Leveraging Cognitive Radio Networks Using Heterogeneous Wireless Channels

by

Yongkang Liu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2013

Abstract

The popularity of ubiquitous Internet services has spurred the fast growth of wireless communications by launching data hungry multimedia applications to mobile devices. Powered by spectrum agile cognitive radios, the newly emerged cognitive radio networks (CRN) are proposed to provision the efficient spectrum reuse to improve spectrum utilization. Unlicensed users in CRN, or secondary users (SUs), access the temporarily idle channels in a secondary and opportunistic fashion while preventing harmful interference to licensed primary users (PUs). To effectively detect and exploit the spectrum access opportunities released from a wide spectrum, the heterogeneous wireless channel characteristics and the underlying prioritized spectrum reuse features need to be considered in the protocol design and resource management schemes in CRN, which plays a critical role in unlicensed spectrum sharing among multiple users.

The purpose of this dissertation is to address the challenges of utilizing heterogeneous wireless channels in CRN by its intrinsic dynamic and diverse natures, and build the efficient, scalable and, more importantly, practical dynamic spectrum access mechanisms to enable the cost-effective transmissions for unlicensed users. Note that the spectrum access opportunities exhibit the diversity in the time/frequency/space domain, secondary transmission schemes typically follow three design principles including 1) *utilizing local free channels within short transmission range*, 2) *cooperative and opportunistic transmissions*, and 3) *effectively coordinating transmissions in varying bandwidth*. The entire research work in this dissertation casts a systematic view to address these principles in the design of the routing protocols, medium access control (MAC) protocols and radio resource management schemes in CRN.

Specifically, as spectrum access opportunities usually have small spatial footprints, SUs only communicate with the nearby nodes in a small area. Thus, multi-hop transmissions in CRN are considered in this dissertation to enable the connections between any unlicensed users in the network. CRN typically consist of intermittent links of varying bandwidth so that the decision of routing is closely related with the spectrum sensing and sharing operations in the lower layers. An efficient opportunistic cognitive routing (OCR) scheme is proposed in which the forwarding decision at each hop is made by jointly considering physical characteristics of spectrum bands and diverse activities of PUs in each single band. Such discussion on spectrum aware routing continues coupled with the sensing selection and contention among multiple relay candidates in a multi-channel multi-hop scenario. An SU selects the next hop relay and the working channel based upon location information and channel usage statistics with instant link quality feedbacks. By evaluating the performance of the routing protocol and the joint channel and route selection algorithm with extensive simulations, we determine the optimal channel and relay combination with reduced searching complexity and improved spectrum utilization.

Besides, we investigate the medium access control (MAC) protocol design in support of multimedia applications in CRN. To satisfy the quality of service (QoS) requirements of heterogeneous applications for SUs, such as voice, video, and data, channels are selected to probe for appropriate spectrum opportunities based on the characteristics and QoS demands of the traffic along with the statistics of channel usage patterns. We propose a QoS-aware MAC protocol for multi-channel single hop scenario where each single SU distributedly determines a set of channels for sensing and data transmission to satisfy QoS requirements. By analytical model and simulations, we determine the service differentiation parameters to provision multiple levels of QoS.

We further extend our discussion of dynamic resource management to a more practical deployment case. We apply the experiences and skills learnt from cognitive radio study to cellular communications. In heterogeneous cellular networks, small cells are deployed in macrocells to enhance link quality, extend network coverage and offload traffic. As different cells focus on their own operation utilities, the optimization of the total system performance can be analogue to the game between PUs and SUs in CRN. However, there are unique challenges and operation features in such case. We first present challenging issues including interference management, network coordination, and interworking between cells in a tiered cellular infrastructure. We then propose an adaptive resource management framework to improve spectrum utilization and mitigate the co-channel interference between macrocells and small cells. A game-theory-based approach is introduced to handle power control issues under constrained control bandwidth and limited end user capability. The inter-cell interference is mitigated based upon orthogonal transmissions and strict protection for macrocell users.

The research results in the dissertation can provide insightful lights on flexible network deployment and dynamic spectrum access for prioritized spectrum reuse in modern wireless systems. The protocols and algorithms developed in each topic, respectively, have shown practical and efficient solutions to build and optimize CRN.

Acknowledgements

The past five years of my graduate and research life in Waterloo is truly the most unique, precious and awarding time of mine. Along this way, there are so many people who deserve to be thanked for their friendship, help and supports offered to me. Within this limited space, I cannot hope to thank them properly.

First and foremost, my deepest and sincerest gratitude goes to my supervisor Professor Xuemin (Sherman) Shen. It is his guidance, support, encouragement, patience and genuine expertise in the past five years that have made this dissertation possible. What I appreciate the most of Professor Shen is his great patience and understanding to students. His enthusiasm and dedication to his work, his students and his family are really inspiring to me.

I would also like to thank the honorable members of my thesis committee, Professor Xianbin Wang, Liang-liang Xie, Sagar Naik and Xinzhi Liu, for serving as my thesis readers. Their insightful comments have significantly affected the substance and presentation of my work.

Being honored to study and work in BBCR group, I would like to express my deep gratitude to Professor Weihua Zhuang and Professor Jon W. Mark for immensely helpful comments and suggestion on my research. Doing research and spending day and night in the lab can be boring and daunting. My colleagues in BBCR group have made my life at the University of Waterloo a colorful and enjoyable experience. I wish to thank Zhongmin Zheng, Kuan Zhang, Xinsheng Zhou, Ning Lu, Ning Zhang, Nan Cheng, Yong Zhou, Yujie Tang, Khadige Abboud, Amila P. K. Tharaperiya Gamage, Qinghua Shen, Xiaohui Liang, Miao Wang, Ran Zhang, Jian Qiao, Xiaoxia Zhang, Tingting Yang, Hongbin Liang, Sanying Li, Wei Cai, Dr. Lin X. Cai, Dr. Hao Liang, Dr. Tom Luan, Dr. Shimin Gong, Dr. Rongxing Lu, Dr. Yuanguo Bi, Dr. Zhiguo Shi, Dr. Yipin Sun, Dr. Hangguan Shan, Dr. Bin Lin, Dr. Fen Hou and many others for their continuous encouragement, selfless help and all the good times we spent together. In particular, I wish to especially thank Dr. Lin X. Cai, Dr. Zhiguo Shi, Hongwei Luo, and Dr. Juncheng Jia for their inspiring discussions and invaluable suggestions on my research.

The thesis is dedicated to my father Shixiang Liu and my mother Xinhua Hao. I would not be completing my studies if they did not teach me the value of hard work and dedication. The thesis is also dedicated to my girl friend Luning (Grace) Fu. I owe them everything, and fear I cannot love them enough in return for that. Thanks to them all for their continuous and ever-caring support which made me always feel their presence so near to me.

Table of Contents

List of Tables x

List of Figures xi

List of Abbreviations xiii

List of Symbols xv

1 Introduction 1

1.1 Motivation and Background . . . 1
1.2 Research Challenges . . . 2
1.3 Our Approach and Contributions . . . 5
1.4 Thesis Outline . . . 8

2 Background 9

2.1 Overview of Cognitive Radio Networks . . . 9
2.2 Dynamic Spectrum Access . . . 11
  2.2.1 Spectrum Access Models . . . 11
  2.2.2 Spectrum Access Opportunities . . . 13
  2.2.3 Spectrum Sensing . . . 15
  2.2.4 Opportunistic Unlicensed Transmissions . . . 17
2.3 Summary . . . 19

3 System Model 20
3.1 Network Model 20
3.2 Channel Model 21
3.3 Traffic Model 23

4 Opportunistic Cognitive Routing in Heterogeneous Wireless Channels 26
4.1 Introduction 27
4.2 Related Works 30
4.3 Opportunistic Routing in CRN 35
4.3.1 Sensing Sequence 37
4.3.2 Relay Selection 38
4.4 Performance Evaluation 40
4.4.1 Simulation Settings 40
4.4.2 End-to-end Delay in Secondary Link 41
4.4.3 Performance of SU Density 42
4.5 Summary 43

5 Joint Resource Allocation for Multi-hop Opportunistic Transmissions 44
5.1 Introduction 44
5.2 Related Work 45
5.3 Spectrum Aware Opportunistic Routing Protocol 46
5.3.1 System Description 47
5.3.2 Protocol Overview 48
5.3.3 Routing Protocol Analysis 49
5.4 Joint Channel and Relay Selection 54
5.4.1 Novel Routing Metric 54
5.4.2 Heuristic Algorithm 55
5.5 Performance Evaluation 59

5.5.1 Simulation Settings 59
5.5.2 PU Activities 62
5.5.3 Multi-user Diversity 63
5.5.4 Effectiveness of Routing Metric 64
5.6 Summary 67

6 QoS-aware Cognitive (QC) MAC for Multimedia Services 70
6.1 Introduction 70
6.2 The QC MAC Scheme 75
6.2.1 Channel Sensing 76
6.2.2 Service Differentiation 78
6.3 Performance Analysis 78
6.4 Numerical and Simulation Results 80
6.4.1 Simulation Settings 80
6.4.2 Delay of Homogeneous Traffic 80
6.4.3 Delay of Heterogeneous Traffic 81
6.4.4 Performance of Service Differentiation 82
6.5 Summary 83

7 Cognitive Cellular Networks under Dynamic Resource Management 85
7.1 Introduction 86
7.2 Trends in Cellular Networks 87
7.2.1 Heterogeneous Cells and Tiered Network Deployment 87
7.2.2 Green Cellular Networks and Sustainable Communications 89
7.2.3 Secure Communications and User Privacy 90
7.3 Cognitive Cellular Networks 91
7.3.1 Interference Management in Tiered Networks 93
7.3.2 Network Bottleneck Mitigation 93

7.3.3 Coordination in Cognitive Cellular Networks 94
7.4 Dynamic Resource Allocation in Tiered Cellular Networks 95
7.5 Summary 99
7.6 Appendix: Proof of Nash Equilibrium 99

8 Conclusion and Future Work **101**
8.1 Conclusion 101
8.1.1 Opportunistic Spectrum Reuse 101
8.1.2 End User Adaption 102
8.1.3 Network Coordination 102
8.2 Future Work 103
8.2.1 Cooperative Communications in CRN 103
8.2.2 Energy Efficiency in CRN 104
8.2.3 Heterogeneous Network Design 105

References **106**

List of Tables

4.1 Simulation parameters in performance evaluation of Chapter 4 40
4.2 Settings of channel usage patterns in performance evaluation of Chapter 4 41

5.1 Summary of notations in Chapter 5 68
5.2 Simulation parameters in performance evaluation of Chapter 5 69
5.3 Average neighbor density under different SU densities 69

6.1 Sensing window design for multimedia services in CRN 78
6.2 Settings of channel usage patterns in performance evaluation of Chapter 6 81

7.1 Specifications of different cells in cellular networks 88

List of Figures

2.1 Channel usage pattern. 13

3.1 CRN topology without infrastructure. 21
3.2 CRN topology with infrastructure. 22
3.3 Dynamic multi-channel access in CRN 23
3.4 Heterogeneous wireless channels. 24

4.1 Secondary transmissions in an ad hoc CRN 27
4.2 Multi-hop CRN over heterogeneous wireless channels. 35
4.3 Relay selection. 39
4.4 Comparison of end-to-end delay under different PU activities. 41
4.5 CDF of path hop counts under different PU activities. 42
4.6 Comparison of end-to-end delay under different SU densities. 43

5.1 The opportunistic cognitive routing timeline 47
5.2 Performance comparison between OCR and SEARCH under different channel conditions (Number of SUs: 200, flow rate: 10/40 pps) 61
5.3 Performance comparison under different traffic loads and PU activities (Number of SUs: 200, flow rate: 10 pps) 62
5.4 Performance comparison of end-to-end delay under different SU densities (flow rate: 10 pps, $E[T_{OFF}] = 200\ ms$) 63
5.5 Performance comparison under different traffic loads (Number of SUs: 200, $E[T_{OFF}] = 200\ ms$) 65

5.6 Performance comparison under different SU densities (flow rate: 10 pps, $E[T_{OFF}] = 200\ ms$) 66

6.1 Multi-channel transmissions in CRN. 76
6.2 Average delay of voice/video flows (Greedy). 82
6.3 Comparison of channel sensing policies (low traffic load). 83
6.4 Comparison of channel sensing policies (high traffic load). 84
6.5 Performance comparison with different sensing windows. 84

7.1 Tiered framework of cognitive cellular networks 90
7.2 The study framework of cognitive cellular networks 91
7.3 Interference management in cognitive cellular networks 96
7.4 Aggregated femtocell utility under intra-femtocell channel gains 98

List of Abbreviations

AODV	Ad hoc on-demand distance vector
AP	Access point
ARQ	Automatic repeat request
CCC	Common control channel
CCRN	Cooperative cognitive radio networks
CR	Cognitive radio
CRN	Cognitive radio networks
CRAHN	Cognitive radio ad hoc network
CSMA	Carrier sense multiple access
CTT	Cognitive transport throughput
DCF	Distributed coordination function
DSA	Dynamic spectrum access
DySPAN	Dynamic spectrum access network
DSDV	Destination-sequenced distance-vector
DSR	Dynamic source routing
ETT	Expected Transmission Time
ETX	Expected Transmission Count
ExOR	Extremely Opportunistic Routing
FBS	Femtocell base station
FCC	Federal communications commission
FH	Frequency hopping
GPS	global positioning system
GPSR	Greedy perimeter stateless routing
ITU	International Telecommunication Union

LBT	Listen-before-transmission
MAC	Medium access control
MJD	Multiband joint detection
MRV	Multiple rendezvous
OCR	Opportunistic cognitive routing
PMP	Point-to-multipoint
POMDP	Partially observable Markov decision process
PU	Primary user
QCMAC	Quality of service cognitive medium access control
QoS	Quality of service
QP	Quiet period
RREP	Route reply
RREQ	Route request
SIFS	Short interframe space
SINR	Signal to interference-plus-noise ratio
SDR	Software-defined radio
SRV	Single rendezvous
SU	Secondary user
TCP	Transmission control protocol
UHF	Ultra high frequency
UWB	Ultra-wideband
VBR	Variable bit rate

List of Symbols

$A_D(S,R)$	Relay advancement of the link SR for D
B_i	Bandwidth of Channel i
C_N	Number of channels
$\mathcal{C}$	Set of orthogonal channels
D_k	Transmission distance on channel k
d_{rD}	Distance between SU r and the destination gateway
d_{SD}	Distance between the source and destination
$d(S,D)$	Euclidian distance between S and D
$E[T]$	Expected transmission delay of an SU
$E[T_i^k]$	Expected transmission delay on channel k of SU i
$E[T_i]$	Expected transmission delay of SU i
$E[T^k]$	Average time an SU spends on one transmission on channel k
$E[T_{OFF}^{c_j}]$	Mean duration of an idle channel, c_j
$E[T_{ON}^{c_j}]$	Mean duration of a busy channel, c_j
$F_k(x)$	CDF of usage pattern of channel k by the PUs
$f_{T_{OFF}^k}(x)$	PDF of the sojourn time of an OFF period of channel k
$f_{T_{ON}^k}(y)$	PDF of the sojourn time of an ON period of channel k
$\mathcal{F}_{OFF}^{c_j}(t)$	CDF of the OFF duration of c_j
$I_R^{c_j}$	SU R detects c_j to be idle
$\overline{I_R^{c_j}}$	SU R detects c_j to be busy
N_d	Number of data channels selected by the SU for transmissions
N_k	Number of SUs contending in channel k
$\mathbf{N}_S$	The SU set of the sender S's neighbors
P_{CO}^k	Probability that channel k has been occupied by a PU

P_i^k	Probability that the current frame can be successfully transmitted on channel k
$P_{OFF,R}^{c_j}(t_0, t_1)$	The probability that c_j is idle at t_1, $t_1 > t_0$
P_{PU}^k	Probability that channel k is idle but a PU turns on during the sensing interval
$P_R^{c_j}(t_1, t_2)$	The probability that c_j is idle during $[t_1, t_2]$ at R
$P_r(d)$	Received radio power at a distance of d
$P_{relay,R_i}^{c_j}$	The probability that the relay via R_i succeeds in c_j
$P_{RSfail}^{c_j}$	The probability that relay selection fails in c_j
P_{ss}^k	Probability that an SU succeeds in sensing and attempts a transmission on channel k
P_{SU}^k	Probability that an SUs sensing fails due to other SUs transmission
P_{s_r}	Opportunistic transmission capability of user r obtained from the transmission history
P_t	Transmission power
P_{TS}^k	Probability that a transmission succeeds only if no PU turns on during the total sensing and transmission time of the SU on channel k
P_{TF}^k	Probability that the transmission fails due to the disruption from PU on channel k
$Ps(r)$	Probability that a transmission succeeds in the r-th attempt
$\mathbf{R}_D$	The set of relay candidates for the destination D
$\mathbf{R}_D^{c_j}$	The set of relay candidates for the destination D in the channel c_j
R_i	Data rate on channel i
SW_{data}	Sensing window size of data traffic
SW_i	Sensing window Size
SW_{video}	Sensing window size of video traffic
SW_{voice}	Sensing window size of voice traffic
t_{AS_i}	Arbitrary sensing period selected by user i
t_i	Transmission time of the current frame by user i
T_{OFF}^k	Sojourn time of an OFF period of channel k
T_{ON}^k	Sojourn time of an ON period of channel k
T_{detc}	Per channel energy detection delay
T_{DTX}	Per hop data packet transmission delay

T_{init}	Sensing initialization delay
T_{relay}	Per hop transmission delay in OCR
T_{RREQ}	RREQ message transmission delay
T_{RRSP}	RRSP message transmission delay
T_{RS}	Per hop relay selection delay
T_{SNS}	Per hop sensing delay
T_{switch}	Transceiver switching time
t_0	The latest channel status observation time
V_{R_i}	The priority of R_i in the relay selection
$X^{c_j}_{R_tR_r}$	SUR_t and SU R_r are affected by the same PU in c_j with value 1
α	Weighting factor
α_k	Mean rate of transition from idle to busy on channel k
α_p	Path-loss coefficient parameter
β	Weighting factor
β_k	Mean rate of transition from idle to busy on channel k
γ	The maximum channel propagation delay
γ_k	Metric to characterize the forwarding capability in each channel
τ_i	Maximum tolerable one hop delay
μ	Backoff mini-slot
ν_r	Relay capability of SU r
ν_{min}	Maximum relay capability of SU r
ν_{max}	Minimum relay capability of SU r
ρ_{c_j}	The chance for an idle state in c_j

Chapter 1

Introduction

1.1 Motivation and Background

The growing popularity of mobile Internet services have vigorously motivated the exploration and exploitation of wider spectrum bands to support higher data rate transmissions over the air [1]. To solve spectrum scarcity resulting from static near-full allocation of spectrum, cognitive radio is proposed as one promising approach to keep promoting wireless communications. In cognitive radio networks (CRN), unlicensed users, or secondary users (SUs), reside in the same area with licensed users, or primary users (PUs), and explore for and exploit the frequency bands where PUs are inactive, i.e., spectrum access opportunities, to transmit their own data.

In CRN, spectrum access opportunities inherently exhibit diversity in time, frequency and space dimensions, and vary at individual SUs. To protect the licensed transmissions of PUs, SUs need to opportunistically access the time-frequency resource blocks where PUs are inactive and operate with strict constraints to protect the active PUs' quality of service (QoS), e.g., the maximum interference allowed at the PU's receiver. As a result, the channel notably shows the ON-OFF usage pattern to SUs and the average time window length is affected by the intensity of PUs' activities in the channel. Meanwhile, spectrum access opportunities, which are distributed in a wide spectrum, show unique transmission characteristics in individual channels including PU's intensity and radio propagation features. Since SUs either coordinate with PUs by spectrum leasing or spectrum sensing, which is cost-sensitive, then the available free bandwidth also relies on the capability of individual SUs, such as PU detection, scanning range per sensing, and the scheduling strategy of on-demand and periodic sensing operations. Therefore, heterogeneous wireless channels in

CRN consist of PU activities, channel transmission profiles and user specific bandwidth, which jointly shape the transmissions of unlicensed traffic in an opportunistic way.

Identifying such heterogeneity in wireless channels and characterizing spectrum access opportunities can significantly boost SUs' transmissions in CRN, which in turn further improves spectrum utilization and the whole network performance. Since CRN are expected to carry multimedia services with various QoS requirements, appropriately accommodating different traffic flows in the idle channels is necessary, which can satisfy SUs' QoS while performing better protection for nearby PUs. Besides spectrum efficiency, an intelligent spectrum exploration and reusage can reduce energy consumption for SUs as well. Since the bandwidth supply for secondary transmissions is recharged in an on-demand manner, it is energy efficient to predict the size of the available spectrum and select the channels with best potentials based upon the priori knowledge of traffic and channel conditions, which reduces the cost in sensing and delay in secondary transmissions.

Under a hierarchical access model, SUs strictly yield to PUs' priority in the accessed channels. Therefore, the bandwidth supply for unlicensed transmissions relies on two main factors: the channel usage pattern and SU's capability of harvesting the free spectrum. Since the former one is determined by PUs in each single frequency band and normally predictable, manipulating SUs with effective coordination and spectrum selection is meaningful for the performance leverage of CRN, which motivates the discussions in this dissertation. By recoganizing the promising potentials in performance improvement, this dissertation addresses the intelligent protocol design and resource management schemes for CRN serving secondary services in heterogeneous wireless channels.

1.2 Research Challenges

To fully utilize heterogeneous spectrum access opportunities in CRN, the major research concerns are on the design of spectrum aware protocols to coordinate SUs' transmissions over multiple channels, upon which efficient radio resource management schemes are expected to cope with dynamic channel conditions in a wide working spectrum and reduce interference to PUs in each channel. Reasons to address this challenging topic are as follows. Specifically, in CRN, SUs need to yield to the priority of the returning PUs during the unlicensed transmissions and prevent harmful interference onto the licensed users. Such layered access structure distinguishes CRN from conventional wireless networks or multi-channel access schemes. In other words, it means we can not directly use the existing schemes to address the challenging issues in CRN. Moreover, SUs use the temporarily unused spectrum bands to feed the demand of unlicensed communications. Thus, detecting and utilizing the varying spectrum resources determines the performance, which requires the communication protocols to coordinate multiple SUs in the network to sense and access multiple available channels within short time windows. The other important factor is that PUs and SUs usually define their own utilities of transmissions, respectively. Then, radio resource management (RRM) in CRN is expected to manage the objectives for individual user groups while keeping improving spectrum utilization.

In general, the challenging research issues in utilizing heterogeneous spectrum access opportunities can be categorized as three aspects: dynamic resource searching, opportunistic transmissions and effective coordination/interworking.

Dynamic Resource Searching

To find the desired spectrum access opportunities, spectrum sensing is performed in the channels periodically and upon request ahead of each secondary transmission attempt. The detection of spectrum access opportunities largely depends on the accurate depiction of channel usage patterns acquired by spectrum sensing. Intuitively, more sensing samples would help update the channel status for more accurate sensing results. But too frequent sensing would occupy the network resources and lead to long delay in transmissions. Moreover, to prevent the pollution to sensing results by ongoing secondary transmissions, the medium access control (MAC) layer protocol in CRN schedules the sensing nodes and transmission nodes working in different channels without collision. The major challenge of designing the sensing policy, normally considered in MAC, lies in adapting to the varying channel usage pattern at different frequencies and locations. Thus, the scheduling algorithm needs to balance the exploration and exploitation of the channel opportunities. Besides, individual SUs usually have limited sensing capability in terms of limited number of cognitive radios, fixed working frequency range at one sensing attempt, etc. To scan a wide spectrum containing multiple channels in a short sensing window, SUs are usually grouped for cooperative sensing and sharing the sensed channel status with each other. It is obvious that clustering the sensing nodes can speed up the sensing process and in turn improve the detection performance, but the tremendous computational and communication loads are then introduced in the node coordination and resource management. Usually, to facilitate practical deployment, these operations are preferred in a distributed way given the limited control bandwidth and challenges the protocol design, which further increases the design difficulty in the sensing scheduling and other MAC operations.

Opportunistic Transmissions

Right after available spectrum is detected, CRN should schedule secondary transmissions in these idle channels. SUs needs to adopt an opportunistic manner in their transmissions with short occupation period, frequency hopping and recovery in noncontinuous resource blocks, which enable SUs to fast respond to incumbent PUs in the working channels. Several challenging issues lie in SUs' opportunistic transmissions. First, the definition and measurement of the utility of SUs working in heterogeneous wireless channels is still an open issue. In CRN, spectrum access opportunities are normally distributed in different frequency bands showing various capabilities of serving the unlicensed traffic. Since the utility also varies with the required QoS in different SU deployment cases, the definition needs to consider both the channel condition and the user service requirement while integrating them into the major goal of improving spectrum utilization. Second, when multiple SU pairs are active in the network, balancing spectral efficiency and fairness among SUs in opportunistic transmissions also should take heed with practical considerations. In the detected idle channels, SUs highly intend to transmit in the resource blocks with the maximum utility, which is highly possible to introduce congestion and longer delay in some most popular channels while wasting the opportunities in the other less popular channels. Third, packing different traffic into different size of resource blocks largely affects the total spectrum efficiency by recognizing that SUs may have vast different views of local spectrum opportunities. To design realtime allocation schemes in CRN, the adaptive design is essential but challenging, which intelligently combines separate spectrum bands for a large bandwidth request or allocates the spectrum for multiple small traffic to share.

Coordination/interworking

In the control panel of CRN, the coordination between SUs plays an important role to achieve global optimality. Equipped with sophisticated spectrum sensors and cognitive radios, SUs have the knowledge of local channel environment and adjust their operations to increase their utility. However, transmissions led by local utility at individual users would possibly make the aggregated utility for the whole network deviate from the global optimality. Given limited bandwidth for the control purpose between SUs and the channel hopping nature of secondary transmissions, it is a difficult task for the operators of CRN to conduct global control and optimization over SUs in dynamic spectrum resources. Therefore, distributed resource management is preferred, especially when channel conditions changes rapidly in a dense SU deployment case. However, performing effective coordination for distributed resource management exhibits unique challenging issues in CRN. On one hand, SUs can not afford too frequent coordination or global information exchange because the network control is usually sensitive to the cost of bandwidth supply in CRN. Usually, the designated common control channels (CCC) for SUs have limited bandwidth or SUs totally communicate with each other in a frequency hopping way. When SUs are working in multiple channels, the protocol design for accessing CCC becomes a critical problem, especially focusing on avoiding the congested channel and the blind node as the network size increases. On the other hand, for the scalable protocol design, i.e., with more spectrum and larger network size, SUs may vary in detecting local spectrum opportunities and have different capabilities in sensing and opportunistic transmissions. As a result, the coordination scheme needs to adapt to the heterogeneity in the spectrum resources and individual users given the various constraints in control frequency and accuracy.

To summarize, the challenging issues identified in this dissertation are usually not stand-alone, which require to treat a systematic way for the design of network protocols and resource management schemes with specified frequency ranges, supported service types and deployment scenarios.

1.3 Our Approach and Contributions

Recognizing the challenges along with the introduction of exploiting heterogeneous wireless channels in CRN, our approach follows three design principles including 1) local spectrum sensing and small area utilization, 2) cooperative and opportunistic transmissions and 3) effective and efficient coordination, for the potential solutions to achieve better PU protection, higher SU utility and improving total spectrum utilization.

With respect to better exploiting spectrum access opportunities with small spatial footprints, small cell secondary transmissions are urgent for more efficient utilization based upon local detection. Therefore, multi-hop transmissions are necessary to support data transmissions for any connection pair in CRN. First, link quality between SUs depends on the local wireless environment and the capability of relay nodes in the search of spectrum access opportunities. In a network with intermittent links of varying bandwidth, routing in CRN is inherently bonded with the adopted MAC layer protocol. Second, building and maintaining the routing table relies on knowledge of instant channel availability and the statistics of spectrum access opportunities. For example, if an active PU is detected in the selected route, then the ongoing secondary transmission needs to find another route to detour the detected PU or stay in current route silently waiting for the next idle window, which relies on the adopted forwarding strategy. Most importantly, since spectrum access opportunities are valid in the much shorter time window compared with the globally established forwarding route, then building and maintaining a flexible routing table responding to channel variations is the major challenge in the design of CRN routing.

To make opportunistic transmissions within dynamic free channels, the major design principle of the MAC layer protocol is to agree with the varying while limited spectrum bandwidth for effective and efficient data transmissions including 1) the cost-effective resource allocation for spectrum sensing and spectrum usage, i.e., dynamic spectrum access, 2) the collaboration between SUs in sensing and transmissions in the multi-channel multi-hop network with limited control overhead, 3) the QoS provisioning by assigning the available spectrum access opportunities to heterogeneous individual user traffics.

With respect to effective network control over SUs and the interworking between PUs and SUs, CRN need to focus on efficient coordination mechanisms. Unlike conventional wireless networks where radio resources are defined universally over the whole network, spectrum access opportunities are defined as some local resources appreciated by nearby SUs for individual utility optimization.

Therefore, a given available resource block is usually shared within a group of nodes in vicinity so that common control channels only carry the coordination locally, which greatly relieves the signaling congestion. Furthermore, other than to simply report channel states without any processing, SUs can manipulate the sending of transmission requests to the central controller or broadcasting in the common control channel so that the communication frequency can adapt to the bandwidth demand with the reduced signaling overhead. Last not the least, contention for request needs to be well designed by introducing some regulation rules and penalty schemes such that no SU intends to deviate from the decision made by the operators, which further reinforce the coordination effect and distributed control.

In this dissertation, we present our proposed schemes according to the proposed design principles to exploit heterogeneous wireless channels for performance improvement in CRN.

In Chapter 4, an efficient opportunistic cognitive routing (OCR) scheme is proposed to support multi-hop transmissions in CRN. The forwarding decision at each hop is made by jointly considering physical characteristics of spectrum bands and diverse activities of PU in each band. Assuming that CRN working in a wide spectrum, the average hop count and transmission power level vary from channel to channel. To effectively explore the spectrum opportunities, SUs with limited cognitive radio capability follow a proper channel sensing sequence for fast and reliable message delivery in a distributed way. We then develop a greedy forwarding scheme that SUs can select the next hop relay based on the geometry information and channel access opportunity of their one hop neighbors. For the proposed

OCR, as routing control messages are locally exchanged, SUs can efficiently make the routing decision and opportunistically access the available channels. Via extensive simulations, it is shown that our proposed scheme outperforms existing opportunistic routing schemes in CRN by exploiting the heterogeneity of spectrum bands for opportunistic channel access.

In Chapter 5, the discussion on cognitive routing proceeds with the coupled spectrum sensing and sharing in a multi-channel multi-hop CRN. The candidate relay selection is investigated. Based on location information and channel usage statistics, an SU distributedly selects the next hop relay and adapts its transmission to the dynamic spectrum opportunities in its neighborhood. Assisted by a common control channel, node coordination among neighboring SUs in relay selection can be performed in a timely way to reduce the selection delay and improve spectrum utilization. In addition, we introduce a novel metric, namely cognitive transport throughput (CTT), to capture the unique properties of CRN and evaluate the potential relay gain of each relay candidate. A heuristic algorithm is proposed to reduce the searching complexity of the optimal selection of channel and relay. Simulation results demonstrate that our proposed OCR well adapts to the spectrum dynamics and outperforms existing routing protocols in CRN.

In Chapter 6, to support multimedia applications in CRN, we propose a distributed QoS-aware MAC protocol for multi-channel single hop scenario. Specifically, based on the channel usage patterns of PUs, SUs determine a set of channels for channel sensing and data transmissions to satisfy their QoS requirements. We further enhance the QoS provisioning of the proposed cognitive MAC by applying differentiated arbitrary sensing periods for various types of traffic. An analytical model is developed to study the performance of the proposed MAC, taking the activities of both PUs and SUs into consideration. Extensive simulations validate our analysis and demonstrate that our proposed MAC can achieve multiple levels of QoS provisioning for various types of multimedia applications in CRN.

In Chapter 7, we further exploit the great opportunities provided by the fast developing cognitive radio technology in conventional cellular communications. Smartphone fever along with roaring mobile traffic pose great challenges for cellular networks to provide seamless wireless access to end users. Operators and vendors realize that new techniques are required to improve spectrum efficiency to meet the ever increasing user demand. In this chapter, specifically, we first present challenging issues including interference management, network coordination, and interworking between access networks in a tiered cognitive cellular network with both macrocells and small cells. Taking into consideration the different network characteristics of macrocells and small cells, we then propose an adaptive resource management framework to improve spectrum utilization efficiency and mitigate the co-channel interference between macrocell and small cell users. A game-theory-based approach for efficient power control has also been provided.

1.4 Thesis Outline

The rest of the dissertation is organized as follows: Chapter 2 introduce the background knowledge and basic concepts of CRN. Chapter 3 gives a brief system description we address in the research. Chapter 4 presents an opportunistic routing scheme utilizing heterogeneous wireless channels. Chapter 5 presents the joint design of channel selection and route selection to support multi-hop transmissions in CRN under dynamic channel conditions. Chapter 6 investigates on the MAC layer protocol design and analysis in support of multimedia services with QoS provisioning. To highlight the impact of CRN research on current wireless communications, we introduce the idea of dynamic spectrum access in cellular heterogeneous networks in Chapter 7. Finally, Chapter 8 concludes the thesis, and points out our future research directions.

Chapter 2

Background

In this chapter, we introduce the basic concepts and background knowledge of the emerging cognitive radio networks and the dynamic spectrum access technology. We focus on the fundamental procedures in secondary transmissions including spectrum sensing and sharing developed in the existing works, which shed light on our research work in the design of MAC layer spectrum sharing, network layer routing and radio resource management schemes.

2.1 Overview of Cognitive Radio Networks

CRN are proposed to provide economic spectrum access and flexible network deployment for the ever growing wireless traffic demand. Under current static spectrum allocation policy, spectrum is divided into isolated frequency bands for exclusive access. Appropriate spectrum for wireless communications are scarce resources, and bidding for extra spectrum resource usually means billions of dollars to pay for the license. Spectrum scarcity resulting from the near-full allocation of spectrum becomes the notable bottleneck to further spur wireless communications. But the allocated frequency bands are unevenly deployed with varying utilizations from 15% to 85% in time and space domains [2]. Even in crowded deployment cases, e.g., in urban area, there are still many spectrum access opportunities for unlicensed transmissions in time/frequency/space domains. For instance, spectrum occupancy is only 13.1% between 30 MHz and 3.0 GHz in New York City [3]. As we are far away to building a dynamic spectrum allocation policy which allows to allocate spectrum on-demand at realtime pricing, CRN introducing spectrum reuse into current underutilized bandwidth has been recognized as one important approach to feed the spectrum hunger [4].

In CRN, the wireless nodes are divided into two separate user groups according to the channel access privilege in the frequency bands: PUs and SUs. And each user group form a network layer in CRN with its own services and transmission utility. PUs are the licensed users who have the exclusive priority when they access the channels. For SUs, on the contrary, they do not have licensed channels. But thanks to cognitive radios, SUs can still identify the temporarily free channels without active PUs, and use them for unlicensed transmissions. The free channels usually span over the spectrum and open a limited time window for unlicensed transmissions, which are referred to as spectrum access opportunities. The terms, e.g., spectrum holes and "white space" coined by the FCC [5], are also commonly used to denote the temporally available channels. In the rest of this dissertation, they are used exchangeably.

CRN work in a prioritized channel access model which enables SUs to be able to access the licensed spectrum bands but requires SUs to quickly respond to the channel dynamics and properly manage the node operations to prevent harmful interference to PUs. Therefore, SUs are designed to detect free channels before the transmissions and keep monitoring the channels in case PUs turn on. When multiple SUs are active in the network, the coordination and networking among SUs largely impacts the utilization of dynamic spectrum resources as well as the service utility of SUs. The hierarchical channel access model is the distinguished feature of CRN and has the practical importance in nowadays wireless communications [6, 7].

A centrally controlled CRN architecture has been standardized to exploit the vacated TV bands in a wireless regional area network. The named IEEE 802.22 system specifies the physical and MAC layers for a fixed point-to-multipoint (PMP) air interface, in which a base station (BS) manages spectrum sensing in addition to the traditional cellular BS features, such as resource allocation and call admission control. In parallel, there exists the need for distributed solutions which enables the secondary transmissions in mesh or ad hoc networks [8]. Since spectrum opportunity is usually valid in a short time window, distributed control and local identification of spectrum resources can make flexible spectrum access and reduce the deployment cost of the infrastructure.

CRN also shed light on designing more efficient wireless systems to satisfy the ever increasing demand of high data rates and lower power consumptions in mobile wireless networks [9]. For example, the performance of the femtocells deployed for indoor cellular signal enhancement is limited by available spectrum resources [10]. Recently, a practical CRN architecture is introduced into the femtocell networks to provide extra out-of-band spectrum resources for fixed and mobile subscribers [11]. By deploying short-range low-cost low-power femtocell base stations (FBSs), or named as home eNode B (HeNb) in 3GPP LTE specifications, the end users of femtocells can use the sensed spectrum opportunities to meet the traffic demands of Internet-based services.

Building CRN is a complex task including a wide range of research issues, such as smart antenna design, wireless signal processing, spectrum sensing and measurement, medium access control (MAC), network discovery and routing, self-organizing and learning, as well as software-defined radio (SDR) platform design. Among them, the exploration and exploitation of spectrum access opportunities is at the heart of the CRN study. Modern signal processing technologies have already proposed promising solutions to detect and trace channel usage patterns. Currently, how to find the desired spectrum opportunities in an efficient way is still an open challenging question, which is also the major concern in this dissertation.

2.2 Dynamic Spectrum Access

In CRN, the unlicensed transmissions of SUs are constrained by the behaviors of licensed users residing in the same frequency bands. But, with the introduction of sophisticated channel sensing schemes, e.g. wide-band sensing, the capability of SUs to sense and detect wide spectrum space is enhanced. At one moment, it is highly possible to find abundant white bandwidth to exploit for CRN. Then, the bottleneck then moves to the aspect how to timely detect the existing opportunities and transmit data in the limited time window. Furthermore, the varying bandwidth supply caused by the channel usage pattern and the capability of cognitive radios requires SUs to utilize more intelligent and aggressive transmission strategy to exploit the sensed resources. As a result, the dynamic spectrum access (DSA) technology was proposed in CRN to perform opportunistic transmissions between unlicensed users given the sensing results of the channels.

2.2.1 Spectrum Access Models

DSA focus on the spectrum reuse by unlicensed users. Identifying and characterizing spectrum access opportunities are the base of enabling DSA transmissions in the layered spectrum access architecture. According to the way of exchanging information between PUs and SUs for the spectrum reuse, there are three channel access models in CRN: interweave, underlay and overlay [6, 7, 12].

In the interweave mode, PUs and SUs are operating independently and there is no signaling interaction between the two layers. SUs identify the free channels through spectrum sensing. And the channel status is labelled as either “ON” or “OFF” given that PUs are active or not. SUs are allowed to transmit in a channel only when it is "OFF". Therefore, the spectrum holes show the clear bound in time and frequency, which is an ideal case on the assumption of the interference model. In spectrum sensing, SUs use cognitive radios to physically detect the existence of PU's signal, e.g., energy detection, feature detection, etc. The cost of performing the sensing is significant but inevitable in an SU's operation. Because the interweave mode requires no modification at PUs and protects them in the time-frequency resource blocks, SUs in the interweave mode are transparent to primary systems and can be deployed in any licensed spectrum bands open for CRN, which is promising. Thus, the interweave mode is widely used in the research works for CRN [13].

In the underlay mode, PUs acknowledge SUs a criterion named as the "interference temperature" for each licensed channel, which indicates the maximum allowable interference at the PU's receiver [14]. Correspondingly, SUs adjust their transmission power to prevent from violating the threshold set by the nearby PUs. Compared with the interweave mode, SUs are allowed to transmit in the same frequency bands with PUs at the same time, which can significantly improve spectrum utilization. For example, ultra-wideband (UWB) communication is a typical underly transmission. However, SUs need to keep monitoring the link gains with the neighboring PUs and estimate signal-to-interference-noise-ratio (S-INR) at the PU's receiver, which means that SUs needs a strong inter-layer connection to make the transmission decision. Therefore, the underlay mode is generally used in the discussion of the theoretical performance boundaries of CRN.

The third approach is the overlay mode in which PUs proactively participate into the access decision of SUs in CRN, or named as cooperative CRN (CCRN) [15, 16]. With the purpose of gaining transmission opportunities, SUs negotiate with PUs for secondary transmissions by relaying PUs traffic through cooperative communication techniques, such as advanced coding or cooperative relaying. Usually, sensing error is inevitable in practical network operations, such as estimation, quantization and delayed feedback, which has negative effect on the network performance. Cooperation can leverage CRN to avoid the sensing error problem and gain spectrum access opportunity for SUs. However, the cooperation requires the significant changes in PUs' protocol stack to enable the interworking with SUs.

Many Internet services usually require various requirements on the quality of services (QoS) for the wireless networks, such as the realtime connection with the servers residing in the wired core network. According to the information available to secondary transmissions on channel usage, the exploration of spectrum access opportunities can adaptively selects the access mode from the three access modes above to agree with the availability and the amount of information between the layered networks in different deployment cases. In this dissertation, the discussions in Chapter 4-6 are focused on the interweave mode since PUs have the strictest interference protection with least requirement on the protocol stack to open the spectrum for SUs. While in Chapter 7, CRN are extended to the heterogeneous cellular networks in the underlay mode because the nodes have default link measurement procedures and backhaul connections.

2.2.2 Spectrum Access Opportunities

Spectrum Opportunity Usage Pattern

Since CRN usually work in multiple spectrum bands for opportunistic transmissions, the spectrum access opportunities exhibit variations in the spectrum usage patterns in terms of radio environment and the statistics in the PU behaviors. Generally, from the point of view of an SU, the channel availability for the secondary transmission can be modeled by an alternating ON/OFF process as shown in Fig. 2.1. As a general definition [17], the sojourn time of an OFF period, when the channel k is free of PUs and available for secondary channel access, can be modeled as a random variable T^k_{OFF} with the probability density function (PDF) $f_{T^k_{OFF}}(x),\ x > 0$. Similarly, the PDF of the duration of an ON period, when the channel is busy, is given as $f_{T^k_{ON}}(y),\ y > 0$.

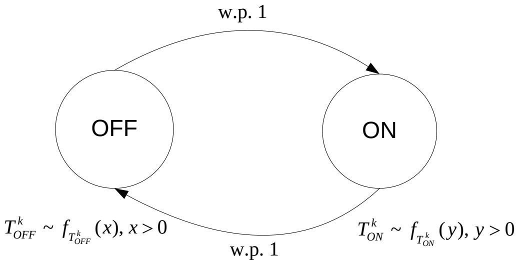


Figure 2.1: Channel usage pattern.

According to the field measurement [18, 19], although most licensed channels appear to have the idle duration from time to time, the statistics of the open window are not the same. The ON/OFF period varies over different bands according to the specific PU behaviors. It can be intuitively assumed that the spectrum opportunity released from a TV band would be different from the one from a busy cellular network band. Such difference shows the channel diversity in the support of traffic demand of SUs in CRN.

Heterogenous Spectrum Access Opportunity

In addition to channel occupancy of PUs, the heterogeneity of frequency channels provides extra diversity gain in secondary transmissions. Features of channel frequency diversity are addressed in the propagation models in the wireless communications [20, 21]. The channel impulse response to the radio signals is different over frequency bands, due to the frequency selectivity of small-scale fading, as well as possibly through the frequency dependence of the path loss. Literature [22] lists important research issues in the heterogeneous multi-channel networks.

It is prospected that the SDR technique will enable the SUs to search the valuable spectrum opportunities and multiple CRs can tune to the totally isolated spectrum bands to serve for the same SU. In this scenario, the spectrum holes in different frequency bands will show the heterogeneity in their transmission characteristics. Currently characterizing the heterogeneous spectrum resources in CRN is still an open problem. As a simple example, the SU would view different neighborhood in the secondary links since the PU activity and the channel transmission characteristics vary with the channel bands tuned on. And for a single neighbor node, there are several channels with different transmission quality to enable the link. The authors in [23] propose a new metric to measure the weights of the heterogeneous resources in the resource scheduling and the route selection. The upper layer performance, e.g., the transmission control protocol (TCP) performance for the end-to-end transmission, is evaluated in the literature [24], and the results show that exploiting heterogeneous channels can improve the performance with carefully defining the heterogeneous interfaces. Literature [25] concerns the impact of the secondary transmission introducing interference on the PU transmissions in the adjacent frequency bands. To mitigate such cross-channel interference, the packet delivery of the SU takes a detour to avoid the PU reservation regions with the account of the cross channel interference fading with the frequency space.

In general, the heterogeneity of the spectrum opportunities can be categorized as follows:

- ▷ Transmission characteristics of the channel: the channel capacity, the transmission distance on the given power level, and the supported modulation & coding schemes.
- ▷ Channel operation policy: channels divided with fixed bandwidths or the flexible bandwidth schemes which combine the adjacent idle channels to improve the bandwidth

- ▷ PU activities: the average channel availability, the average time held for the spectrum opportunity, and the maximum power allowed for the SU.
- ▷ Topology Related Weights: different weights for the same channel in different links

2.2.3 Spectrum Sensing

SUs keep monitoring the frequency bands to trace the channel status and look for the spectrum holes through spectrum sensing. Spectrum sensing is to detect the active PUs in the sensed channels. Depending on the signal processing schemes used in the sensing process [6], the sensing techniques can be classified into three categories: energy detection, matched filtering, and feature detection. Among them, energy detection with the threshold-based decision is widely applied because of its low computational and implementation complexities [26]. With energy detection, SUs sense the presence/absence of PUs based on the energy level of the received signals. In addition, it does not require any knowledge on the PUs' signals. A detailed introduction on the existing spectrum sensing techniques can be found in [13].

Most research activities in spectrum sensing focus on how to efficiently and accurately detect the spectrum opportunities for medium access in the physical and MAC layers with or without user cooperations. In [17], to maximize the discovery of spectrum opportunities, a sensing-period adaptation mechanism is proposed as well as an optimal channel sequencing design. A channel usage pattern estimation technique is also proposed being aware of the environment on the channels, which provides a good reference efficient link layer scheduling.

Individual SUs have limited sensing capability, it is difficult and costly to scan the whole frequency range consisting of multiple channels at one sensing effort. Given the fact that SUs in the neighborhood usually experience the similar primary activities, sharing the sensing results in the local area is promising to improve the sensing accuracy, extend the sensing spectral range and reduce the delay. In [27], multiple SUs are scheduled to sense the same set of channels in a sensing period before the transmission is performed. In a fixed time window for spectrum sensing and transmission, the channel status would be slowly varying in PU activity and the spectrum access decision is made by the soft combination of multiple sensing samples. The optimal sensing period among these channels is determined to achieve the maximum throughput performance which maintains the sensing accuracy. In the literature [28], the authors propose a cooperative sensing scheme to allow multiple SUs who share the identical channel status to sense different parts of the spectrum bands and exchange the sensing results to identify more spectrum opportunities and increase the decision accuracy. To cover as many spectrum bands in one sensing slot, SUs are scheduled to sense different sets of channels. According to the number of SUs taking part in the cooperative sensing, the optimization is performed in the sensing strategy. Such enhanced sensing scheme with node cooperation is incorporated in the discussion of the call admission control (CAC) schemes in the literature [29]. Moreover, in literature [30], the authors separate the sensing function from the SUs and form an independent spectrum sensing networks with spectrum sensors. The main research issue in this work is to optimize the placement of the spectrum sensors to minimize the hidden terminal effect in the PU detection.

Recently, researchers extend the discussion to the impact of the traffic QoS on the sensing strategies in CRN. In an underlay mode, a queue-aware spectrum sensing in [31] employs the queueing status of SUs to indicate the statistical QoS requirements. By dynamically adjust the energy threshold to decide the presence of the PUs' signals, CRN can achieve different sensing strategies. When the queue length is small, the traffic demand of SUs is not so heavy compared with the capability of secondary transmission by the channels. In such case, a more conservative sensing strategy, i.e., using a lower threshold in the energy detection of the spectrum occupancy, can be used by SUs to mitigate the interference imposed on the primary system. On the contrary, when a large queue length is observed which means more packets are waiting to transmit at SUs, SUs should apply a more aggressive sensing strategy to enlarge the supply of spectrum opportunities while introduce more interference on PUs. The results of the work show that the proposed dynamic sensing scheme can support much higher data traffic loads for SUs than the traditional sensing, however, the interference constraints on the licensed users should be taken. On the other side, in an overlay mode, the spectrum opportunities are defined in time-frequency blocks with clear bound. Therefore, SUs cannot increase the supply of additional opportunities for secondary transmission by introducing more interference on PUs.

Spectrum sensing is normally operating in a narrowband sensing, i.e., the sensor works on one single channel for each sensing attempt. Under the Listen-before-Transmission (LBT) scheme for the opportunistic transmission, the narrowband sensing requires a considerable delay to perform a sequential scanning over multiple spectrum bands to look for spectrum holes. To reduce the detection delay while maintain the accuracy, researchers have turn to study for solutions using multi-channel sensing. For example, to detect the spectrum holes in a wide frequency band, e.g., the ultra high frequency (UHF) bands from 470 to 890 MHz with each channel bandwidth of 6 MHz, a multiband joint detection (MJD) scheme is introduced in [32]. The MJD scheme detects the PU signals over multiple frequency bands rather than just in one band at a time as the narrowband sensing.

A bank of multiple narrowband detectors of an SU are jointly optimized to improve the aggregate opportunistic throughput of CRN under the constraints of individual bands on the interference to PUs. To prevent hidden terminal problem, this work is further extended to incorporate the spatially distributed SUs to jointly detect the spectrum utilization of the individual bands in the wideband sensing which is called spatial-spectrum joint detection in [33, 34]. Such cooperative sensing strategy can improve the sensing reliability over multiple spectrum bands.

2.2.4 Opportunistic Unlicensed Transmissions

Based upon the collected sensing results, SUs first evaluate the spectrum access opportunities and decide secondary bandwidth. Then, the detected free channels are shared by SUs to transmit data in an opportunistic way.

Spectrum Opportunity Evaluation and Decision

Once spectrum access opportunities are identified, SUs first evaluate the utilities and select the most appropriate frequency channels by the QoS demands. One of the criteria in the evaluation is to minimize the interruption probability in the secondary transmission if some service will be carried on in the candidate spectrum opportunities. It is closely related to the expected duration of the spectrum opportunity and the traffic packet pattern. Furthermore, the statistical behaviors of the PUs should be calculated into the priority weights of the spectrum opportunity selection. By modeling the spectrum opportunities, the SU can choose the feasible spectrum holes for the packet transmission, or timely switch current transmission to another available spectrum hole if the PU is detected.

Researchers have been searching for the spectrum prediction and decision algorithms in the field of the artificial intelligence. The trade-off is considered between the cost and benefit in the design of spectrum sensing strategies. Using self-learning algorithm to trace the channel status, [35] proposes a channel predication scheme using Partially Observable Markov Decision Process (POMDP). In this framework, the approach considers the spectrum sensing and spectrum access jointly. Since no information exchange is required between the SUs, the individual SU learns the spectrum status by its own sensing process at its position. In a timely slotted channel model, the spectrum opportunity decision is generally a Markov decision process using partially observable spectrum sensing results in each sensing slot. The learning process of the primary activity is reinforced over a sufficiently long time totally in self cognitive way. The systematic optimization relies on sensing policy, reliance policy of sensing vector and action policy.

In [36], the spectrum opportunity management in the multi-hop transmission is formulated as a decentralized optimization problem in CRN, and the learning process is addressed as a fictitious play problem. The utility of each potential spectrum holes are evaluated based on the observation at the node and by the information exchanged from its neighborhood. The distance in the unit of hops is analyzed with the value of the information, and a concept "information cell" is proposed to calculate the value of the status report from the neighbor nodes.

Spectrum Sharing

In CRN, the detected spectrum opportunities are collected in a pool for unlicensed transmissions. Since multiple SUs may request the bandwidth at the same time, the spectrum sharing scheme should be installed to coordinate the simultaneous SU pairs in the spectrum resource pool as well as to avoid interference to PUs. Currently, the coexistence of several SUs has received little attention as compared to the effort on mitigating interference to PUs. But spectrum sharing largely affects the performance of CRN, especially in a dense deployed network. Thus, this feature requires the MAC protocol in CRN to schedule each single SU's behavior in spectrum sensing, access request/contention and transmissions [6, 37].

Specifically, in time, in the same frequency bands, due to the appearances of licensed users, the transmissions are needed to interrupted and resume after licensed user's transmission or in another free frequency band. In frequency bands, as the intensity of PUs' activities, they may show different spectrum access opportunities for unlicensed transmission. Since the transmissions of PU are affecting the covering area, so that the spectrum access opportunities also show the diversity in different locations. Therefore, the secondary transmissions have the rules to agree with these features of transmission. Accordingly, to fight against the time interruption, the estimation of time window and the multichannel to provide continuous bandwidth for unlicensed transmissions are proposed. to protect the PU transmission, the listen-before-transmission schemes are widely used in cognitive radio transmissions. Frequency hopping is used to take the spectrum opportunities appearing in different frequency bands. While to take the local spectrum hole in a frequency band, the unlicensed users sense and transmission in locally found free bands. Thus, multi-hop transmission will be considered in the transmission scenario for the nodes not within the transmission range.

Recently, researchers pay more attention on QoS provisioning in the MAC design for real time multimedia applications in CRN. In [38], voice capacity, the maximum number of voice connections that can be supported with QoS guarantee, is analytically derived in a voice only CR network, assuming there is only one available spectrum band shared by both PUs and SUs. In [39], multimedia content are distributed over multiple unused spectrum bands based on digital fountain codes.

2.3 Summary

In this chapter, we have surveyed the basic concepts of CRN and the related work on the discussions of spectrum sensing and sharing. Extensive research has been done for secondary transmissions in CRN. However, these works cannot be directly applied to solve the challenges in this dissertation. Based upon these works, we explore the solutions from routing protocol design, MAC protocol design and resource management schemes.

Chapter 3

System Model

In this chapter, we list the major assumptions on the system model to study the protocol design, analysis and algorithm development to improve spectrum utilization in CRN.

3.1 Network Model

In this dissertation, we consider that CRN consist of PUs and SUs residing in the same geographical area on orthogonal wireless channels. SU works in the interweave mode, i.e., it opportunistically transmits when it determines that no PU's signal is present in the interested sensing channel through energy detection [12]. Each SU is equipped with one half-duplex cognitive transceiver which can switch to one channel at one time for spectrum sensing or opportunistic transmissions. SU collects the channel availability individually in its vicinity with the specific sensitivity determined by the PU system operating in the same channel. In CRN, spectrum access opportunities are shown as the idle frequency channels with limited open time windows in a region unaffected by PU activities.

The network topology can be with or without infrastructure support for SUs. In terms of the network scale, it can be further divided into single hop and multi-hop scenarios.

Fig. 3.1 illustrates an infrastructureless CRN in which SUs, either fixed or mobile, are randomly distributed and are autonomously operating in ad hoc mode without a central controller. SUs can communicate with each other only when they are located within the limited transmission range of each other determined by the sensing and operating in the same free channel. Different spectrum access opportunities may appear at different

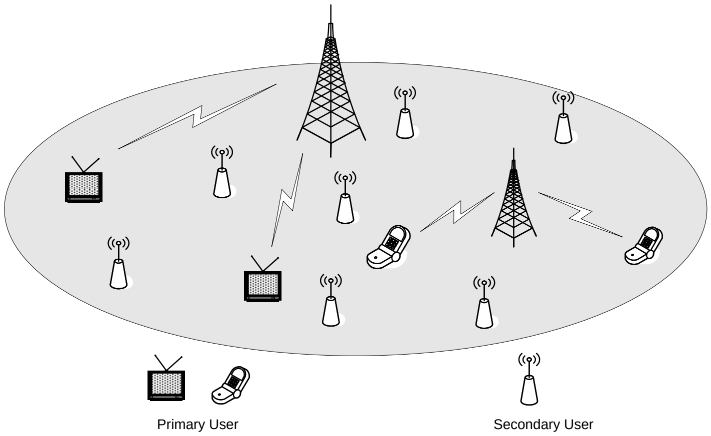


Figure 3.1: CRN topology without infrastructure.

positions by the spectrum sensing of individual SUs, which results in a potential network heterogeneity. The connectivity among SUs may vary with the time in different channels.

Fig. 3.2 illustrates an infrastructure-based CRN. A number of access points (APs) are deployed in CRN serving as the Internet portals so that SUs can connect to the wired backbone network for global communications through these APs. Since the APs also working in the unlicensed bands at low power, only SUs within the coverage of an AP have direct links. Such deployment case can be found in the cognitive radio femtocell Networks [11]. The SUs out of the AP cells would transmit their packets to APs either by multi-hop transmissions via relay or hold the packet until they move into an AP cell.

3.2 Channel Model

In CRN, PUs and SUs share a set of orthogonal wireless channels at different frequencies, $C = \{c_1, c_2, ..., c_m\}$. An hierarchical channel access model is applied in CRN, i.e., PUs are granted with higher and exclusive channel access priorities when they are active and SUs can not disturb PUs' transmissions. Once a PU is turned on in a channel, SUs nearby are not allowed to transmit in the same channel. And the affected ongoing SU pairs should pause to wait for the channel becomes available again or resume their transmissions in other free channels at the moment.

For each single channel c_i in C, an alternating ON (PU is active)/OFF (PU is inactive) process is applied to model the channel availability for SUs. We assume that the ON and

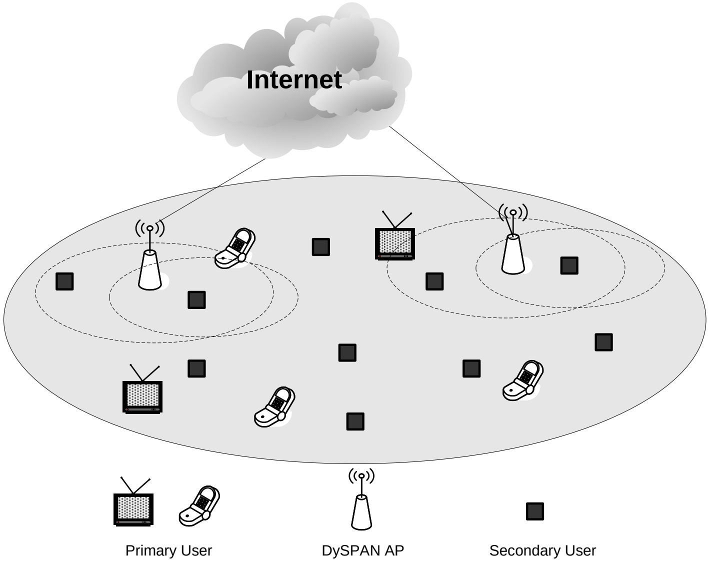


Figure 3.2: CRN topology with infrastructure.

OFF parameters of the channel are slowly time varying and largely dependent on the PU activity and the node density of the primary networks in the channel. The sojourn time of an OFF period, when the channel k is free for secondary channel access, is modeled as a random variable T^k_{OFF} with the PDF $f_{T^k_{OFF}}(x),\ x > 0$. Similarly, the PDF of the duration of an ON period, when the channel is busy, is given as $f_{T^k_{ON}}(y),\ y > 0$. The transitions between the two channel states are shown in Fig. 2.1. At different locations, SUs may experience different available channel sets. Therefore, each SU is enabled to distributively trace the channel states by periodic sensing with energy detection in the similar way as [17]. As Fig. 3.3 is shown, it is a typical view of an SU on the channel status. By sensing the channels independently or collaboratively, SUs make sure that no PUs are active at times/locations/frequencies in the spectrum access opportunities they detect.

In this dissertation, C consists of frequency channels spanning over a wide spectrum range. Different channels exhibit different propagation characteristics, e.g., transmission distances, achievable data rates on the channels. We assume the radio propagation model as

$$P_r(d) = \frac{P_t}{d^{\alpha_p}} \tag{3.1}$$

where P_t is the transmission power of node and the received radio power at a distance is

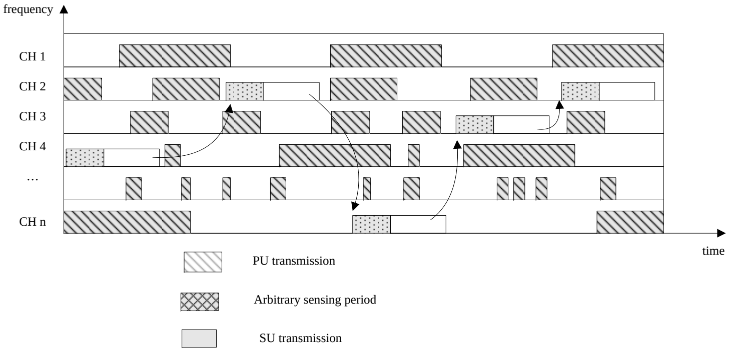


Figure 3.3: Dynamic multi-channel access in CRN

denoted by d.

In free space, the value of the path-loss coefficient parameter, α_p, is selected in the range of $[2, 6]$ [20, 40]. In practise, $P_r(d)$ can be specified at each frequency band using empirical models to indicate the path loss effect in real deployment scenarios, e.g., COST-231 Hata models [20]. Channel c_i has the bandwidth B_i, so according to Shannon's channel capacity, the data rate for a link in a single channel can be represented by

$$R_i = B_i \log(1 + \frac{P_r(d)}{N_0}) \tag{3.2}$$

Each channel has a set of supported transmission patterns in terms of typical data rates, which are fixed and determined by the regulations for unlicensed transmissions. We assume that the control messages are transmitted using the most robust pattern which supports the lowest data rate while the longest transmission distance.

In the multi-hop transmission, as shown in Fig. 3.4, the SU's communication coverage varies in each channel, which results in the neighboring SUs and PUs change with the frequency. On one hand, with more SUs in the neighborhood, it is very likely to find a better relay. On the other hand, more PUs in the communication coverage may reduce the transmission opportunity of the tagged SU since more PU activities may be detected.

3.3 Traffic Model

With the ever-increasing demand of wireless multimedia services, e.g, voice of IP (VoIP) and video conference, SUs may carry various applications, such as voice, video, and data,

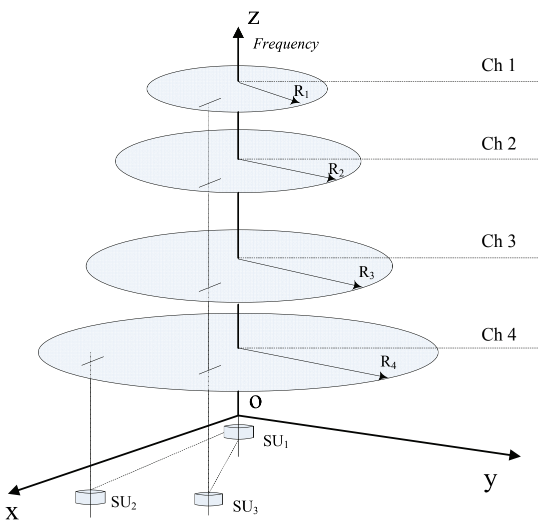


Figure 3.4: Heterogeneous wireless channels.

which have different QoS requirements in terms of throughput and delay. In this study, we consider multiple services supported in CRN.

For voice traffic, we consider VoIP application here. In a VoIP system, the sampled voice signals in analogue are formed into constant-bit-rate (CBR) flows after the compression and encoding. For example, iLBC voice codec is a very bandwidth efficient codec with a rate of 15.2 kbps and has been used in Internet soft-phone applications, e.g., Skype. The voice traffic is modeled as a CBR traffic flow with a packet interval of 20 ms and a packet size of 38 Bytes [41]. To satisfy the QoS demand for voice traffic, according to International Telecommunication Union (ITU) standards [42], the one-way end-to-end delay of voice traffic should be no greater than 150 ms for good voice quality and up to 400 ms for acceptable voice quality, with an echo canceler. Also, the voice packet loss rate should be no more than 1% to maintain satisfactory voice quality.

Compared with voice traffic, realtime video traffic usually demands higher throughput and is modeled as a variable bit rate (VBR) flow with different compression ratios and various payload formats in the codec. In this study, we consider the codec of H.263 and H.264/MPEG-4 AVC, which support very efficient video compression and are applied in a broad range of video applications from low rate Internet video streaming to high definition video (HDV), such as Flash video contents as used on sites as YouTube, Google Video. For example, a H.263 video "Star Trek - First Contact" trace file has the mean rate of 256 kbps and the peak rate of 1.5 Mbps [43]. The mean frame rate and frame size are 25 frame per second and 4420 Bytes, respectively. The general QoS metrics of video applications include throughput, delay, jitter, and packet loss.For interactive applications, e.g., video telephony, the normal tolerable delay should be less than 100 ms and packet loss rate should be below 1%.

We also consider the background data service which is modeled as a simple saturated data flow in the performance study of CRN. Currently, there are many data applications in the Internet services including email, web browsing, file transfer, etc. Each SU is assumed to have a data packet in the queue ready for transmission. The saturated data model is applicable for large volume bulk data transfer applications. Although data traffic is usually delay-insensitive, generally it requires no transmission error. Transmission errors can be improved by the schemes applied in the link layer and the reliable transmission control protocols in the transport layer. In this study, we evaluate the QoS metrics for data traffic in terms of throughput and fairness.

Chapter 4

Opportunistic Cognitive Routing in Heterogeneous Wireless Channels

In CRN, SUs usually operate at low power level to protect nearby PUs from severe interference and only communicate with the peer nodes within a close distance to better utilize local spectrum access opportunities. Therefore, it is highly possible that the destination of the secondary traffic is out of the transmission range of the source in CRN. Multi-hop transmissions would be one typical deployment case for unlicensed users. In this chapter, the routing protocol for multi-hop transmissions in CRN is studies by using the spectrum access opportunities located in heterogeneous wireless channels. Specifically, the proposed opportunistic cognitive routing (OCR) scheme first introduces a novel metric to effectively characterize the forwarding capability of an SU in each single channel, which identifies diverse channel usage patterns and propagation characteristics in the channels. Then, each transmitting SU determines a proper spectrum sensing sequence by the metric to effectively explore the instant spectrum access opportunities. To further improve the reliability and efficiency of secondary data delivery, relay candidates residing in the same channel distributively evaluate the relaying capability based on location information and transmission history. Generally, an SU exhibiting the higher distance gain (i.e., closer towards the destination) and larger success transmission capability are preferred by the sender to be the next hop relay. Since the relay node at each hop is appropriately selected and operates in an opportunistic way, OCR adapts well to the network dynamics, e.g., PUs' activities, user mobility, etc., in a multi-hop multi-channel CRN.

4.1 Introduction

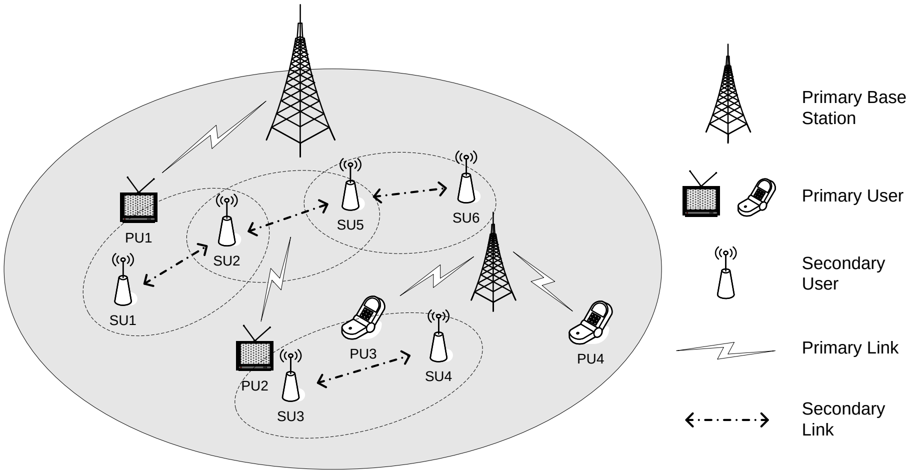


Figure 4.1: Secondary transmissions in an ad hoc CRN

CRN has been emerging as a prominent solution to improve the efficiency of spectrum usage to meet the increasing user demand on broadband wireless communications. In CRN, SUs can utilize spectrum access opportunities for unlicensed transmissions when PUs do not occupy the licensed spectrum. Therefore, the most critical issue in CRN is the exploration and exploitation of the spectrum access opportunities for SUs' transmissions and in the meantime preventing harmful interference to PUs' transmissions [44, 45]. While most research in CRN has focused on a single-hop wireless access network, the research community has recently realized that cognitive paradigm can be applied in multi-hop networks to provide great potential for unexplored services and enable a wide range of multimedia applications with the extended network coverage.

The end-to-end path in CRN is composed of the concatenated links residing in the free channels, which enables the opportunistic transmission. For instance, in an infrastructureless CRN as shown in Fig. 4.1, SUs, $(SU_i, i = 1, ..., 6)$, independently trace spectrum access opportunities in the channels, make access decisions and conduct pairwise transmissions with the neighboring users. For the traffic generated at SU_1 for SU_6, it requires the forwarding path through SU_2 and SU_5 in the network. Since the PU activity varies with locations, the selected route for SUs is intermittent in time. Moreover, although SUs explore a given set of channels, at a moment each single node may also have different views of the channel availability and utility for secondary transmissions. As it is shown in Fig. 4.1, SU_2 needs to get aware of the activity of the nearby PU_1 in its licensed channel. Once PU_1 turns on, SU_2 has to keep silence in PU_1's channel and look for other free channels to link with SU_5. Apparently, a valid link for secondary transmissions requires the SU pair reside in the same spectrum access opportunity and connects to each other within the transmission range.

The key issue for multi-hop transmissions in CRN is to identify the spectrum access opportunities in the channels over the network and coordinate SUs to relay the packets in these intermittent links.

Therefore, the routing decision affects the strategy of spectrum sensing and dynamic resource management at the SUs in the route, and vice versa. In unlicensed transmissions, CRN routing schemes have much closer links with MAC layer protocols. Since MAC protocol in CRN shapes the channel access behaviors of individual SUs, the multi-hop CRN requires MAC to support not only the pairwise transmissions between SUs, but also the multi-hop transmissions for end-to-end traffic flows. Meanwhile, due to the short time window given in the highly dynamic environment, the packet forwarding scheme needs to operate in an opportunistic and efficient way to satisfy the end-to-end performance demand for the stringent traffic QoS. Therefore, we investigate the following issues in the multi-hop transmissions in CRN with both challenges and opportunities.

Multi-hop Transmissions in Opportunistic Channels

In CRN, a link is active only if two SUs are connected in the same idle channel. The path for a source-destination pair is formed by an ordered sequence of the active links in the designated spectrum opportunity over the channels. Thus, the link quality is closely related with spectrum sensing as well as the MAC protocol for multiple user access. As SUs reside in different channels with spectrum opportunities, it results in the network topology which is always varying as different links are enabled at different moments.

Meanwhile, heterogeneous spectrum bands affects the routing protocol design. The spectrum bands may exhibit different propagation characteristics and constraints for spectrum utilization, which results in different achievable data rates in the channels, communication coverage, etc. For instance, in geographic routing, two neighbors with equal forwarding distance gain, the one with better link reliability and higher data rate would be referred by the sender to be the next relay. To route packets in CRN, the routing algorithm should be aware of the PU occupancy in the channels, such as the cost caused by the multi-channel switching and the delay caused by the detour on the channel to avoid the PU reservation region.

The maintenance of routing tables in CRN introduces huge control overhead. Compared with the conventional wireless networks, probing and updating the routes in CRN largely rely on the resources harvested by sensing operations. The frequent communications for routing table update would consume a large amount of spectrum access opportunities which could have contributed to the payload transmissions. Therefore, the efficient routing scheme with less signaling overhead but effective route selection is in a unique position in the design of routing protocols in CRN.

Opportunistic Forwarding in Heterogeneous Channels

The conventional routing schemes usually predetermine the routing tables based upon the link statistics from the long term observations. However, in CRN, such predefined route can not fully exploit the spectrum access opportunities with short time open windows. Moreover, due to the partial knowledge of channel conditions in a wide spectrum, it becomes more difficult for individual SUs to select the right transmission settings to fight against the fast varying channel conditions, which would degrade the performance. Opportunistic forwarding can mitigate the negative effects by forwarding the packets in an opportunistic way. Specifically, in opportunistic routing, the sender broadcasts the packet to multiple neighboring nodes in the direction to the destination and selects the node who has successfully received the packet and shows the most promising capability to forward the packet at the next hop.

Based on the exact packet reception at the neighboring nodes, opportunistic forwarding introduces the multi-user diversity gain and benefits the transmissions when the channel is fast varying. The diversity gain is calculated based upon the per hop forwarding advancement and the link reliability. As most existing opportunistic routing protocols are studied in a single channel scenario, how to extend opportunistic routing in a multi-channel CRN is an open research issue. In a multi-channel structure, transmissions can take place in multiple channels simultaneously using opportunistic routign. Hidden terminals should take heed in the multi-channel case. Since the control communications and the data transmissions are in the same channels, the sender tells the transmission quality in the links from the receivers' feedbacks to verify the relay selection. Thus, if any unrelated transmissions occur in the scheduled feedback period, the sender would be mislead and make the wrong match in the forwarding decision, which negatively impacts the end-to-end performance, such as delay and throughput. The procedures should be carefully designed and verified for the protection of potential hidden terminal problem.

4.2 Related Works

A bunch of existing literatures have discussed the multi-hop transmissions in conventional wireless networks in routing protocol design and algorithm. Recently, researchers have identify the importance of the routing for leveraging performance of CRN, especially the close relationship between route selection and the opportunistic channel access for unlicensed users. In [46], the authors have verified that CRN can achieve the full connectivity, i.e., any pair of SUs can be connected by a sequence of available links, if the deployment densities of PUs and SUs satisfy the given constraints, which validates the multi-hop transmissions in CRN from information theory. [8] specifies the challenging issues associated with the cognitive radio ad hoc networks (CRAHN) or the infrastructureless CRN including optimizing the dynamic network topology for SUs in multi-hop transmissions. In [47], it has further proved that both PUs and SUs in a hierarchical access channel can achieve the same throughput scaling law as in the stand-alone case [48]. It suggests that the performance of the multi-hop transmissions in CRN can be manipulated with the trade-off between the delay and the throughput. Consequently, in the multi-hop CRN, the design of the routing protocol should adapt to the various design goals. Compared to the traditional routing schemes in the wireless networks, routing in the multi-hop CRN is still an open problem gaining more attentions. Hereafter, the major research issues has been addressed in related works as follows.

Joint Route and Channel Assignment

The design and evaluation of the routing schemes in CRN is not an independent issue in the network layer. The route state should be considered in a vector consisting of link connections, working frequency and the available time window. The pool of links largely depends on available spectrum opportunities at the moment, which are determined by spectrum sensing and spectrum management. Therefore, in CRN, the route selection and the channel assignment are usually considered together [6].

The idea that cross-layer solutions jointly optimizing the route and the operating spectrum at each hop has been well received by many papers [49, 50, 51, 52]. In literature [49], the authors confirm that the joint design of routing and MAC scheduling notably outperforms the decoupling design scheme under provisioned settings. The coupling scheme is presented in an interference graph. Given any possible pattern of route and channel allocation, the best pattern is found via the exhaustive search. In literature [50], the authors propose a layered graph model for the joint design. The links in each channel are modeled as in one horizontal plane with the forwarding cost, while the vertical links represent the channel switching operations at the same SU with the predefined switching cost. Then, the optimal solution in the layered graph is the shortest path with minimum cost. In [51], the end-to-end delay for multi-hop transmissions is accumulated by the operation delay during the spectrum sensing and the transmission delay at each hop. The authors formulate the routing problem as the optimization problem to minimize the total propagation delay along the end-to-end path.

These joint design solutions are usually based on two assumptions: 1) the entire network topology is known, and 2) the channels are relatively static. Then the central server or any SU in the network can perform the search for the optimal solution and have enough time to update the routing tables along the route to guide the forwarding. However, as the network size grows, it would become more difficult for CRN to perform such optimal selection considering the rapid communication and computational overhead. Furthermore, to make the routing more adaptive to the dynamic spectrum resources in CRN, we still have to design efficient schemes for the route establishment, the route selection and the coordination within the heterogeneous resources. The first issue is related with the efficiency of the end-to-end path set-up while the other two are about building an evaluation framework and utilizing the diversity in the spectrum opportunities to fight against the link fluctuation.

Route Establishment

To establish a route in CRN, SUs first need to know its location and the surrounding network topology. The routing schemes can be normally categorized as proactive routing or reactive routing. In the proactive routing protocols, every node keeps its own routing table recording the topology tree reaching to every other node in the network, and it updates periodically no matter it has packet to send or not, such as DSDV. When it is used in CRN, the Cogmesh protocol in [53] conducts such operation among the SUs in a small area forming a cluster. The frequent interaction in the neighborhood can help the cluster members keep the connections using the dynamic resources. However, this method is inefficient if the long distance route is made. The considerable delay would be expected to get the comprehensive route update for all SUs, and it needs the huge bandwidth consumption no matter there is transmission or not, which limits the scalability of CRN.

Reactive routing, on the contrary, provides the latest link status for the packet transmission ready to go. Literatures [54, 55, 52, 56] are the recent research efforts to use the modified reactive routing scheme to update the knowledge of the topology for the SUs in CRN. The authors show their preference for the destination-oriented approach following the modified AODV protocol. In their proposals, according to the reception quality of the

RREQ messages, the destination node determines the optimal route according to the design objective. The nodes in the route update their routing tables based on the feedback by the RREP messages. Although it has been widely applied in the wireless ad hoc networks, the on-demand routing via the RREQ-RREP communication does not recognize some unique features of CRN: 1) the confirmation delay is comparable with the coefficient time for the link to be stable; 2) the RREQ-RREP transmission in the common control channel does not indicate the link quality in the data channels; 3) the flooding-based RREQ-RREP transmission is costly to operate in CRN. In [57], the authors address the possibility of the utilization of the source-oriented reactive routing schemes, such as the DSR. But it does not satisfy the features either.

The routing schemes mentioned above are all built upon the priori knowledge of the network topology and link quality, and they are the topology-based schemes, i.e., SUs exchange the route probing messages, such as RREQ/RREP, to update the knowledge of the link quality, then the topology. Because the topology-based routing introduce large overhead, it suits to the case with the relatively stable link quality and guaranteed bandwidth. In [58], the authors have proved that these static routing scheme mainly work well when spectrum opportunities is stable and the time window is relative long compared with the time for the route formation, such as the deployment in the TV bands. In some more dynamic channels, however, these schemes would waste a lot of spectrum opportunities and reduce the spectrum utilization. Furthermore, as the channels are affected by both channel fading and PU activity, the widely used assumption in previous works does not hold any more that the route once formed would be rarely changed in the following data transmission along the path. Thus, how to route the SU's transmission is still under investigation with challenges.

The position-based routing scheme shows some potentials to agree with the dynamic links. In position-based routing, or geographic routing, an SU only needs to know its position information, its neighboring SUs' in the channels and the destination node's to make the routing decision. The location service to help the SU get the position information can be performed out of band. And the sender only concerns the selection of the next hop relay in the direction to the destination for the one hop opportunistic transmission so that the round-trip communication to set up the route is not required. Literature [59] provides a survey on the position-based routing in the wireless networks. Currently, with the hardware development in the positioning devices, such as the GPS, it is easy to embed such function into SUs. Combined with the link-level spectrum sensing and management, SUs can pick up the best opportunities at the moment to push a packets closer to its destination.

[25] makes the first attempt to utilize geographic routing in CRN. But the discussion

is limited in the stable channel case. This scheme depends on the RREQ message to probe the path before the data follows. The holding time for the spectrum opportunities is expected to be larger than the route establishing time. Otherwise, the cost to maintain the route would reduce its benefits on the guidance of the data transmission. Admittedly, geographic routing has its own limitation. There exists the loop route problem, and the neighborhood discovery totally depends on the distance calculation while no appropriate link metric to present the PU consideration has been proposed. The simple but efficient route establishment in geographic routing shows the possibility to improve the end-to-end transmission performance while keeping the minimum overhead in the routing decision process. Geographic routing would be an option in the case of dynamic links with the frequent interruption by the PU activity.

Routing Metric & Opportunistic Route Selection

No matter in topology-based routing or position-based routing, it is essential to decide the end-to-end route or the next hop according to some predefined routing metrics [60]. In CRN, the routing metric shows the designer's interest in the end-to-end performance, such as the maximum end-to-end throughput, the minimum end-to-end opportunistic transmission delay and the minimum interruption probability due to PU activity.

In literatures, the characteristics of dynamic links have been addressed in the design. In [61], the author capture the impact of PU activities on the links for secondary transmissions and propose a probabilistic path metric to indicate the reliability of transmissions. In [55], the routing metric connects the path stability with the PU activity. The interference of the wireless links working in the same channel are considered in [62, 63]. The expected transmission count (ETX) is proposed in [62]. Later, [63] is aimed at maximizing the throughput with another metric called the expected transmission time (ETT). In [23], ETT is further extended in a more complicated case with more diversity of the controllable resources. These metrics account for the effect of the PU activity and the maximum duration of the opportunistic transmissions in CRN. In [64], from the perspective of protecting PUs, the throughput degradation of PUs caused by the potential route selected by SUs is considered, and the routing metric incorporates the inter-layer factors in the route selection, e.g., the mutual interference, to make the trade-off of the design objectives in CRN.

Additional diversity gain lies in utilizing multiple candidate routes for one source-destination pair. Opportunistic routing exploit the broadcast nature of wireless communications to assist the route selection process [65, 66]. The source node first broadcasts the packet to some candidates in its neighborhood, and waits for the feedbacks to select the next hop relay from the successful receivers. Geographic information was later introduced in the opportunistic route selection [67, 68]. Since geographic information requires less communication overhead in the route update, in a dense network, it turns out that opportunistic routing combined with geographic routing metric is a good alternative for multi-hop transmission in dynamic links since it does not rely on the global topology knowledge. The end-to-end throughput performance of such combination is analyzed, which confirms such performance improvement [68, 69, 70, 71].

The existing geographic and opportunistic routing schemes are normally designed for the single channel case. In CRN, since the candidate routes are distributed in different channels, it is not technically sound to broadcast packets in each single channel before selecting the relay. And current three-way handshaking in opportunistic routing is designed to get the receivers' feedback in the same channel. If a PU turns on in the working channel, the feedback needs to find another free channel to report the previous packet reception to the sender.

Selfish Behavior in the Relay

The time variance of available resources limits the interactions among SUs, especially in the multi-hop scenario. It largely depends on the SU itself to make the decision to access the licensed bands. In this case, constrained by the traffic requirement (e.g. delay deadline) or the power consumption (e.g. battery lifetime), SU preferably conducts rational (selfish) behavior to maximize its own utility for every detected resource. In the distributed networks, the selfish behavior will cause the poor spectrum utilization due to the collision or unfairness among SUs. In the multi-hop case, the selfish relay nodes result in the low packet rate or long delay for the multi-hop transmissions [72]. To analyze and solve this problem, game theory is a powerful tool which design the coordination among SUs and punish the violations [73]. To solve the selfish node problem in the multi-hop CRN, one of the promising solutions is to design a good incentive scheme for the data forwarding. One feasible incentive scheme is as follows. As SUs can benefit from the local information exchange in spectrum sensing and sharing, they are willing to help relay others' packets if some useful information can be knowledged in the operations. Thus, the local information can be embedded in the data packet header, then the intermediate SU can update its status vector of its neighborhood through the relay operation. Every destination node for the single hop receives the local information from its source node, and inserts its local information before forwarding the packet to the next hop which can maintain the validity of the information for the next hop. By multiple flows in the neighborhood, the SU can acquire a good amount of local information. The incentive scheme design in CRN is still an open problem. The packet header should be carefully designed, and the benefit through the relay incentive should be compared with normal local information exchange. And the security issues are also raised in case of the cheating.

4.3 Opportunistic Routing in CRN

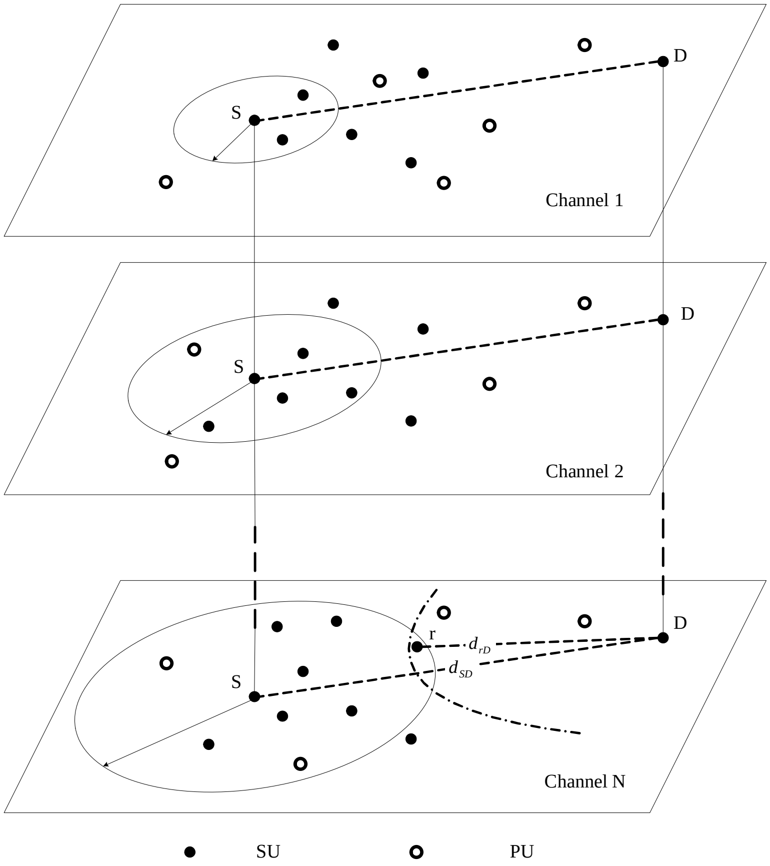


Figure 4.2: Multi-hop CRN over heterogeneous wireless channels.

We propose an opportunistic routing scheme to support multi-hop transmissions in CRN under dynamic channel conditions. We consider a densely deployed infrastructureless multi-hop CRN consisting of N orthogonal channels, $C_1, C_2, \cdots, C_N$, as shown in Fig. 4.2. SUs are randomly distributed over the network. Each SU is equipped with a single half-duplex cognitive radio transceiver which can perform the spectrum sensing or transmission/reception on one frequency channel at one time.

One pairwise coordination scheme is enabled in CRN, e.g., one common control channel to acknowledge the pair of SUs to meet on the detected idle channel. As a dedicated universal spectrum band may not be always available in the licensed bands for SUs, ultra-wideband (UWB) is considered as an ideal option for unlicensed common control channel [74] [1]. Exploiting the spreading technique in the UWB system, senders use their receivers' spreading codes to initiate handshake transmissions over the unlicensed UWB control channel. After a successful handshake, or namely transceiver synchronization, both the sender and receiver tune to the same data channel for communications.

The sensing distance is two times of the transmission distance. In CRN, SUs have the location information of the destination D, i.e., the gateway, and themselves through GPS or other localization services. We assume the the channel hopping sequence of the destination gateway is known to all its one-hop neighbors so that there is no transceiver synchronization problem between the SUs and the destination gateway. As the frequency channels may span over a wide spectrum range, different channels exhibit different propagation characteristics, which result in various transmission distances, and diverse achievable data rates in these channels. As shown in Fig. 4.2, the SU's communication coverage differs in each channel, which results in different sets of neighboring SUs and PUs. On the one hand, with more SUs in the neighborhood, it is very likely to find a better relay. On the other hand, more PUs in the communication coverage may reduce the transmission opportunity of the tagged SU as more PU activities may be detected. An independent and identically distributed two-stage ON/OFF model is applied to all PUs in each channel. SUs distributively trace the ON/OFF parameters by periodic sensing. How to assure accurate channel sensing and parameter estimation has been extensively studied in [17], which is beyond the scope of this work.

We first introduce a channel sensing metric that characterizes the dynamic properties of CRN, based on which a secondary sender can determine a proper spectrum sensing sequence and probe the spectrum opportunities to broadcast a Route REQuest (RREQ) message. We then propose a novel routing metric for selecting the best SU that has the highest opportunistic relay capability, using a distributed medium access mechanism. Basically, the available neighboring SUs operating in the same channel as the sender receive the RREQ message and determine their capability to forward the data towards the destination. An SU with a higher forwarding capability will select a smaller backoff window and transmit a

---

[1] UWB is especially attractive as it is power limited and causes little interference to other communication networks

Route REPly (RREP) messages, and thus is more likely to be selected as a relay in the next hop. After the successful exchange of RREQ and RREP messages, the sender will forward the data to the selected SU. Otherwise, the sender will attempt to re-broadcast the RREQ in the next channel. The procedure repeats until the data message successfully reaches the destination. The pseudo codes of the sender and receiver algorithms are described in Algorithms 1 and 2. More details of the metric design and opportunistic routing protocol will be presented in the following subsections.

Algorithm 1: Sender

```
1: if (Source user $S$ has a packet for Destination user $D$ ) then
2:     $S$ calculates $\gamma_k$, determines the spectrum sensing sequence and starts sensing in the first
       channel;
3:     if Channel is sensed idle then
4:         Broadcast an RREQ message;
5:         if Receive an RREP message correctly then
6:             Transmit data to the SU that responds with an RREP in current channel;
7:         else if Receive multiple RREP messages then
8:             Retransmit an RREQ to notify collided SUs;
9:             Go to Step 5;
10:        end if
11:    else
12:        Select the next channel and start sensing;
13:    end if
14:    Update $\gamma_k$ and reordering the spectrum sensing sequence;
15:    Go to Step 3;
16: end if
```

4.3.1 Sensing Sequence

To effectively probe the spectrum opportunities in a highly dynamic CRN, it is critical to determine a set of channels with an appropriate spectrum sensing sequence for each SU. As radio waves at different frequency bands propagate differently, various channels may exhibit diverse propagation characteristics. For example, some frequency bands can support very high data rate at very short distance, e.g., millimeter wave bands, while others are used for low rate communications over medium and long ranges. In addition, the activities of SUs in CRN is heavily dependent on those of PUs. If the channel is occupied by a large number of active PUs, it is less likely that an SU can find an opportunity in this channel.

Recognizing the diversity in different channels, we propose a new metric to characterize the forwarding capability in each channel, which is given by

$$\gamma_k = \alpha \frac{T_{OFF}}{T_{ON} + T_{OFF}} R_k + (1 - \alpha) D_k \tag{4.1}$$

where T_{OFF} and T_{ON} are the average time durations that the channel is idle and busy, respectively, $\frac{T_{OFF}}{T_{ON}+T_{OFF}}$ is the probability that the channel is sensed idle, R_k is the achievable data rate in channel k with the corresponding transmission distance D_k, normalized by $\max_k D_k$ for $k \in C_N$, and α is a weighting factor. Each SU updates T_{OFF} and T_{ON} via periodic sensing. In each channel, the SU may achieve different data rates at different distances with adaptive modulation technique. Generally, a fewer number of hops is required to forward the data with a larger transmission distance. In this scheme, we choose the largest D_k to calculate γ_k and sort channels in the descending order of γ_k for channel sensing. How to choose the most proper modulation scheme to balance D_k and R_k is under further investigation. SUs select the channel with the highest forwarding capability (i.e., more available network resources and a longer transmission distance), and attempts data communication when opportunity appears. We can balance the available channel resources and the transmission coverage by adjusting the value of α. The impacts of different α on the routing performance will be studied in Section 4.4.

Algorithm 2: Receiver

```
Listening on the current channel;
if An RREQ message is recieved then
    Calculate $\nu_r$ and select a backoff timer;
    while Backoff timer ! = 0 do
        if Overhear other RREP messages then
            Stop the backoff timer and go to Step 1;
        end if
    end while
    Send an RREP message to Sender and go to Step 1;
end if
```

4.3.2 Relay Selection

Once a sender discovers a spectrum opportunity in one channel via spectrum sensing, it broadcasts a RREQ message over the channel. All SUs operating over the same channel

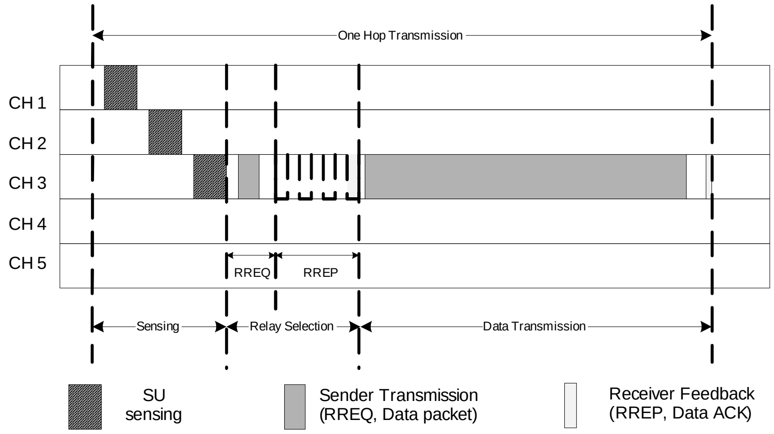


Figure 4.3: Relay selection.

and in the transmission range of the sender are relay candidates which can help forward the data in the next hop. To facilitate opportunistic routing in CRN, each SU needs to respond a RREP message to the sender, so that the sender can select the best candidate to relay the message towards the destination.

In our proposed protocol, an SU, e.g., SU r, estimates its relay capability by calculating

$$\nu_r = \beta P_{s_r} + (1-\beta)\frac{d_{SD}}{d_{rD}}, \tag{4.2}$$

where the successful transmission probability P_{s_r} indicates the opportunistic transmission capability of user r in a multi-hop CRN, which is obtained from the transmission history of SU r, d_{rD} is the distance between SU r and the destination gateway, and d_{SD} is the distance between the source and destination, as shown in Fig. 4.2. β is a weighting factor to balance the user transmission capability and the forward distance gain. SU r is qualified in the relay candidate selection only when ν_r is within $[\nu_{min}, \nu_{max}]$ indicated in the RREQ message. To reduce the possible collisions among multiple relay candidates, we use a contention based MAC for relay selection process. As shown in Fig. 4.3, each SU selects a backoff timer $W_r \in$[0, W] based on the estimated ν_r. The backoff timer reduces by one for every idle mini-slot and an SU can transmit only when its backoff timer reaches zero. Therefore, the SU with the highest ν_r will transmit first and thus be selected as the next hop relay. Other SUs stop their own timer upon overhearing a RREP from other SUs. It is also possible that multiple SUs may select the same backoff timer which causes collisions. If multiple copies of RREP are received at the sender, the sender will re-transmit a RREQ message, and only those collided SUs in the previous round contention will enter the second round

Number of SUs	200
Number of PUs per channel	4
Channel number	6
Delay Threshold	150 ms
λ_{ON}	0.08 /ms
λ_{OFF}	0.04 /ms
Source-destination distance	360 m
Spectrum Sensing Time	1 ms
A backoff mini-slot	4 μs

Table 4.1: Simulation parameters in performance evaluation of Chapter 4

contention by randomly selecting a backoff timer in [0, W]. The process repeats until only one relay is selected. After successfully delivering the data to the next hop, the sender switches to the listening state over the same channel.

4.4 Performance Evaluation

4.4.1 Simulation Settings

We evaluate the performance of the proposed routing protocol under the network parameter settings in Table 4.1 if no other specification is made in the individual study. The simulation model is built in C++ using the reference of the Cognitive Radio Cognitive Network Simulator [75]. The heterogeneity wireless channel patterns, such as the transmission distance and the transmission data rate, are listed in Table 4.2. PUs and SUs are randomly distributed in an area of 300×300 m^2. One connection is initiated in CRN, and the distance between the source-destination pair is arbitrarily set to 360 m. We use a saturated flow with a package size of 500 bytes, and packet delay bound of 150 ms. We model the PU activity as an exponential ON-OFF process with parameters λ_{ON} and λ_{OFF}. SUs individually estimate the ON/OFF durations in each channel. The channel switch time plus the minimum sensing duration with energy detection is 1 ms. Each mini slot is 4 μs. We run each experiment for 120 s, and the first 20 s in each trial run is for SUs to trace PU activities in the channels. We repeat them 50 times to calculate the average value.

Our OCR scheme is compared with the "GPSR+" protocol which is an extensive version of the classical GRSR protocol [76] in a multi-channel CRN. In GPSR+, the sender senses the channels in descending order by the channel maximum transmission distance, and

channel	PU Tx range	SU Tx range	SU data rate
CH1, CH2	50 m	20 m	6 Mbps
CH3, CH4	100 m	60 m	4 Mbps
CH5, CH6	500 m	100 m	0.5 Mbps

Table 4.2: Settings of channel usage patterns in performance evaluation of Chapter 4

greedy forwarding is undertaken by the sender in the relay selection. The relay at the next hop is determined by the sender with the closest position to the destination among all relay candidates in the same channel as the sender.

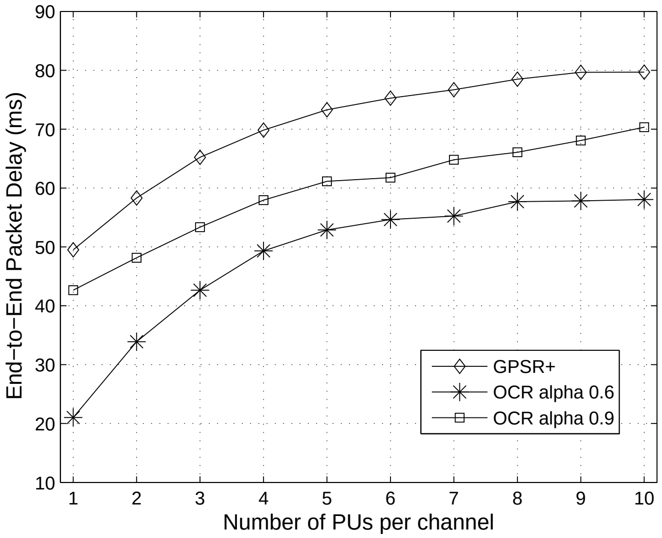


Figure 4.4: Comparison of end-to-end delay under different PU activities.

4.4.2 End-to-end Delay in Secondary Link

We first evaluate the end-to-end delay performance of the proposed OCR scheme under different PU activities. We set the weighting factor α to be 0.6 and 0.9, respectively, and compare the performance with GPSR+. GPSR+ prefers the channel which has the larger SU transmission range in the sensing sequence. As shown in Fig. 4.4, it can be seen that our proposed OCR scheme significantly outperform GPSR+ in terms of end-to-end packet delay, especially in the case $\alpha = 0.6$. Using a greedy forwarding algorithm to select a relay with the highest distance gain, GPSR+ can achieve shortest path with minimum hops in most cases. However, in CRN, more PUs' activities may be detected over a longer communication coverage, and less transmission opportunity can be exploited for SUs' transmissions. In our proposed OCR, by jointly considering the channel characteristics and PU activities, SUs can efficiently exploit the channel access opportunities for data forwarding. By well balancing the two factors, a much lower delay can be achieved. Optimal parameter setting is beckon for further investigation.

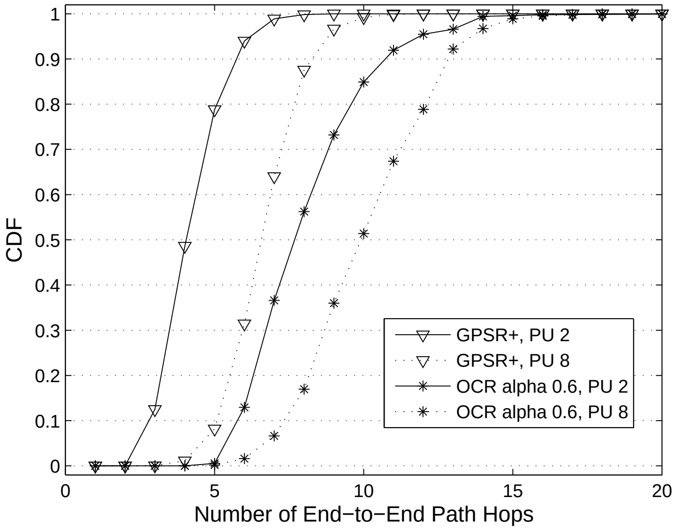


Figure 4.5: CDF of path hop counts under different PU activities.

We then study the impacts of heterogeneity wireless channels on the performance of the proposed routing scheme. With a shorter communication coverage, there is less likely that an SU will be interfered by a PU, and thus the SU is able to explore more spectrum opportunities for data transmissions. Fig. 4.5 shows the cumulative density function of the hop counts for the end-to-end transmission in CRN. There is a marked shift of the average hop count to the higher value which means the transmission distance per hop shrinks due to PU activity.

4.4.3 Performance of SU Density

We also investigate the performance under different node density. The number of PU is 4 in each channel. As shown in Fig. 4.6, the delay drops drastically as the number of SUs

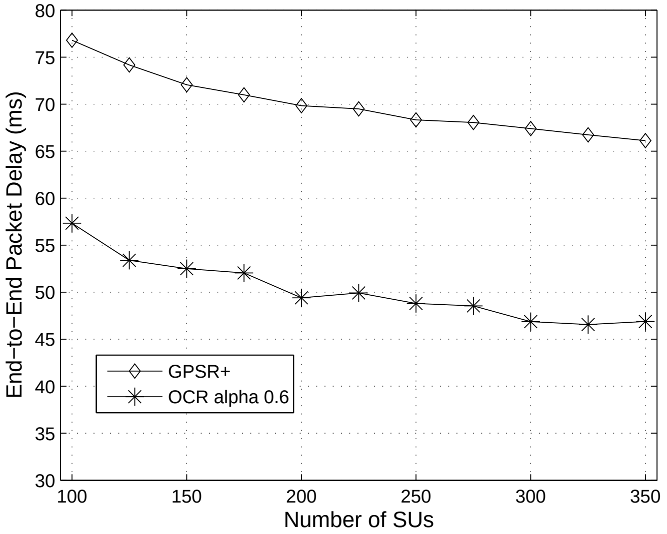


Figure 4.6: Comparison of end-to-end delay under different SU densities.

increases from 100 to 200. Then the decrease rate slows down as the SU number further increases above certain level like 250 SUs here. In a sparse network condition, more SUs can improve the network connectivity, and it is more likely to find a proper relay, thus the number of channel switching and sensing can be significantly reduced and a better delay performance can be achieved. When the network is sufficiently dense, the performance improvement mainly relies on the diversity in the opportunistic relay selection. In this case, further increasing the node density of CRN has little impact on the delay performance of SUs.

4.5 Summary

In this chapter, we have analyzed the features of multi-hop transmissions in CRN and survey the related work on routing secondary traffic over dynamic links. We have proposed an opportunistic routing scheme for multi-channel CRN, by jointly exploiting heterogeneity wireless channel characteristics and geometry information of SUs. By taking spectrum opportunity and selecting the best relay at each hop, our proposed routing scheme can adapt well to the network dynamics and achieve better performance than existing cognitive routing protocols.

Chapter 5

Joint Resource Allocation for Multi-hop Opportunistic Transmissions

In this chapter, we continue with the study of cognitive routing in a multi-channel multi-hop CRN, by utilizing channel usage statistics in the discovery of spectrum access opportunities to improve transmission performance of SUs. The joint channel selection and route selection is considered to optimize the per hop relay performance.

5.1 Introduction

To fully explore the potentials of the multi-hop CRN in support of multimedia applications, it is crucial to study routing in CRN, taking into account the unique properties of the cognitive environment. Existing research efforts mainly focus on effective spectrum sensing and sharing schemes in the physical and MAC layers. Some recent studies indicate that the next major breakthrough in CRN lies in utilizing the diversity gain of spare spectrum in the time, frequency, and space domains to enhance transmissions among SUs [77]. However, in multi-hop CRN, SUs distributed at different locations may have different views of the usage patterns of PUs over multiple frequency channels, which makes it extremely challenging for SUs to coordinate with each other and to exploit the multi-channel and multi-user diversity gain. Some preliminary works on spectrum-aware routing have been proposed for joint channel assignment and route establishment [50, 78, 25, 79]. However, these routing algorithms are based on a pre-determined end-to-end routing table, which is more suitable for static spectrum access system where the channel conditions do not change frequently, e.g., in a CRN operating in TV bands [80]. In CRN, spectrum opportunities of mobile SUs may change over hops from time to time, which leads to expensive maintenance of the routing table. Some recent research extends the work in a wide spectral band under highly dynamic channel conditions other than TV bands [58, 81]. A QoS differentiation scheme and an opportunistic relay forwarding scheme are proposed in our previous works [81, 82], respectively, considering heterogeneous channel usage patterns. These works either mainly focus on the QoS provisioning in a multi-channel scenario or only exploit the diversity of channel propagation characteristics in multi-hop transmissions, which do not specify the impact of the channel usage statistics on SUs' transmissions, especially in a multi-hop CRN.

The remainder of this chapter is organized as follows. The related work is presented in Section 5.2. A multi-channel opportunistic cognitive routing protocol is proposed in Section 5.3. To maximize the relay performance of the OCR, a novel routing metric is designed and the practical implementation issues are discussed in Section 5.4, followed by performance evaluation in Section 5.5.

5.2 Related Work

Routing in CRN can be formulated as a global optimization problem with the channel-link allocation for data flows in the network [83]. Xin et al. [50] propose a layered graph to depict the topology of CRN in a snapshot and allocate multiple links over orthogonal channels to enhance the traffic throughput by establishing a near-optimal topology. Pan et al. [78] propose a joint scheduling and routing scheme according to the long term statistics of the link transmission quality for SUs. Gao et al. [84] develop a flow routing scheme which mitigates the network-wide resource for multicast sessions in multi-hop CRN. These works on cognitive routing pre-determine an end-to-end relay path in CRN based on the global network information. However, the channel conditions of secondary links are highly dependent on PUs' activities in CRN. SUs usually need to track the channel status by periodic sensing [17] or field measurements [77]. When the channel status changes, source nodes need to re-calculate a path. Khalif et al. [58] show that the involved computation and communication overhead for re-building routing tables for all flows is nontrivial, especially when the channel status changes frequently.

Compared with centralized scheduling, distributed opportunistic routing is more suitable for a dynamic CRN since SUs can select the next hop relay to adapt to the variations of local channel/link conditions [70, 66]. Instead of using a fixed relay path, a source node broadcasts its data to neighboring nodes, and selects a relay based on the received responses under current link conditions [70]. Liu et al. [85] propose to apply an opportunistic routing algorithm in cognitive networks where the forwarding decision is made under the locally identified spectrum opportunities. So far, most opportunistic routing protocols have been studied in a single channel scenario. In a multi-channel system, the channel selection and relay link negotiation may introduce extra delay, which degrades the performance of the network. How to extend opportunistic routing in a multi-channel CRN is still an open research issue.

It is also recognized that with available localization services, geographic routing can achieve low complexity and high scalability under dynamic link conditions in various wireless networks, such as wireless mesh networks [76], ad hoc networks [86] and vehicle communication networks [87]. With geographic routing, a node selects a relay node that is closer to the destination for achieving distance advances in each hop. Chowdhury and Felice [25] introduce geographic routing in CRN to calculate a path with the minimal latency. However, their work still focuses on building routing tables and thus is not suitable for dynamic CRN. Considering the unique features of CRN, it is essential to design a distributed opportunistic routing algorithm by tightly coupling with physical layer spectrum sensing and MAC layer spectrum sharing to adapt to the network dynamics in CRN.

5.3 Spectrum Aware Opportunistic Routing Protocol

In this section, an opportunistic cognitive routing (OCR) protocol is proposed where SUs forward the packets in the locally identified spectrum access opportunities. To adapt to the channel dynamics, SUs opportunistically select the relay nodes from multiple candidates according to the distance gain and the channel usage statistics. The main contributions of this work are four-fold: (1) we propose an opportunistic cognitive routing (OCR) protocol in which forwarding links are selected based on the locally identified spectrum access opportunities. Specifically, the intermediate SU independently selects the next hop relay based on the local channel usage statistics so that the relay can quickly adapt to the link variations; (2) the multi-user diversity is exploited in the relay process by allowing the sender to coordinate with multiple neighboring SUs and to select the best relay node with the highest forwarding gain; (3) We design a novel routing metric to capture the unique properties of CRN, referred to as cognitive transport throughput (CTT). Based on the novel metric, we propose a heuristic algorithm that achieves superior performance with reduced computation complexity. Specifically, CTT represents the potential relay gain over

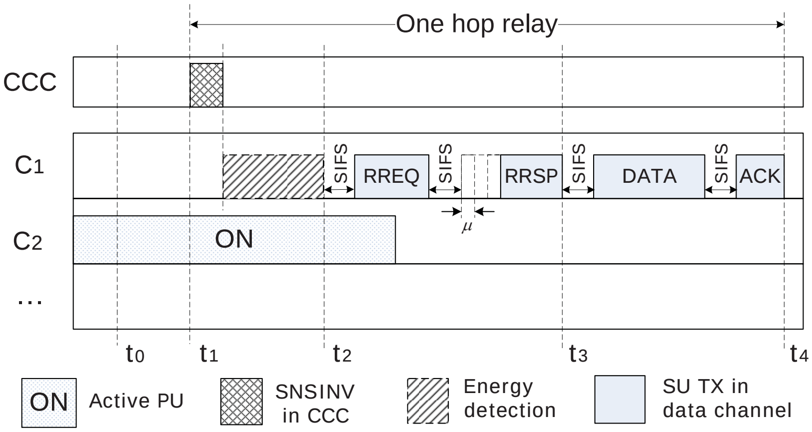


Figure 5.1: The opportunistic cognitive routing timeline

the next hop, which is used in the channel sensing and relay selection to enhance the OCR performance; and (iv) we evaluate the performance of the proposed OCR in a multi-hop CRN. Simulation results show that the proposed OCR protocol adapts well to the dynamic channel/link environment in CRN.

5.3.1 System Description

We consider a multi-hop CRN where multiple PUs and SUs share a set of orthogonal channels, $C = \{c_1, c_2, ..., c_m\}$. SUs can exchange messages over a common control channel (CCC)[1]. Each SU is equipped with two radios: one half-duplex cognitive radio that can switch among C for data transmissions and the other half-duplex normal radio in CCC for signaling exchange.

When a source SU communicates with a destination node outside its transmission range, multi-hop relaying is required. As shown in Fig. 5.1, at each hop, the sender first senses for a spectrum access opportunity and selects a relay node in the detected idle channel[2]. We model the occupation time of PUs in each data channel as an independent and identically distributed alternating ON (PU is active) and OFF (PU is inactive) process. SUs track the channel usage pattern, i.e., ON or OFF, and obtain the channel usage statistics through periodic sensing operations. Generally, the statistics of channel usage time change slowly.

---

[1] The CCC can be implemented by bidding on a narrow spectrum band [57] or accessing the temporarily spare spectrum bands in a predefined frequency hopping sequence [88].

[2] In some extreme case when geographic routing fails to reach the destination, we can apply the right-hand rule for route recovery as proposed in GPSR [76].

The parameter estimation is beyond the scope of this work and the details can be found in [77, 17]. With GPS or other available localization services, SUs can acquire their own location information, and the source nodes have the corresponding destinations' location information, e.g., an edge router or a gateway in the network. A summary of main notations used in the paper is given in Table 5.1 for easy reference.

5.3.2 Protocol Overview

As shown in Fig. 5.1, the per hop relay in OCR includes three steps, i.e., channel sensing, relay selection, and data transmission.

In the channel sensing step, the sender searches for a temporarily unoccupied channel in collaboration with its neighbors using energy detection technique. Before sensing the data channel, the sender broadcasts a short message, i.e., *sensing invitation* (SNSINV), in the CCC to inform neighboring nodes of the selected data channel, and the location information of the sender and the destination. The transmission of SNSINV message in the CCC follows the CSMA/CA mechanism as specified in IEEE 802.11 MAC. Upon receiving the SNSINV, the neighboring SUs set the selected data channel as non-accessible so that no SU will transmit in the selected data channel during the sensing period of the sender. In this way, the co-channel interference from concurrent secondary transmissions can be mitigated. Using the location information in SNSINV, the neighboring SUs evaluate whether they are eligible relay candidates, e.g., whether a relay node is closer to the destination than the sender and thus can provide a relay distance gain. Eligible relay candidates will collaborate with the sender in channel sensing and relay selection. Other SUs cannot transmit in the selected data channel during the reserved time period specified in SNSINV. When the channel is sensed idle, i.e., no PU activity is detected, the sender will initiate a handshake with relay candidates in the relay selection step. Otherwise, the sender selects another channel and repeats the channel sensing process.

In the relay selection step, the sender selects the next hop relay from the relay candidate SUs. Specifically, when the channel is sensed idle, the sender first broadcasts a routing request (RREQ) message to the relay candidates. Eligible candidates reply routing response (RRSP) messages in a sequence specified by the sender. A relay candidate is assigned a higher priority to transmit RRSP after a shorter backoff window if it has a larger link throughput [70], a greater relay distance advancement [76], or a higher link reliability [61]. A candidate SU keeps listening to the data channel until it overhears an RRSP or it transmits an RRSP when its backoff timer reaches zero. The sender selects the first replying relay candidate as the next hop relay. If the sender receives no RRSP message, which implies no relay candidate is available in the selected channel, it will repeat the channel sensing and relay selection steps. After a successful RREQ-RRSP handshake, the sender transmits data to the selected relay node in the data transmission step.

5.3.3 Routing Protocol Analysis

We study the impacts of PUs' activities on the performance of the proposed OCR protocol. In CRN, when PUs appear in a channel, an SU needs to stop its current transmission, update its record of the channel status, and reselect a data channel. Thus, PUs' appearance will result in a larger transmission delay, and involve extra overhead for channel sensing and relay selection. To evaluate the impacts of PUs' activities on the protocol performance, we first introduce the main performance metrics, namely, relay distance advancement and per hop transmission delay. Based on the introduced metrics, we then analyze the success probability in each step, i.e., channel sensing, relay selection, and data transmission.

1) Performance Metrics

We first introduce the relay distance advancement and the per hop delay for performance evaluation. The relay advancement is measured by the geographic distance gain. For a sender S in CRN, $\mathbf{N}_S$ is the set of SUs within its transmission range. The neighboring relay candidate set for the relay to the destination D is denoted by $\mathbf{R}_D \subseteq \mathbf{N}_S$. If an SU $R \in \mathbf{N}_S$ is selected as the relay, the relay advancement $A_D(S,R)$ in terms of the difference in the distances between the SU pairs, (S,D) and (R,D) can be expressed by

$$A_D(S,R) = d(S,D) - d(R,D), \tag{5.1}$$

where $d(S,D)$ and $d(R,D)$ are the Euclidian distances between (S,D) and (R,D), respectively.

The per hop transmission delay T_{relay} is comprised of three parts: sensing delay (T_{SNS}), relay selection delay (T_{RS}), and packet transmission delay (T_{DTX}).

The sensing delay T_{SNS} includes the transmission time of an SNSINV message, T_{init}, and the energy detection time, T_{detc},

$$T_{SNS} = T_{init} + T_{detc}. \tag{5.2}$$

Based on the relay capability, candidate SUs are sorted in a given prioritized order. In the relay selection, the i-th relay candidate R_i sends an RRSP message only when the first $i-1$ higher-priority candidates are not available. Therefore, the relay selection delay $T_{RS}(i)$ is given by

$$T_{RS}(i) = T_{RREQ} + (i-1)\mu + T_{RRSP} + 2\ SIFS, \tag{5.3}$$

where T_{RREQ} and T_{RRSP} are the transmission time of an RREQ message and an RRSP message, respectively, and μ is the duration of one mini-slot in the backoff period. According to [89], the length of a mini-slot can be calculated as $\mu = 2 \cdot \gamma + t_{switch}$, where γ is the maximum channel-propagation delay within the transmission range, and t_{switch} is the time duration that the radio switches between the receiving mode and the transmitting mode.

Once R_i is selected, the packet transmission delay T_{DTX} is

$$T_{DTX} = T_{DATA} + T_{ACK} + 2\ SIFS, \tag{5.4}$$

which includes the packet transmission delay (T_{DATA}) and the ACK transmission time (T_{ACK}).

The transmission delay $T_{relay}(R_i)$ via the relay at R_i is the delay sum

$$T_{relay}(R_i) = T_{SNS} + T_{RS}(i) + T_{DTX}. \tag{5.5}$$

2) Channel Sensing

Denote $I_R^{c_j}$ ($\overline{I_R^{c_j}}$) as the event that c_j is sensed to be idle (busy) by an SU R in the channel c_j. A channel is determined to be idle given that it is sensed idle at the starting time of t_1 and remains idle until sensing completes at t_2, as shown in Fig.5.1. According to the renewal theory, the channel status can be estimated by the distribution of the channel state duration and the sensing history [90]. Specifically, given the channel status (idle or busy) observed at an earlier time, e.g., t_0, we have $P_{OFF,R}^{c_j}(t_0, t_1)$, the probability that c_j is idle (OFF) at t_1, $t_1 > t_0$. Assume ON and OFF durations follow exponential distributions with mean $1/E[T_{ON}^{c_j}]$ and $1/E[T_{OFF}^{c_j}]$[3],

$$\begin{aligned} & P_{OFF,R}^{c_j}(t_0, t_1) \\ = & \begin{cases} \rho_{c_j} + (1-\rho_{c_j}) e^{-\Delta_{c_j}(t_1 - t_0)}, & \text{if } c_j \text{ is OFF at } t_0, \\ \rho_{c_j} - \rho_{c_j}\ e^{-\Delta_{c_j}(t_1 - t_0)}, & \text{if } c_j \text{ is ON at } t_0, \end{cases} \\ & \text{where} \begin{cases} \rho_{c_j} = \frac{E[T_{OFF}^{c_j}]}{E[T_{ON}^{c_j}] + E[T_{OFF}^{c_j}]}, \\ \Delta_{c_j} = \frac{1}{E[T_{ON}^{c_j}]} + \frac{1}{E[T_{OFF}^{c_j}]}. \end{cases} \end{aligned} \tag{5.6}$$

---

[3] which are commonly used in other works [77, 17]

Note that ρ_{c_j} indicates the chance for an idle state in c_j.

We then calculate the likelihood of the channel staying idle during the sensing period. According to the renewal theory, the residual time of a state in an alternating process truncated since the time origin can be expressed by the equilibrium distribution of the state duration [90]. Thus, the probability that the channel at R stays in the idle state during the sensing period $[t_1, t_2]$ can be calculated as

$$P_R^{c_j}(t_1, t_2) = \int_{t_2-t_1}^{\infty} \frac{\mathcal{F}_{OFF}^{c_j}(u)}{E[T_{OFF}^{c_j}]} du, \tag{5.7}$$

where $\frac{\mathcal{F}_{OFF}^{c_j}(t)}{E[T_{OFF}^{c_j}]}$ is the probability density function (PDF) of the residual time of an idle channel since the time origin when it is observed as idle. $\mathcal{F}_{OFF}^{c_j}(t)$ is the cumulative distribution function (CDF) of the duration of the OFF state in c_j with mean $E[T_{OFF}^{c_j}]$, i.e., $\mathcal{F}_{OFF}^{c_j}(t) = \int_0^t f_{T_{OFF}^k}(x)dx$. Then, the probability that R detects a spectrum access opportunity in c_j is given by

$$Pr\{I_R^{c_j}\} \; = \; P_{OFF,R}^{c_j}(t_0, t_1) \cdot P_R^{c_j}(t_1, t_2). \tag{5.8}$$

For the OCR protocol, $Pr\{I_S^{c_j}\}$ denotes the probability of sensing success when the sender S detects c_j as an idle channel. Once the sender finds an idle channel, it will move to the relay selection step. Otherwise, the sender will switch to another channel and initiate the channel sensing process.

3) Relay Selection

After detecting an idle channel, the sender needs to select a relay for data forwarding. In OCR, the prioritized RRSP transmission enables the relay candidate of the highest relay priority to notify the sender its availability for data forwarding. However, active PUs may interrupt the handshaking process and cause the failures in the relay selection when an SU candidate cannot reply due to the detection of active PUs. Such case is very rare, and it happens only when a nearby PU turns on during the selection period. Since the relay selection is very short in time, usually less than 1 millisecond, we mainly consider the case when a candidate SU detects the selected channel which is occupied by an active PU in the sensing. In this case, the candidate will not respond to the RREQ. If no relay candidate responds to the RREQ message at the moment, the relay selection fails. Therefore, we have

$$P^{c_j}_{RSfail} = Pr\{I^{c_j}_S\} \cdot Pr\Big\{ \bigcap_{R_i \in \mathbf{R}^{c_j}_D} \overline{I^{c_j}_{R_i}} \Big| I^{c_j}_S \Big\}, \tag{5.9}$$

where $Pr\{I^{c_j}_S\}$ indicates the probability that the sender initiates the relay selection when it detects an idle channel as defined in Eq. (5.8). In c_j, one feasible relay selection $\mathbf{R}^{c_j}_D = \{R_1, R_2, \ldots, R_n\}$ contains a set of SUs in $\mathbf{R}_D$ with the size of $n = |\mathbf{R}^{c_j}_D|$. Denote V_{R_i} as the priority of R_i in the RRSP transmission. $\mathbf{R}^{c_j}_D$ is sorted in the descending order of V_{R_i}, i.e., $V_{R_1} > V_{R_2} > \ldots > V_{R_n}$. The event that no relay candidate replies in the relay selection step, is equivalent to the event that all SUs in $\mathbf{R}^{c_j}_D$ sense the channel busy in the previous sensing with the probability $Pr\Big\{ \bigcap_{R_i \in \mathbf{R}^{c_j}_D} \overline{I^{c_j}_{R_i}} \Big| I^{c_j}_S \Big\}$.

In the CRN, we assume that an SU is affected by at most one active PU in one frequency band. Such assumption holds in the frequency bands such as the downstream bands in cellular network where the adjacent cells/sectors are usually assigned with different working frequencies to avoid the co-channel interference [91]. Thus, the channel usage pattern is mainly determined by the PU activity at the spot of the individual SU. Let $X^{c_j}_{R_t R_r} = 1$ if a pair of SUs, R_t and R_r, are affected by the same PU in c_j, and $X^{c_j}_{R_t R_r} = 0$ otherwise. $X^{c_j}_{R_t R_r}$ can be acquired and maintained by the periodic exchange of the channel status in the SU's neighborhood. A cognitive transmission is successful only if both ends of the link are not influenced by active PUs. For example, if the channel utilities of c_j at R_t and R_r are $\rho^{c_j}_{R_t}$ and $\rho^{c_j}_{R_r}$, respectively, the link quality of the link l_{tr} can be expressed by $P^{c_j}_{l_{tr}} = \rho^{c_j}_{R_t} \cdot \rho^{c_j\,(1 - X^{c_j}_{R_t R_r})}_{R_r}$. Therefore, $Pr\Big\{ \bigcap_{R_i \in \mathbf{R}^{c_j}_D} \overline{I^{c_j}_{R_i}} \Big| I^{c_j}_S \Big\}$ in Eq. (5.9) is given by

$$\begin{aligned} & Pr\Big\{ \bigcap_{R_i \in \mathbf{R}^{c_j}_D} \overline{I^{c_j}_{R_i}} \Big| I^{c_j}_S \Big\} \\ = \; & Pr\Big\{ \overline{I^{c_j}_{R_1}} \Big| I^{c_j}_S \Big\} \cdot \prod_{i=2}^{n} Pr\Big\{ \overline{I^{c_j}_{R_i}} \Big| \Big\{ \bigcap_{k=1}^{i-1} \overline{I^{c_j}_{R_k}} \Big\} \cap I^{c_j}_S \Big\} \\ = \; & (1 - X^{c_j}_{SR_1}) Pr\Big\{ \overline{I^{c_j}_{R_1}} \Big\} \\ & \cdot \prod_{i=2}^{n} \Big[(1 - X^{c_j}_{SR_i}) Pr\Big\{ \overline{I^{c_j}_{R_i}} \Big\}^{\prod_{k=1}^{i-1}(1 - X^{c_j}_{R_k R_i})} \Big]. \end{aligned} \tag{5.10}$$

Suppose that the i-th relay candidate R_i in the selected relay selection order $\mathbf{R}_D^{c_j}$ is available, R_i will be selected as the next hop relay with the probability $P_i^{c_j}$, given that previous $i-1$ candidates are not available,

$$P_i^{c_j} = \begin{cases} Pr\{I_S^{c_j}\} \cdot Pr\{I_{R_1}^{c_j} \big| I_S^{c_j}\}, & \text{for } i = 1, \\ Pr\{I_S^{c_j}\} \cdot Pr\Big\{\big\{\bigcap_{k=1}^{i-1} \overline{I_{R_k}^{c_j}}\big\} \cap \{I_{R_i}^{c_j}\} \Big| I_S^{c_j}\Big\}, & \\ & \text{for } 2 \le i \le n, \end{cases} \tag{5.11}$$

where $Pr\Big\{\big\{\bigcap_{k=1}^{i-1} \overline{I_{R_k}^{c_j}}\big\} \cap \{I_{R_i}^{c_j}\} \Big| I_S^{c_j}\Big\}$ can be expressed as

$$\begin{aligned} & Pr\Big\{\Big\{\bigcap_{k=1}^{i-1} \overline{I_{R_k}^{c_j}}\Big\} \cap \{I_{R_i}^{c_j}\} \Big| I_S^{c_j}\Big\} \\ = \ & Pr\Big\{\overline{I_{R_1}^{c_j}} \Big| I_S^{c_j}\Big\} \cdot \prod_{u=2}^{i-1} Pr\Big\{\overline{I_{R_u}^{c_j}} \Big| \Big\{\bigcap_{r=1}^{u-1} \overline{I_{R_k}^{c_j}}\Big\} \cap I_S^{c_j}\Big\} \\ & \cdot Pr\Big\{I_{R_i}^{c_j} \Big| \Big\{\bigcap_{k=1}^{i-1} \overline{I_{R_k}^{c_j}}\Big\} \cap I_S^{c_j}\Big\} \\ = \ & (1 - X_{SR_1}^{c_j}) Pr\Big\{\overline{I_{R_1}^{c_j}}\Big\} \\ & \cdot \prod_{u=2}^{i-1} \Big[(1 - X_{SR_l}^{c_j}) Pr\Big\{\overline{I_{R_u}^{c_j}}\Big\}^{\prod_{r=1}^{u-1}(1-X_{R_rR_u}^{c_j})}\Big] \\ & \cdot \Big[\prod_{k=1}^{i-1}(1 - X_{R_kR_i}^{c_j})\Big] Pr\{I_{R_i}^{c_j}\}^{(1-X_{SR_i}^{c_j})}. \end{aligned} \tag{5.12}$$

4) Data Transmission

Once R_i is selected, the data transmission in the link l_{SR_i} succeeds when no active PU appears during the transmission period $[t_3, t_4]$ in c_j. Thus, the successful relay probability at current hop via R_i can be expressed by

$$\begin{aligned} P^{c_j}_{relay,R_i} &= P^{c_j}_i \cdot P^{c_j}_{l_{SR_i}}(t_3,t_4) \\ &= P^{c_j}_i \cdot P^{c_j}_S(t_3,t_4) \cdot P^{c_j}_{R_i}(t_3,t_4)^{(1-X^{c_j}_{SR_i})}. \end{aligned} \tag{5.13}$$

5.4 Joint Channel and Relay Selection

We then jointly consider the selection of the sensing channel and relay node to improve the performance of the proposed OCR. As many factors, including channel usage statistics, the relay distance advances, and transmission priority of relay candidates, may affect the relay performance, we introduce a new metric to capture these factors and apply it in a heuristic algorithm to select the best relay in one data channel at a reduced computation complexity.

5.4.1 Novel Routing Metric

We design a new metric, the cognitive transport throughput (CTT), $CTT(c_j, \mathbf{R}^{c_j}_D)$, to characterize the one hop relay performance of OCR in the selected channel c_j with the selected relay candidate set $\mathbf{R}^{c_j}_D$, in unit of bit·meter/second.

$$\begin{aligned} CTT(c_j, \mathbf{R}^{c_j}_D) &= E\big[L \cdot \frac{A^{c_j}_D}{T^{c_j}_{relay}}\big] \\ &= \sum_{R_i \in \mathbf{R}^{c_j}_D} P^{c_j}_{relay,R_i} \frac{L \cdot A_D(S,R_i)}{T_{relay}(R_i)} \end{aligned} \tag{5.14}$$

The physical meaning of the CTT defined in Eq. (5.14) is the expected bit advancement per second for one hop relay of a packet with the payload L in the channel c_j. To improve the OCR performance, we should maximize the one hop relay performance along the path as one hop performance improvement contributes to the end-to-end performance. In addition, as the multi-user diversity is implicitly incorporated in the relay selection process, we can also achieve a high multi-user diversity gain by maximizing CTT. From Eq. (5.14), we can jointly decide channel c_j and the corresponding relay selection order $\mathbf{R}^{c_j}_D$ to maximize CTT.

5.4.2 Heuristic Algorithm

To obtain c^* and $\mathbf{R}_D^{c^*}$ for the largest CTT, we can exhaustively search for all possible combinations of the sensing channel and the subset of the relay candidate set. Given m channels and up to n relay candidates, an exhaustive search needs to find the locally optimal one in each channel by comparing the value of CTT under all possible permutations of the set of relay candidates. Since the CTT value is sensitive to the set size as well as the permutation, given that k candidate nodes are incorporated in the relay selection, $1 \leq k \leq n$, there are $P(n, k)$ types of opportunistic forwarding patterns. Therefore, over m channels, the exhaustive search should take $m \cdot \sum_{k=1}^{n} P(n, k)$ times of the CTT calculation to return the global optimum. If n goes to infinity, we can get $\lim_{n\to\infty} m \cdot \sum_{k=1}^{n} P(n, k) = \lim_{n\to\infty} m \cdot \sum_{k=0}^{n} \frac{n!}{(n-k)!} = \lim_{n\to\infty} m \cdot n! \cdot \left[\sum_{k=0}^{n} \frac{1}{k!} - 1\right]$. Thus the exhaustive search running time is $O(m \cdot n! \cdot e)$, where e is the base for natural logarithms. We can see that once n becomes very large, the exhaustive search becomes infeasible in real implementations.

To reduce the complexity, we propose an efficient heuristic algorithm to reduce the searching space yet achieve similar performance of the optimal solution. The performance comparison will be given in the following section.

Given independent channel usage statistics in different channels, we can decompose the optimization problem into two phases. First, we compare all possible relay selection orders in each channel and find the optimal one which maximizes the CTT. Then, we choose the relay selection order with the largest CTT value over all channels and select the corresponding channel as the sensing channel. Since the number of channels is usually limited, it is more important to reduce the searching complexity for the best relay selection order in a single channel.

To find the optimal relay selection order, the sender should decide both the number of the relay candidates and the relay priority of each candidate. According to Eq. (5.14), a neighboring SU, R_i, is an eligible relay candidate if it contributes to a positive relay distance advancement, $A_D(S, R_i)$. One feasible relay selection order $\mathbf{R}_D^{c_j}$ in c_j is an ordered subset of $\mathbf{R}_D$ in the descending order of relay priority V_{R_i}. A larger size of $\mathbf{R}_D^{c_j}$ include more relay candidates and achieves a higher diversity gain, which improves the per hop throughput at the cost of increased searching complexity.

To reduce the searching space and improve the algorithm efficiency, we have the following Lemma.

Lemma 5.4.1 *Given a feasible relay selection set* $\mathbf{R}_D^{c_j}$*,* $\exists R_{i_1}, R_{i_2} \in \mathbf{R}_D^{c_j}$*, if* $V_{R_{i_1}} > V_{R_{i_2}}$*,* $X_{R_{i_1}R_{i_2}}^{c_j} = 1$*, then* $CTT(c_j, \mathbf{R}_D^{c_j} \setminus \{R_{i_2}\}) \geq CTT(c_j, \mathbf{R}_D^{c_j})$*.*

Proof 1 *Suppose* $\mathbf{R}_D^{c_j} = \{R_1, \ldots, R_{i_1}, \ldots, R_{i_2}, \ldots\}$. *According to Eq. (5.11), if* $V_{R_{i_1}} > V_{R_{i_2}}$, $X_{R_{i_1}R_{i_2}}^{c_j} = 1$ *and* $X_{R_{i_1}R_{i_2}}^{c_j} = 1$, $P_{i_2}^{c_j} = 0$. *Thus,* $P_{relay,R_{i_2}}^{c_j} = 0$. *From Eq. (5.14),*

$$\begin{aligned} CTT(c_j, \mathbf{R}_D^{c_j}) &= \sum_{r=1}^{i_2-1} P_{relay,R_r}^{c_j} \frac{L \cdot A_D(S, R_r)}{T_{relay}(R_r)} \\ &\quad + \sum_{r=i_2+1}^{|\mathbf{R}_D^{c_j}|} P_{relay,R_r}^{c_j} \frac{L \cdot A_D(S, R_r)}{T_{relay}(R_r)} \\ &\leq \sum_{r=1}^{i_2-1} P_{relay,R_r}^{c_j} \frac{L \cdot A_D(S, R_r)}{T_{relay}(R_r)} \\ &\quad + \sum_{r=i_2+1}^{|\mathbf{R}_D^{c_j}|} P_{relay,R_r}^{c_j} \frac{L \cdot A_D(S, R_r)}{T_{relay}(R_r) - \mu} \\ &= CTT(c_j, \mathbf{R}_D^{c_j} \setminus \{R_{i_2}\}), \end{aligned}$$

which shows that the CTT performance does not drop when R_{i_2} *is deleted from* $\mathbf{R}_D^{c_j}$.

Lemma 5.4.1 indicates that we can reduce the size of the relay selection by excluding the relay candidates that are affected by the same PU. The reduced set of relay candidates will not degrade CTT. Specifically, for a given set of relay candidates, the sender groups the SUs that are affected by the same PU, selects the SU with the highest relay priority, and deletes other SUs in a group from the set.

We observe the following property which can be used to further reduce the searching space.

Property 5.4.2 (Tail Truncation Rule) *Given a feasible relay selection* $\mathbf{R}_D^{c_j}$, $\exists R_i \in \mathbf{R}_D^{c_j}$, $X_{SR_i}^{c_j} = 1$, *then* $CTT(c_j, \mathbf{R}_D^{c_j}) = CTT(c_j, \mathbf{R}_D^{c_j} \setminus \{R_k \big| R_k \in \mathbf{R}_D^{c_j}, V_{R_k} < V_{R_i}\})$.

Proof 2 *If* S *and* R_i *are affected by the same PU,* $Pr\left\{\overline{I_{R_i}^{c_j}} \middle| I_S^{c_j}\right\} = 0$. *According to E-*

q. (5.11), $P_k^{c_j} = 0, \forall R_k \in \mathbf{R}_D^{c_j}, V_{R_k} < V_{R_i}$*. Thus,*

$$\begin{aligned} CTT(c_j, \mathbf{R}_D^{c_j}) &= \sum_{r=1}^{i} P_{relay,R_r}^{c_j} \frac{L \cdot A_D(S, R_r)}{T_{relay}(R_r)} \\ &\quad + \sum_{r=i+1}^{|\mathbf{R}_D^{c_j}|} 0 \cdot \frac{L \cdot A_D(S, R_r)}{T_{relay}(R_r)} \\ &= CTT(c_j, \mathbf{R}_D^{c_j} \setminus \{R_k \big| R_k \in \mathbf{R}_D^{c_j}, \\ &\quad V_{R_k} < V_{R_i}\}), \end{aligned}$$

which shows that the CTT performance does not change when the relay candidates are removed from $\mathbf{R}_D^{c_j}$ *with lower priority than* R_i*.*

Property 5.4.2 indicates that the size of the relay candidate set can be further reduced by deleting SUs whose relay priorities are lower than the SU that is affected by the same PU as the sender. In other words, we can reduce the searching set without degrading the performance of the current flow while the deleted candidates can also participate in other transmissions, which further improve the network performance.

As discussed above, the relay priority plays a critical role in relay selection. It is well known that in geographic routing, users closest to the destination is the best next hop relay as it provides the greatest distance gain. It is also proved that the geographic routing approaches the shortest path routing with the distance advance metric [92]. Therefore, we also apply the distance advance and verify its efficiency in the proposed OCR.

Thus, the CTT metric can be approximated as

$$\begin{aligned} CTT(c_j, \mathbf{R}_D^{c_j}) &\simeq \frac{L}{T_{relay}} \cdot \sum_{i=1}^{|\mathbf{R}_D^{c_j}|} P_{relay,R_i} A_D(S, R_i) \\ &= \frac{L}{T_{relay}} \cdot E[A_D^{c_j}], \end{aligned} \tag{5.15}$$

where $E[A_D^{c_j}]$ is the estimated relay advancement in c_j, and T_{relay} is the estimated one hop transmission delay in Eq. (5.5). To maximize the CTT in each channel, we need to find an optimal relay selection to maximize $E[A_D^{c_j}]$. When opportunistic routing over independent links uses $E[A_D^{c_j}]$ as a routing metric, [92] has proved that the optimal relay priority should be set according to the distance of the relay candidate to the destination. In addition, the

Algorithm 3: The MAXCTT algorithm

```
Input: the channel set $C$, the relay candidate set $\mathbf{R}_D$, $r_{max}$
Output: the selected channel $c^*$, the relay selection order $\mathbf{R}_D^{c^*}$
$c^* \leftarrow 0$;   $\mathbf{R}_D^{c^*} \leftarrow \emptyset$;   $CTT_{max} \leftarrow 0$;
for each $c_j$ do
    $\mathbf{N} \leftarrow \mathbf{R}_D$; $\mathbf{R}_E \leftarrow \emptyset$; $\mathbf{R}_D^{c_j} \leftarrow \emptyset$; $R_p \leftarrow \emptyset$; $CTT_{c_j} \leftarrow 0$;
    while ($\mathbf{N} \neq \emptyset$) do
        $\mathbf{R}_E \leftarrow$ insert an SU $R_i \in \mathbf{N}$ that has max $A_D(S, R_i)$;
        Remove $R_j \in \mathbf{N}$ with $X_{R_iR_j} = 1$ from $\mathbf{N}$;
    end while
    while ($\mathbf{R}_E \neq \emptyset$ && $|\mathbf{R}_D^{c_j}| < r_{max}$ && $X_{SR_p} \neq 1$) do
        for each SU $R_i \in \mathbf{R}_E$ do
            $\mathbf{R}_T \leftarrow \mathbf{R}_D^{c_j} + R_i$; Sort $\mathbf{R}_T$ in the descending order of $A_D(S, R)$;
            Get $CTT$ on $\mathbf{R}_T$ according to Eq. (5.14);
            if ($CTT > CTT_{c_j}$) then
                $CTT_{c_j} \leftarrow CTT$; $R_p \leftarrow R_i$;
            end if
        end for
        $\mathbf{R}_D^{c_j} \leftarrow$ insert $R_p$ in the descending order of $A_D(S, R)$; $\mathbf{R}_E \leftarrow \mathbf{R}_E - R_p$;
    end while
    if ($CTT_{c_j} > CTT_{max}$) then
        $c^* \leftarrow c_j$; $\mathbf{R}_D^{c^*} \leftarrow \mathbf{R}_D^{c_j}$; $CTT_{max} \leftarrow CTT_{c_j}$;
    end if
end for
return  $(c^*, \mathbf{R}_D^{c^*})$;
```

maximum $E[A_D^{c_j}]$ increases with the number of relay candidates. Therefore, we can assign the relay priority in the descending order of $A_D(S, R)$.

We then propose a heuristic algorithm, MAXCTT, as shown in Algorithm 3. The inputs are the channel set C, the set of relay candidates $\mathbf{R}_D$, and the maximum number of relay candidates in relay selection r_{max}. MAXCTT selects the SUs from $\mathbf{R}_D$ to form the relay selection order $\mathbf{R}_D^{c_j}$ and calculates the achieved CTT_{c_j} in each c_j. By comparing CTT_{c_j} over the channels, MAXCTT returns the channel c^* that has CTT_{max} and the corresponding relay selection order $\mathbf{R}_D^{c^*}$ as the algorithm output.

Specifically, an eligible relay candidate set $\mathbf{R}_E$ is first formed by excluding the SUs affected by the same PU in c_j according to Lemma 5.4.1, which is a subset of $\mathbf{R}_D$ (line 4–line 6). A recursive searching [70] is then applied to obtain $\mathbf{R}_D^{c_j}$(line 8–line 14). At the beginning of the searching step, $\mathbf{R}_D^{c_j}$ contains no SU. Each time, $\mathbf{R}_D^{c_j}$ includes one more relay candidate out of the remaining SUs in $\mathbf{R}_E$ which provides the best CTT improvement. The selected relay candidates are sorted in the descending order of $A_D(S, R)$ in $\mathbf{R}_D^{c_j}$. The formed $\mathbf{R}_D^{c_j}$ contains all relay candidates from $\mathbf{R}_E$, and it satisfies the requirements of r_{max} and Property 5.4.2 (line 8). The search ends when 1) all relay candidates are included, or 2) the set size reaches the upper boundary, i.e, r_{max}, or 3) the relay selection needs to be truncated according to Property 5.4.2. The recursive searching obtains the optimal $\mathbf{R}_D^{c_j}$ in c_j when the size of the selection order is at most 2, and it achieves almost the same performance as the optimal solution when the final order contains more than 2 candidates according to Lemma 5.1 in [70]. Suppose that the largest size of $\mathbf{R}_E$ over the channels is n, at most $m \cdot \sum_{k=1}^{n} k$ times of the CTT calculations are required to find CTT_{max}. Thus, the time complexity of MAXCTT is $O(m \cdot n^2)$, which is much lower than exhaustive search.

5.5 Performance Evaluation

In this section, we evaluate the performance of the OCR protocol by simulation under different network settings, e.g., channel conditions, number of SUs, and traffic loads, using an event-driven simulator coded in C/C++ [82, 75]. The network parameter settings are shown in Table 5.2 if no other specification is made in the individual study.

5.5.1 Simulation Settings

The PU activity in each channel is modeled as an exponential ON-OFF process with parameters $1/E[T_{ON}]$ and $1/E[T_{OFF}]$, and the idle rate $\rho = E[T_{OFF}]/(E[T_{ON}] + E[T_{OFF}])$

is selected accordingly. The channel status is updated by periodic sensing and on-demand sensing before data transmissions. We set up a CRN with multiple PUs and SUs randomly distributed in an $800 \times 800\ m^2$ area. We set a pair of SUs as source and destination with a distance of 700 m, and a constant bit rate (CBR) flow is associated with the SU pair with packet size 512 bytes and flow rate of 10 packets per second (pps). The unit disc model is applied for the data transmission. The channel switch time is 80 μs [93], the minimum sensing duration with energy detection is 5 ms, and a mini-slot is 4 μs [89]. We evaluate the performance of the proposed OCR protocol in terms of the end-to-end delay, the packet delivery ratio (PDR) and the hop count, i.e., the total number of transmission hops between the source and destination SUs. We run each experiment for 40 s and repeat it 500 times to calculate the average value.

We then compare the performance of the OCR protocol with that of SEARCH [25], based on different metrics for the channel and relay selection, which are listed as follows.

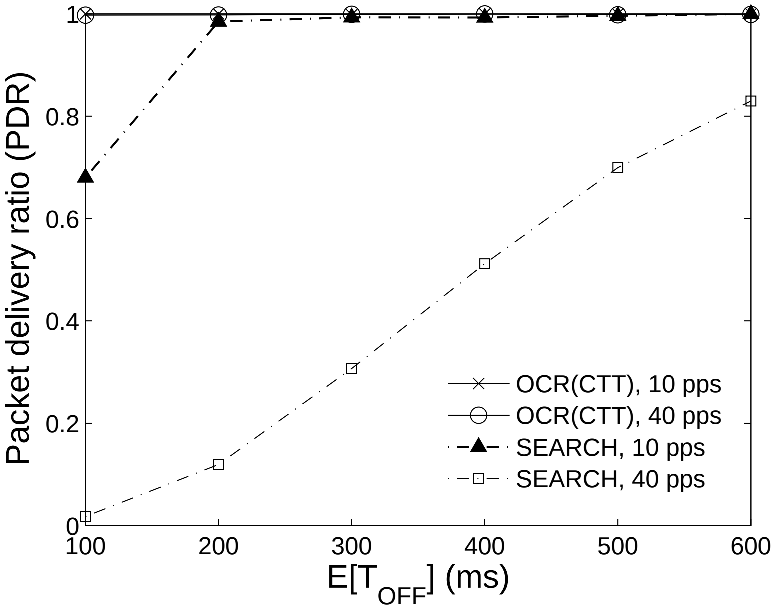
1. SEARCH: SEARCH [25] is a representative geographic routing protocol in CRN. It sets up a route with the minimal latency before data transmissions. If an active PU is detected which blocks the route, SEARCH pauses the transmissions and recalculates the route. We modify SEARCH by updating route periodically to adapt to the dynamic changing spectrum access opportunities along the route.

2. OCR (CTT): For OCR (CTT), the channel and the relay candidate set are jointly selected by using the proposed CTT metric and heuristic algorithm proposed in Section 5.4.

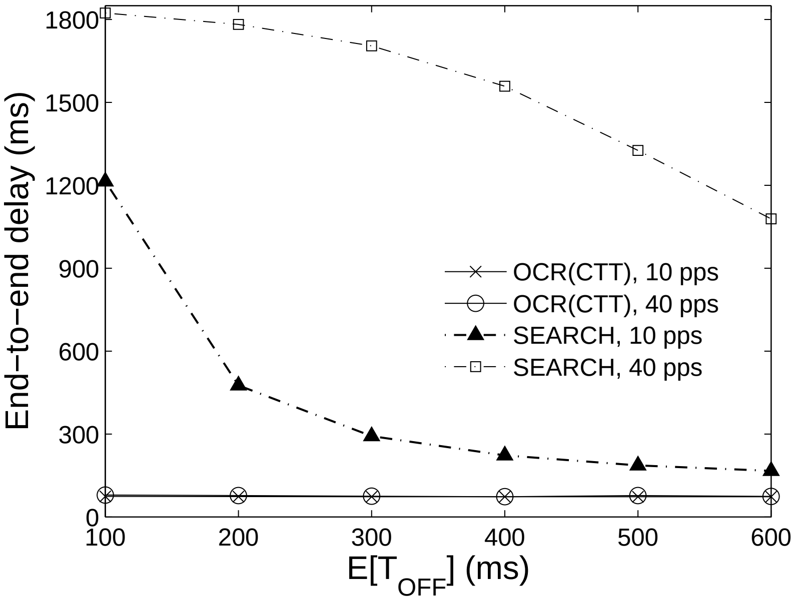
3. OCR (OPT): For OCR (OPT), the channel and the relay candidate set are determined by exhaustively searching for the biggest CTT over all possible channel-relay sets.

4. GOR: For geographic opportunistic routing (GOR) algorithm, the SU first selects the channel with the greatest success probability of packet transmissions; if the channel is sensed idle, the SU then select a relay SU over the channel. The relay selection order is based on the location information and the relay capability of SUs [85].

5. GR: For geographic routing (GR), an SU first selects the channel for sensing as in GOR. If the selected channel is sensed idle, the SU then selects the SU closest to the destination as the next hop relay.



(a) Packet delivery ratio (PDR)



(b) End-to-end delay

Figure 5.2: Performance comparison between OCR and SEARCH under different channel conditions (Number of SUs: 200, flow rate: 10/40 pps)

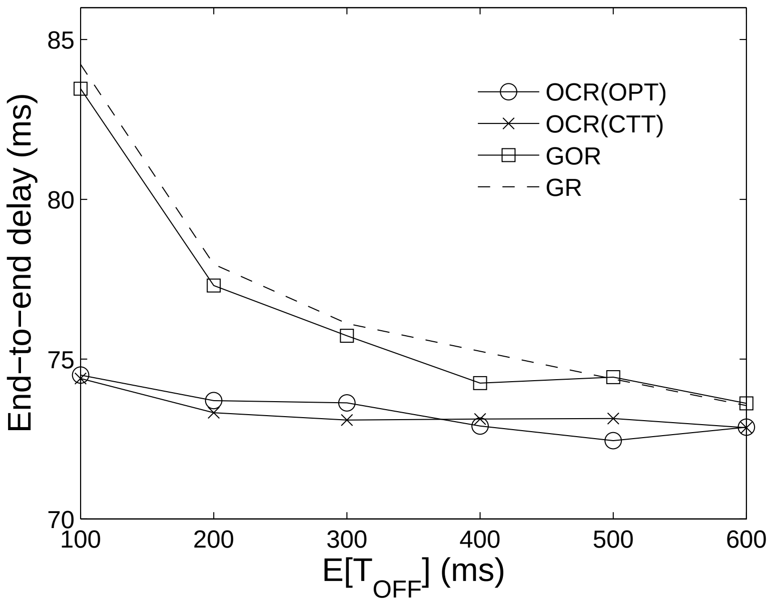


Figure 5.3: Performance comparison under different traffic loads and PU activities (Number of SUs: 200, flow rate: 10 pps)

5.5.2 PU Activities

We first evaluate the performance of OCR under different PU activity patterns. The average PU OFF duration $E[T_{OFF}]$ varies from 100 ms (high channel dynamics) to 600 ms (low channel dynamics). The PDR performance of OCR and SEARCH are compared under different traffic loads in Fig. 5.2(a). A smaller $E[T_{OFF}]$, e.g., 100 ms, indicates the available time window is shorter and thus SUs' transmissions are more likely to be interrupted by PUs. We can see a marked PDR improvement under dynamic channel conditions for the per hop relay schemes, e.g., OCR (CTT), compared with SEARCH which is based on the global route establishment. In OCR (CTT), SUs are allowed to locally search and exploit spare spectrum and select the available links to data forwarding. Thus, OCR (CTT) can adapts well in the dynamic data channels. On the contrary, SEARCH uses a pre-determined routing table. Once an active PU is detected along the relay path, intermediate SUs should defer the packet relay until they update their routing tables according to the current channel availabilities in CRN. Since more SUs are involved in the route establishment, the handshakes between SUs in the network to establish the relay path introduce a large overhead and results in a longer delay.

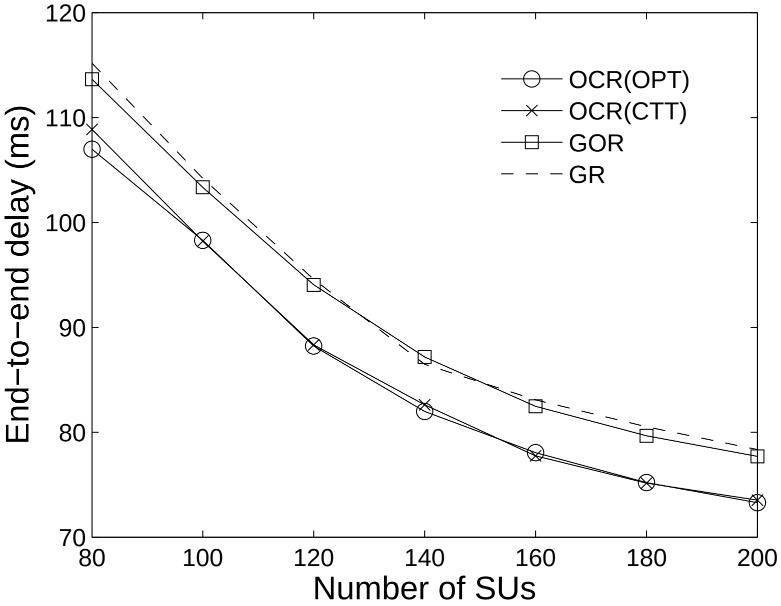


Figure 5.4: Performance comparison of end-to-end delay under different SU densities (flow rate: 10 pps, $E[T_{OFF}] = 200\ ms$)

Fig. 5.2(b) and Fig. 5.3 compare the end-to-end delay performance. All routing protocols achieve a better delay performance when the idle channel state becomes longer, e.g., from 100 ms to 600 ms, as more packets can be transmitted during the idle state. When the channel state change frequently, SEARCH needs to update routing tables accordingly which involves a long delay for route recovery. Our proposed OCR protocols are opportunistic routing algorithms that quickly adapt to the dynamic channel environment and achieve better delay performance compared with SEARCH. OCR (CTT) also outperforms GR and GOR since the latter two protocols perform the channel and relay selection separately while OCR (CTT) jointly consider the channel selection and relay selection.

5.5.3 Multi-user Diversity

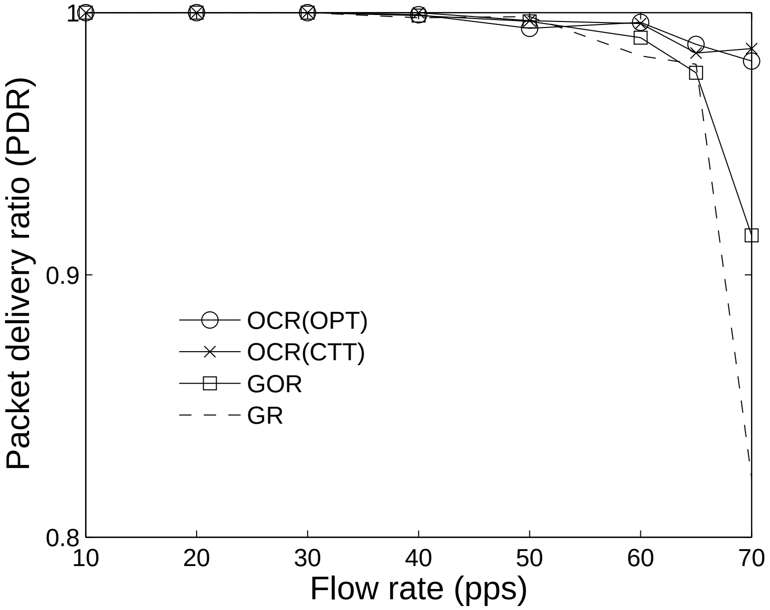
We investigate the impacts of node density on the relay performance. The number of SUs in CRN varies from 100 to 200. When the number of SUs is large, the sender has more neighbors as shown in Table 5.3. With more SUs in the neighborhood, the relay is more likely to find a feasible relay link with better relay distance advance, which reduces the hop count number. The relay performance increases with the number of SUs due to the larger diversity gain. As a result, for all protocols, the hop count of the end-to-end relay decreases and the PDR increases with SU density by exploiting the multi-user diversity in CRN. The end-to-end delay performance under different SU densities is compared in Fig. 5.4. For GR and GOR, a channel is selected first, and then SUs coordinates to serve as relay. The coordination overheads increase with the number of SUs, which also degrades the PDR performance. The proposed OCR (CTT) jointly considers the channel and relay selection, and SU coordination overhead is minimized as sender determines the relay selection order based on the relay priority.

We also compare the performance of the heuristic algorithm for the channel-relay selection in OCR (CTT) with the optimal one in OCR (OPT) where the selection is based on exhaustive search. Fig. 5.4 shows that OCR (CTT) achieves almost the same performance as OCR (OPT), even when the returned number of the selected relay candidates is over 2, according to the value of r_{max} in Table 5.2. Table 5.3 indicates that as the SU density increases in the network, the number of neighbors along the forwarding direction of the sender will increase accordingly. For example, given 160 SUs over 6 channels, the average number of neighbors of an SU is around 11. OCR (OPT) takes over 6.5×10^8 times of the CTT calculation to find the globally optimal solution which is infeasible for real time implementation. In the simulated scenario, although at most 4 neighbors are under independent PU coverage which significantly reduces the searching space, OCR (OPT) still takes 384 runs while OCR (CTT) only needs 60 runs, which achieves the marked reduction at the computational expense.

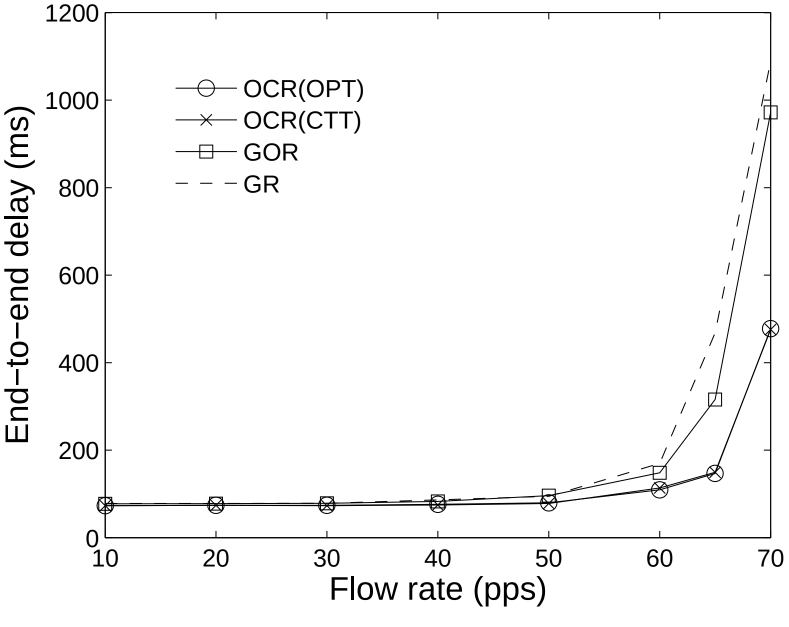
5.5.4 Effectiveness of Routing Metric

We further compare the performance of OCR (CTT) with that of GR and GOR to evaluate the effectiveness of routing metrics used in the channel and relay selection. We first compare the performance under different traffic loads. We change the traffic load by varying the flow rate from 10 pps (light load) to 70 pps (heavy load). As shown in Fig. 5.5(a) and Fig. 5.5(b), when the traffic load increases, the PDR and delay performance degrade. However, the decreasing rate of OCR (CTT) is much lower than that of GR and GOR. This is because OCR (CTT) jointly considers the optimal channel and link selection, while the other two OCR protocols select the channel and relay separately.

We define P_{ef} to be the ratio of the number of successful relay transmissions to the number of the sensing operations performed in the data channels. P_{ef} indicates the effectiveness of the routing metrics since the transmission relies on detection of an idle channel and an available relay node. If P_{ef} approaches to 1, the selected channel for each hop relay



(a) Packet delivery ratio (PDR)



(b) End-to-end delay

Figure 5.5: Performance comparison under different traffic loads (Number of SUs: 200, $E[T_{OFF}] = 200\ ms$)

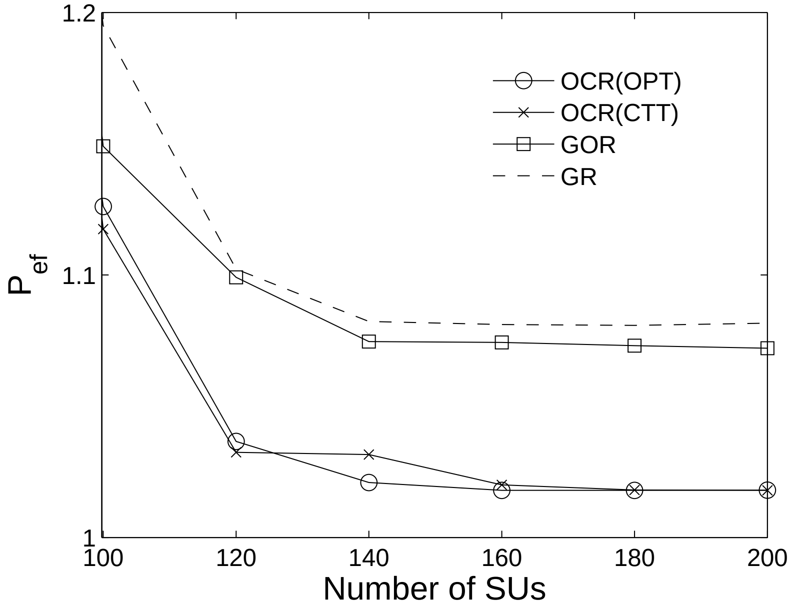


Figure 5.6: Performance comparison under different SU densities (flow rate: 10 pps, $E[T_{OFF}] = 200\ ms$)

almost surely is available for data transmission. Fig. 5.6 shows the performance of P_{ef} under different node densities. In all network scenarios, OCR (CTT) outperforms GR and GOR, because CTT metric jointly considers the channel access and relay selection.

5.6 Summary

In this chapter, we have proposed an opportunistic cognitive routing (OCR) protocol to improve the multi-hop transmission performance in CRN. We have studied the impact of PU activities on the operation of OCR in channel sensing, relay selection and data transmission. Furthermore, we have proposed a novel metric, CTT, for the channel and relay selection. Based on the metric, we have proposed a heuristic channel and relay selection algorithm which approaches optimal solution. We have compared the performance of OCR (CTT) with that of the existing routing approaches, e.g., SEARCH, GR and GOR and show that the proposed OCR achieves the highest PDR and the lowest delay. In our future work, we will study protocol design with uncertain channel usage statistics and study the impacts of the measurement errors on the protocol performance.

Table 5.1: Summary of notations in Chapter 5

Symbol	Definition
$C = \{c_j\}$	Channel set, $j = \{1, 2, ..., m\}$
$\mathbf{N}_S$	The SU set of the sender S's neighbors
$\mathbf{R}_D$	The set of relay candidates for the destination D
$\mathbf{R}_D^{c_j}$	The set of relay candidates for the destination D in the channel c_j
$A_D(S, R)$	Relay advancement of the link SR for D
$(c^*, \mathbf{R}_D^{c^*})$	The transmission channel and the ordered relay set selected by MAXCTT
$CTT(c_j, \mathbf{R}_D^{c_j})$	The cognitive transport throughput (CTT)
$d(S, D)$	Euclidian distance between S and D
$E[T_{ON}^{c_j}](E[T_{OFF}^{c_j}])$	Mean duration of a busy(idle) c_j
$\mathcal{F}_{OFF}^{c_j}(t)$	CDF of the OFF duration of c_j
$I_R^{c_j}$ $(\overline{I_R^{c_j}})$	SU R detects c_j to be idle (busy)
T_{detc}	Per channel energy detection delay
T_{DTX}	Per hop data packet transmission delay
T_{init}	Sensing initialization delay
T_{relay}	Per hop transmission delay in OCR
$T_{RREQ}(T_{RRSP})$	RREQ (RRSP) message transmission delay
T_{RS}	Per hop relay selection delay
T_{SNS}	Per hop sensing delay
T_{switch}	Transceiver switching time
t_0	The latest channel status observation time
$P_i^{c_j}$	The probability that R_i is selected as the relay in c_j
$P_{OFF,R}^{c_j}(t_0, t_1)$	The probability that c_j is idle at t_1, $t_1 > t_0$.
$P_R^{c_j}(t_1, t_2)$	The probability that c_j is idle during $[t_1, t_2]$ at R
$P_{relay,R_i}^{c_j}$	The probability that the relay via R_i succeeds in c_j
$P_{RSfail}^{c_j}$	The probability that relay selection fails in c_j
V_{R_i}	The priority of R_i in the relay selection.
$X_{R_tR_r}^{c_j} = 1(0)$	SUR_t and SU R_r are (not) affected by the same PU in c_j
ρ_{c_j}	The chance for an idle state in c_j
μ	Backoff mini-slot
γ	The maximum channel propagation delay

Table 5.2: Simulation parameters in performance evaluation of Chapter 5

Number of channels	6
$\{\rho_{c_1}, \rho_{c_2}, \rho_{c_3}, \rho_{c_4}, \rho_{c_5}, \rho_{c_6}\}$	{0.3, 0.3, 0.5, 0.5, 0.7, 0.7}
Number of PUs per channel	11
PU coverage	250 m
$E[T_{OFF}]$	[100 ms, 600 ms]
Number of SUs	[100, 200]
SU transmission range	120 m
Source-destination distance	700 m
SU CCC rate	512 kbps
SU data channel rate	2 Mbps
CBR delay threshold	2 s
Mini-slot time, μ	4 μs
Per channel sensing time	5 ms
Channel switching time	80 μs
PHY header	192 μs
r_{max}	2

Table 5.3: Average neighbor density under different SU densities

Number of SUs	100	120	140
Average number of neighbors	7.0686	8.4823	9.8960
Number of SUs	160	180	200
Average number of neighbors	11.3097	12.7235	14.1372

Chapter 6

QoS-aware Cognitive (QC) MAC for Multimedia Services

In this chapter, we study how to efficiently probe the spectrum opportunities to satisfy the QoS requirements of heterogeneous applications, such as voice, video, and data. We propose a cognitive MAC protocols with QoS provisioning for multimedia services in CRN. Based on the characteristics and QoS demands of the traffic and the statistics of channel usage patterns, each SU distributively determines a set of channels for sensing and data transmission. By distributing multiple SUs over different channels, the contention level among SUs on each channel can be effectively reduced and the resource utilization is improved accordingly. We then present a priority spectrum access scheme to further enhance the QoS provisioning in CRN with heterogeneous traffic. By applying differentiated arbitrary sensing periods for SUs with various types of traffic, the higher priority traffic, e.g., voice, will have more opportunity to access the available channels when competing with other SUs. By carefully designing the service differentiation parameters, multiple levels of QoS provisioning can be achieved.

6.1 Introduction

In the past few decades, the popularity of wireless communications has contributed to personal and business benefits in a prosperous market of mobile services along with the explosive growth of traffic demands over the air. Various Internet services are now penetrating from the desk into the pocket, which has launched data hungry multimedia applications on smartphones and tablets thanks to more powerful computational and communication capabilities. As the result, today's wireless devices are consuming more bandwidth than ever before, e.g., a single smartphone monthly consumes 35 times the bandwidth of a conventional voice-only cellphone [1]. Furthermore, the aggressive improvement on hardware and software platforms of mobile devices are keeping offloading more data traffic from the Internet wirelessly. The global mobile data traffic thus has been increasing exponentially in the past five years and such trend shows no fading in the expectable future [94].

In CRN, different SUs may carry various applications with different traffic characteristics and QoS demands. In addition, the traffic patterns in each spectrum bands may also be different depending on the types of PUs [95]. For example, the OFF periods in TV bands are relatively long when programs are terminated, while they could be very small in cellular bands where a large number of cellular customers carry voice traffic with a very low rate. To efficiently explore the spectrum opportunity to provide QoS for SUs, it is essential to consider the characteristics of both the traffic and the channel usage in the MAC design [96, 97]. Various traffic characteristics, diverse QoS requirements, and heterogeneity channel usage patterns should be taken into consideration in the MAC design and analysis for the DySPAN.

The exploration of the spectrum opportunities should be efficient to satisfy the QoS requirements of heterogeneous applications, such as voice, video, and data. For instance, each SU distributively determines a set of channels for sensing and data transmission based on the characteristics and QoS demands of the traffic, and the statistics of channel usage patterns on its spot. Multiple SUs are distributed over different channels so that the contention level among SUs over each channel can be effectively reduced and the resource utilization is improved accordingly.

For heterogeneous traffics, a priority spectrum access scheme should be incorporated to further enhance the QoS provisioning in the DySPAN. For example, by applying differentiated arbitrary sensing periods for SUs with various types of traffic, SUs with a higher priority traffic, e.g., voice, have more opportunity to access the available channels when competing with other SUs. Meanwhile, the service differentiation parameters should be carefully designed so that multiple levels of QoS provisioning can be achieved.

In CRN, the transmission power of the SUs should be managed below the interference limitation of the primary networks on the frequency channels [14]. The source-destination pairs in CRN may be out of the transmission coverage of each other, so the MAC layer protocol in CRN should support the multi-hop architecture for the opportunistic transmission. To improve the performance of CRN, there are many interesting topics for further investigation, such as the hidden terminal problem, the selfish behavior, the incentive scheme for the relay nodes, and the route selection. Since spectrum opportunities are distributed among orthogonal frequency channels, the multi-channel MAC protocols become the candidates for CRN. In conventional wireless networks, the multi-channel MAC protocol is well studied to coordinate the node behaviors using multiple channels, especially in the mesh networks [98, 99]. In CRN architecture, the spectrum opportunities are spatio-temporally available, which brings more uncertainty in the secondary transmission through multiple hops.

Node Coordination

The time variance of the spectrum opportunities limits the validity of the information exchange, hence the conventional centralized MAC protocol can not adapt to CRN efficiently and effectively, especially in the multi-hop transmission. Meanwhile, the fully distributed MAC protocol usually lacks of the cooperation scheme for nodes to share the local information regularly, which cannot guarantee the gain of the cooperative communication. Some proposed MAC protocols of CRN clusters the distributed SUs in the local region as cooperative entities in spectrum management [53].

More open issues are still on this aspect. In the cluster based MAC protocol, the life cycle of a cluster needs carefully analysis on the aspect of the available resources. Currently the operations intra and inter clusters cost large overhead which limit the efficient utilization of the spectrum opportunities, because some short duration ones cannot accommodate one whole frame. Because the common control channel in the cluster is unlicensed, the procedure of the band switching is an interesting topic here. Unlike the simple case for a couple of transceivers to take the switching action, the cluster head should predict the PU activity, select the candidate control channel, make the decision and inform every member in its cluster timely. All actions should be conducted in a limited time and be evaluated by the new performance in the new bands, which requires a sophisticated and efficient support from the MAC layer protocol design.

Common Control Channel

Once the sensing and decision policy is made, the ongoing connection of the cognitive transmission may avoid using the bands with high appearance probability and switch to other detected idle bands. In the spectrum switching process, which is also called spectrum mobility in [6], the notification between the communication pairs need careful design, especially in the cluster based application.

CRN transmission also has the multi-channel hidden terminal problem which is due to the failure of all nodes to meet on the same channel when transmission reservation messages are broadcasting [98]. The hidden terminal problem would severely impact the collision probability in the wireless channels when the nodes within the transmission range of each other would introduce significant interference if the coordination control messages were missed due to the lost and the synchronization problems.

In [100], the authors propose a solution to the multi-channel hidden terminal problem in CRN using periodic beacon to synchronize the channels for the secondary transmission. By extending the work in [98, 99], In a fixed beacon period, all SUs have to switch to a default channel to listen to controlling messages. Each transmission pair exchanges the selection message about current idle channels they would use. At the rest of the beacon period, the pairs switch to their negotiated channel to perform single channel carrier sense multiple access (CSMA) the same as the IEEE 802.11 DCF does. So on each idle channel, the traffic load is mitigated by the division onto different channels while the synchronization window is a fixed overhead for the networks. A quiet period (QP) is set to coordinate the sensing period among SUs on the same channel to improve the PU detection. However, the drawback of this scheme is that one transmission pair can only use one idle channel in one period, not able to extend to one pair over multiple channel case.

One key issue in the multi-channel MAC is the establishment of the common control channel(s) (CCC). Since the SUs have various candidate channel set using in a frequency hopping way, it is necessary to design a rendezvous scheme for the SUs to meet on the new CCC band after the switching process. Currently the released proposals suggest several solutions to this problem. The type of the common control channel can be dedicated or randomized; the ownership can be network sharing or pairwise [37]. Generally, we can distinguish single rendezvous (SRV) and multiple rendezvous (MRV) schemes. In SRV protocols, exchange of control information occurs only on one channel at any time, whereas the MRV schemes use several channels in parallel for this purpose. Within the SRV schemes, we can further distinguish three different classes: one class using a common control channel, another class using common hopping, and a last class using a split-phase approach.

For the dedicated control channel [101], the SUs in CRN are licensed to use a given frequency band for signaling. It is easier for the SUs to trace the control channel and be aware of the networks status. However, the capacity of the control channel limits the size of the CR networks. For the randomized CCC case, [102] discusses on the pair of the transceivers synchronize in an available channel by the defined frequency hopping (FH) sequence. FH sequences are required to design to enlarge the rendezvous probability [103, 104]. There are also several hybrid solutions shown in [101]. For the multiple control channel case, each pair of transceivers first meet in the initial common control channel which is dedicated, then both switch to another randomized one for further control communication. In this way, it can achieve both the quick rendezvous and the large control channel capacity. Usually, the switching process also need consider the channel capacity for the candidate bands. In [105] the authors compare different CCC switching policy and the effect on the system performance, especially the effect of the error during the candidate selection.

There are open topics, such as the disagreement problem in multi channel case in the multi-hop CRN. Because the availability of spectrum opportunities is spatially specified by the individual SUs, the transmission node pair may have different view of current available channel set. If the transmission channel index is mismatched at the transceivers, not only the secondary transmission would be blocked, but also it would introduce severe interference to the primary systems on the channels. So it needs the exchange of the control messages before the data transmission. The temporary availability of the resources limits the comprehensive control information exchange, the design of signaling and the way to synchronization. FH sequence for rendezvous should be carefully designed.

Performance Analysis

The performance analysis of wireless networks can be divided as three main aspects: the traffic performance, the security and the energy efficiency. In the first part, the multi-hop transmission of various applications are evaluated via secondary links in the system parameters, such as the end-to-end delay and the throughput. Literature [106] addresses some important concepts in the design of CRN with the consideration of the performance. The authors suggest that the proactive optimization on the network performance is more appropriate to apply in CRN than the reactive adaptation schemes in conventional wireless networks. In [88], the authors propose a MAC protocol for the multi-hop transmission in CRN by introducing multiple users into the identification of PU activity on the channels. In their discussion of the throughput performance, they find that the maximum throughput occurs in the direct link in the neighborhood while it would result into low scalability. The opportunistic forwarding scheme should be used to overcome this problem. The performance of the secondary transmission is largely dependent on the amount of spectrum opportunities both in time (duration of each spectrum hole) and frequency (the number of licensed channels). To improve the throughput in secondary links, wideband sensing and concurrent transmission in multi-channels are implied in [107]. In literature [108], a performance analysis on current released typical multi-channel MAC for the DSA transmission is conducted.

Since CRN coexists with primary systems in the hierarchical channel access, the security issue here is to guarantee the spectrum sensing process is not interrupted by the misbehavior. Literature [109] addresses the issue in the incumbent emulation (IE) attack. A malicious SU may try to occupy extra spectrum opportunities by transmitting signals the same as PUs, so that it takes advantage of the higher priority in accessing spectrum resources as a PU. Because current CRN widely use the energy detection technique to detect the PU activity, it is more vulnerable to resist such attack. Therefore, this problem should be carefully considered.

Energy consideration shows its unique impact on CRN performance with close relation in the aspects of the DSA transmission. Literature [88] identifies that the scanning cost would be a considerable contributor to the total energy cost for the nonactive SU which has no packet to transmit. In order to reduce the energy waste, it requires the sophisticated node management to shut down the SU if no traffic is on. On the other hand, another major factor in the energy consumption is the overhearing on the channels, i.e., the received packet which is not for the receiver would also cost the energy for the unrelated nodes. Thus, the densely deployed CRN can achieve high scalability while it should pay the cost in the energy consumption where the designer should make a trade-off here. Moreover, unlike conventional wireless networks with fixed number of channel, CRN can enlarge the channel set by incorporating more spectrum opportunities distributed on more channels which would reduce the energy cost mentioned above. However, the number of channels to sense should make a bargaining so that the introduction of additional energy consumption in the scanning can be compensated by the improvement in energy efficiency.

6.2 The QC MAC Scheme

In the proposed QC MAC scheme, by exploiting diverse channel usage patterns, SUs select appropriate channels that satisfy their QoS requirements for channel sensing and data transmissions. The transmission procedure of an SU is shown in Fig. 6.1. Without loss of generality, an SU senses the first channel and starts data transmission if the channel is sensed idle for a sensing interval. To reduce possible collisions among SUs, each SU will sense the channel for an arbitrary sensing period (ASP), which consists of the basic sensing period that assures satisfactory sensing accuracy plus some random slots selected from a sensing window $[0, SW_i]$. If the channel is sensed busy, SU switches to the second channel. The pseudo code of the SU transmission procedure is shown in Algorithm 4. Note that at the beginning of the channel sensing period, the sender will initiate a handshake with its receiver over the control channel for transceiver synchronization. It is also possible that a

PU may appear during SU's data transmission, in which case the transmission fails and the SU will switch to the next channel to retransmit the data. Therefore, the key research issue in the cognitive MAC design is how to effectively explore the spectrum opportunities and select an appropriate set of channels to assure the QoS performance of SUs.

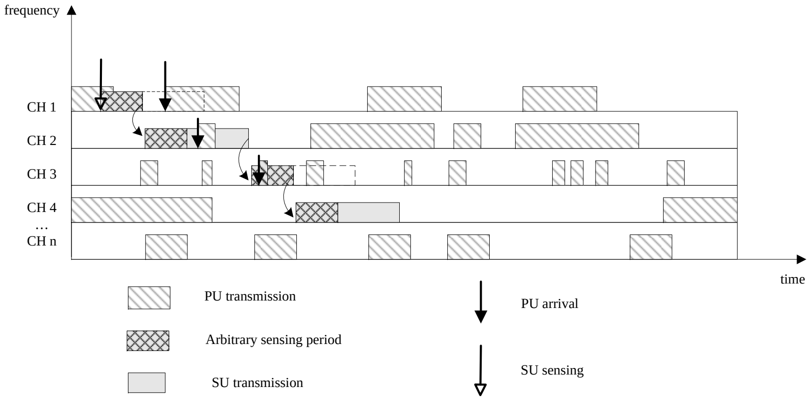


Figure 6.1: Multi-channel transmissions in CRN.

6.2.1 Channel Sensing

Denote the transmission time of the current frame of user i as t_i, which is constant for voice traffic but may vary for video traffic. To fully utilize the spectrum opportunity, SUs calculate the probability that the current frame can be successfully transmitted over a free channel k according to the equilibrium distribution of the idle channel window [90], which is given by

$$\begin{aligned} P_i^k &= Pr(\text{Channel k is idle, Transmission is successful}) \\ &= \frac{\beta_k}{\alpha_k + \beta_k} F_k(t_{AS_i} + t_i), \end{aligned} \tag{6.1}$$

where $1/\alpha_k$ and $1/\beta_k$ are the mean idle and busy periods of channel k, t_{AS_i} is an arbitrary sensing period selected by user i, t_i is the transmission time of the current frame, and $F_k(t)$ is estimated by $P_R^{c_j}(0,t)$ in Eq.(5.7) where $E[T_{OFF}^{c_j}] = 1/\alpha_k$, the derivation of which can be found in [17]. Notice that with an arbitrary sensing period, it is less likely that two or more SUs transmit simultaneously and collide with each other. Therefore, the transmission is successful if no PU appears during the following $t_{AS_i} + t_i$ interval.

Algorithm 4: SU Transmission Procedure

```
Begin:
 1: if (User $i$ has a data for User $j$) then
 2:    Calculate $P_s^i$ and decide the sensing sequence;
 3:    Select the first channel;
 4:    Sense the channel and initiate a handshake with User $j$ over the control channel;
 5:    if (Channel is sensed busy during $ASP_i$ then
 6:       Select the next channel;
 7:       Go to line 4;
 8:    else
 9:       Transmit data;
10:       if (Transmission is successful) then
11:          Go to End;
12:       else
13:          Go to line 6;
14:       end if
15:    end if
16: end if
End
```

Based on the calculated successful transmission probability in each channel, SUs can determine the channel sensing sequence with two different policies: greedy and ascending. For the greedy policy, SUs simply sort channels in a descending order and always use channels with the highest success probability for achieving a low delay and high throughput. However, the channel with less PUs' activity is more likely to be selected by SUs, which causes high contention level among SUs sharing the same radio resources and degrade the performance accordingly. Therefore, we propose the second sensing policy that allows different SUs select various channels based on the QoS requirements of their applications. For instance, each realtime frame is associated with the maximum tolerable one hop delay τ_i. Given P_i^k, the expected transmission time over channel k is estimated as $E[T_k^i] = (t_{AS_i} + t_i)/P_i^k$. Therefore, an SU first selects a group of channels that satisfy

$$E[T_i^k] = (t_{AS_i} + t_i)/P_i^k < \tau_i, \tag{6.2}$$

and senses these channels in the ascending order of P_i^k. That is, each SU first selects a channel with the minimum P_k^i that satisfies its delay requirement. If the channel is sensed busy, the SU will switch to the next channel with second lowest P_k^i. As the channel is sorted in the ascending order of P_k^i, the expected delay of SU i, $E[T_i] \leq E[T_k^i] < \tau_i$. Therefore, the delay performance can be guaranteed with the ascending policy. Notice that although

Table 6.1: Sensing window design for multimedia services in CRN

	Strict Priority	Statistical Priority	No Priority
Voice	[0,31]	[0,31]	[0,31]
Video	[32,63]	[0,63]	[0,31]
Data	[64,127]	[0,127]	[0,31]

SU can estimate the channel usage pattern by PUs, it is difficult if not impossible for an SU to accurately estimate the number of SUs currently sharing the spectrum bands. For a simple yet robust MAC design, an SU can set a stringent delay bound τ_i and select a channel sets with more opportunities to incorporate the impacts of other SUs. The impacts of contention level among SUs will be analytically studied in Section 6.3, which can provide important guideline for an SU to set the parameter τ_i and select a set of channels for opportunistic transmissions.

6.2.2 Service Differentiation

We further enhance the QoS provisioning of the proposed cognitive MAC by introducing service differentiation in the arbitrary sensing periods of different traffic flows. Basically, a smaller sensing window is applied for a higher priority real time applications so that they have a higher chance to access data channels when opportunity appears, i.e., $SW_{voice} < SW_{video} < SW_{data}$. In addition, by carefully determining the sensing windows for different types of traffic, multiple levels of QoS provisioning can be achieved for multimedia applications in CRN. As shown in Table 6.1, a statistical priority is provided by simply doubling the sensing windows for various types of traffic, while a strict priority can be achieved when non-overlapped sensing windows are used. The performance of the differentiated service provisioning using different settings will be evaluated in Section 6.4.

6.3 Performance Analysis

In this subsection, we develop an analytical model to study the delay performance of the proposed QC MAC.

An SU senses channel k and attempts to transmit if the channel is sensed idle for t_{AS_i}. In other words, an SU's sensing fails if 1) the channel has been occupied by a PU with probability

$$P^k_{CO} = Pr(\text{Channel occupied}) = \frac{\alpha_k}{\alpha_k + \beta_k}; \tag{6.3}$$

2) the channel is idle but a PU turns on during the sensing interval with probability

$$P^k_{PU} = Pr(\text{PU on}) = \frac{\beta_k}{\alpha_k + \beta_k} F_k(t_{AS_i}) \tag{6.4}$$

or 3) the channel becomes busy due to any other SU's transmissions. We consider a homogeneous case that all SUs use a constant sensing window for channel access. Let the maximum sensing window $SW_i = W$. Given there are N_k SUs contending in channel k, the probability that the tagged SU wins the contention, i.e., all of the remaining SUs select a larger sensing window than the tagged SU, is given by

$$Pr(\text{SU } i \text{ wins contention}) = \sum_{j=1}^{W} \frac{1}{W} (\frac{W-j}{W})^{N_k}. \tag{6.5}$$

As the SUs contend for channel access only when no PU activity is detected, the probability that an SU's sensing fails due to other SU's transmission is given by

$$P^k_{SU} = \frac{\beta_k}{\alpha_k + \beta_k} (1 - F_k(t_{AS_i}))(1 - \sum_{j=1}^{W} \frac{1}{W} (\frac{W-j}{W})^{N_k}). \tag{6.6}$$

As it is very complicated to track the number of SUs in each data channel due to highly dynamic spectrum access in CRN, we use the average number of SUs to estimate the contention level in each channel, $N_k = (N-1)/N_d$, where N is the total number of SUs in the system, N_d is the number of data channels selected by the SU for transmissions. Therefore, the probability that an SU succeeds in sensing and attempts a transmission over channel k is

$$P^k_{ss} = Pr(\text{sensing succeeds}) = 1 - P^k_{CO} - P^k_{PU} - P^k_{SU}. \tag{6.7}$$

An SU transmits data when its sensing succeeds, or it switches to the next channel when sensing fails. The average time an SU spends on one transmission over channel k is

$$E[T^k] = P^k_{ss}(t_{AS_i} + t_i) + (1 - P^k_{ss}) t_{AS_i}. \tag{6.8}$$

A transmission succeeds only if no PU turns on during the total sensing and transmission time of the SU,

$$\begin{aligned} P^k_{TS} &= Pr(\text{No PU transmit and SU } i \text{ wins}) \\ &= \frac{\beta_k}{\alpha_k + \beta_k} \sum_{j=1}^{W} \frac{1}{W} (\frac{W-j}{W})^{N_k} F_k(t_{AS_i} + t_i). \end{aligned} \tag{6.9}$$

Or the transmission fails due to the disruption from PU with the probability

$$P_{TF}^{k} = P_{ss}^{k} F_k(t_i). \tag{6.10}$$

Without loss of generality, an SU checks the set of selected channels in a round robin sequence, $\{CH_1, CH_2, ..., CH_{N_d}, CH_1, ...\}$, until the packet is successfully transmitted. The probability that a transmission succeeds in the r-th attempt is

$$Ps(r) = P_{TS}^{r} \prod_{j=1}^{r-1} (1 - P_{TS}^{j}), \tag{6.11}$$

where P_{TS}^{r} corresponds to the probability of a successful transmission over the channel in the r-th attempts. We obtain the average transmission delay of an SU as

$$E[T] = \sum_{r=1}^{\infty} E[T^r] P_{TS}^{r} \prod_{j=1}^{r-1} (1 - P_{TS}^{j}). \tag{6.12}$$

6.4 Numerical and Simulation Results

6.4.1 Simulation Settings

In this chapter, we evaluate the performance of the proposed QC MAC via extensive simulations written in C. Three types of flows are simulated in a single hop network, i.e., voice, video, and data, as described in Section 3.3. The initial arrival time of each flow is uniformly distributed in [0, 5ms]. The channels are modeled as exponential ON/OFF models with the parameters listed in Table 6.2. The capacity of each channel is 10 Mbps. The channel switch time plus the basic sensing duration is 1 ms. SUs add an arbitrary number of mini-slots after the basic sensing duration and each mini-slot is 4 μs. The arbitrary sensing window setting is tabulated in Table 6.1. The delay bound of realtime voice and video traffic is set to be 20 ms. We run each experiment for 100s and repeat them 50 times to calculate the average value.

6.4.2 Delay of Homogeneous Traffic

We first study the delay performance of the proposed QC MAC in support of homogeneous traffic, i.e., voice or video flows, in Fig. 6.2. We compare the voice and video delay

Channel	α	β	$\frac{\alpha}{\alpha+\beta}$	Channel	α	β	$\frac{\alpha}{\alpha+\beta}$
CH1	0.215	0.4	0.351	CH6	0.1	0.1	0.5
CH2	0.054	0.1	0.35	CH7	0.653	0.4	0.62
CH3	0.278	0.4	0.41	CH8	0.163	0.1	0.62
CH4	0.069	0.1	0.409	CH9	1.2	0.4	0.75
CH5	0.4	0.4	0.5	CH10	0.3	0.1	0.75

Table 6.2: Settings of channel usage patterns in performance evaluation of Chapter 6

performance of the proposed QC with that of fractional (FRC) scheme which senses the channel in the descending order of the average channel available time. All SUs use the same sensing window [0, 31] without service differentiation. It can be seen that the average delay of voice/video traffic increases with the number of SUs. The delay of voice packets using both QC and FRC schemes are low because small voice packets are more likely to be transmitted opportunistically when PUs are inactive. For video traffic with much larger payloads, the probability of transmission failure becomes high as a PU is more likely to turn on and interfere with the SU during a longer transmission time of a video packet. When a transmission fails, an SU will switch to the next channel for sensing and retransmission, which results in a longer delay. It is also shown in Fig. 6.2 that the proposed QC MAC achieves much lower delay compared with FRC. This is because, in QC MAC, SUs always select a proper set of channels that assure high probability of successful frame transmissions, while only the average channel utilization is considered in FRC. As shown in the figure, the analytical results approximate those obtained by simulations well.

6.4.3 Delay of Heterogeneous Traffic

We then study the performance of QC MAC supporting heterogeneous traffic with different channel sensing policies, i.e., greedy and ascending, under various traffic loads in Fig. 6.3 and Fig. 6.4. When traffic load is low, e.g., there are 2 video flows and 1 to 5 voice flows in the network, greedy scheme achieves better delay performance than ascending scheme. Using the greedy scheme, SUs always select the channels with the highest success probabilities so that channels with good condition, e.g., fewer PU activities, will be efficiently utilized. When more voice flows joins in the network, ascending scheme slightly outperforms greedy for video traffic. For the greedy scheme, all SUs are likely to select the best channels for their transmissions and the contention level becomes high in those channels as the number of SU increases. In the ascending scheme, different users may select various channels that satisfy their QoS requirements, and thus the contentions among SUs are distributed over multiple channels. As shown in Fig. 6.4, when there are 10 video flows and

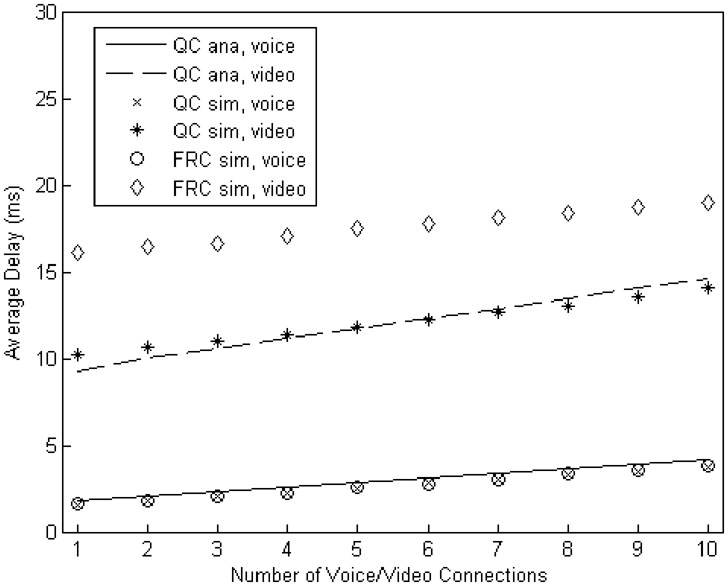


Figure 6.2: Average delay of voice/video flows (Greedy).

up to 30 voice flows, ascending scheme achieves better delay performance for video traffic. As small voice packets are more likely to take any opportunity for transmission and less likely to be interrupted by PUs, the delay performance of voice flows are low in all cases. Overall, greedy scheme is suitable for CRN with light traffic loads, while ascending scheme is more efficient when there are multiple types of SUs with different QoS requirements. Our proposed schemes achieves a much lower delay than FRC under different traffic loads.

6.4.4 Performance of Service Differentiation

We also investigate the performance of the service differentiation scheme, using the sensing window setting listed in Table 6.1. We have 10 voice and 10 video flows in the network. To study the impacts of background data transmissions, a saturated data flow is set up in each channel. As shown in Fig. 6.5, the delay of voice traffic does not change much with the traffic loads in the network; the delay of video traffic slightly increases; while the data throughput decreases when more video SUs join in the network. By applying different sensing windows for voice, video, and data, multimedia traffic have a higher priority to be transmitted when opportunity appears. It can be seen that the voice delay is around 7 ms

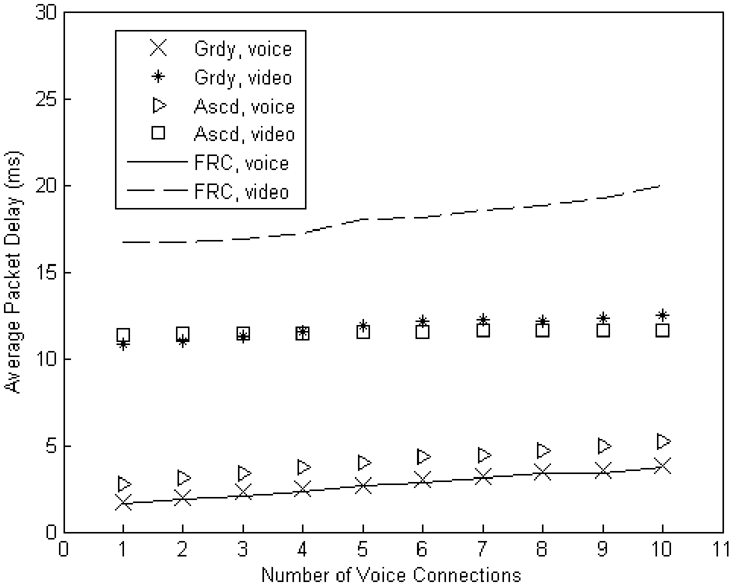


Figure 6.3: Comparison of channel sensing policies (low traffic load).

using strict sensing window setting, 9 ms when statistical window is applied, and 16 ms for constant window setting. Similarly, the video delay is around 27 − 30 ms, 30 − 34 ms, and 50 − 55 ms for strict, statistical, and constant sensing window settings, respectively. When a strict priority setting is applied, data packets have a lower probability to access the channel, and thus multimedia applications achieves a better delay at the cost of a lower throughput of data flows. In CRN in support of different types of multimedia applications, differentiated service is required to provision QoS for delay-sensitive real time applications.

6.5 Summary

In this chapter, we have proposed a distributed QoS-aware MAC with service differentiation for cognitive radio networks supporting heterogeneous multimedia applications. An analytical model has been developed to study the QoS performance of the proposed MAC, considering the activities of both PUs and SUs. Simulation results validate our analysis, and demonstrate that the proposed MAC provides satisfactory QoS support for multimedia applications.

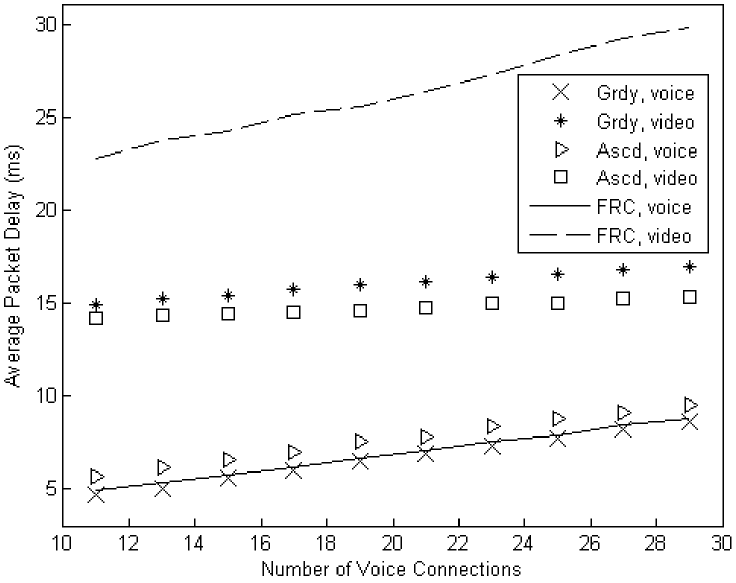


Figure 6.4: Comparison of channel sensing policies (high traffic load).

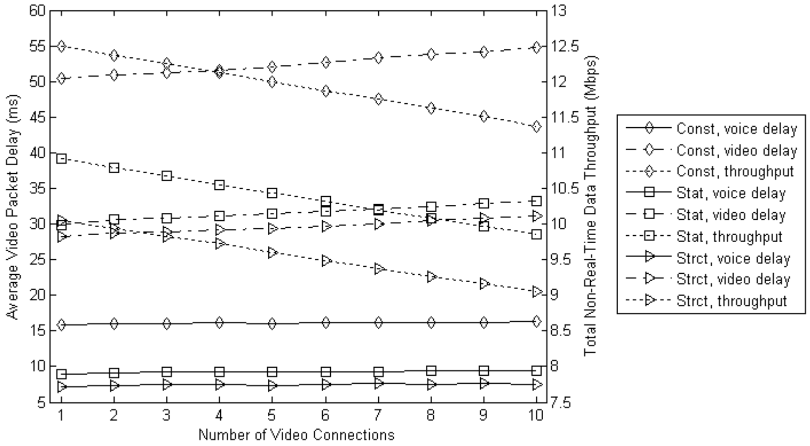


Figure 6.5: Performance comparison with different sensing windows.

Chapter 7

Cognitive Cellular Networks under Dynamic Resource Management

In this chapter, we propose a new framework of cognitive cellular networks which applies cognitive radio techniques in resource management and network coordination of cellular networks. We first study the trends in current cellular networks which greatly shape the challenges and exhibit the design potentials in improving network capacity and transmission quality. Specifically, the tiered network structure, energy awareness along with the security issues are identified as the main trends. Next, we treat the arising challenges in the trends by introducing the cognitive radio techniques on the aspects of network coexistence, dynamic spectrum access and the coordination mechanism design under limited bandwidth. Therefore, we propose a new framework of cognitive cellular networks in which we list three major research issues: interference management in the tiered network structure, mitigation of network bottlenecks, and coordination in resource management. For each research issue, we specify the features, compare the candidate solutions and verify the impact over the system performance. With a case study of cognitive cellular networks, we further discuss on the design of interference management in femtocell deployment where we apply the game theory to model the operations with heterogeneous network entities and limited inter-network coordination capability.

7.1 Introduction

The fast development of cellular communications corroborates the success of the mobile Internet which penetrates into our daily lives by connecting end user devices to the Internet with diverse services of qualities. Thanks to the powerful computation and communication hardware platform on the mobile devices, individual users generate more data than ever before. For example, a smartphone user generates as much as 35 times data of that of a voice-only cellphone. The global mobile data traffic has been doubled for the past four years in a row [1].

As the result, today's wireless devices are consuming more bandwidth than ever before, e.g., a single smartphone monthly consumes 35 times the bandwidth of a conventional voice-only cellphone [1]. Furthermore, the aggressive improvement on hardware and software platforms of mobile devices are keeping offloading more data traffic from the Internet wirelessly. The global mobile data traffic thus has been increasing exponentially in the past five years and such trend shows no fading in the expectable future [94].

To improve the network capacity and meet the explosively growing data demand, more frequency bands have been assigned for 4G cellular networks. Meanwhile, efficient communications techniques, such as MIMO and smart antennas, have been applied working with scheduling schemes designed for multi-dimensional resource allocation, such as in OFDMA systems, to effectively improve the network performance. Furthermore, to accommodate more users in the serving area, frequency reuse and network splitting are also introduced with interference management. However, these evolutional efforts cannot fully solve the bandwidth shortage. Besides, in today's ecosystem of cellular networks, which are comprised of both operators and users, end users are more actively participating into the networking and resource allocation to ensure their perceived quality of service (QoS). For example, to improve the communication quality in an indoor environment, end users could deploy femtocells which operate in the licensed spectrum [110]. With the launch of new mobile services, such as e-health and personal financial services, critical QoS, security and privacy issues are arising at both operators and end users [111].

Therefore, a re-visit to the network deployment and resource management issues is necessary at both operators and end users in radio access networks and the backhaul. To improve the overall system performance, the operators should not only optimize the resource utilization within the traditional domain of radio access, but also steer the usage patterns of end users. Some new opportunities are emerging, such as offloading traffic from macrocells to user deployed femtocells [110], or to operator deployed WiFi networks [112]. However, under such heterogeneous network deployment, spectrum sharing becomes complicated as multiple users attached with different network access portals generate mutual interference. Dynamic spectrum access in cognitive radio study has shown the potentials to further enlarge the pool of available resources for the users while reducing the access cost, e.g., the sensing delay. In a layered network structure with prioritized spectrum access, the users with lower priority trace the temporal and spatial distribution of spectrum access opportunities and adapt their transmission to the activities of prioritized user group [113]. A significant gain is expected by applying cognitive spectrum sharing in femtocells with efficient coordinations between the femtocells and the macrocells for resource allocation. Until now, few literatures have addressed this critical issue, and the specifications on the network deployment and operation are still open issues.

7.2 Trends in Cellular Networks

In cellular networks, two methods are usually used to meet the ever-increasing bandwidth demands of mobile users. The first method is to add more spectrum bands at the expense of billions of dollars. Since the spectrum resources are inherently limited and very expensive, the operators usually turn to the second method, i.e., deploying new physical and link layer techniques, such as MIMO, high order modulation and smart antennas, to further improve the spectrum utilization efficiency [114, 115]. However, these advanced techniques usually have high operational complexity and maintenance cost. To this end, a simple yet efficient solution is required to improve the spectrum utilization and dimension the future system design and management.

Huge investigations have been put into wireless infrastructure and the adoption of advanced link enhancement techniques, e.g., building more cellular towers with multiple antennas, using high performance radio amplifiers and higher order modulation and coding schemes, etc. However, as suggested by Shannon channel capacity, in bandwidth-limited case, These in-band approaches becomes more difficult as approaching to the theoretical limit since the marginal cost to achieve the incremental spectrum utilization increases tremendously, which makes the resource management complicated and inefficient [116].

7.2.1 Heterogeneous Cells and Tiered Network Deployment

Compared with macrocell base stations, base stations of small cells have a smaller communication range with a lower transmission power. Small cells are usually deployed as

Table 7.1: Specifications of different cells in cellular networks

	Macrocell	Microcell	Picocell
Transmit Power	50 W	a few Watts	>200 mW
Range	$1 \sim 5$ km	$300 \sim 1000$m	< 200 m
Deployment	Operator	Operator	Operator
Operating Bands	Operator's	Operator's	Operator's
Coverage	Outdoor/Indoor	Outdoor	Outdoor/Indoor
	Femtocell	WiFi	
Transmit Power	$10 \sim 100$ mW	$100 \sim 200$mW	
Range	$20 \sim 30$m	$100 \sim 200$m	
Deployment	User	User	
Operating Bands	Operator's	Unlicensed	
Coverage	Indoor	Indoor	

the complement of macrocells with different purposes, e.g., to improve network capacity in hotspots, to compensate the long distance loss for users at network edge and to provide coverage in the blind zone, as shown in Fig. 7.1. According to the working frequency and the deployment and control schemes, small cells can be roughly categorized into two types:

Out-of-band small cells: these small cells operate in the frequency bands other than the licensed frequency bands of macrocells, e.g., WiFi cells in the unlicensed 2.4 GHz ISM bands. Out-of-band small cells are usually deployed by end users. Nowadays, the operators become interested in deploying WiFi access to offload mobile data from cellular cells to WiFi hotspots. Such operator-deployed WiFi networks are usually open to their own subscribers only.

In-band small cells: these small cells operate in the same frequency bands as the macrocells and are usually deployed and managed by the operators, which are referred to as microcells and picocells. Recently, a new type of small cell, i.e., femtocell, is introduced to enhance the indoor cellular signal with a simplified cellular base station (BS) connected to the cellular core network via the third party Internet cable service [110]. Because these femtocells can be deployed by end users, the operators only have limited control over these femtocells, which makes it challenging to mitigate the co-channel interference and manage the radio network resources.

A brief summary of the specifications of small cells is listed in Table 7.1. Compared with macrocells, small cells have unique advantages in some usage scenarios, such as capacity enhancement in hotspots, and coverage expansion into the houses and workplaces. Basically, small cells allow for flexible BS deployment and simple transceiver design due to the limited communication coverage. As the radio environment is becoming more and more complicated, using small cells is beneficial for the operators to deal with the localized coverage and link enhancement while offloading the traffic from the macrocells to femtocells or WiFi [1, 117].

Recognizing the differences between macrocells and small cells, it is necessary to revisit the network planning and management issues in a tiered network integrating both macrocells and small cells. First, different cells may have different capabilities of serving users. Small cells can provide a higher throughput for local users while macrocells provide mobile services with reduced link capacity for remote users. Second, it is critical to determine the cell size and the number of cells to achieve the maximum network capacity in the serving area. Users can select the best access cell among multiple visible cells to achieve a high diversity gain, however, more visible cells will cause more burden on the network coordination and energy management [118]. Last, but not the least, the implementation encounters more challenges. For the in-band small cells, severe co-channel interference from and into the macrocells may degrade the performance of the whole cellular networks.While for out-of-band small cells operating independently from macrocells, operators may not be able to rely on a centralized control architecture to help schedule traffic offloading from macrocells to small cells. Distributed traffic offloading remains an open research issue.

7.2.2 Green Cellular Networks and Sustainable Communications

Green radio communication networks have attracted great attention recently as the information and communication society realized the necessity for achieving energy efficiency and being environmental friendly [119, 120]. Operators and users are resort to efficient power management solutions to reduce the energy consumption of BSs and extend the battery life of mobile devices.

Although cellular networks are widely deployed to provide ubiquitous wireless access worldwide, some users in developing countries have very restricted access especially when they roam in off-grid suburb areas where power supplies rely on transported fuels such as diesel, which are very expensive. To reduce the cost of off-grid BS deployment, researchers and engineers work on the development of green base stations, i.e., base stations that are powered by sustainable power supplies, such as solar, wind and tides [121]. Unlike

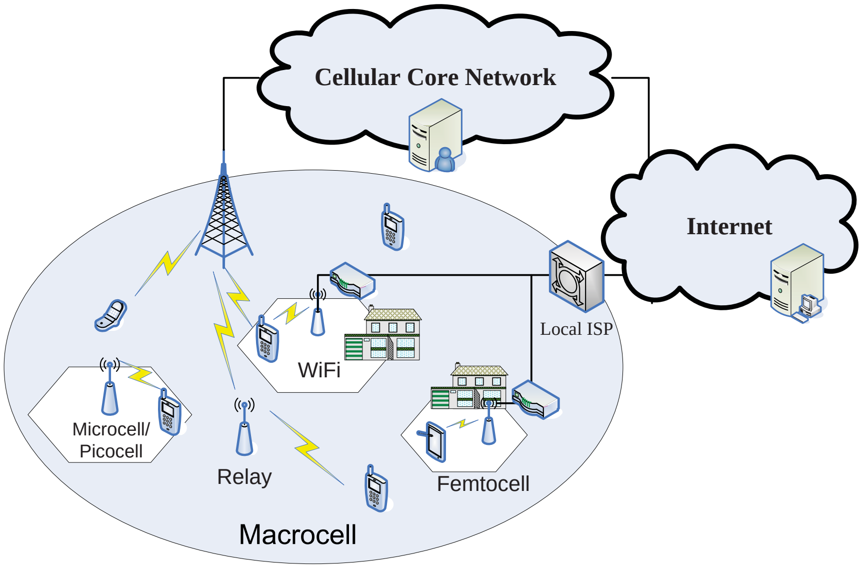


Figure 7.1: Tiered framework of cognitive cellular networks

traditional electricity power supply, the renewable power supply is inherently variable in its availability and capacity, which poses great challenges in the network resource management. Considering dynamic characteristics of sustainable energy sources, in a green cellular network powered by renewable energy, the fundamental design criterion and the main performance metric should shift from energy efficiency to energy sustainability [122]. Under such a new green network paradigm, network planning and resource management issues should be thoroughly re-visited.

7.2.3 Secure Communications and User Privacy

In a tiered network, users can set up small cells for offloading data which will traverse the cellular core network to the Internet. This opens a door for malicious attackers who can easily set up a femtocell and eardrop or even change the information traversed over the cell. Although there are some existing attack models and analysis in general computer networks, these models do not properly capture the openness and flexibility in the spectrum access of femtocells. How to ensure secure data transmission and preserve user's privacy in the new tiered network with secure macrocell and open femtocell is still an open issue.

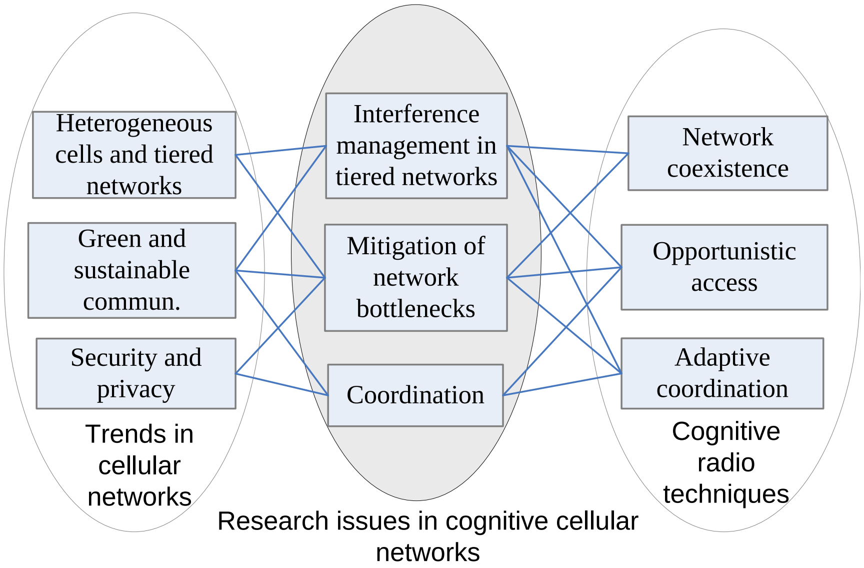


Figure 7.2: The study framework of cognitive cellular networks

7.3 Cognitive Cellular Networks

To address the challenges brought by the trends emerging in cellular networks, we should revisit the research issues in network deployment and resource management from multiple aspects. In general, it requires flexible network deployment methods over diverse spectral environments and dynamic resource management schemes in heterogeneous networks. We propose a new framework, called *cognitive cellular networks*, as shown in Fig. 7.2. We apply the cognitive radio techniques and investigate in cellular networks in the following issues: 1) heterogeneous network coexistence, 2) spectral diversity and opportunistic access, and 3) adaptive interworking with constrained network coordination.

Specifically, a strong candidate solution in cognitive cellular networks should first allow and facilitate the coexistence of small cells of the same or different types, e.g., in-band and out-of-band cells, with macrocells under a spectrum access strategy. In cognitive radio networks, the coexistence problem is formulated between two independent user groups, i.e., primary users (PUs) and secondary users (SUs), respectively. SUs coexist with PUs under a predefined spectrum sharing method which specifies the visibility of nodes intra- and inter- user groups, priority in spectrum access and conflict resolution [113]. Usually, the spectrum sharing methods can be categorized as overlay, underlay or interweave to agree with the requirements in different deployment scenarios. In overlay mode, for instance, PUs actively participate in the spectrum sharing and release some bandwidth in exchange for SUs' relay assistance during PUs' transmissions to mitigate the interference from SUs. While in interweave mode, spectrum sensing and the database of channel usage pattern are the primary solutions to deal with the coexistence issues since SUs are transparent to PUs.

The second question is how to identify the available resources for the transmissions with different priorities so that it can improve the spectrum utilization efficiency. In cellular networks, users are usually scheduled for data transmission in the time, frequency, code and space domains by a central controller. In a tiered network with heterogeneous network environments, however, the centralized approach may not be available or will be costly from both computational and communication aspects. In cognitive radio networks, the available spectrum resources have been finely identified at different locations and times. The transmission pairs select the spectrum access opportunities which can satisfy the required transmission qualities, e.g., length and/or bandwidth. And the traffic flows are routed according to the distribution of spectrum resources at the nodes [113]. Introducing adaptive resource management in cognitive cellular networks can improve the resource utilization efficiency via making opportunistic transmission decisions based on the local traffic and channel conditions.

It is also critical to design efficient interworking schemes for heterogeneous cells in cognitive cellular networks. In the tiered network architecture, the nodes have diverse capabilities in transmissions. The coordinations between end users and cells or inter-cells greatly affect the network performance since the mismatch of the operations or inappropriate transmission settings would generate severe interference. When the coordination has constraints in the network topology and limited bandwidth for the control panel, the case becomes worse. For example, the coordination between the femtocell and the macrocell is limited since the femtocell BS is indirectly connected to the cellular core network through a local Internet cable, which prohibits the operators to perform integrated network operations. Distributed decision making has been shown as a promising solution in the cognitive radio study [123]. Based upon partial and/or delayed network information, e.g., channel gains, the distributed decision making process can be modeled to capture the interworking between the femtocells and the parent macrocell. To achieve efficient spectrum sharing among a large number of distributed users, game theoretical approach is usually used for resource management of heterogeneous cells [124, 125, 126].

In cognitive cellular networks, the aforementioned issues are considered on the following aspects including interference management over the tiered network architecture, mitigating the network bottleneck for opportunistic and energy efficient spectral access, and coordination schemes to optimize the network utility. A good candidate solution to the dynamic resource management usually integrates the cognitive radio techniques to optimize the performance and concerning about the implementation issues, as shown in Fig. 7.2.

7.3.1 Interference Management in Tiered Networks

In cognitive cellular networks, small cells are employed to enhance the link quality and network capacity. As small cells are operating in the same frequency as the macrocells, severe co-interference exists among macrocell and small cell users. As shown in Fig. 7.1, the mobile users served by macrocells may move to the edge of cell where they may experience strong signals from the private femtocells. Similarly, the low power transmissions in small cells are also likely interfered with macrocell users. To mitigate the co-channel interference, some candidate approaches have been proposed, including:

Spectrum splitting approach refers to the resource allocation by assigning orthogonal resources, e.g., subcarriers, to the transmission pairs with strong interference.In the tiered network, the operator can split the spectrum into subbands and assign them to the neighboring small cells to reduce the interference between the neighboring cells. However, such static allocation may cause waste of spectrum and reduce the adaption to the varying traffic demands.

Power control approach is to adjust the transmit power of nodes in the network to secure the reception quality at the receivers. It is a good candidate to reduce the interference in the network and encourage the energy efficient transmissions. However, the central controller needs to acquire actual channel conditions and nodes' operational parameters to optimize the performance, which introduces heavy coordination cost, especially in the tiered architecture.

Offloading approach tries to reduce the strong interference source by arbitrarily handover these users to the cells with better link qualities to mitigate their interference over the neighbors. In this approach, both link qualities and the resource allocation needs to be considered before the handover [127]. The availability of such cell is another issue when the targeted femtocell is of closed access for its private user only.

7.3.2 Network Bottleneck Mitigation

In cognitive cellular networks, as small cells becomes more likely to be deployed by users, it is very difficult for operators to determine the available network resources in realtime operations. In addition, the capacity of access links (e.g., the links between the users and the femtocell BS), and backhaul links (e.g., the one between the femtocell BS and its parent macrocell BS) may vary. The cellular downlink throughput can achieve 100 Mbps, while the backhaul of femtocell has limited capacity provided by the Internet service providers, normally up to 10 Mbps according to the data plan by regions and price. Therefore, the smaller bandwidth of femtocell backhaul becomes network bottleneck that limits the quality of service of users. To tackle this problem, a possible solution is to allow multi-path data transmissions through different network interfaces, e.g., using WiFi and cellular networks [112], for the throughput aggregation at the end users.

In a communication network where the wireless backhaul is the bottleneck, opportunistic data forwarding is an efficient solution for cognitive cellular networks, by jointly considering the forwarding capability of femtocell BSs and the traffic loads, as proposed in [113]. Specifically, the femtocell BS evaluates its forwarding capability based upon the expected relay advancement in the forwarding direction as well as the interference in the transmission channels, which determines the order of relay candidates along the forwarding path. To fight against the fading in wireless channels, the proposed forwarding scheme incorporates multiple nodes at each transmission so that the successful receiver, if there is any, can continue with the data forwarding if the nodes with higher forwarding capability fail. Such opportunistic forwarding scheme well adapts to the dynamic channel conditions and significantly reduces the link failures and the resulting retransmissions in the backhaul.

7.3.3 Coordination in Cognitive Cellular Networks

In cognitive cellular networks, a user senses the channel conditions and makes the best strategy for its own utility. The egocentricity of individual operations may impair the whole network performance when effective coordination mechanisms are missing. Overall, the resource management in cognitive cellular networks can be formulated as a network utility maximization problem as follows,

$$\begin{aligned} &\max_{\mathbf{a}} \sum_{i\in C}\sum_{j\in C_i} U_{\mathbf{a_j}} && (7.1)\\ \textbf{s.t.}\quad & \mathbf{I_a} \le \mathbf{\Gamma} \\ & U_{\mathbf{a_j},\mathbf{a_{-j}}} \ge U_{\mathbf{a'_j},\mathbf{a_{-j}}}, \quad \forall\ \mathbf{a_j}\in\mathbf{a},\ \mathbf{a'_j}\neq\mathbf{a_j}. \end{aligned}$$

Specifically, under a transmission strategy, denoted by $\mathbf{a}$, which specifies the operation parameters of each node, e.g., cell selection, transmit power, etc., the objective is to maximize the aggregated utility functions of all links in the network, i.e., $\max_{\mathbf{a}} \sum_{i\in C}\sum_{j\in C_i} U_{\mathbf{a_j}}$

where C is the set of cells including all macrocells and small cells in the network, and C_i represents the set of active wireless access connections in cell i. Given the other nodes' transmissions, $\mathbf{a_{-j}}$, each node selects its transmission strategy, $\mathbf{a_j}$, to best respond to $\mathbf{a_{-j}}$, i.e., $U_{\mathbf{a_j},\mathbf{a_{-j}}} \geq U_{\mathbf{a'_j},\mathbf{a_{-j}}}, \ \forall \, \mathbf{a_j}, \mathbf{a'_j} \in \mathbf{a}, \ \mathbf{a'_j} \neq \mathbf{a_j}$. Furthermore, one candidate transmission strategy should not violate the network coexistence rules Γ, which determines the maximum allowable interference in the links, i.e., $\mathbf{I_a} \leq \mathbf{\Gamma}$. The operators manipulate the decision making of individuals from the network aspect, such as load balance, interference management, and security. Candidate approaches include introducing incentive schemes [124], defining new utility functions for players [128], etc.

Besides the competitions in the zero-sum game for radio resources, users and small cells can also cooperate for the channel condition monitoring, handover management and relay transmission. The cooperation can benefit the users who have limited capability to acquire the necessary network or channel conditions to make decisions. No matter competition or cooperation, the participating users require the knowledge of all possible moves of other players or the required coordination information in cooperative communication. In cognitive cellular networks, the design of the coordination connections is critical by considering the overhead and performance.

7.4 Dynamic Resource Allocation in Tiered Cellular Networks

We investigate the power management problem in a tiered network with both macrocells and femtocells. Self-deployed femtocells may cause severe interference with nearby macrocell users. As shown in Fig. 7.3, Femtocell 2 is located at the edge of the macrocell. In the downlink, the leaked signal from Femtocell 2 to the nearby macrocell user, UE2, may be stronger than UE2's received signal from the macrocell base station as UE2 is located at the edge of the macrocell. Therefore, Femtocell 2 introduces significant interference on UE2's transmission. Many literatures on femtocells have addressed such problem [110]. However, existing solutions mainly focus on the centralized resource management, which may not be suitable for a tiered cognitive cellular network where a robust distributed approach is more desirable due to the random deployment of femtocells.

In reality, the central controller of the macrocell can hardly fully control the affiliated femtocells because these femtocells may not follow the scheduling information but like to aggressively compete for network resources to maximize their own utility. For example, femtocell base stations can increase their transmit power for achieving higher throughput

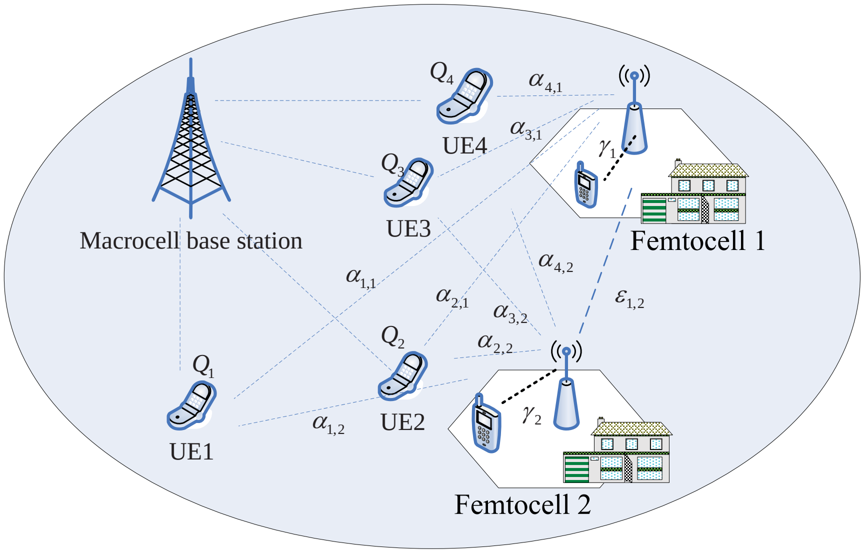


Figure 7.3: Interference management in cognitive cellular networks

while causing greater interference to the neighboring users. In current cellular network, the central controller may not be able to specify the violation behaviors of individual femtocells, even when the neighboring macrocell users report the reception failure caused by such violation. In other words, the central controller can not effectively to eliminate the co-interference resulting from self-deployed femtocells.

To analyze the motivation and behavior of femtocells in the violation of the centralized scheduling, we apply a game theory approach to study power management, which is widely used for resource allocation among PUs and SUs in cognitive radio networks [128]. We derive the downlink interference, and the uplink analysis can be obtained in a similar way.

As shown in Fig. 7.3, a group of closed access femtocells are located in the serving area of a macrocell, all cells operating in the same frequency band. In the macrocell, there are M active macrocell users in the downlink, denoted by $U = \{u_1, u_2, \ldots, u_M\}$. u_i has a threshold Q_i, which indicates the maximum tolerable interference level in the downlink. The macrocell users are scheduled to transmit in non-overlapping resource blocks so that there is no interference among macrocell users. The active femtocells in the downlink form a set $F = \{f_1, f_2, \ldots, f_N\}$ with the size of N. In a femtocell, the femto BS schedules one user for transmission at one time. Therefore, there is one active link in each femtocell at any time. The transmit power of the femtocell f_j is denoted by P_j. We assume that the channel gains, $\alpha_{i,j}$ of the link between u_i and f_j, $\epsilon_{j,k}$ of the link between f_j and f_k, and γ_j of the link within f_j are known to each femtocell as well as Q_i of each u_i, and the links are symmetrical.

The aggregated interference at u_i should satisfy $P_1\alpha_{i,1} + P_2\alpha_{i,2} + \cdots + P_N\alpha_{i,N} \leq Q_i$, otherwise, u_i is blocked. The capacity function c_j for femtocell f_j is defined as

$$c_j = log_2(1 + \frac{P_j\gamma_j}{N_0 + \Sigma_{k\neq j}P_k\gamma_k}) \tag{7.2}$$

It is obvious that the femtocell BSs prefer to using the maximal transmission power to achieve the highest link capacity. Therefore, femtocells may like to violate the power control strategy made by the central controller, which may cause co-channel interference with macrocell users. To address this issue, we apply randomized silencing policy proposed in cognitive radio networks [124]. The policy is very straightforward, i.e., if any macrocell user u_i experiences the interference greater than its limit Q_i, the central controller will randomly select one active femtocell from F and force it to turn off in current transmission period. Such silencing process continues for several rounds until no macrocell user reports the block case[1].

Given a power allocation strategy of femtocells, $\mathbf{P} = \{P_1, P_2, \ldots, P_N\}$, once the interference requirement is met at each macrocell user, the utility of macrocell is determined. Therefore, the power control problem in this case study can be formulated as

$$\begin{aligned} &\max_{\mathbf{P}} \sum_{j\in F} E[c_j \cdot \mathbf{1}_j] && (7.3)\\ \mathbf{s.t.}\;\; &P_1\alpha_{i,1} + P_2\alpha_{i,2} + \cdots + P_N\alpha_{i,N} \leq Q_i, \forall i \in U \\ &\mathbf{P} \in NE \end{aligned}$$

where the objective of resource allocation is to find the maximum aggregated utility of femtocells, which can be denoted as $\max_{\mathbf{P}} \sum_{j\in F} E[c_j \cdot \mathbf{1}_j]$ with the function $\mathbf{1}_j = 1$ if f_j is not shut down after the silencing process, and 0 otherwise. As each femtocell intends to maximize its utility by selecting the transmission power best responding to the transmission powers of other nodes, a candidate $\mathbf{P}$ would be a power allocation of Nash Equilibrium (NE).

Here, we present some preliminary results to explore the NEs for the optimal value. Using the theorems in [124], we can easily prove the orthogonal power allocation $\mathbf{P}_{OR} = \{\min_i\{\frac{Q_i}{\alpha_{i,1}}\}, \min_i\{\frac{Q_i}{\alpha_{i,2}}\}, \ldots, \min_i\{\frac{Q_i}{\alpha_{i,N}}\}\}, i \in U$ is an NE.

---

[1] The shutdown process is valid in cellular networks where macrocell users are protected with higher priority because they have been admitted in the serving macrocell. When the self deployed femtocells register at the cellular operator, they are required to yield to the priority of macrocells if confilictions occur.

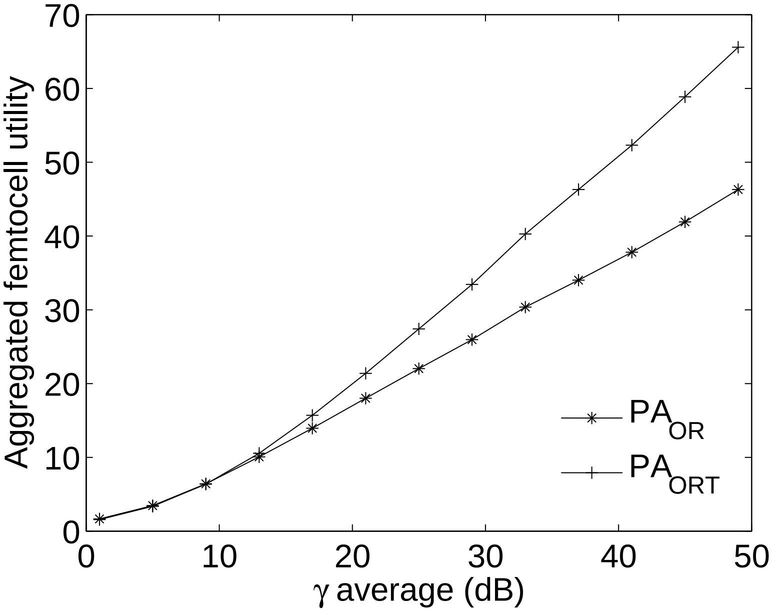


Figure 7.4: Aggregated femtocell utility under intra-femtocell channel gains

Moreover, we notice that a macrocell user causes significant interference when it is closer to some active femtocells than others. Therefore, the femtocell needs to avoid transmission in the resource block assigned to the nearby macrocell user if it is detected by the femtocell[2]. The femtocell can learn the allocation of resource blocks of each macrocell user by listening to the allocation message broadcasted by the macrocells at the beginning of each transmission period. In such case, we can also prove that the orthogonal power allocation $\mathbf{P}_{ORT} = \{\min_{i \in U \setminus S_1}\{\frac{Q_i}{\alpha_{i,1}}\}, \min_{i \in U \setminus S_2}\{\frac{Q_i}{\alpha_{i,2}}\}, \ldots, \min_{i \in U \setminus S_N}\{\frac{Q_i}{\alpha_{i,N}}\}\}$ is an NE where S_j is the set of macrocell users near f_j with the channel gain $\alpha_{i,j}$ greater than a predefined threshold[3].

To evaluate the performance of the proposed power allocation scheme, we simulate the network with Rayleigh fading channels where all $\alpha_{i,j}$ fade independently with average $\bar{\alpha}$, all

---

[2] It is valid in cellular networks because the users measure the signal strength from visible cells and the femtocell can detect such nearby macrocell user through this process although it may rejects the user's access requst if the user is not a private member.

[3] Comparable with macrocell users, femtocells have also the predetermined QoS threshold. For example, in a femtocell f_j, it has the SINR threshold μ_j, then the selected transmission power should be greater than $\frac{\mu_j}{\gamma_j}$ for successful reception at the receiver.

$\epsilon_{j,k}$ fade independently with average $\bar{\epsilon}$, and all γ_j fade independently with average $\bar{\gamma}$ [124]. We set $M = 20$ and $N = 5$ in the macrocell, and $Q = 1$, $\bar{\alpha} = 1dB$, $\bar{\epsilon} = 1dB$, and $\bar{\gamma}$ ranges from 1 to 50 dB. In each experiment, we randomly select two femtocells with a significant neighboring macrocell user, i.e, $\bar{\alpha} \approx \bar{\gamma}$. It can be seen in Fig. 7.4 that the recognition of the significant interference sources in femtocells can improve the interference management in cellular network and achieve a higher throughput for femtocells.

Based upon the discussion above, a good candidate solution is to integrate the considerations of the performance requirements and the corresponding techniques. Specifically, we treat the interference management problem as the major design objective by identifying the link interference conditions in the macrocell and the femtocells. Meanwhile, we also take into considerations the performance requirements of limited interworking bandwidth and efficient coordination. On the network bottleneck mitigation, regarding the limited bandwidth from the femtocell to the core network, we design the solution using game theory to analyze the nodal behavior in favor of its own utility. On the coordination mechanism design, we develop the random shutdown scheme as the penalty for misbehavior which generates less control overhead while maintaining effective regulation effect.

7.5 Summary

In this article, we have studied network deployment and resource management in a cognitive cellular network with both macrocells and small cells in a tiered architecture. We have discussed the main research trends and challenging research issues in cellular networks and proposed a framework of cognitive cellular network to address the challenges. A game theory based approach for efficient power control has also been provided for studying resource allocation in a cognitive cellular network where femtocells are deployed.

7.6 Appendix: Proof of Nash Equilibrium

Proposition 7.6.1 $\mathbf{P}_{OR} = \{\min_i\{\frac{Q_i}{\alpha_{i,1}}\}, \min_i\{\frac{Q_i}{\alpha_{i,2}}\}, \ldots, \min_i\{\frac{Q_i}{\alpha_{i,N}}\}\}, i \in U$ *is an NE under random silencing.*

Proof 3 *With orthogonal power allocation, each femtocell user generates an interference at each macro user equal to Q. According to the violation penalty, if any macro user reports* *the interference violation, the random silence will proceed. Therefore, from the point view of each single femtocell user, it only has the chance to transmit with the power equal to the minimum allowed power by any active macro user nearby, i.e.,* $\min_i\{\frac{Q_i}{\alpha_{i,j}}\}$ *for femtocell user* j*. On the one hand, no user can increase its utility by decreasing its power. On the other hand, if any user increases its power (thus violating the* Q*), then it is shut down with certainty and its utility is always zero. This shows that* $\mathbf{P}_{OR}$ *is always a NE.*

Next, if the most significant interference sources have been allocated in orthogonal resource slots, then, for each femtocell, it can increase its instant transmission utility by increasing the power in the active resource slots. However, the total resource slots have been reduced.

Proposition 7.6.2 *the orthogonal power allocation*

$$\mathbf{P}_{ORT} = \{\min_{i\in U\setminus S_1}\{\frac{Q_i}{\alpha_{i,1}}\}, \min_{i\in U\setminus S_2}\{\frac{Q_i}{\alpha_{i,2}}\}, \ldots, \min_{i\in U\setminus S_N}\{\frac{Q_i}{\alpha_{i,N}}\}\}$$

is a NE under random silencing.

Proof 4 *Once the set of the significant interference sources is determined by each femtocell, then the resource slots for femtocell is determined by* $\frac{M-m_j}{M}$ *where* $m = |S_j|$*. On the one hand, no femtocell can increase its utility by decreasing its power. On the other hand, in the allocated resource blocks, no femtocell user is motivated to increase its transmission power because it would definitely violate the interference at the most vulnerable macro user in* $U \setminus S_j$ *and be shutdown without no gain.*

Chapter 8

Conclusion and Future Work

8.1 Conclusion

This dissertation reports our research work on enabling secondary spectrum reuse for CRN in heterogeneous wireless channels. With the distinguished dynamic and diverse features of spectrum access opportunities, our target is to explore the efficient, scalable and, more importantly, practical solutions to accommodate secondary traffic, protect primary services and improve spectrum utilization. To this end, a comprehensive and systematic study which redesigns the communication layer protocols and optimizes the resource management schemes to address those new features and challenges is necessary. In this chapter, we summarize our research outcome from the perspectives of opportunistic spectrum reuse, end user adaption and network coordination, respectively.

8.1.1 Opportunistic Spectrum Reuse

In CRN, SUs operate at the low power level to exploit locally unused channels and prevent harmful interference to licensed transmissions. In order to provide the connection between any two SUs out of the direct transmission range, the multi-hop transmission is necessary. This, however, is a daunting task due to the intense route establishment and maintenance cost. In Chapter 4, our goal is to develop a practical and cost-effective solution on building the path dedicated for secondary multi-hop transmissions in CRN. To this end, we design an opportunistic cognitive routing scheme in which SUs update the routing tables based upon physical characteristics of spectrum bands and diverse activities of PUs in each band. The forwarding decision is made by jointly estimating the remaining distance to the destination and the reliability of transmissions in each candidate channel. Since SUs only need to update limited global information, i.e., the geographical information of the source-destination pair, it provides the efficient routing as SUs mainly communicate with the close neighboring nodes to exchange routing messages. In Chapter 5, based upon the idea of local opportunistic forwarding, we further examine the impact of multiple relay candidates in per hop relay selection and the joint design of channel selection and route selection. Recognizing the short time window of available spectrum, a prioritized feedback mechanism is designed to reduce the contentions among multiple relay candidates. Via extensive simulations, we show the effectiveness of the proposed scheme, which outperforms existing routing schemes in CRN by adapting to the dynamic and diverse wireless channels for opportunistic channel access.

8.1.2 End User Adaption

The basic goal of any communication systems is to provide the users with their *desired* service quality of applications. Therefore, to investigate on the specific user requirements and provide the corresponding service guarantee is the fundamental issue in the communication system design. This, however, is a very challenging issue for CRN due to the nature of unlicensed transmissions. In Chapter 6 of the dissertation, we focus on the channel selection in spectrum sensing and sharing to satisfy QoS of different types of secondary traffics. Recognizing the statistics of traffic in time and packet size given the required QoS, we propose a distributed QoS-aware MAC protocol so that SUs perform sensing and access based upon the channel usage pattern in each single channel. Moreover, we can manipulate the channel access priority of various types of traffic by applying differentiated arbitrary sensing periods in the detected spectrum holes, which further enhances the QoS provisioning of the proposed cognitive MAC. The analytical model and extensive simulations show that our proposed MAC can achieve multiple levels of QoS provisioning for various types of multimedia applications in CRN.

8.1.3 Network Coordination

In the presence of intense network dynamics, how to fully utilize the varying network resources towards the global welfare of the system is a fundamental issue of the tiered networks. In order to conduct effective network control and resource allocation, CRN provide abundant candidate solutions to exchange control messages between SUs with/without coordination with PUs. However, in a more general tiered communication network infrastructure, e.g., cellular heterogeneous networks, effective and efficient network coordination is still an open problem. This motivates our research in Chapter 7 on exploring solutions in cellular communications based upon the lessons learnt in CRN. In Chapter 7, we first address challenging issues including interference management, network coordination, and cell interworking in a tiered cellular network with both macrocells and small cells. We identify the different network characteristics of macrocells and small cells, which exhibit the similar features as the ones between PUs and SUs. We then propose an adaptive resource management framework to integrate the operation objectives of different cells into the global one of improving spectrum utilization and mitigating the co-channel interference between cells. In this manner, a game-theory-based approach for efficient power control is provided. We verify the efficiency of the coordination in the proposed schemes and the effectiveness of protecting the transmissions of macrocell users.

8.2 Future Work

Towards the efficient, low-cost and ubiquitous spectrum reuse for unlicensed services, there are still many open issues remaining to be solved. Next we outline several interesting directions for future work.

8.2.1 Cooperative Communications in CRN

Most discussions in CRN are currently build upon the fundamental assumption that PUs and SUs are operating independently with exclusive utility of the access spectrum. However, to gain more acceptance of open spectrum reuse, a good business model which can promote CRN is necessary to consider the welfare of license holders in the spectrum sharing game. Moreover, in practical deployment, spectrum sensing errors are inevitable. As an alternative, cooperation has been leveraged in CRN to gain spectrum access opportunity and eliminate harmful interference to PUs caused by sensing errors. In cooperative cognitive radio networks (CCRN), SUs relay PUs' traffic using cooperative communication techniques in return for spectrum access opportunities in the remaining resource blocks owned by PUs. In this manner, PUs with poor link qualities are motivated to proactively participate into spectrum reuse given the incentives of improving their own transmissions through SUs' help. It is promising but challenging for SUs to cooperative with PUs. First, SUs are obliged to apply more aggressive transmission schemes to support both primary relay and secondary traffics in the spectrum. Given the spectrum resources and the demanded QoS of primary services, SUs would achieve their own chances to transmit only when they complete the assist primary transmissions as soon as possible. Therefore, to improve their own utility, SUs need to improve PUs' transmission performance at the first hand, which is the distinguished feature in CCRN. Second, secondary transmissions are cost-sensitive operations for both PUs and SUs. On the one hand, PUs need to select the appropriate relaying SU set from neighboring SUs by evaluating the link quality. Different SUs may exhibit different capabilities of forwarding traffic, which in turn determines the amount of remaining resources for secondary transmissions. On the other hand, SUs need to take the forwarding operations into account as the expected transmission cost to evaluate the cooperation gain for their own transmissions. Since PUs and SUs have their own services, effective coordination in relay selection and relay volunteering is critical in CCRN. Third, as PUs always prefer the set of SUs who can enhance the transmissions and reward them with bandwidth, the fairness among SUs needs to be considered in case some SUs may starve simply due to the inferior position for relaying. Some incentive mechanisms should be introduced into SUs.

8.2.2 Energy Efficiency in CRN

Most research works in CRN mainly focus on the spectrum efficient merits in the system design and optimization, such as the throughput and delay. Since secondary transmissions are based upon spectrum exploration, spectrum sensing and negotiation with licensed holders consume a notable piece of energy in CRN operations. One of the biggest impediments of future wireless communications systems is the need to limit the energy consumption of battery-driven devices so as to prolong the operational times and to avoid active cooling. In fact, without effective energy saving, there is a significant threat that SUs will be searching for power outlets rather than available spectrum. In CRN, energy aware system design lies in almost all aspects of SUs' operations, such as efficient secondary transmissions, optimized topology control, effective coordination and cooperation. Among these efforts, for wireless devices, temporarily turning off idle radios and scheduling devices in sleep mode can significantly reduce the energy consumption and prolong the working period. However, there are several challenging issues to introduce power saving mode in CRN. First, since SUs need to keep tracing the channel usage patterns, new mechanisms are necessary to acknowledge awaken SUs with the latest channel status to help them to quickly resume to work. Information exchange within neighboring SUs can be an option in the design of network/cluster entry procedures. Moreover, in infrastructureless CRN, SUs not only transmit their own packets, but also help forward packets from other nodes. When power saving is introduced, the individual sleep decision should also incorporate the impact on the entire network utility. For example, some SUs may act as the “anchor nodes” to maintain network connectivity. Therefore, some incentive mechanisms are also expected to keep balance between the individual energy efficiency and global performance.

8.2.3 Heterogeneous Network Design

A totally realtime spectrum trade market is expected as the promising future for dynamic spectrum reuse to fully exploit the potentials of current spectrum bandwidth for emerging wireless services. Since the implementation of CRN is still in its infancy, enabling frequency reuse among the heterogeneous network entities in the same operator’s network is a good trial test for the proposed CRN technologies. As proposed in Chapter 7, the cellular heterogeneous networks (HetNet) still have many challenging issues to study using cognitive radio technologies. In HetNet, users can deploy their own network access points to build the private small cells for exclusive cellular access, which is out of scope of the conventional network planning and optimization. As it is expected that there would be hundreds of these small cells installed in a macrocell, interference management in HetNet becomes a severe problem. Furthermore, small cells are in charge of offloading cellular traffic from macrocells to allow macro base stations to better serve mobile public users. A strong backhaul is necessary to connect the access points of small cells to the cellular core network. The robust control over the backhaul is essential. For example, the collapse of backhaul may push all private users’ traffic back to macrocells, which results in chain effect and threatens the cellular safety.

References

[1] Cisco, "Cisco visual networking index: global mobile data traffic forecast update, 2011-2016," Feb. 2012.

[2] NTIA, "U.S. frequency allocations," http://www.ntia.doc.gov/osmhome/allochrt.pdf, [Online; accessed 11-Feb-2013].

[3] D. A. Roberson, C. S. Hood, J. L. LoCicero, and J. T. MacDonald, "Spectral occupancy and interference studies in support of cognitive radio technology deployment," *Proc. IEEE SDR'06*, pp. 26–35, 2006.

[4] PCAST, "Report to the President: realizing the full potential of government spectrum to spur economic growth," http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast_spectrum_report_final_july_20_2012.pdf, July, 2012, [Online; accessed 11-Feb-2013].

[5] C. Cordeiro, K. Challapali, D. Birru, and S. Shankar, "IEEE 802.22: The first worldwide wireless standard based on cognitive radios," *Proc. IEEE DySPAN*, pp. 328–337, 2005.

[6] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Comput. Netw.: Int. J. Comput. Telecommun. Netw.*, vol. 50, no. 13, pp. 2127–2159, Sept. 2006.

[7] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, 2007.

[8] I. F. Akyildiz, W. Y. Lee, and K. Chowdhury, "Crahns: Cognitive radio ad hoc networks," *Ad Hoc Netw. (Elsevier)*, vol. 7, no. 5, pp. 810–836, July 2009.

[9] Y. Liu, L. X. Cai, H. Luo, and X. Shen, "Deploying cognitive cellular networks under dynamic resource management," *IEEE Wireless Commun. Mag.*, to appear.

[10] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: a survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, 2008.

[11] M. M. Buddhikot, I. Kennedy, F. Mullany, and H. Viswanathan, "Ultra-broadband femtocells via opportunistic reuse of multi-operator and multi-service spectrum," *Bell Labs Tech. J.*, vol. 13, no. 4, pp. 129–144, 2009.

[12] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[13] T. Yücek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 116–130, Jan. 2009.

[14] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Select. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[15] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, "Spectrum leasing to cooperating secondary ad hoc networks," *IEEE J. Select. Areas Commun.*, vol. 26, no. 1, pp. 203–213, Jan. 2008.

[16] B. Cao, L. X. Cai, H. Liang, J. W. Mark, Q. Zhang, H. V. Poor, and W. Zhuang, "Cooperative cognitive radio networking using quadrature signaling," *Proc. IEEE INFOCOM*, Mar. 2012.

[17] H. Kim and K. G. Shin, "Efficient discovery of spectrum opportunities with MAC-layer sensing in cognitive radio networks," *IEEE Trans. Mobile Comput.*, vol. 7, no. 5, pp. 533–545, May 2008.

[18] D. Willkomm, S. Machiraju, J. Bolot, and A. Wolisz, "Primary users in cellular networks: A large-scale measurement study," *Proc. IEEE DySPAN*, pp. 1–11, Oct. 2005.

[19] D. Chen, S. Yin, Q. Zhang, M. Liu, and S. Li, "Mining spectrum usage data: a large-scale spectrum measurement study," *Proc. ACM/IEEE MOBICOM*, Sept. 2009.

[20] T. S. Rappaport, *Wireless communications principles and practices.* Prentice-Hall, 2002.

[21] A. F. Molisch, L. J. Greenstein, and M. Shafi, "Propagation issues for cognitive radio," *Proc. IEEE*, vol. 97, no. 5, pp. 787–802, May 2009.

[22] V. Bhandari and N. H. Vaidya, “Heterogeneous multi-channel wireless networks: Scheduling and routing issues,” *Technical Report, UIUC*, Oct. 2007.

[23] H. Wu, F. Yang, K. Tan, J. Chen, Q. Zhang, and Z. Zhang, “Distributed channel assignment and routing in multiradio multichannel multihop wireless networks,” *IEEE J. Select. Areas Commun.*, vol. 24, no. 11, pp. 1972–1983, Nov. 2006.

[24] W. Yoon and N. Vaidya, “Routing exploiting multiple heterogeneous wireless interfaces: A tcp performance study,” *Computer Communications (Elsevier)*, vol. 33, no. 1, pp. 23–34, Jan. 2010.

[25] K. R. Chowdhury and M. D. Felice, “SEARCH: A routing protocol for mobile cognitive radio ad-hoc networks,” *Computer Communications (Elsevier)*, vol. 32, no. 18, pp. 1983–1997, 2009.

[26] F. F. Digham, M. S. Alouini, and M. K. Simon, “On the energy detection of unknown signals over fading channels,” *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 3575–3579, Jan. 2007.

[27] R. Fan and H. Jiang, “Optimal multi-channel cooperative sensing in cognitive radio networks,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 1128–1138, Mar. 2010.

[28] A. Alshamrani, X. Shen, and L. Xie, “A cooperative MAC with efficient spectrum sensing algorithm for distributed opportunistic spectrum networks,” *J. Communications*, vol. 4, no. 10, p. 728C740, 2009.

[29] A. Alshamrani, L. Xie, and X. Shen, “Adaptive admission control and channel allocation policy in cooperative ad hoc opportunistic spectrum networks,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1618–1629, Apr. 2010.

[30] Z. Han, R. Fan, and H. Jiang, “Replacement of spectrum sensing in cognitive radio,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2819–2826, June 2009.

[31] Q. Du and X. Zhang, “Queue-aware spectrum sensing for interference-constrained transmissions in cognitive radio networks,” *Proc. IEEE ICC*, 2010.

[32] Z. Quan, S. Cui, A. H. Sayed, and H. Poor, “Wideband spectrum sensing in cognitive radio networks,” *Proc. IEEE ICC*, pp. 901–906, May 2008.

[33] Z. Quan, S. Cui, H. V. Poor, and A. H. Sayed, “Collaborative wideband sensing for cognitive radios,” *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 60–73, Nov. 2008.

[34] Z. Quan, S. Cui, A. H. Sayed, and H. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1128–1140, Mar. 2009.

[35] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive mac for opportunistic spectrum access in ad hoc networks: A pomdp framework," *IEEE J. Select. Areas Commun.*, vol. 25, no. 3, pp. 589–600, Apr. 2007.

[36] H. P. Shiang and M. van der Schaar, "Ditributed resource management in multihop cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 941–953, Feb. 2009.

[37] T. V. Krishna and A. Das, "A survey on mac protocols in OSA networks," *Computer Netw. (Elsevier)*, vol. 53, no. 9, pp. 1377–1394, June 2009.

[38] P. Wang, D. Niyato, and H. Jiang, "Voice service capacity analysis for cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1779–1790, May 2010.

[39] H. Kushwaha, Y. Xing, R. Chandramouli, and H. Heffes, "Reliable multimedia transmission over cognitive radio networks using fountain codes," *Proc. IEEE*, vol. 96, no. 1, pp. 155 –165, 2008.

[40] Q. Chen, C. T. Chou, S. S. Kanhere, W. Zhang, and S. k. Jha, "Performance of multi-hop whisper cognitive radio networks," *Proc. IEEE DySPAN*, 2010.

[41] L. X. Cai, X. Shen, J. W. Mark, L. Cai, and Y. Xiao, "Voice capacity analysis of WLAN with unbalanced traffic," *IEEE Trans. Veh. Technol.*, vol. 55, no. 5, pp. 752–761, May 2006.

[42] International Telecommunicaion Union, "General characteristics of international telephone connections and international telephone circuits one-way trans mission time," Feb, 1996.

[43] Video Traces Research Group, "Video traces for network performance evaluation," http://trace.kom.aau.dk/TRACE/pics/FrameTrace/h263/indexc366.html, [Online; accessed 11-Feb-2013].

[44] H. T. Cheng and W. Zhuang, "Simple channel sensing order in cognitive radio networks," *IEEE J. Select. Areas Commun.*, vol. 29, no. 4, Apr. 2011.

[45] A. Alshamrani, X. Shen, and L. Xie, "QoS provisioning for heterogeneous services in cooperative cognitive radio networks," *IEEE J. Select. Areas Commun.*, vol. 29, no. 4, pp. 819–830, Apr. 2011.

[46] W. Ren and Q. Zhao, "Connectivity of cognitive radio networks: Proximity vs. opportunity," *Proc. ACM MobiCom Workshop on Cognitive Radio Networks*, 2009.

[47] C. Yin, L. Gao, and S. Cui, "Scaling laws for overlaid wireless networks: A cognitive radio network vs. a primary network," *Proc. IEEE GLOBECOM*, Dec. 2008.

[48] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Info. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[49] Q. Wang and H. Zheng, "Route and spectrum selection in dynamic spectrum networks," *Proc. IEEE CCNC/ARCON*, Jan. 2006.

[50] C. Xin, B. Xie, and C.-C. Shen, "A novel layered graph model for topology formation and routing in dynamic spectrum access networks," *Proc. IEEE DySPAN*, 2005.

[51] R. Pal, "Efficient routing algorithms for multi-channel dynamic spectrum access networks," *Proc. IEEE DySPAN*, pp. 288–291, Apr. 2007.

[52] A. Sampath, L. Yang, L. Cao, H. Zheng, and B. Zhao, "High throughput spectrum-aware routing for cognitive radio based ad hoc networks," *Proc. IEEE CrownCom'08*, May 2008.

[53] T. Chen, H. Zhang, G. M. Maggio, and I. Chlamtac, "Cogmesh: a cluster-based cognitive radio network," *Proc. IEEE DySPAN*, p. 168C178, Apr. 2007.

[54] G. Cheng, W. Liu, Y. Li, and W. Cheng, "Spectrum aware on-demand routing in cognitive radio networks," *Proc. IEEE DySPAN*, pp. 571–574, Apr. 2007.

[55] J. Cheny, H. Li, J. Wu, and R. Zhang, "Starp: A novel routing protocol for multi-hop dynamic spectrum access networks," *Proc. ACM Mobicom-MICNET'09*, Sept. 2009.

[56] K.-C. Chen, B. K. Cetin, Y.-C. Peng, N. Prasad, J. Wang, and S. Lee, "Routing for cognitive radio networks consisting of opportunistic links," *Wireless Communications and Mobile Computing (Wiley)*, vol. 10, no. 4, pp. 451–466, Apr. 2010.

[57] I. Pefkianakis, S. H. Y. Wong, and S. Lu, "SAMER: Spectrum aware mesh routing in cognitive radio networks," *Proc. IEEE DySPAN*, pp. 1–5, Oct. 2008.

[58] H. Khalif, N. Malouch, and S. Fdida, “Multihop cognitive radio networks: To route or not to route,” *IEEE Network*, vol. 23, no. 4, pp. 20–25, 2009.

[59] M. Mauve, J. Sidmer, and H. Hartenstein, “A survey on position based routing in mobile ad-hoc networks,” *IEEE Netw. Mag.*, vol. 15, no. 6, pp. 30–39, Nov. 2001.

[60] M. Cesana, F. Cuomo, and E. Ekici, “Routing in cognitive radio networks: Challenges and solutions,” *Ad Hoc Netw. (Elsveir)*, vol. 9, no. 3, p. 228C248, May 2011.

[61] H. Khalife, S. Ahuja, N. Malouch, and M. Krunz, “Probabilistic path selection in opportunistic cognitive radio networks,” *Proc. IEEE GLOBECOM*, pp. 4861–4865, 2008.

[62] D. D. Couto, D. Aguayo, J. Bicket, and R. Morris, “A high throughput path metric for multi-hop wireless routing,” *Proc. ACM/IEEE MOBICOM*, 2003.

[63] R. Draves, J. Padhye, and B. Zill, “Routing in multi-radio, multi-hop wireless mesh networks,” *Proc. ACM/IEEE MOBICOM*, pp. 114–128, 2004.

[64] G. Lei, W. Wang, T. Peng, and W. Wang, “Routing metrics in cognitive radio networks,” *Proc. IEEE ICCSC’08*, pp. 265–269, May 2008.

[65] S. Biswas and R. Morris, “Opportunistic routing in multi-hop wireless networks,” *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 1, pp. 69–74, 2004.

[66] ——, “ExOR: Opportunistic multi-hop routing for wireless networks,” *ACMSIGCOMM Computer Communication Review*, vol. 34, no. 1, pp. 133–144, 2005.

[67] M. Zorzi and R. R. Rao, “Geographic random forwarding (GeRaF) for ad hoc and sensor networks: Energy and latency performance,” *IEEE Trans. Mobile Comput.*, vol. 2, no. 4, pp. 349–365, Oct.-Dec. 2003.

[68] ——, “Geographic random forwarding (GeRaF) for ad hoc and sensor networks: Multihop performance,” *IEEE Trans. Mobile Comput.*, vol. 2, no. 4, pp. 337–348, Oct.-Dec. 2003.

[69] K. Zeng, W. Lou, and H. Zhai, “On end-to-end throughput of opportunistic routing in multirate and multihop wireless networks,” *Proc. IEEE INFOCOM*, Apr. 2009.

[70] K. Zeng, Z. Yang, and W. Lou, “Location-aided opportunistic forwarding in multirate and multihop wireless networks,” *IEEE Trans. Veh. Technol.*, vol. 58, no. 6, pp. 3032–3040, July 2009.

[71] ——, “Opportunistic routing in multi-radio multi-channel multi-hop wireless networks,” *IEEE Infocom Mini-Conference’10*, Mar. 2010.

[72] X. Y. Li, Y. Wu, P. Xu, G. H. Chen, and M. Li, “Hidden information and actions in multi-hop wireless ad hoc networks,” *Proc. ACM MobiHoc*, pp. 283–292, May 2008.

[73] Z. Ji and K. J. Liu, “Dynamic spectrum sharing: A game theoretical overview,” *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 88–94, June 2007.

[74] P. Pawelczak, R. V. L. Xia, and I. G. M. M. Niemegeers, “Cognitive radio emergency networks - requirements and design,” *Proc. IEEE DySPAN*, 2005.

[75] Michigan Tech. Univ., “Cognitive Radio Cognitive Network Simulator,” http://stuweb.ee.mtu.edu/~ljialian/index.htm, [Online; accessed 11-Feb-2013].

[76] B. Karp and H. T. Kung, “GPSR: Greedy perimeter stateless routing for wireless networks,” *Proc. ACM/IEEE MOBICOM*, 2000.

[77] M. Wellens, J. Riihijarvi, and P. Mahonen, “Evaluation of adaptive MAC-layer sensing in realistic spectrum occupancy scenearios,” *Proc. IEEE DySPAN*, 2010.

[78] M. Pan, C. Zhang, P. Li, and Y. Fang, “Joint routing and link scheduling for cognitive radio networks under uncertain spectrum supply,” *Proc. IEEE INFOCOM*, 2011.

[79] ——, “Spectrum harvesting and sharing in multi-hop cognitive radio networks under uncertain spectrum supply,” *IEEE J. Select. Areas Commun.*, vol. 30, no. 2, pp. 369–378, Feb. 2012.

[80] R. Murty, R. Chandra, T. Moscibroda, and P. Bahl, “SenseLess: A database-driven white spaces network,” *IEEE Trans. Mobile Comput.*, vol. 11, no. 2, pp. 189–203, 2012.

[81] L. X. Cai, Y. Liu, X. Shen, J. W. Mark, and D. Zhao, “Distributed qos-aware mac for multimedia over cognitive radio networks,” *Proc. IEEE GLOBECOM*, Dec. 2010.

[82] Y. Liu, L. X. Cai, X. Shen, and J. W. Mark, “Exploiting heterogeneity wireless channels for opportunistic routing in dynamic spectrum access networks,” *Proc. IEEE ICC*, June 2011.

[83] M. Cesana, F. Cuomo, and E. Ekici, “Routing in cognitive radio networks: Challenges and solutions,” *Ad Hoc Networks (Elsevier)*, vol. 9, no. 3, pp. 228–248, May 2011.

[84] C. Gao, Y. Shi, Y. T. Hou, H. D. Sherali, and H. Zhou, “Multicast communications in multi-hop cognitive radio networks,” *IEEE J. Select. Areas Commun.*, vol. 29, no. 4, pp. 784–793, 2011.

[85] Y. Liu, L. X. Cai, and X. Shen, “Joint channel selection and opportunistic forwarding in multi-hop cognitive radio networks,” *Proc. IEEE GLOBECOM*, Dec. 2011.

[86] A. Abdrabou and W. Zhuang, “A position-based QoS routing scheme for UWB ad hoc networks,” *IEEE J. Select. Areas Commun.*, vol. 24, no. 4, Apr. 2006.

[87] ——, “Statistical QoS routing for IEEE 802.11 multihop ad hoc networks,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1542–1552, March 2009.

[88] M. Timmers, S. Pollin, A. Dejonghe, L. V. der Perre, and F. Catthoor, “Distributed multichannel MAC protocol for multihop cognitive radio networks,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 446–459, Jan. 2010.

[89] Y. Bi, L. X. Cai, X. Shen, and H. Zhao, “Efficient and reliable broadcast in intervehicle communications networks: A cross layer approach,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2404–2417, 2010.

[90] D. R. Cox, *Renewal Theory.* Butler and Tanner, 1967.

[91] J. P. Castro, *The UMTS Network and Radio Access Technology.* John Wiley & Sons, 2001.

[92] K. Zeng, W. Lou, J. Yang, and D. R. Brown, “On geographic collaborative forwarding in wireless ad hoc and sensor networks,” *Proc. WASA‘07*, Aug. 2007.

[93] A. Sampath, L. Yang, L. Cao, H. Zheng, and B. Y. Zhao, “High throughput spectrum-aware routing for cognitive radio based ad hoc networks,” *Proc. CROWNCOM’08*, 2008.

[94] Ericsson, “Traffic and market report,” June 2012.

[95] Y. Liu, L. X. Cai, X. Shen, and J. W. Mark, “Spectrum-aware QoS MAC for multimedia services in cognitive radio networks,” *IEEE COMSOC MMTC E-Letter*, vol. 7, no. 6, pp. 13–16, July 2012.

[96] C. Cormio and K. R. Chowdhury, “A survey on MAC protocols for cognitive radio networks,” *Ad Hoc Networks*, vol. 7, no. 7, pp. 1315–1329, 2009.

[97] J. Jia, Q. Zhang, and X. Shen, “HC-MAC: A hardware-constrained cognitive MAC for efficient spectrum management,” *IEEE J. Select. Areas Commun.*, vol. 26, no. 1, pp. 106–117, 2008.

[98] J. So and N. Vaidya, “Multi-channel mac for ad hoc networks: Handling multi-channel hidden terminals using a single transceiver,” *Proc. ACM MobiHoc*, pp. 222–233, May 2004.

[99] J. Wang, Y. Fang, and D. Wu, “A power-saving multi-radio multi-channel mac protocol for wireless local area networks,” *Proc. IEEE INFOCOM*, Apr. 2006.

[100] C. Cordeiro and K. Challapali, “C-MAC: A cognitive mac protocol for multi-channel wireless networks,” *Proc. IEEE DySPAN*, pp. 147–157, Apr. 2007.

[101] J. Mo, H. So, and J. Walrand, “Comparison of multichannel mac protocols,” *IEEE Trans. Mobile Comput.*, vol. 7, no. 1, pp. 50–65, Jan. 2008.

[102] P. Bahl, R. Chandra, and J. Dunagan, “SSCH: Slotted seeded channel hopping for capacity improvement in IEEE 802.11 ad-hoc wireless networks,” *Proc. ACM/IEEE MOBICOM*, 2004.

[103] R. Zheng, J. Hou, and L. Sha, “Optimal block design for asynchronous wakeup and its applications in multi-hop wireless networks,” *IEEE Trans. Mobile Comput.*, vol. 5, no. 9, pp. 1228–1241, Sept. 2006.

[104] K. Bian, J. M. Park, and R. Chen, “A quorum-based framework for establishing control channels in dynamic spectrum access networks,” *Proc. ACM/IEEE MOBICOM*, pp. 20–25, Sept. 2009.

[105] L. Yang, L. Cao, and H. Zheng, “Proactive channel access in dynamic spectrum networks,” *Elsevier Physical Communication*, vol. 1, no. 2, pp. 103–111, Jun. 2008.

[106] R. W. Thomas, D. H. Friend, L. A. DaSilva, and A. B. Mackenzie, “Cognitive networks: Adaptation and learning to achieve end-to-end performance objectives,” *IEEE Commun. Mag.*, vol. 44, no. 12, pp. 51–57, Dec. 2006.

[107] J. Mwangoka, K. Letaief, and Z. Cao, “Robust end-to-end qos maintenance in non-contiguous OFDM based cognitive radios,” *Proc. IEEE ICC*, 2008.

[108] P. Pawelczak, S. Pollin, H. So, A. Bahai, R. Prasad, and R. Hekmat, “Performance analysis of multichannel medium access control algorithms for opportunistic spectrum access,” *IEEE Trans. Veh. Technol.*, vol. 58, no. 6, pp. 3014–3031, Jul. 2008.

[109] R. Chen, J.-M. Park, Y. T. Hou, and J. H. Reed, "Toward secure distributed spectrum sensing in cognitive radio networks," *IEEE Commun. Mag.*, Apr. 2008.

[110] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: past, present, and future," *IEEE J. Select. Areas Commun.*, vol. 30, no. 497-508, p. 3, April 2012.

[111] X. Liang, X. Li, M. Barua, L. Chen, R. Lu, X. Shen, , and H. Luo, "Enable pervasive healthcare through continuous remote health monitoring," *IEEE Wireless Commun. Mag.*, vol. 19, no. 6, pp. 10–18, 2012.

[112] W. Song and W. Zhuang, "Multi-service load sharing for resource management in the cellular/WLAN integrated network," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 725–735, Feb. 2009.

[113] Y. Liu, L. X. Cai, and X. Shen, "Spectrum-aware opportunistic routing in multi-hop cognitive radio networks," *IEEE J. Select. Areas Commun.*, vol. 30, no. 10, pp. 1958–1967, Nov. 2012.

[114] M. Awad, V. Mahinthan, M. Mehrjoo, X. Shen, and J. W. Mark, "A dual decomposition-based resource allocation for OFDMA networks with imperfect CSI," *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2394–2403, May 2010.

[115] J. Yun and K. Shin, "Adaptive interference management of OFDMA femtocells for co-channel deployment," *IEEE J. Select. Areas Commun.*, vol. 29, no. 6, pp. 1225–1241, June 2011.

[116] E. Dahlman, S. Parkvall, and J. Skold, *4G LTE LTE-Advanced for Mobile Broadband.* Oxford, UK: Academic Press, 2011.

[117] J. Buoccuzzi and M. Ruggiero, *Femtocells design & application.* New York: McGraw Hill, 2011.

[118] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of K-tier downlink heterogeneous cellular networks," *IEEE J. Select. Areas Commun.*, vol. 30, no. 3, pp. 550–560, Apr. 2012.

[119] L. X. Cai, Y. Liu, H. Luan, X. Shen, J. W. Mark, and H. V. Poor, "Adaptive resource management in sustainable energy powered wireless mesh networks," *Proc. IEEE GLOBECOM*, Dec. 2011.

[120] ——, “Sustainability analysis and resource management of wireless mesh networks,” *IEEE J. Select. Areas Commun.*, to appear.

[121] Z. Zheng, L. X. Cai, R. Zhang, and X. Shen, “RNP-SA: Joint relay placement and sub-carrier allocation in wireless communication networks with sustainable energy,” *IEEE Trans. Wireless Commun.*, vol. 1, no. 10, p. 3818C3828, Oct. 2012.

[122] L. X. Cai, H. V. Poor, Y. Liu, H. Luan, X. Shen, and J. W. Mark, “Dimensioning network deployment and resource management in green mesh networks,” *IEEE Wireless Commun. Mag.*, vol. 18, no. 5, pp. 58–65, Oct. 2011.

[123] L. Saker, S. E. Elayoubi, R. Combes, and T. Chahed, “Optimal control of wake up mechanisms of femtocells in heterogeneous networks,” *IEEE J. Select. Areas Commun.*, vol. 30, no. 3, pp. 664–672, Apr. 2012.

[124] R. D. Taranto, P. Popovski, O. Simeone, and H. Yomo, “Efficient spectrum leasing via randomized silencing of secondary users,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3739–3749, Dec. 2010.

[125] D. Li, Y. Xu, X. Wang, and M. Guizani, “Coalitional game theoretic approach for secondary spectrum access in cooperative cognitive radio networks,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 844–856, Mar. 2011.

[126] X. Kang, R. Zhang, and M. Motani, “Price-based resource allocation for spectrum-sharing femtocell networks: a Stackelberg game approach,” *IEEE J. Select. Areas Commun.*, vol. 30, no. 3, pp. 538–549, Apr. 2012.

[127] Z. Niu, Y. Wu, J. Gong, , and Z. Yang, “Cell zooming for cost-efficient green cellular networks,” *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 74–79, Nov. 2010.

[128] J. Zhang and Q. Zhang, “Stackelberg game for utility-based cooperative cognitive radio networks,” *Proc. ACM MobiHoc*, 2009.