. *IEEE Trans. Commun.*, 1(20):15–22, February 1972.

[36] B. Vucetic. An adaptive coding scheme for time-varying channels. *IEEE Trans. Commun.*, 39(5):653–663, May 1991.

[37] W. T. Webb and R. Steele. Variable rate QAM for mobile radio. *IEEE Trans. Commun.*, 43(7):2223–2230, July 1995.

[38] S. M. Alamouti and S. Kallel. Adaptive trellis-coded multiple-phase-shift keying for Rayleigh fading channels. *IEEE Trans. Commun.*, 42(6):2305–2314, June 1994.

[39] A. J. Goldsmith and S. Chua. Variable-rate variable-power MQAM for fading channels. *IEEE Trans. Commun.*, 45(10):1218–1230, October 1997.

[40] S. M. Alouini and A. J. Goldsmith. Adaptive modulation over Nakagami fading channels. *Springer Wirel. Pers. Commun.*, 13(1-2):119–143, May 2000.

[41] P. Ormeci, X. Liu, D. L. Goeckel, and R. D. Wesel. Adaptive bit-interleaved coded modulation. *IEEE Trans. Commun.*, 49(9):1572–1581, September 2001.

[42] L. Lin, R. D. Yates, and P. Spasojevic. Adaptive transmission with discrete code rates and power levels. *IEEE Trans. Commun.*, 51(12):2115–2125, December 2003.

[43] J. Block and M. Pursley. A protocol for adaptive transmission in direct-sequence spread-spectrum packet radio networks. *Communications, IEEE Transactions on*, 52(8):1388–1396, August 2004.

[44] T. Ji and W. Stark. Rate-adaptive transmission over correlated fading channels. *IEEE Trans. Commun.*, 53(10):1663–1670, October 2005.

[45] L. Lin, R. D. Yates, and P. Spasojevic. Adaptive transmission with finite code rates. *IEEE Trans. Inform. Theory*, 52(5):1847–1860, May 2006.

[46] R. Liu, J. Luo, and P. Spasojevic. Adaptive transmission with variable-rate Turbo bit-interleaved coded modulation. *IEEE Trans. Wireless Commun.*, 6(11):3926–3936, November 2007.

[47] M. B. Pursley and T. C. Royster. Low-complexity adaptive transmission for cognitive radios in dynamic spectrum access networks. *IEEE J. Select. Areas Commun.*, 26(1):83–94, January 2008.

[48] M. S. Alouini and A. J. Goldsmith. Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques. *IEEE Trans. Veh. Technol.*, 48(4):1165–1181, July 1999.

[49] S. Catreux, V. Erceg, D. Gesbert, and Jr. Heath, R. W. Adaptive modulation and MIMO coding for broadband wireless data networks. *IEEE Commun. Mag.*, 40(6):108–115, June 2002.

[50] R. K. Mallik, M. Z. Win, J. W. Shao, M. S. Alouini, and A. J. Goldsmith. Channel capacity of adaptive transmission with maximal ratio combining in correlated Rayleigh fading. *IEEE Trans. Wireless Commun.*, 3(4):1536–1276, July 2004.

[51] Z. Zhou, B. Vucetic, M. Dohler, and Y. Li. MIMO systems with adaptive modulation. *IEEE Trans. Veh. Technol.*, 54(5):1828–1842, September 2005.

[52] FCC. Notice of proposed rule making and order. ET Docket no. 03-222, December 2003.

[53] SDR Forum. Cognitive Radio Work Group. http://www.sdrforum.org/pages/aboutTheForum/currentProjects.asp.

[54] M. A. McHenry. NSF spectrum occupancy measurements projects summary. Technical report, Shared Spectrum Company, August 2005.

[55] IEEE. IEEE 802.11 working group on wireless local area networks. http://www.ieee802.org/11/.

[56] IEEE. IEEE 802.16 working group on broadband wireless access standards. http://www.ieee802.org/16/.

[57] IEEE. IEEE 802.22 working group on wireless regional area networks. http://www.ieee802.org/22/.

[58] J. H. Reed. An introduction to cognitive radio and some research trends in cognitive radio. In *2005 Wireless Symposium at Virginia Tech*, Blacksburg, VA, June 2005.

[59] W. Lane. Topic 9: Cognitive radio potential for public safety. http://www.fcc.gov/pshs/techtopics/techtopic9.html.

[60] R. Thomas. Cognitive networks. Phd dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA, June 2007.

[61] D. Niyato and E. Hossain. Spectrum trading in cognitive radio networks: A market-equilibrium-based approach. *IEEE Wireless Commun.*, 15(6):71–80, December 2008.

[62] Ubiquisys. Ubiquisys and percello unveil next generation femtocell platform. http://www.ubiquisys.com/ub3b/pressreleases.php?id=115, June 22 2009.

[63] C. Lu, F. H. P. Fitzek, and P. C. F. Eggers. A cooperative scheme enabling spatial reuse in wireless networks. *Springer Wireless Commun. Mag.*, 15(6):457–471, September 2007.

[64] P. Greenis. Smart communications for smart grids. In *2009 Nanotech Conference and Expo*, Houston, TX, May 3-7 2009.

[65] J. Neel and A. Amanna. Introduction to cognitive radio technology for transportation systems. In *2009 SDR Forum Workshop on Smart Communications in Transportation Systems (SDR 2009)*, Houston, TX, June 18 2009.

[66] J. Neel. Synthetic symmetry in cognitive radio networks. In *2007 SDR Forum Technical Conference and Product Exposition (SDR 2007)*, Denver, Co, November 5-9 2007.

[67] T. W. Rondeau, B. Le, C. J. Rieser, and C. W. Bostian. Cognitive radios with genetic algorithms: Intelligent control of software defined radios. In *2004 SDR Forum Technical Conference and Product Exposition (SDR 2004)*, pages C–3–C–8, Phoenix, AZ, November 15-18 2004.

[68] J. Mitola. Cognitive radio for flexible mobile multimedia communications. In *1999 IEEE International Workshop on Mobile Multimedia Communications (MoMuC 1999)*, pages 3–10, San Diego, CA, November 15-17 1999.

[69] H. Mehdi and K. Feher. Power and spectrally efficient modulation for PCS and satellite TV broadcasting. *IEEE Trans. Consumer Electron.*, 42(1):92–100, February 1996.

[70] A. He, S. Srikanteswara, J. H. Reed, X. Chen, W. H. Tranter, K. K. Bae, and M. Sajadieh. Minimizing energy consumption using cognitive radio. In *2008 IEEE International Performance, Computing and Communications Conference (IPCCC 2008)*, pages 372–377, Austin, TX, December 7-9 2008.

[71] J. G. Proakis. *Digital Communications*. Mc Graw Hill, New York, NY, fourth edition, 2001.

[72] Frederick H. Raab. Average efficiency of power amplifiers. In *Proc. RF Technol. Expo*, pages 474–486, Anaheim, CA, Jan. 30 - Feb. 1 1986.

[73] D. A. Neamen. *Electronic Circuit Analysis and Design*. McGraw Hill Higher Education, New York, NY, 2 edition, 2000.

[74] Texas Instruments. *TRF4002 PCF RF power amplifier*, 1999.

[75] H. S. Kim and B. Daneshrad. Energy-aware link adaptation for MIMO-OFDM based wireless communication. In *2008 IEEE Military Communications Conference (MILCOM 2008)*, pages 1–7, San Diego, CA, November 17-19 2008.

[76] A. He, S. Srikanteswara, K. K. Bae, T. R. Newman, J. H. Reed, W. H. Tranter, M. Sajadieh, and M. Verhelst. System power consumption minimization for multichannel communications using cognitive radio. In *2009 IEEE International Conference on Microwaves, Communications, Antennas and Electronics Systems (COMCAS 2009)*, Tel Aviv, Isreal, November 9-11 2009.

[77] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, New York, NY, 1991.

[78] J. M. Cioffi. *Data Transmission Theory*. in preparation.

[79] I E. Telatar. Capacity of multi-antenna gaussian channels. *European Trans. Telecom.*, 10(6):585–595, November 1999.

[80] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath. Capacity limits of MIMO channels. *IEEE J. Select. Areas Commun.*, 21(5):684–702, June 2003.

[81] D-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn. Fading correlation and its effect on the capacity of multielement antenna systems. *IEEE Trans. Commun.*, 48(3):502–513, March 2000.

[82] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen. A stochastic MIMO radio channel model with experimental validation. *IEEE J. Select. Areas Commun.*, 20(6):1211–1226, August 2002.

[83] K. Yu, M. Bengtsson, B. Ottersten, D. McNamara, and P. Karlsson. A wideband statistical model for NLOS indoor MIMO channels. In *2002-Spring IEEE 55th Vehicular Technology Conference (VTC2002-Spring)*, pages 370–374, Birmingham, AL, May 6-9 2002.

[84] V. A. Aalo. Performance of maximal-ratio diversity systems in a correlated Nakagami-fading environment. *IEEE Trans. Commun.*, 43(8):2360–2369, August 1995.

[85] A. He, J. Gaeddert, K. Bae, T. R. Newman, J. H. Reed, L. Morales, and C. Park. Development of a case-based reasoning cognitive engine for IEEE 802.22 WRAN applications. *ACM Mobile Computing and Communications Review*, 13(2):37–48, April 2009.

[86] 3rd Generation Partnership Project. Self-configuring and self-optimizing network use cases and solutions (release 9). *Evolved Universal Terrestrial Radio Access Network (E-UTRAN)*, September 2009.

[87] FCC. FCC adopts rules for unlicensed use of television white spaces. `http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-286566A1.pdf`, November 8 2008.

[88] QinetiQ. Cognitive radio technology a study for Ofcom - volume 1. `http://www.ofcom.org.uk/research/technology/research/emer_tech/cograd/`, February 12 2007.

[89] R. Lo. Moving towards cognitive NetOps. In *JTRS Science and Technology Forum, JSTeF 3*, San Diego, CA, January 26-29 2009.

[90] FCC. Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies. Notice of proposed rulemaking and order, Federal Communications Commission, Washington, D.C., December 2003.

[91] IEEE. Standard Terms, Definitions, and Concepts for Spectrum Management, Policy Defined Radio, Adaptive Radio, and Software Defined Radio. `http://grouper.ieee.org/groups/emc/emc/ieee_emcs_-_sdcom/P1900-1/IEEE\%20P1900-1\%20Draft\%20-\%20Current.doc`.

[92] AAAI. Artificial intelligence topics. `http://www.aaai.org/AITopics/pmwiki/pmwiki.php/AITopics/HomePage`.

[93] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice-Hall, Upper Saddle River, NJ, USA, 1999.

[94] A. Fehske, J. Gaeddert, and J. H. Reed. A new approach to signal classification using spectral correlation and neural networks. In *2005 IEEE 1st International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN 2005)*, pages 144–150, Baltimore, MD, November 8-11 2005.

[95] A. F. Cattoni, M. Ottonello, M. Raffetto, and C. S. Regazzoni. Neural networks mode classification based on frequency distribution features. In *2007 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2007)*, pages 251–257, Orlando, FL, August 1-3 2007.

[96] X.-L. Zhu, Y.-A. Liu, W.-W. Weng, and D-M. Yuan. Channel sensing algorithm based on neural networks for cognitive wireless mesh networks. In *2008 4th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM 2008)*, pages 1–4, Dalian, China, October 12-14 2008.

[97] M. Hasegawa, T. Ha Nguyen, G. Miyamoto, Y. Murata, and S. Kato. Distributed optimization based on neurodynamics for cognitive wireless clouds. In *2007 IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2007)*, pages 1–5, Athens, Greece, September 3-7 2007.

[98] Z. Zhang and X. Xie. Intelligent cognitive radio: Research on learning and evaluation of cr based on neural network. In *2007 ITI 5th International Conference on Information and Communications Technology (ICICT 2007)*, pages 33–37, Dhaka, Bangladesh, December 16-18 2007.

[99] MPRG. OSSIE. `http://ossie.wireless.vt.edu/`.

[100] C. R. Aguayo Gonzalez, C. B. Dietrich, F. E. Kragh, S. Sayed, H. I. Volos, J. D. Gaeddert, P. M. Robert, and J. H. Reed. Open-source SCA-based core framework and rapid development tools enable software-defined radio education and research. *IEEE Commun. Mag.*, 47(10):48–55, October 2009.

[101] N. Baldo and M. Zorzi. Learning and adaptation in cognitive radios using neural networks. In *2008 IEEE 5th Consumer Communications and Networking Conference (CCNC 2008)*, pages 998–1003, Las Vegas, NV, January 10-12 2008.

[102] B. Ustundag and O. Orcay. Pattern based encoding for cognitive communication. In *2008 3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2008)*, pages 1–6, Singapore, May 15-17 2008.

[103] O. Orcay and B. Ustundag. Pattern recognition in cognitive communication. In *2008 23rd International Symposium on Computer and Information Sciences (ISCIS 2008)*, pages 1–6, Istanbul, Turkey, October 27-29 2008.

[104] C. Blum and A. Roli. Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Comput. Surv.*, 35(3):268–308, September 2003.

[105] F. Glover. Future paths for integer programming and links to artificial intelligence. *Comput. Oper. Res.*, 13(5):533–549, May 1986.

[106] H. Robbins and S. Monro. A stochastic approximation method. *Ann. Math. Stat.*, 22(3):400–407, 1951.

[107] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addsson-Wiley, Reading, MA, 1989.

[108] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, May 1983.

[109] F. W. Glover and M. Laguna. *Tabu Search*. Kluwer Academic Publishers, Boston, MA, 1997.

[110] M. Dorigo and C. Blum. Ant colony optimization theory: a survey. *Theoretical Computer Science*, 344(2-3):243–278, November 2005.

[111] T. R. Newman, B. A. Barker, A. M. Wyglinski, A. Agah, J. B. Evans, and G. J. Minden. Cognitive engine implementation for wireless multicarrier transceivers. *Wirel. Commun. Mob. Comput. (Wiley InterScience)*, 7(9):1129–1142, 2007.

[112] J. F. Hauris. Genetic algorithm optimization in a cognitive radio for autonomous vehicle communications. In *2007 International Symposium on Computational Intelligence in Robotics and Automation (CIRA 2007)*, pages 427–431, Jacksonville, FL, June 20-23 2007.

[113] S. K. Park, Y. Shin, and W. C. Lee. Goal-pareto based nsga for optimal reconfiguration of cognitive radio systems. In *2007 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2007)*, pages 147–153, Orlando, FL, August 1-3 2007.

[114] D. Thilakawardana and K. Moessner. A genetic approach to cell-by-cell dynamic spectrum allocation for optimising spectral efficiency in wireless mobile systems. In *2007 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2007)*, pages 367–372, Orlando, FL, August 1-3 2007.

[115] Z. Zhang and X. Xie. Application research of evolution in cognitive radio based on ga. In *2008 IEEE 3rd Conference on Industrial Electronics and Applications (ICIEA 2008)*, pages 1575–1579, Singapore, June 3-5 2008.

[116] Y. Liu, H. Jiang, and Y. Q. Huang. Design of cognitive radio wireless parameters based on multi-objective immune genetic algorithm. In *2009 WRI International Conference on Communications and Mobile Computing (CMC 2009)*, pages 92–96, Kunming, Yunnan, China, January 6-8 2009.

[117] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proc. IEEE*, 77(2):257–286, February 1989.

[118] T. W. Rondeau, C. J. Rieser, T. M. Gallagher, and C. W. Bostian. Online modeling of wireless channels with hidden markov models and channel impulse responses for cognitive radios. In *2004 IEEE MTT-S International Microwave Symposium Digest*, pages 739–742, Fort Worth, TX, June 6-11 2004.

[119] K. Kim, I. A. Akbar, K. K. Bae, J.-s. Urn, C. M. Spooner, and J. H. Reed. Cyclostationary approaches to signal detection and classification in cognitive radio. In *2007 IEEE 2nd International Symposium*

on New Frontiers in Dynamic Spectrum Access Networks (DySPAN 2007), pages 212–215, Dublin, Ireland, April 17-20 2007.

[120] C. Ghosh, C. Cordeiro, D. P. Agrawal, and M. B. Rao. Markov chain existence and hidden markov models in spectrum sensing. In *2009 IEEE International Conference on Pervasive Computing and Communications (PerCom 2009)*, pages 1–6, Galveston, TX, March 9-13 2009.

[121] I. A. Akbar and W. H. Tranter. Dynamic spectrum allocation in cognitive radio using hidden markov models: Poisson distributed case. In *2007 IEEE Southeast Conference*, pages 196–201, Richmond, VA, March 22-25 2007.

[122] C.-H. Park, S.-W. Kim, S.-M. Lim, and M. S. Song. HMM based channel status predictor for cognitive radio. In *2007 Asia-Pacific Microwave Conference (APMC 2007)*, pages 1–4, Bangkok, Thailand, December 11-14 2007.

[123] J. Paine. Expert systems. `http://www.j-paine.org/students/lectures/lect3/lect3.html`.

[124] J. H. Reed et al. Development of a cognitive engine and analysis of WRAN cognitive radio algorithms: phase II. Report submitted to ETRI, Mobile and Portable Radio Research Group (MPRG), Wireless@Virginia Tech, Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, December 2006.

[125] Y. Zhao, J. Gaeddert, L. Morales, K. Bae, J-S. Um, and J. H. Reed. Development of radio environment map enabled case- and knowledge-based learning algorithms for IEEE 802.22 wran cognitive engines. In *2007 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2007)*, Orlando, FL, August 1-3 2007.

[126] C. Clancy, J. Hecker, E. Stuntebeck, and T. O'Shea. Applications of machine learning to cognitive radio networks. *IEEE Wireless Commun.*, 14(4):47–52, August 2007.

[127] T. Weingart, D. C. Sicker, and D. Grunwald. A statistical method for reconfiguration of cognitive radios. *IEEE Wireless Commun.*, 14(4):34–40, August 2007.

[128] T. R. Gruber. Toward principles for the design of ontologies used for knowledge sharing. *Int. J. Human-Comput. Studies*, 43(5-6):907–928, December 1995.

[129] N. Guarino. Formal ontology in information systems. In *1998 Formal Ontology in Information Systems (FOIS 1998)*, pages 3–15, Trento, Italy, June 6-8 1998.

[130] XTM. XML Topic Maps (XTM) 1.0. `http://www.topicmaps.org/xtm/`.

[131] W3C. Resource Description Framework. `http://www.w3.org/RDF/`.

[132] W3C. OWL Web Ontology Language Overview. `http://www.w3.org/TR/owl-features/`.

[133] W3C. W3C Semantic Web Activity. `http://www.w3.org/2001/sw/`.

[134] BBN Technologies. XG policy language framework - request for comments. `http://www.autonomic-communication.org/web/bodies/DARPA_XG_rfc_policylang.pdf`, April 2004.

[135] M. M. Kokar and L. Lechowicz. Language issues for cognitive radio. *IEEE J. Select. Areas Commun.*, 97(4):689–707, April 2009.

[136] J. Wang, M. M. Kokar, K. Baclawski, and D. Brady. Achieving self-awareness of SDR nodes through ontology-based reasoning and reflection. In *2004 SDR Technical Conference and Product Exposition (SDR 2004)*, Phoenix, AZ, November 15-18 2004.

[137] J. D. Poston, W. D. Horne, M. G. Taylor, and F. Z. Zhu. Ontology-based reasoning for context-aware radios: insights and findings from prototype development. In *2005 IEEE 1st International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN 2005)*, pages 634–637, Baltimore, MA, November 8-11 2005.

[138] A. Ginsberg, J. D. Poston, and W. D. Horne. Toward a cognitive radio architecture: Integrating knowledge representation with software defined radio technologies. In *2006 Military Communications Conference (MILCOM 2006)*, pages 1–7, Washington, DC, October 23-25 2006.

[139] J. L. Kolodner and D. Leake. A tutorial introduction to case-based reasoning. In David Leake, editor, *Case-Based Reasoning: Experiences, Lessons and Future Directions*, chapter 2, pages 31–65. MIT Press, Cambridge, MA., 1996.

[140] J. H. Reed et al. Development of a cognitive engine and analysis of WRAN cognitive radio algorithms: phase III. Report submitted to ETRI, Mobile and Portable Radio Research Group (MPRG), Wireless@Virginia Tech, Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, December 2007.

[141] M. Khedr and H. Shatila. CogMAX - A cognitive radio approach for WiMAX systems. In *2009 IEEE/ACS International Conference on Computer Systems and Applications (AICCSA 2009)*, pages 550–554, Rabat, Morocco, May 10-13 2009.

[142] B. Le, T. W. Rondeau, and C. W. Bostian. Cognitive radio realities. *Wirel. Commun. Mob. Comput.*, 7(9):1037–1048, 2007.

[143] Soar. Soar. `http://sitemaker.umich.edu/soar/home`.

[144] T. M. Mitchell. *Machine Learning*. McGraw-Hill, New York, NY, 1997.

[145] G. Hardin. The tragedy of the commons. *Science*, 162(3859):1243–1248, December 1968.

[146] J. Neel. Analysis and design of cognitive radio and distributed radio resource management algorithms. Phd dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA, September 2006.

[147] J. Neel, R. Buehrer, J. Reed, and R. Gilles. Game theoretic analysis of a network of cognitive radios. In *2002 IEEE 45th International Midwest Symposium on Circuits and Systems (MWSCAS 2002)*, pages 409–412, Stillwater, OK, August 4-7 2002.

[148] E. Altman and Z. Altman. S-modular games and power control in wireless networks. *IEEE Trans. Automat. Contr.*, 48(5):839–842, May 2003.

[149] Y. Zeng and Z. Zhang. Joint transmit beamforming and power control in multi-user MIMO downlink using a game theoretic approach. *The Journal of China Universities of Posts and Telecommunications*, 14(2):14–18, June 2007.

[150] V. Srivastava and L. DaSilva. Game theoretic analysis of a network of cognitive radios. In *2006 IEEE International Conference on Communications (ICC 2006)*, pages 3850–3855, Istanbul, Turkey, June 11-15 2006.

[151] J. Huang, Z. Han, M. Chiang, , and H. V. Poor. Auction-based distributed resource allocation for cooperation transmission in wireless networks. In *2007 IEEE Global Telecommunications Conference (GLOBECOM 2007)*, pages 4807–4812, Washington DC, November 26-30 2007.

[152] T. Roughgarden. *Algorithmic Game Theory*, chapter Routing Games, pages 31–65. Cambridge University Press, 2007.

[153] R. Gilles R. Komali, A. MacKenzie. Effect of selfish node behavior on efficient topology design. *IEEE Trans. on Mobile Comput.*, 7(9):1057–1070, September 2008.

[154] M. van der Schaar and F. Fu. Spectrum access games and strategic learning in cognitive radio networks for delay-critical applications. *Proc. IEEE*, 97(4):720–740, April 2009.

[155] Y. Zhao, S. Mao, J. O. Neel, and J. H. Reed. Performance evaluation of cognitive radios: Metrics, utility functions, and methodology. *Proc. IEEE*, 97(4):642–659, April 2009.

[156] A. He, S. Srikanteswara, K. K. Bae, J. H. Reed, and W. H. Tranter. Energy consumption minimization for mobile and wireless devices - a cognitive approach. *IEEE Trans. Consum. Electron.*, 56(3):1814–1821, August 2010.

[157] A. He, S. Srikanteswara, K. K. Bae, T. R. Newman, J. H. Reed, W. H. Tranter, M. Sajadieh, and M. Verhelst. Power consumption minimization for MIMO systems - a cognitive radio approach. *IEEE J. Sel. Areas Commun.*, 29(2):469–479, February 2011.

[158] Analog Device. *ADM1191 Digital power monitor with convert pin and ALERTB output*, 2010.

[159] A. He, K. K. Bae, T. R. Newman, J. Gaeddert, K. Kim, R. Menon, L. Morales-Tirado, J. Neel and Y. Zhao, J. H. Reed, and W. H. Tranter. A survey of artificial intelligence for cognitive radios. *IEEE Trans. Veh. Technol.*, 59(4):1578–1592, May 2010.

[160] A. He, S. Srikanteswara, K. K. Bae, T. R. Newman, J. H. Reed, W. H. Tranter, M. Sajadieh, and M. Verhelst. Power consumption minimization for MIMO systems using cognitive radio. In *2009 SDR Technical Conference and Product Exposition (SDR 2009)*, Washington D.C., December 1-4 2009.

[161] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics, Philadelphia, PA, 2001.

[162] T. S. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall, Upper Saddle River, NJ, 2nd edition edition, 2002.

[163] G. McLachlan and D. Peel. *Finite Mixture Models*. Wiley-Interscience, Hoboken, NJ, 2000.

[164] V. A. Marcenko and L. A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4), 1967.

[165] C. Cordeiro, K. Challapali, D. Birru, and S. Shankar N. IEEE 802.22: the first worldwide wireless standard based on cognitive radios. In *2005 IEEE 1st International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN 2005)*, pages 328 – 337, Baltimore, MD, 8-11 November 2005.

[166] S. Shiu and S. K. Pal. *Foundations of Soft Case-Based Reasoning*. Wiley Series on Intelligent Systems. Wiley-Interscience, Hoboken, NJ, 2004.

[167] T. Weingart, D. C. Sicker, and D. Grunwald. A method for dynamic configuration of a cognitive radio. In *2006 IEEE 1st Workshop on Networking Technologies for Software Defined Radio Networks (SDR 2006)*, pages 93–100, Reston, Virginia, USA, September 2006.

[168] J. H. Reed et al. Applying artificial intelligence to the development of a cognitive radio engine. Report submitted to ARO, Mobile and Portable Radio Research Group (MPRG), Wireless@Virginia Tech, Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, July 2006.

[169] Y. Zhao, L. Morales, J. Gaeddert, K. K. Bae, J-S. Um, and J. H. Reed. Applying radio environment maps to cognitive wireless regional area networks. In *2007 IEEE 2nd International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN 2007)*, pages 115–118, Dublin, Ireland, April 17-20 2007.