

การดำเนินการตรวจจับและจัดสรรคลื่นความถี่วิทยุบนระบบวิทยุรู้คิด

THE IMPLEMENTATION OF SPECTRUM SENSING AND SPECTRUM
ALLOCATION ON COGNITIVE RADIO

พิชชา รุ่งสว่าง
PITCHA RUNGSAWANG

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ.2560

KMITL-2017-EN-M-070-080

การดำเนินการตรวจจับและจัดสรรคลื่นความถี่วิทยุบนระบบวิทยุรู้คิด

THE IMPLEMENTATION OF SPECTRUM SENSING AND SPECTRUM ALLOCATION ON COGNITIVE RADIO



พิชชา รุ่งสว่าง
PITCHA RUNGSAWANG

เลขที่ 148263
เลขทะเบียน 148263
พ.ศ. 2560

148263986

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ.2560

KMITL-2017-EN-M-070-080

THE IMPLEMENTATION OF SPECTRUM SENSING AND SPECTRUM
ALLOCATION ON COGNITIVE RADIO

PITCHA RUNGSAWANG

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF ENGINEERING IN COMPUTER ENGINEERING
FACULTY OF ENGINEERING
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2017
KMITL-2017-EN-M-070-080

COPYRIGHT 2017

FACULTY OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ การดำเนินการตรวจจับและจัดสรรคลื่นความถี่วิทยุบนระบบรับรู้คิด
Thesis Title The Implementation of Spectrum Sensing and Spectrum Allocation on Cognitive Radio
นักศึกษา นางสาวพิชชา รุ่งสว่าง
รหัสประจำตัว 58601097
ปริญญา วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชา วิศวกรรมคอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์ ดร.อำนาจ ขาวเน
หมายเลขวิทยานิพนธ์ KMITL-2017-EN-M-070-080

คณะกรรมการสอบวิทยานิพนธ์		ลายมือชื่อ
ผศ.ดร.ภุชงค์	อุทโยภาส	
ผศ.ดร.สุรินทร์	กิตติธรรกุล	
ดร.ธัญชัย	ตรีภาค	
ผศ.ดร.ศักดิ์ชัย	ทิพย์จักขุรัตน์	
ดร.อำนาจ	ขาวเน	

วัน / เดือน / ปี ที่สอบ วันอังคารที่ 18 กรกฎาคม พ.ศ. 2560 เวลา 17.00-19.00 น.
สถานที่สอบ ณ อาคาร A ชั้น 5 ห้องประชุม 2

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

คณะวิศวกรรมศาสตร์ รับรองแล้ว



(รองศาสตราจารย์ ดร. คมสัน มาลีสี)

คณบดี คณะวิศวกรรมศาสตร์

วันที่ 18 กรกฎาคม พ.ศ. 2560

หัวข้อวิทยานิพนธ์

การดำเนินการตรวจจับและจัดสรรคลื่นความถี่วิทยุบน
ระบบวิทยุรู้คิด

นักศึกษา

นางสาวพิชชา รุ่งสว่าง

รหัสประจำตัว

58601097

ปริญญา

วิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชา

วิศวกรรมคอมพิวเตอร์

พ.ศ.

2560

อาจารย์ที่ปรึกษาวิทยานิพนธ์

ดร.อำนาจ ขาวเน

บทคัดย่อ

ระบบวิทยุรู้คิดเป็นระบบที่ได้รับการนำเสนอเพื่อใช้ความถี่บนคลื่นวิทยุอย่างมีประสิทธิภาพ เนื่องจากในปัจจุบันมีความต้องการในการใช้คลื่นความถี่วิทยุที่มากขึ้น โดยมีหลักการคือผู้ใช้ที่ไม่มีใบอนุญาต (ผู้ใช้รอง) จากผู้ที่มีสิทธิอนุญาตจากประเทศนั้นๆ เข้าใช้ช่องสัญญาณบนแถบความถี่ของผู้มีใบอนุญาต (ผู้ใช้หลัก) ให้ใช้ความถี่นั้นๆ ซึ่งไม่ส่งผลกระทบต่อผู้มีใบอนุญาตแต่อย่างใด ระบบวิทยุรู้คิดต้องการความสามารถด้านการตรวจจับสเปกตรัมเพื่อตรวจสอบว่ามีผู้ใช้หลักใช้ความถี่นั้นๆ อยู่หรือไม่ และเพื่อเพิ่มประสิทธิภาพในการจัดสรรช่องสัญญาณ งานวิจัยชิ้นนี้จึงได้นำความรู้เรื่องการเรียนรู้แบบเสริมแรงมาช่วยในการเรียนรู้สิ่งแวดล้อมที่เปลี่ยนไปตลอดเวลาอีกด้วย

Thesis	The Implementation of Spectrum Sensing and Spectrum Allocation on Cognitive Radio
Student	Miss. Pitcha Rungsawang
Student ID.	58601097
Degree	Master of Engineering
Program	Computer Engineering
Year	2017
Thesis Advisor	Dr. Amnach Khawne

ABSTRACT

Cognitive Radio is a system designed to use radio spectrum in the most efficient way possible—as demands for radio spectrum have been heavily increased nowadays. The system works by allowing unlicensed users to use radio channel on the licensed users' spectrum without interfering with the licensed users. The system needs spectrum sensing to verify the usage of each specific spectrum. In ever-shifting environment, Reinforcement Learning is also used for establishing spectrum allocation as well.

กิตติกรรมประกาศ

วิทยานิพนธ์เล่มนี้สำเร็จลุล่วงลงได้ เนื่องจากคำแนะนำคำปรึกษา ความช่วยเหลือ คำสั่งสอนที่ดีจาก ดร.อำนาจ ขาวเน ผู้เป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์ ขอขอบพระคุณท่านอาจารย์เป็นอย่างสูงที่ได้อบรม สั่งสอน ติดตาม ให้ความเมตตาแก่ลูกศิษย์คนนี้เป็นอย่างดี

ขอขอบคุณ อาจารย์สมเกียรติ แสดงฤทธิ์ หัวหน้าบริษัทไซบิ้นส์จำกัด ที่ข้าพเจ้าได้ทำงานอยู่อีกทั้งอนุญาตให้ข้าพเจ้าได้มาเรียน เคียวยุติให้ทำตามกรอบของเวลาอันมีค่าให้สุดความสามารถ ให้ความรู้ให้กำลังใจ ให้คำแนะนำที่ดี

ขอขอบคุณ อาจารย์ภาควิชาโทรคมนาคม และนงนุช ที่ให้อนุญาตให้ยืมเครื่อง USRP และให้คำแนะนำ วิธีใช้ ให้ความเอาใจใส่ในการเรียนรู้ของข้าพเจ้าเป็นอย่างดี

ขอขอบคุณบทความของผู้เขียนวารสารวิชาการ การประชุมวิชาการ ที่ข้าพเจ้าได้อ่านถึงแม้ว่าไม่ได้อ้างอิงมาในเอกสารอ้างอิง แต่ก็ทำให้ข้าพเจ้าเปิดมุมมอง รวบรวมเป็นความรู้ ดัดแปลงต่อยอดได้เป็นอย่างดี

ขอขอบพระคุณท่านอาจารย์ทุกท่านที่ให้ความรู้สั่งสอนในช่วงการเรียน course work ในหลักสูตรของแผน ก 2 ของภาคคอมพิวเตอร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังที่เปิดมุมมอง เปิดความคิด เพิ่มเติมในสิ่งที่ข้าพเจ้าไม่เคยคิดไม่เคยรู้มาก่อน

และขอขอบคุณเพื่อนร่วม course work ที่อยู่ร่วมกันมา 8 วิชา ที่คอยเตือน คอยส่งข่าวทุกเรื่องทั้งวันเวลาเรียน วันเวลาไม่เรียน ไปเซนต์ชื่อ ส่งงาน สัมมนา ประชุมทางวิชาการ

สุดท้ายนี้ ข้าพเจ้าขอกราบขอบพระคุณบิดา มารดา ที่คอยให้กำลังใจ ในเวลาที่ข้าพเจ้าเหนื่อยทั้งจากการทำงานและเรียน ขอบพระคุณที่อนุญาตให้ทำตามที่ข้าพเจ้าต้องการ และคอยห้ามในสิ่งผิด สิ่งไม่ดีที่จะทำให้ไม่เกิดวิทยานิพนธ์ฉบับนี้ขึ้น

ข้าพเจ้าขอขอบคุณงามความดีที่ได้จากการวิทยานิพนธ์ฉบับนี้ให้กับบิดามารดาที่เป็นที่รักยิ่งของข้าพเจ้า ครูอาจารย์ทุกท่านที่ให้คำสั่งสอนประสิทธิ์ประสาทวิชาความรู้และประสบการณ์อันมีค่าที่หาที่ไหนไม่ได้ให้แก่ข้าพเจ้า

พิชชา รุ่งสว่าง

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VI
สารบัญรูป.....	VII
บทที่ 1	VII
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา.....	2
1.3 สมมติฐานของการศึกษา	2
1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในงานวิจัย	2
1.5 ขอบเขตงานวิจัย.....	3
1.6 ขั้นตอนการศึกษา.....	3
1.7 เครื่องมือและอุปกรณ์ที่ใช้ในงานวิจัย.....	4
1.8 โครงสร้างของวิทยานิพนธ์.....	4
บทที่ 2	5
2.1 หลักการวิทยุที่ดำเนินการด้วยซอฟต์แวร์ (Software Defined Radio – SDR).....	5
2.2 วิทยุรู้คิด (Cognitive Radio).....	6
2.2.1 การตรวจจับสเปกตรัม (Spectrum sensing)	9
2.2.2 การใช้สเปกตรัมร่วมกัน (Spectrum sharing).....	11
2.3 การตรวจจับแบบตรวจจับพลังงานของสัญญาณ (Energy detection).....	12
2.3.1 หลักการการตรวจจับ	12
2.3.2 รูปแบบความน่าจะเป็นที่ใช้สำหรับวิเคราะห์การตรวจจับสเปกตรัม	13
2.4 การเรียนรู้แบบเสริมแรง (Reinforcement Learning).....	13
2.4.1 การเรียนรู้แบบคิว (Q-Learning).....	15
2.5 การประเมินประสิทธิภาพการตรวจจับสเปกตรัม.....	18
2.5.1. กราฟแสดงความสัมพันธ์ระหว่างผลบวกจริงและผลบวกปลอม (Receiver Operating Characteristic curve – ROC curve).....	18

สารบัญ (ต่อ)

	หน้า
บทที่ 3	20
3.1 การตรวจจับสเปกตรัม (Spectrum sensing).....	20
3.1.1 การตรวจจับสเปกตรัมโดยใช้ความรู้เรื่องการเรียนรู้แบบเสริมแรง	20
3.2 การใช้คลื่นความถี่ร่วมกัน (Spectrum sharing)	22
3.3 การดำเนินการจำลองระบบวิทยุรู้คิด (Implementation).....	24
บทที่ 4	27
4.1 กระบวนการตรวจจับสเปกตรัมที่นำเสนอ	27
4.1.1 กระบวนการตรวจจับแบบตรวจจับพลังงาน	27
4.1.2 กระบวนการเรียนรู้แบบคิว	27
4.1.3 กระบวนการปรับปรุงพารามิเตอร์ของวิทยุในวิทยุรู้คิด	28
4.2 การติดตั้งในส่วนของฮาร์ดแวร์	31
บทที่ 5	33
5.1 ผลการทดลองเมื่อจำลองสภาพแวดล้อมเป็นเขตเมืองหนาแน่น	34
5.2 ผลการทดลองเมื่อจำลองสภาพแวดล้อมเป็นเขตชานเมือง.....	36
5.3 ผลการทดลองเมื่อจำลองสภาพแวดล้อมเป็นเขตชนบท	38
5.4 เปรียบเทียบผลการทดลองจากกราฟ ROC curve	40
บทที่ 6	41
เอกสารอ้างอิง	42
ภาคผนวก.....	44
ประวัติผู้เขียน.....	51

สารบัญตาราง

ตารางที่	หน้า
2.1 เมตริกซ์ความสับสน.....	18
4.1 พารามิเตอร์ในการใช้งานการเรียนรู้ของคิว.....	28
4.2 เลขที่อยู่ไอพีของ USRP และ คอมพิวเตอร์.....	32
4.3 พารามิเตอร์ในการส่งของ USRP.....	32
4.2 พารามิเตอร์ในการส่งของ USRP.....	28
5.1 เลขซีกำลังการสูญเสียในเส้นทางสำหรับสภาพแวดล้อมต่างๆ.....	33
5.2 สรุปพารามิเตอร์ทั้งหมดที่ใช้ในการทดลอง.....	34

สารบัญรูป

รูปที่	หน้า
2.1	แผนภาพการทำงานของวิทยุดำเนินการด้วยซอฟต์แวร์.....5
2.2	USRP ยี่ห้อ Ettus รุ่น N210.....6
2.3	จำลองภาพจากอุปกรณ์วิเคราะห์สเปกตรัม.....7
2.4	หลักการทำงานของระบบวิทยุรู้คิด.....8
2.5	องค์ประกอบของระบบวิทยุรู้คิด.....9
2.6	ขั้นตอนของ energy detection.....10
2.7	ขั้นตอนของ cyclostationary feature detection.....10
2.8	ขั้นตอนของ matched filtering detection.....10
2.9	การใช้สเปกตรัมร่วมกันแบบผสมผสานกัน.....11
2.10	การใช้สเปกตรัมร่วมกันแบบเบื้องล่าง.....11
2.11	การใช้สเปกตรัมร่วมกันแบบซ้อนทับ.....12
2.12	ขั้นตอนการตรวจจับพลังงานของสัญญาณ.....12
2.13	รูปแบบของการเรียนรู้แบบเสริมแรง.....14
2.14	Flow chart ของการเรียนรู้แบบคิว.....17
2.15	ตัวอย่างการเขียนกราฟ ROC.....19
3.1	ขั้นตอนการแบ่งสเปกตรัมใช้ร่วมกัน.....23
4.1	กระบวนการตรวจจับสเปกตรัมแบบตรวจจับพลังงาน.....27
4.2	ตัวอย่างสเปกตรัม.....29
4.3	ตัวอย่างค่าเกณฑ์ระดับ.....30
4.4	ตัวอย่างการยกสัญญาณและ noise floor.....30
4.5	แบบจำลองโมเดลในการติดตั้งฮาร์ดแวร์ของระบบวิทยุรู้คิด.....31
5.1	ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตเมืองหนาแน่น, $SNR = 13.7, P_{fa} = 10^{-2}$35
5.2	ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตเมืองหนาแน่น, $SNR = 13.7, P_{fa} = 10^{-4}$35
5.3	ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตเมืองหนาแน่น, $SNR = 13.7, P_{fa} = 10^{-6}$36
5.4	ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตชานเมือง, $SNR = 24.5, P_{fa} = 10^{-2}$37
5.5	ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตชานเมือง, $SNR = 24.5, P_{fa} = 10^{-4}$37
5.6	ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตชานเมือง, $SNR = 24.5, P_{fa} = 10^{-6}$38
5.7	ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตชนบท, $SNR = 30.7, P_{fa} = 10^{-2}$39
5.8	ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตชนบท, $SNR = 30.7, P_{fa} = 10^{-4}$39
5.9	ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตชนบท, $SNR = 30.7, P_{fa} = 10^{-6}$40

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

คลื่นความถี่วิทยุเป็นการติดต่อสื่อสารแบบไร้สายขั้นพื้นฐาน โดยปรกติคลื่นความถี่วิทยุจะได้รับการจัดสรรโดยต้องได้รับการอนุญาตจากองค์กรซึ่งบริหารจัดการคลื่นความถี่วิทยุของประเทศนั้นๆ ทำให้อาจเกิดการจัดสรรที่มากเกินไปจนเกินความจำเป็นต่อการใช้ประโยชน์ ทำให้เสียทรัพยากรคลื่นความถี่วิทยุ ในปัจจุบันมีความต้องการในการใช้คลื่นความถี่วิทยุจำนวนมากทั้งจำนวนของผู้ใช้และจำนวนของอุปกรณ์ที่ใช้ในการรับส่ง การที่บริหารจัดการทรัพยากรคลื่นความถี่วิทยุอย่างมีประสิทธิภาพไม่เพียงพอกับการเพิ่มปริมาณผู้ใช้ขึ้นอย่างต่อเนื่อง ทำให้ต้องมีการศึกษาการจัดสรรทรัพยากรคลื่นความถี่วิทยุขึ้นมาใหม่เพื่อแก้ปัญหาความคลาดแคลนและเพิ่มความอ่อนตัวในการใช้ทรัพยากรดังกล่าวนั้น บางคนกล่าวว่า ทรัพยากรคลื่นความถี่วิทยุนั้นเป็นทรัพยากรที่มีอยู่อย่างจำกัด แต่ระบบวิทยุรู้คิด (Cognitive Radio, CR) นี้จะทำให้เห็นว่า คลื่นความถี่วิทยุสามารถบริหารจัดการได้ดีกว่านี้ และใช้ประโยชน์ได้อย่างชาญฉลาด ระบบวิทยุรู้คิด [1] จะสามารถให้ผู้ใช้ที่ไม่ได้รับใบอนุญาตหรือผู้ใช้สำรอง (Secondary User, SU) ใช้ความถี่ที่ต้องได้รับการอนุญาตได้ โดยไม่ส่งผลกระทบต่อผู้ที่เป็นเจ้าของคลื่นความถี่หรือผู้ใช้ที่ได้รับอนุญาตหรือผู้ใช้หลัก (Primary User, PU) โดยต่อไปนี้จะเรียกผู้ใช้ที่ไม่ได้รับใบอนุญาตว่าผู้ใช้สำรองหรือ SU และจะเรียกผู้ใช้ที่ได้รับใบอนุญาตว่าผู้ใช้หลักหรือ PU ซึ่ง โดยปรกติ ผู้ใช้สำรองจะใช้คลื่นความถี่ของผู้ใช้หลักในขณะที่ผู้ใช้หลักไม่ได้ใช้งาน สิ่งที่จะทำให้ระบบวิทยุรู้คิดมีความฉลาดที่จะเรียนรู้สิ่งแวดล้อม (environment) ว่ามีผู้ใช้งานหลักใช้งานคลื่นความถี่นั้นๆ อยู่หรือไม่ กระบวนการนี้เรียกว่า การตรวจจับสเปกตรัม (spectrum sensing) [2] ซึ่งจะใช้มาตรวัดกำลังของสัญญาณในช่วงที่สนใจเพื่อหาช่องว่างของแถบความถี่ (spectrum hole) ความยากของการตรวจจับสเปกตรัมนี้คือ การระบุและตรวจจับสัญญาณของผู้ใช้หลักท่ามกลางสิ่งแวดล้อมที่มีสัญญาณรบกวนรุนแรง และยังมีความเร็วและความถูกต้องของการตรวจจับก็ต้องให้ความสำคัญด้วย

การเรียนรู้แบบเสริมแรง (Reinforcement Learning) เป็นการเรียนรู้กับสิ่งแวดล้อมที่เปลี่ยนแปลงตลอดเวลา การเรียนรู้แบบเสริมแรงจะใช้วิธีทางคณิตศาสตร์มาหาจุดที่ดีที่สุด แม้ว่าได้คำตอบที่ดีที่สุดแล้ว ก็ยังเรียนรู้ในสิ่งแวดล้อมภายนอก เพื่อเสาะแสวงหาคำตอบที่ดีที่สุดในระยะยาว

โดยการทดลองเพื่อเรียนรู้และพัฒนาความสามารถของระบบวิทยุรู้คิดบนการประยุกต์ใช้หลักการทางการเรียนรู้แบบเสริมแรง เพื่อให้ใช้แถบคลื่นความถี่ที่ว่างอยู่นั้น ผู้ทดลองได้ทำการทดลองจริงบนเครื่องมือ USRP (Universal Software Radio Peripheral) ซึ่งเป็นชุดเครื่องมือเพื่อศึกษา

ระบบการติดต่อสื่อสารบนคลื่นความถี่วิทยุ [3] และใช้ LabVIEW® เป็นโปรแกรมไว้เขียนโปรแกรมจำลองสถานการณ์สิ่งแวดล้อมในระดับกายภาพ และรวมไฟล์ .m ซึ่งใช้ใน MATLAB มาใช้ร่วมกันได้อีกด้วย

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

- 1.2.1 เพื่อให้สามารถศึกษาค้นคว้าและทำความเข้าใจงานวิจัยที่เกี่ยวกับการตรวจจับสเปกตรัมด้วยหลักการตรวจจับพลังงานในระบบวิทยุรู้คิดได้เป็นอย่างดี
- 1.2.2 เพื่อให้มีความรู้ความเข้าใจในทฤษฎีพื้นฐานของการตรวจจับสเปกตรัม (spectrum sensing) และการเรียนรู้ของเครื่องแบบการเรียนรู้แบบเสริมแรง (reinforcement learning)
- 1.2.3 เพื่อใช้งานเครื่องมือ USRP ในการทดลองการตรวจจับสเปกตรัมบนสภาพแวดล้อมจริงได้
- 1.2.4 เพื่อประยุกต์ใช้การเรียนรู้แบบเสริมแรงมาช่วยเพิ่มประสิทธิภาพของการตรวจจับสเปกตรัมได้
- 1.2.5 เพื่อทำการตรวจสอบประสิทธิภาพในการตรวจจับสเปกตรัมบนระบบวิทยุรู้คิดได้

1.3 สมมติฐานของการศึกษา

ในปัจจุบันมีการวิจัยมากมายซึ่งศึกษาเกี่ยวกับการนำความรู้ด้านการเรียนรู้แบบเสริมแรงมาช่วยในการตรวจจับสเปกตรัม แต่ส่วนมากก็เป็นเพียงการทดสอบ จำลองสถานการณ์บนโปรแกรมช่วยต่างๆ บนคอมพิวเตอร์ ในวิจัยนี้จะทดสอบบนเครื่องวิทยุที่ใช้ระบบ SDR (Software Defined Radio) มาช่วยในการทดลองบนสถานการณ์สิ่งแวดล้อมจริง ซึ่งมีความท้าทายในการรบกวนของสัญญาณรบกวนที่ไม่สามารถกำหนดได้ที่การทดลองบนโปรแกรมในเครื่องคอมพิวเตอร์ไม่สามารถสร้างได้ให้เสมือนการทดลองในสิ่งแวดล้อมจริงอย่างแน่นอน และปรับปรุงให้การตรวจจับสเปกตรัมด้วยการเรียนรู้แบบเสริมแรงมีประสิทธิภาพที่ดีในสถานการณ์สิ่งแวดล้อมจริงที่เปลี่ยนแปลงตลอดเวลา

1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในงานวิจัย

- 1.4.1 หลักการวิทยุที่ควบคุมด้วยซอฟต์แวร์ (Software Defined Radio – SDR)
- 1.4.2 หลักการวิทยุรู้คิด (Cognitive Radio – CR)
 - 1.4.2.1. การตรวจจับสเปกตรัม (Spectrum sensing)
 - 1.4.2.2. การใช้สเปกตรัมร่วมกัน (Spectrum sharing)
- 1.4.3 การตรวจจับพลังงาน (Energy Detection – ED)

- 1.4.4 การเรียนรู้แบบเสริมแรง (Reinforcement Learning – RL)
 - 1.4.4.1. การเรียนรู้แบบคิว (Q-learning)
- 1.4.5 การประเมินประสิทธิภาพ
 - 1.4.5.1. กราฟแสดงความสัมพันธ์ระหว่างผลบวกจริงและผลบวกปลอม (Receiver Operating Characteristic curve – ROC curve)

1.5 ขอบเขตงานวิจัย

- 1.5.1. ศึกษาและปรับปรุงประสิทธิภาพการตรวจจับสเปกตรัมแบบการตรวจจับพลังงานบนระบบวิทยุรู้คิด
- 1.5.2. วิจัยมุ่งเน้นไปที่การปรับปรุงประสิทธิภาพการตรวจจับสเปกตรัมแบบการตรวจจับพลังงานในสภาพแวดล้อมจริง โดยใช้อุปกรณ์ USRP คู่กับโปรแกรม LabVIEW® เพื่อจำลองเป็นวิทยุรู้คิด ซึ่งเป็นผู้ใช้งานหลัก 1 ตัว ผู้ใช้สำรอง 1 ตัวแยกกัน และตัวสร้างสัญญาณรบกวนแบบเกาส์เซียนสีขาวแบบบวกอีก 1 ตัว
- 1.5.3. ใช้การกล้ำสัญญาณแบบ 16QAM, QPSK และ BPSK
- 1.5.4. ใช้ความถี่ 2.4GHz
- 1.5.5. โมเดลที่นำเสนอนี้ เป็นการตรวจจับสเปกตรัมแบบการตรวจจับพลังงานโดยใช้ความรู้ด้านการเรียนรู้แบบควมาช่วยในการปรับปรุงประสิทธิภาพ

1.6 ขั้นตอนการศึกษา

- 1.6.1 ศึกษาภาพรวมของวิทยุรู้คิด โดยศึกษาความเป็นมาเป็นไปของระบบ หลักการของระบบ วิจัยที่เกี่ยวกับระบบวิทยุรู้คิด ที่ได้รับการตีพิมพ์จากวารสาร การประชุมวิชาการที่เชื่อถือได้
- 1.6.2 ศึกษาวิจัยเกี่ยวกับการแก้ไขปัญหาของระบบวิทยุรู้คิดด้วยการใช้หลักการของการเรียนรู้ของเครื่อง, ปัญญาประดิษฐ์, เหมืองข้อมูล
- 1.6.3 ตั้งขอบเขต, วัตถุประสงค์ของการวิจัย
- 1.6.4 ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการทำการตรวจจับสเปกตรัมแบบต่างๆ และการแก้ปัญหาคาการตรวจจับสเปกตรัมโดยใช้หลักการของการเรียนรู้แบบเสริมแรง
- 1.6.5 ศึกษาการใช้งาน USRP และโปรแกรม LabVIEW®
- 1.6.6 ทดลองการใช้ USRP และ LabVIEW® เพื่อทำการตรวจจับสเปกตรัมในเวลาจริง
- 1.6.7 ออกแบบวางแผนโดยการใช้หลักการการเรียนรู้แบบเสริมแรงมาช่วยแก้ปัญหาคาการตรวจจับสเปกตรัมแบบการตรวจจับพลังงาน ให้ดีขึ้น
- 1.6.8 ทดลองโมเดลที่ออกแบบไว้กับสภาพแวดล้อมจริง

1.7 เครื่องมือและอุปกรณ์ที่ใช้ในงานวิจัย

- 1.7.1 คอมพิวเตอร์ CPU Core I7 ระบบปฏิบัติการ Windows 2 เครื่อง
- 1.7.2 เครื่อง USRP NI-2920 3 เครื่อง
- 1.7.3 เสาอากาศแบบรอบทิศทาง (omni directional antenna)
- 1.7.4 โปรแกรมแลปวิว (LabVIEW®) 2015
- 1.7.5 โปรแกรมแมทแลป (MATLAB) รุ่น R2015b
- 1.7.6 สายสัญญาณ Gigabit Ethernet, Gigabit Ethernet switch

1.8 โครงสร้างของวิทยานิพนธ์

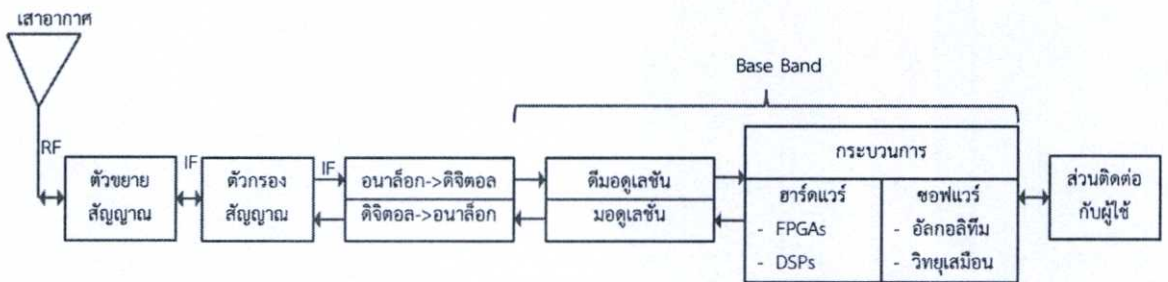
- วิทยานิพนธ์ฉบับนี้แบ่งออกเป็น 6 บท แต่ละบทประกอบด้วยเนื้อหาดังต่อไปนี้
- บทที่ 1 กล่าวถึงความเป็นมาและความสำคัญของปัญหา ความมุ่งหมายและวัตถุประสงค์ของการศึกษา สมมติฐานของการศึกษา ทฤษฎีหรือแนวความคิดที่ใช้ในการศึกษา ขอบเขตของการศึกษา ขั้นตอนของการศึกษา รวมถึงเครื่องมือที่ใช้ในงานวิจัย และโครงสร้างของวิทยานิพนธ์
- บทที่ 2 กล่าวถึงความรู้พื้นฐานที่เกี่ยวข้องกับงานวิจัย ประกอบด้วย หลักการ SDR หลักการวิทยุรู้คิด การตรวจจับแบบตรวจจับพลังงาน การเรียนรู้แบบเสริมแรง การเรียนรู้แบบคิว การเขียนกราฟ ROC
- บทที่ 3 เป็นการทบทวนงานวิจัยที่เกี่ยวข้อง เช่น การตรวจจับสเปกตรัม การใช้คลื่นความถี่วิทยุร่วมกัน การดำเนินการจำลองระบบวิทยุรู้คิด
- บทที่ 4 กล่าวถึงแนวความคิดในกระบวนการตรวจจับสเปกตรัมที่มีการเรียนรู้แบบคิวร่วมด้วย การติดตั้งระบบจำลองในการดำเนินการทดลอง
- บทที่ 5 กล่าวถึงการทดลอง ผลการทดลอง และการวิเคราะห์ผลการทดลอง โดยการทดลองประกอบด้วย การทดลองเปรียบเทียบประสิทธิภาพของการส่งสัญญาณจากผู้ใช้หลักที่แตกต่างกันไป ทั้งวิธีการส่ง และสภาพแวดล้อมที่ใช้ในการส่ง
- บทที่ 6 สรุปผลการทดลอง ข้อเสนอแนะและแนวทางในการปรับปรุงในอนาคต

บทที่ 2

ความรู้พื้นฐานที่เกี่ยวข้องกับงานวิจัย

2.1 หลักการวิทยุที่ดำเนินการด้วยซอฟต์แวร์ (Software Defined Radio – SDR)

วิทยุที่ดำเนินการด้วยซอฟต์แวร์ คือวิทยุที่มีฟังก์ชันส่วนใหญ่ที่ได้รับการดำเนินการด้วยซอฟต์แวร์ และสามารถตั้งค่าการติดต่อสื่อสารได้หลายหลายรูปแบบในวิทยุตัวเดียว เช่น ในสมัยก่อน วิทยุแบบดั้งเดิม ถ้าเป็นวิทยุรับสัญญาณ HF ก็จะได้รับเพียงสัญญาณ HF ไม่สามารถรับสัญญาณ VHF ได้ หรือวิทยุแบบดั้งเดิมซึ่งรับได้เพียงการกล้ำสัญญาณได้ไม่ก็รูปแบบเนื่องจากข้อจำกัดทางฮาร์ดแวร์ แต่วิทยุที่ดำเนินการด้วยซอฟต์แวร์สามารถติดตั้งไลบรารีได้มากขึ้น ใน SDR ประกอบไปด้วยส่วนดำเนินการซึ่งเป็นฮาร์ดแวร์ยกตัวอย่างเช่น ตัวประมวลผลทั่วไป (General Purpose Processor – GPP), ตัวประมวลผลสัญญาณดิจิทัล (Digital Signal Processor – DSP), อุปกรณ์ที่สามารถโปรแกรมสำหรับออกภาคสนาม (Field Programmable Gate Array – FPGA) ซึ่งสัญญาณที่ได้จากการประมวลผลจากอุปกรณ์เหล่านี้จะเรียกว่า Waveform



รูปที่ 2.1 แผนภาพการทำงานของวิทยุดำเนินการด้วยซอฟต์แวร์

รูปที่ 2.1 แสดงแผนภาพการทำงานของวิทยุดำเนินการด้วยซอฟต์แวร์ ในส่วนของวิทยุฝั่งรับ เสาอากาศทำการส่งสัญญาณในอากาศเข้ามาเป็นความถี่วิทยุ (Radio Frequency – RF) ผ่านตัวขยายสัญญาณกลายเป็นความถี่กลาง (Intermediate Frequency – IF) ผ่านตัวกรองเลือกสัญญาณที่ต้องการ แปลงสัญญาณอนาล็อกเป็นสัญญาณดิจิทัลเพื่อนำไปประมวลผลต่อไป ส่วนด้านฝั่งส่งของวิทยุ SDR คือ เมื่อผ่านกระบวนการสัญญาณที่ต้องการจะส่งเรียบร้อยแล้ว ทำการกล้ำสัญญาณ (modulation) เพื่อให้สัญญาณส่งผ่านความถี่ในย่านความถี่ที่ต้องการ เปลี่ยนจากสัญญาณอนาล็อกไปเป็นสัญญาณดิจิทัล กรองสัญญาณ และตัวขยายสัญญาณ ก่อนส่งผ่านเสาอากาศต่อไป

ในปัจจุบัน SDR ได้พัฒนาในส่วนของการเปลี่ยนจากอุปกรณ์ที่ต้องเป็นฮาร์ดแวร์ให้มาใช้ซอฟต์แวร์มากขึ้น เริ่มจากในส่วนกระบวนการทำกับสัญญาณ การมอดูเลชัน การแปลงสัญญาณอนาล็อก ดิจิตอล เรื่อยมาจนถึงภาคแรก จนมาถึงการทำให้เสาอากาศเป็น เสาอากาศอัจฉริยะได้

เมื่อมีการประดิษฐ์หรือสร้างรูปแบบสัญญาณขึ้นมาใหม่ ไม่จำเป็นต้องเปลี่ยนฮาร์ดแวร์อีกต่อไป เพียงแค่เพิ่มโมดูลให้กับซอฟต์แวร์ใน SDR วิทยุที่เป็น SDR ก็จะสามารถใช้ประโยชน์ได้ต่อซึ่งในปัจจุบันมีรูปแบบสัญญาณเพิ่มขึ้นมากมาย

อุปกรณ์การศึกษา SDR ในปัจจุบันมีมากมาย ในส่วนงานวิจัยนี้ได้เลือกนำมาใช้ศึกษาคือเครื่อง USRP (Universal Software Radio Peripheral) เป็น SDR ที่นิยมใช้เพื่อการศึกษาในหัวข้อเกี่ยวกับเรื่องคลื่นความถี่วิทยุ, การติดต่อสื่อสารบนคลื่นความถี่วิทยุ ซึ่งมีมากมายหลายแบบตามความต้องการของระบบ เช่น รองรับคลื่นความถี่ตั้งแต่ 0 เฮิรตซ์ ถึง 6 กิกกะเฮิรตซ์, ส่วนเชื่อมต่อกับคอมพิวเตอร์เป็นแบบ USB หรือ สายแลนิกิกกะบิต, ตัวแปลงอนาล็อกเป็นดิจิตอลความละเอียด 14 บิต, ตัวแปลงดิจิตอลเป็นอนาล็อก 16 บิต, รองรับเสาอากาศแบบ MIMO (Multiple Input Multiple Output) เป็นต้น

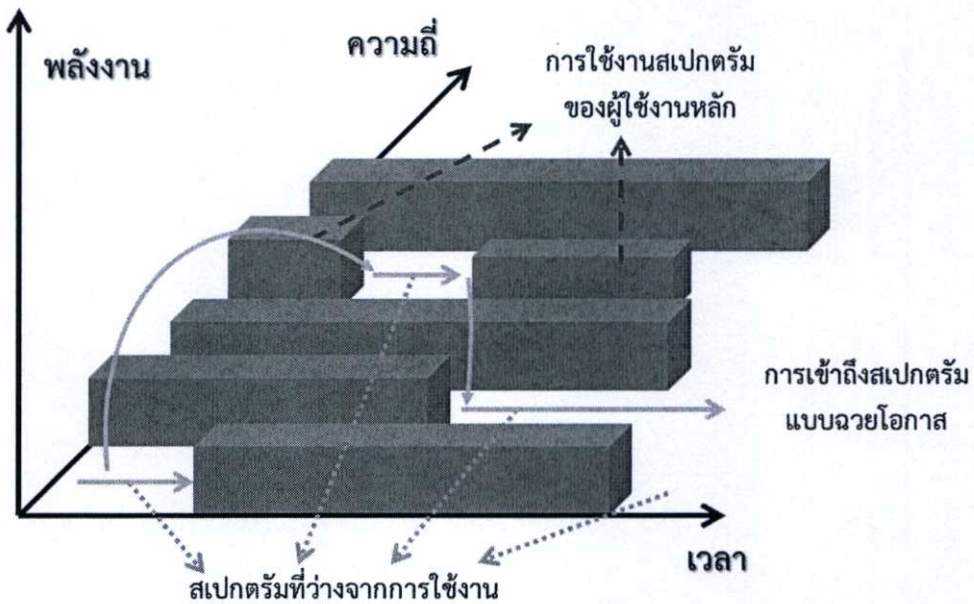


รูปที่ 2.2 USRP ยี่ห้อ Ettus รุ่น N210

2.2 วิทยุรู้คิด (Cognitive Radio)

วิทยุรู้คิดได้รับการนำเสนอครั้งแรกจาก Mitola ในปี 1999 เนื่องจากต้องการแก้ปัญหาการขาดแคลนของคลื่นความถี่วิทยุ และเพื่อใช้ประสิทธิภาพของคลื่นความถี่วิทยุให้มากขึ้น ด้วยการจัดการบริหารคลื่นความถี่วิทยุที่มีความชาญฉลาด ระบบวิทยุรู้คิดนี้จะมีผู้ใช้หลัก (Primary User, PU) คือผู้ใช้งานที่ได้รับใบอนุญาตจากองค์กรที่จัดสรรคลื่นความถี่วิทยุ และผู้ใช้สำรอง (Secondary User, SU หรือ Cognitive Radio, CR) คือผู้ใช้ที่ไม่ได้รับใบอนุญาตจากองค์กรซึ่งมีหน้าที่จัดสรรทรัพยากรทางคลื่นความถี่วิทยุ หลักการของวิทยุรู้คิดคือผู้ใช้สำรองจะตรวจจับคลื่นความถี่รอบตัว ว่ามีความถี่ที่ตนต้องการใช้ติดต่อสื่อสารนั้นมีผู้ใช้งานหลัก ใช้งานความถี่นั้นอยู่หรือไม่ ถ้าไม่มีผู้ใช้สำรอง

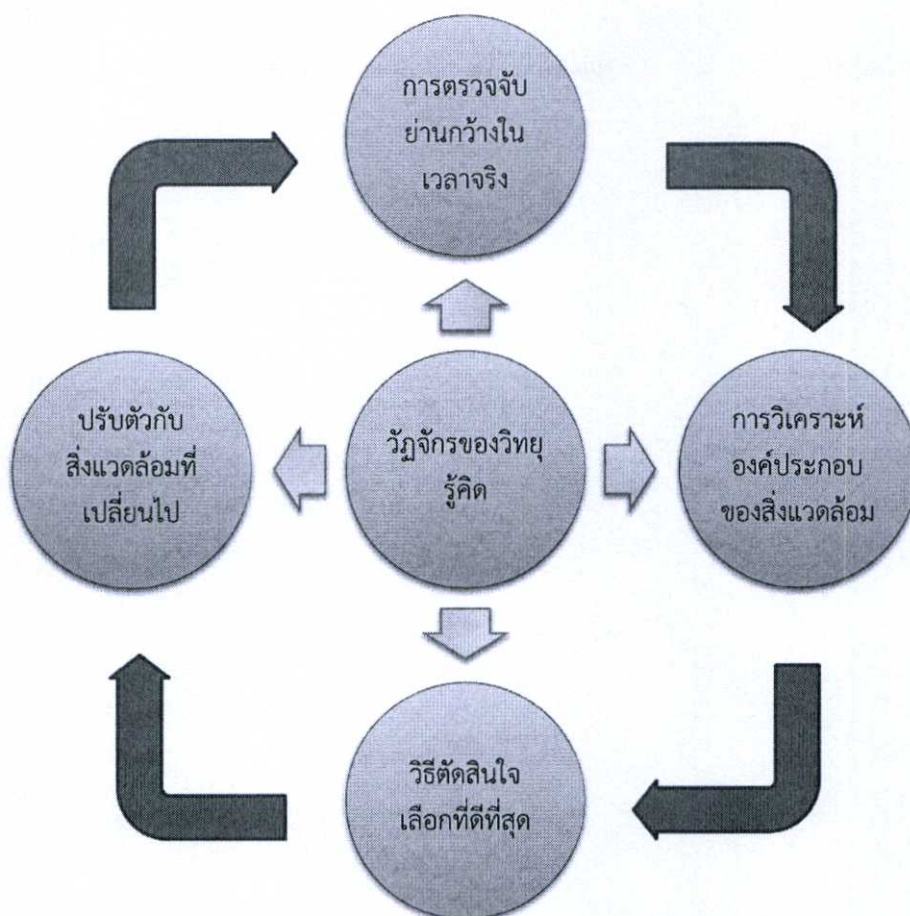
จะใช้ความถี่นั้น ในเวลาใดในทันที แต่ถ้าพบว่าไม่มีผู้ใช้งานหลักใช้งานอยู่ ก็อาจจะไม่ส่งข้อมูล หรือส่งข้อมูลโดยมีการกำหนดรูปแบบการส่งเพื่อไม่ให้เกิดสัญญาณรบกวนกับผู้ใช้งานหลักเด็ดขาด



รูปที่ 2.3 จำลองการเข้าถึงสเปกตรัมแบบฉวยโอกาส (opportunistic access)

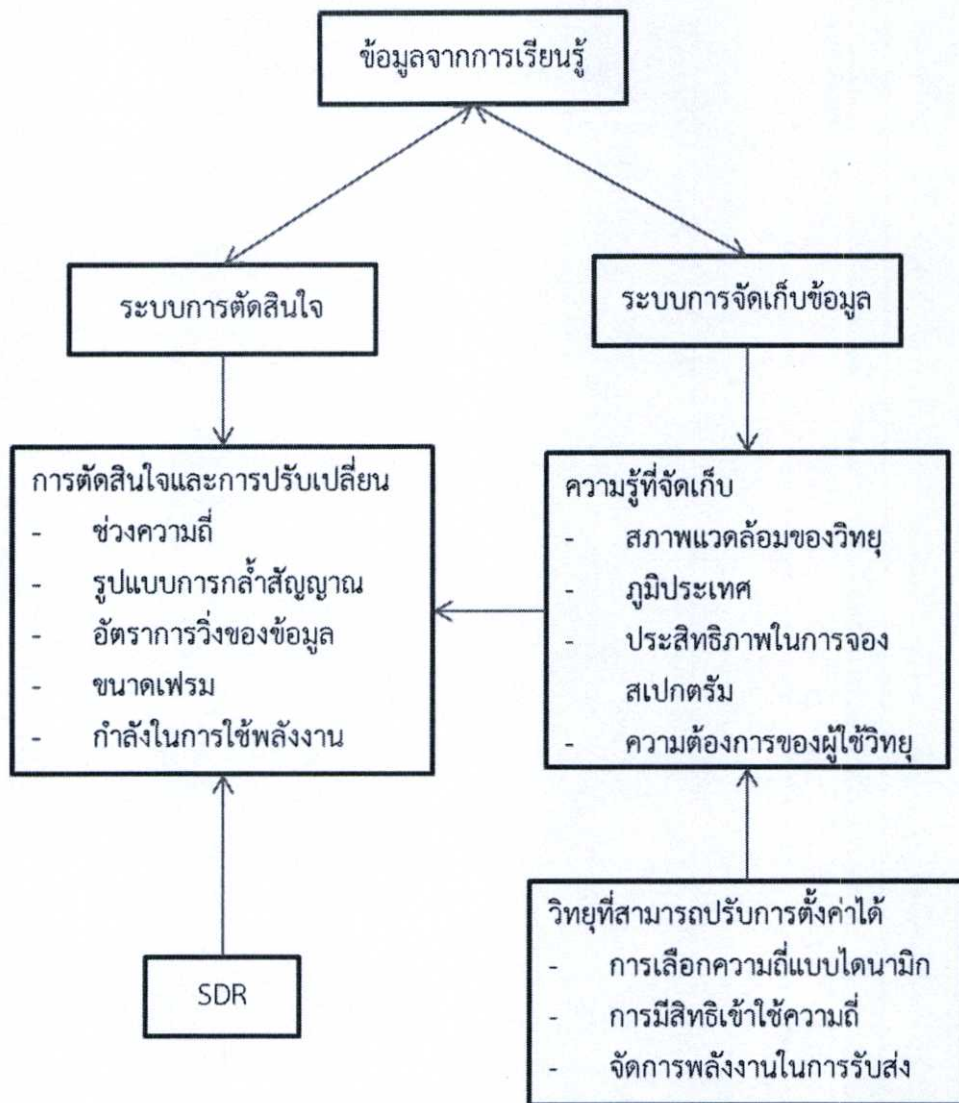
จากรูปที่ 2.3 อธิบายการเข้าถึงสเปกตรัมแบบฉวยโอกาส (opportunistic access) โดยการใช้หลักการตรวจจับสเปกตรัม (spectrum sensing) เพื่อตรวจสอบว่าในช่วงความถี่ตรงไหนสามารถใช้งานได้โดยปราศจากการใช้งานของผู้ใช้หลักและจะคืนความถี่ดังกล่าวเมื่อการใช้งานของผู้ใช้หลักปรากฏขึ้นเพื่อไม่ให้เกิดการรบกวนไปสู่ผู้ใช้งานหลัก ในปัจจุบันมีการศึกษาหลากหลายทางที่ใช้ในการช่วยระบบวิทยุรู้คิดนี้ตรวจสอบและให้ระบบทำงานได้ตามทฤษฎีได้มากขึ้น เช่น การใช้ความรู้ด้านการเรียนรู้ของเครื่อง (machine learning) มาช่วยในการตรวจจับสเปกตรัม, ช่วยในการทำนายการใช้งานของผู้ใช้งานหลัก, เรียนรู้การใช้งานของผู้ใช้งานหลัก เป็นต้น

ระบบวิทยุรู้คิดเป็นระบบที่ต้องเรียนรู้จากสิ่งแวดล้อมรอบตัวและยังสามารถปรับตัวปรับการตั้งค่าของการส่งวิทยุตามสิ่งแวดล้อมรอบตัวได้เพื่อให้ได้คำตอบที่ดีที่สุด อธิบายดังรูปที่ 2.4



รูปที่ 2.4 หลักการทำงานของระบบวิทยุรู้คิด

ก่อนที่ระบบวิทยุรู้คิดจะปรับแต่งปรับตัวตามสิ่งแวดล้อมดังกล่าวข้างต้นนั้น ระบบจะต้องได้รับข้อมูลที่สำคัญจากสิ่งแวดล้อมซึ่งจะดีหรือไม่อยู่ที่ผู้ศึกษาว่าเลือกความสำคัญของข้อมูลที่เข้ามา เช่น ลักษณะของผู้ใช้งานหลัก, คลื่นความถี่วิทยุที่ผู้ใช้งานหลักใช้, การก่อกวนสัญญาณ, ข้อมูลทางภูมิศาสตร์, ข้อมูลทางด้านผู้รับ เป็นต้น และการปรับแต่งวิทยุรู้คิดนั้น สามารถปรับแต่ง กำลังของการส่ง, ความถี่ในการส่ง, การก่อกวนสัญญาณ, ขนาดของสัญญาณในการส่ง เป็นต้น



รูปที่ 2.5 องค์ประกอบของระบบวิทยุรู้คิด

2.2.1 การตรวจจับสเปกตรัม (Spectrum sensing)

จากหลักการของระบบวิทยุรู้คิดในรูปที่ 2.4 มีขั้นตอนหนึ่ง คือ การตรวจจับสเปกตรัม เป็นการตรวจจับว่ามีช่องสัญญาณที่ว่างจากผู้ใช้งานหลักหรือไม่ ผู้ใช้สำรอง หรือ วิทยุรู้คิดจะคอยตรวจจับสเปกตรัมเป็นระยะๆ ว่าถ้าไม่มีการใช้งานหลักจะตั้งสมมติฐานในการตรวจสอบเป็น H_0 (Hypothesis 0) แต่ถ้าตรวจพบที่มีการใช้งานของผู้ใช้หลักอยู่จะตั้งสมมติฐานในการตรวจสอบเป็น H_1 (Hypothesis 1) ซึ่งช่องสัญญาณที่ผู้ใช้สำรองสามารถใช้งานได้ต้องได้รับการตั้งสมมติฐานว่า H_0 เป็นจริง ดังสมการที่ 2.1

$$f(t) = \begin{cases} H_0, & ; \text{กรณีไม่มีการใช้งานหลัก} \\ H_1, & ; \text{กรณีมีการใช้งานหลัก} \end{cases} \quad (2.1)$$

การตรวจจับสัญญาณซึ่งในปัจจุบันมีการศึกษาค้นคว้าและพัฒนาต่อยอดให้มีประสิทธิภาพดีขึ้นเป็นอย่างมาก แต่ในวิทยานิพนธ์นี้จะขอยกตัวอย่างวิธีการดั้งเดิมในยุคแรกๆ ของการศึกษาระบบวิทยุรู้คิดดังต่อไปนี้

1. การตรวจจับพลังงานของสัญญาณ (Energy detection) เป็นการคำนวณพลังงานที่ได้รับการตรวจจับ ณ เวลาใดเวลาหนึ่ง แล้วนำไปเทียบกับเกณฑ์ (threshold) ซึ่งค่าเกณฑ์ที่ยอมรับได้ ก็ยังคงได้รับการศึกษาพัฒนาเพื่อให้ค่าเกณฑ์นั้นถูกต้องแม่นยำมากขึ้น ซึ่งวิธีการตรวจจับพลังงานนี้ ได้นำมาใช้ในการวิจัยครั้งนี้ ซึ่งจะอธิบายอย่างละเอียดในบทถัดไป



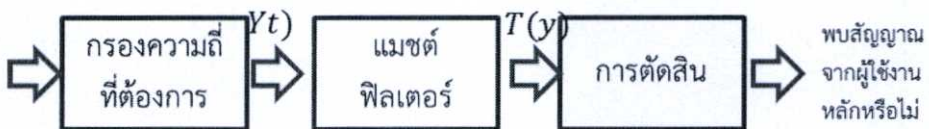
รูปที่ 2.6 ขั้นตอนของการตรวจจับพลังงานของสัญญาณ (energy detection)

2. Cyclostationary Feature Detection วิธีนี้อาศัยข้อมูลในการรับส่งของผู้ใช้งานหลัก เช่น สถิติในการส่งสัญญาณในรูปแบบการส่งที่แตกต่างกัน, อัตราการกล้ำสัญญาณ, ความถี่นำพา เพื่อการตรวจจับที่แม่นยำยิ่งขึ้นกว่า การตรวจจับพลังงานของสัญญาณ



รูปที่ 2.7 ขั้นตอนของ cyclostationary feature detection

3. Matched Filtering Detection เป็นการตรวจจับโดยการทำสหสัมพันธ์กับสัญญาณที่เข้ามา ซึ่งไม่รู้รูปแบบของสัญญาณ ด้วยรูปแบบ (template) ของโมเดลในการตัดสินใจ

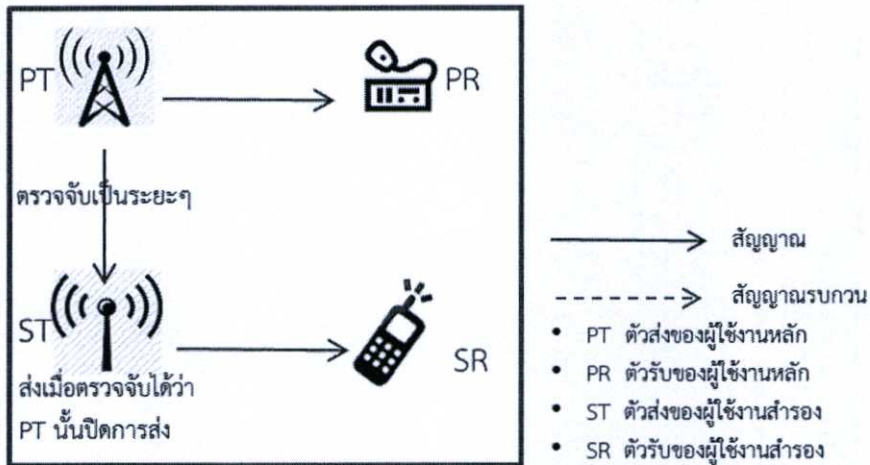


รูปที่ 2.8 ขั้นตอนของ matched filtering detection

2.2.2 การใช้สเปกตรัมร่วมกัน (Spectrum sharing)

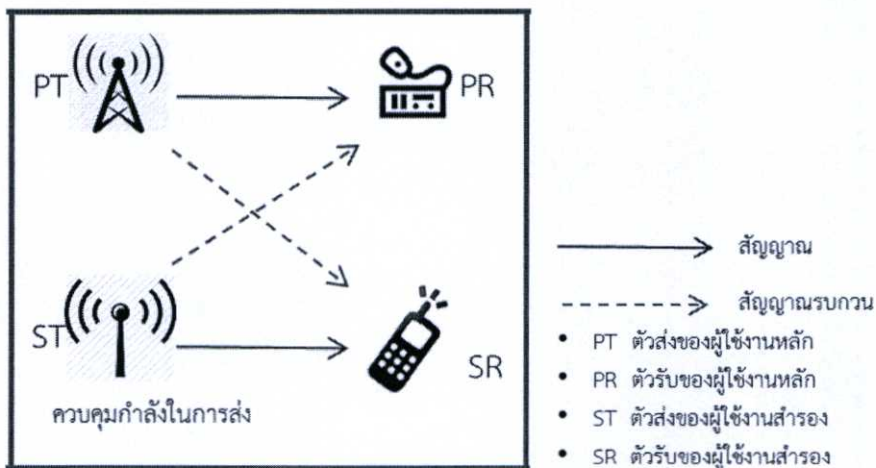
การใช้สเปกตรัมร่วมกัน (Spectrum Sharing) มี 3 ประเภท ได้แก่

1. แบบผสมผสานกัน (Interweave) ผู้ใช้สำรองจะตรวจสอบผู้ใช้งานหลักเป็นระยะๆ ว่ามีการส่งข้อมูลใดๆหรือไม่ ผู้ใช้งานรองเมื่อพบว่าผู้ใช้งานหลักไม่ส่งข้อมูลจึงทำการส่งหรือใช้ความถี่นั้นในการส่งแบบไม่มีสัญญาณรบกวนให้กับผู้ใช้งานหลัก



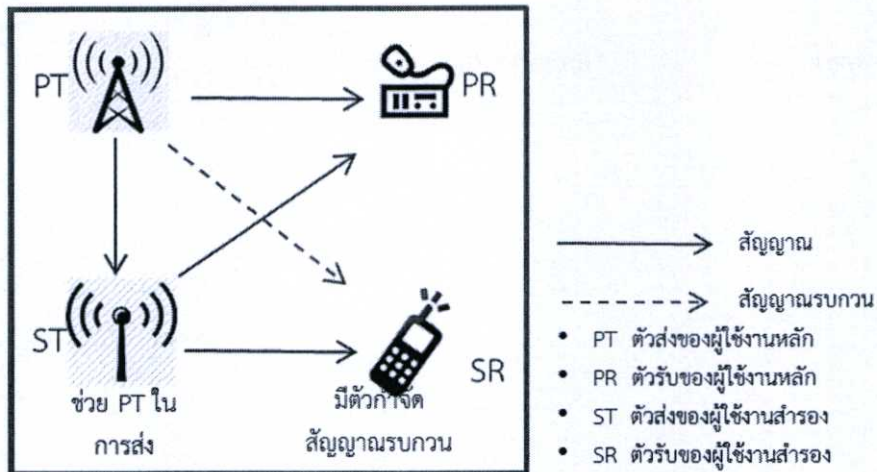
รูปที่ 2.9 การใช้สเปกตรัมร่วมกันแบบผสมผสานกัน

2. แบบเบี่ยงล่าง (Underlay) ผู้ใช้สำรองจะส่งโดยมีการควบคุมกำลัง (Power controlled transmission) ในการส่ง เพื่อไม่ให้ผู้ใช้งานหลักได้รับผลกระทบใดๆ



รูปที่ 2.10 การใช้สเปกตรัมร่วมกันแบบเบี่ยงล่าง

3. แบบซ้อนทับ (Overlay) ผู้ใช้สำรองจะมีการเข้ารหัส หรือกระทำการใดๆ ในการส่ง เพื่อให้แยกออกจากข้อมูลใดเป็นของผู้ใช้งานหลัก ข้อมูลใดเป็นของผู้ใช้สำรอง

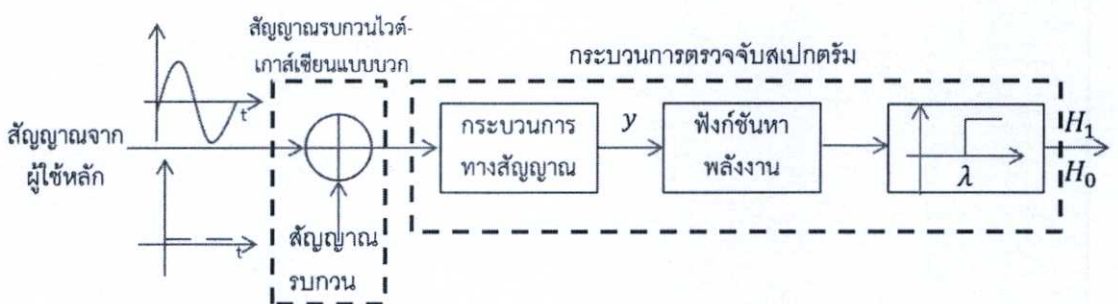


รูปที่ 2.11 การใช้สเปกตรัมร่วมกันแบบซ้อนทับ

2.3 การตรวจจับแบบตรวจจับพลังงานของสัญญาณ (Energy detection)

2.3.1 หลักการการตรวจจับ

การตรวจจับแบบตรวจจับพลังงานของสัญญาณเป็นการตรวจจับของผู้ใช้สำรอง ตรวจจับช่องสัญญาณที่สามารถใช้งานได้ โดยการตรวจว่าช่องสัญญาณนั้นมีการปรากฏพลังงานของผู้ใช้งานหลักอยู่หรือไม่ ถ้ามีพลังงานของผู้ใช้งานหลักอยู่ซึ่งแปลได้ว่า มีการใช้งานของผู้ใช้งานหลักอยู่ ผู้สำรองก็ไม่สามารถใช้งานช่องสัญญาณนั้นได้ โดยข้อดีของการตรวจจับแบบตรวจจับพลังงานนี้คือ ไม่จำเป็นต้องรู้ความรู้เริ่มต้นของผู้ใช้งานหลัก เช่น รูปแบบในการกล้ำสัญญาณ เป็นต้น สิ่งที่ต้องการคือ เกณฑ์ในการตรวจจับว่ามีหรือไม่มีสัญญาณอยู่, ความแรงของสัญญาณ, ขนาดของสัญญาณ, จำนวนของการแซมปลิง และประสิทธิภาพของการตรวจจับ



รูปที่ 2.12 ขั้นตอนการตรวจจับพลังงานของสัญญาณ

สัญญาณที่ส่งจากฝั่งผู้ใช้งานหลักไปยังฝั่งรับของผู้ใช้งานหลัก แทนด้วย สัญญาณเชิงซ้อน s ซึ่งประกอบด้วย ส่วนจริง s_r และส่วนจินตภาพ s_i ตัวอย่างเช่น $s = s_r + js_i$ และถ้าสัญญาณที่ผู้ใช้งานหลักได้รับ, การตรวจจับของผู้ใช้สำรองที่ได้รับสัญญาณ y ได้รับการแซมปลิงจำนวน n แซมปลิง $y(n)$ จะเป็นดังต่อไปนี้

$$y(n) = \begin{cases} w(n) : H_0 \\ x(n) + w(n) : H_1 \end{cases} \quad (2.2)$$

โดยที่สามารถกำหนดสมมติฐานในการตรวจจับได้ 2 ข้อคือ H_0 คือ Hypothesis 0 แปลว่า ไม่พบสัญญาณของผู้ใช้งานหลัก และ H_1 คือ Hypothesis 1 แปลว่า ไม่พบสัญญาณของผู้ใช้งานหลัก

หลักการของการตรวจจับแบบตรวจจับพลังงาน จากรูปที่ 2.12 คือ ผู้ใช้รองรับสัญญาณในอากาศ บนช่องสัญญาณที่ต้องจะใช้งาน เมื่อรับสัญญาณดังกล่าวมา ก็จะกรองสัญญาณรบกวนออกก่อนเป็นอันดับแรก นำสัญญาณที่ได้ไปทำการใส่กำลังสอง และนำไปหาอินทิเกรต จะได้ผลลัพธ์ที่จะ

นำไปเทียบกับค่าเกณฑ์ (threshold) ต่อไปว่าช่องสัญญาณนี้มีการใช้งานอยู่หรือไม่

ฟังก์ชันหาค่าพลังงานก่อนนำไปเปรียบเทียบ

$$T(n) = \frac{1}{N} \sum_{n=0}^{N-1} |y[n]|^2 \quad (2.3)$$

2.3.2 รูปแบบความน่าจะเป็นที่ใช้สำหรับวิเคราะห์การตรวจจับสเปกตรัม

เพื่อเป็นการวิเคราะห์การตรวจจับสเปกตรัม ไม่ว่าจะใช้รูปแบบการตรวจจับสเปกตรัมแบบใด การวิจัยส่วนมากใช้ความน่าจะเป็นในการวิเคราะห์การตรวจจับสเปกตรัมดังต่อไปนี้

- ความน่าจะเป็นในการตรวจจับถูกต้อง (Probability of Detection: P_d) อธิบายเป็นสมการได้ว่า

$$P_d = P(H_1|H_1) \quad (2.4)$$

- ความน่าจะเป็นในการตรวจจับผิดแบบ Missed (Probability of Missed Detection: P_m) คือตรวจจับได้ว่าไม่พบ แต่ความจริงคือมีผู้ใช้หลักอยู่ อธิบายเป็นสมการได้ว่า

$$P_m = P(H_0|H_1) \quad (2.5)$$

ความน่าจะเป็นในการตรวจจับผิดแบบ false alarm (Probability of false alarm: P_{fa}) คือตรวจจับได้ว่าพบ แต่ความจริงคือไม่มีผู้ใช้หลักอยู่ อธิบายเป็นสมการได้ว่า

$$P_{fa} = P(H_1|H_0) \quad (2.6)$$

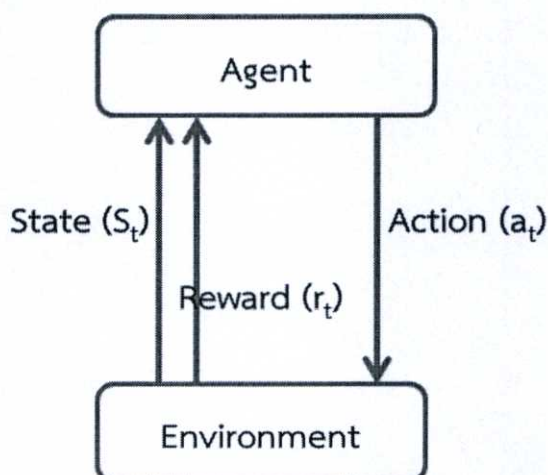
2.4 การเรียนรู้แบบเสริมแรง (Reinforcement Learning)

การเรียนรู้ของเครื่อง (Machine Learning) แบ่งออกเป็น 3 ประเภทคือ

1. Supervised Learning คือ มีเรียนรู้จากผลลัพธ์ที่มีอยู่ก่อนแล้ว เช่น ให้เครื่องเรียนรู้และแบ่งแยกข้อมูลว่า ข้อมูลที่ได้มา เป็นสีแดง หรือสีดำ
2. Unsupervised Learning คือ ไม่มีผลลัพธ์ก่อนหน้า แต่จะใช้การจัดกลุ่มของข้อมูลที่ได้รับเข้ามา ว่าข้อมูลแบบนี้อยู่กลุ่มเดียวกัน และสามารถเลือกจำนวนกลุ่มที่ต้องการแบ่งได้เอง

3. Reinforcement Learning คือ เรียนรู้จากการกระทำและให้คะแนนจากการกระทำนั้น ว่าให้ผลลัพธ์ได้ถูกต้องหรือไม่ ถ้าเป็นการกระทำที่ดีจะให้คะแนนเป็นบวก แต่ถ้าการกระทำนั้นเป็นสิ่งไม่ดี จะให้คะแนนติดลบ

การเรียนรู้แบบเสริมแรง จากที่กล่าวมาข้างต้น คือเป็นการเรียนรู้ของเครื่องชนิดหนึ่งที่แบ่งออกมาจาก supervised learning และ unsupervised learning การเรียนรู้แบบเสริมแรงเป็นการเรียนรู้ที่อาจเรียกได้ว่าเป็นการลองผิดลองถูกจากสถานการณ์หรือสิ่งแวดล้อมที่ agent หรือตัวที่ทำการเรียนรู้พบเจอโดยการเลือกสถานการณ์ถัดไปจากการกระทำในปัจจุบันด้วยรางวัล (Reward) หรือบทลงโทษ (Punishment) ความสามารถในการเรียนรู้แบบเสริมแรงคือจะพยายามสร้างสมดุลระหว่างการสำรวจคำตอบอื่นๆ ที่ดีกว่า หรือการใช้คำตอบที่คิดว่าดีที่สุดที่สุด้นั้นแล้ว (Exploration and exploitation) ซึ่งเหมาะสมกับการแก้ปัญหาที่มีสิ่งแวดล้อมที่หลากหลาย เปลี่ยนแปลงตลอดเวลา หรือสถานการณ์ที่ไม่แน่นอน



รูปที่ 2.13 รูปแบบของการเรียนรู้แบบเสริมแรง

ส่วนประกอบของการเรียนรู้แบบเสริมแรง

1. นโยบาย (Policy) คือ การจับคู่ของการเลือกการกระทำ (Action) ที่และสถานะที่เหมาะสม ซึ่ง agent ได้รับเข้ามาจากสิ่งแวดล้อมที่กำหนด โดยมีสัญลักษณ์เป็น π
2. การกำหนดรางวัล (Reward function) เป็นการเปลี่ยนสถานะที่ได้รับเข้ามา หรือการจับคู่ของสถานะและการกระทำให้เป็นตัวเลขเดียว ซึ่งมีสัญลักษณ์เป็น r
3. การกำหนดมูลค่า (Value function) มีสัญลักษณ์เป็น V จะคล้ายกับ reward เพียงแต่การกำหนดมูลค่าเป็นการกำหนดค่าให้กับปัญหานั้นๆ ในระยะยาว หรืออาจหมายถึงการคำนวณปัญหาในอนาคตได้อีกด้วย

2.4.1 การเรียนรู้แบบคิว (Q-Learning)

การที่ agent จะเรียนรู้สิ่งต่างๆ แบบการเรียนรู้แบบคิวนั้น agent จะต้องมีหน่วยความจำหรือเปรียบเทียบได้เป็นสมองของคน หน่วยความจำของ agent ในการเรียนรู้แบบคิว คือ เมตริกซ์ Q ซึ่งประกอบด้วยแถวหมายถึงสถานะปัจจุบันของ agent และสดมภ์หมายถึงการกระทำเพื่อไปสู่สถานะถัดไป ซึ่งความหมายของเมตริกซ์ Q มีความหมายเหมือน เมตริกซ์ R ซึ่งได้รับการสมมติให้เป็น ระบบรางวัล ดังสมการที่ 2.6

ในขั้นแรก agent จะถูกสมมติให้ว่าไม่รู้อะไรมาก่อน ดังนั้น เมตริกซ์ Q จึงมีค่าเป็น เมตริกซ์ศูนย์ การเรียนรู้แบบคิวสามารถเพิ่มแถวหรือสดมภ์ขึ้นมาเมื่อใดก็ได้ ถ้าพบว่ามิสถานะใหม่เกิดขึ้น ทำให้ยืดหยุ่นต่อการเรียนรู้ของ agent

เริ่มต้นเมตริกซ์ Q ให้กับระบบด้วยเมตริกซ์ศูนย์

$$Q = \begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array} \begin{array}{c} A \\ B \\ C \\ D \\ E \\ F \end{array} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (2.7)$$

สมมติระบบรางวัลเป็นเมตริกซ์ R

$$R = \begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array} \begin{array}{c} \text{สถานะ} \\ A \\ B \\ C \\ D \\ E \\ F \end{array} \begin{array}{c} \text{การกระทำ} \\ A \\ B \\ C \\ D \\ E \\ F \end{array} \begin{bmatrix} - & - & - & - & 0 & - \\ - & - & - & 0 & - & 100 \\ - & - & - & 0 & - & - \\ - & 0 & 0 & - & 0 & - \\ 0 & - & - & 0 & - & 100 \\ - & 0 & - & - & 0 & 100 \end{bmatrix} \quad (2.8)$$

สมการการเปลี่ยนแปลงของการเรียนรู้แบบคิว เป็นดังนี้

$$Q(s_t, a_t) = \alpha \cdot (R(s_t, a_t) + \gamma \text{Max}[Q(s_{t+1}, a_{all})]) \quad (2.9)$$

เมื่อ

s_t คือ สถานะปัจจุบัน

a_t คือ การกระทำที่ทำให้เกิดสถานะปัจจุบัน

α คือ อัตราการเรียนรู้ (Learning Rate)

γ คือ อัตราการลด (discount factor)

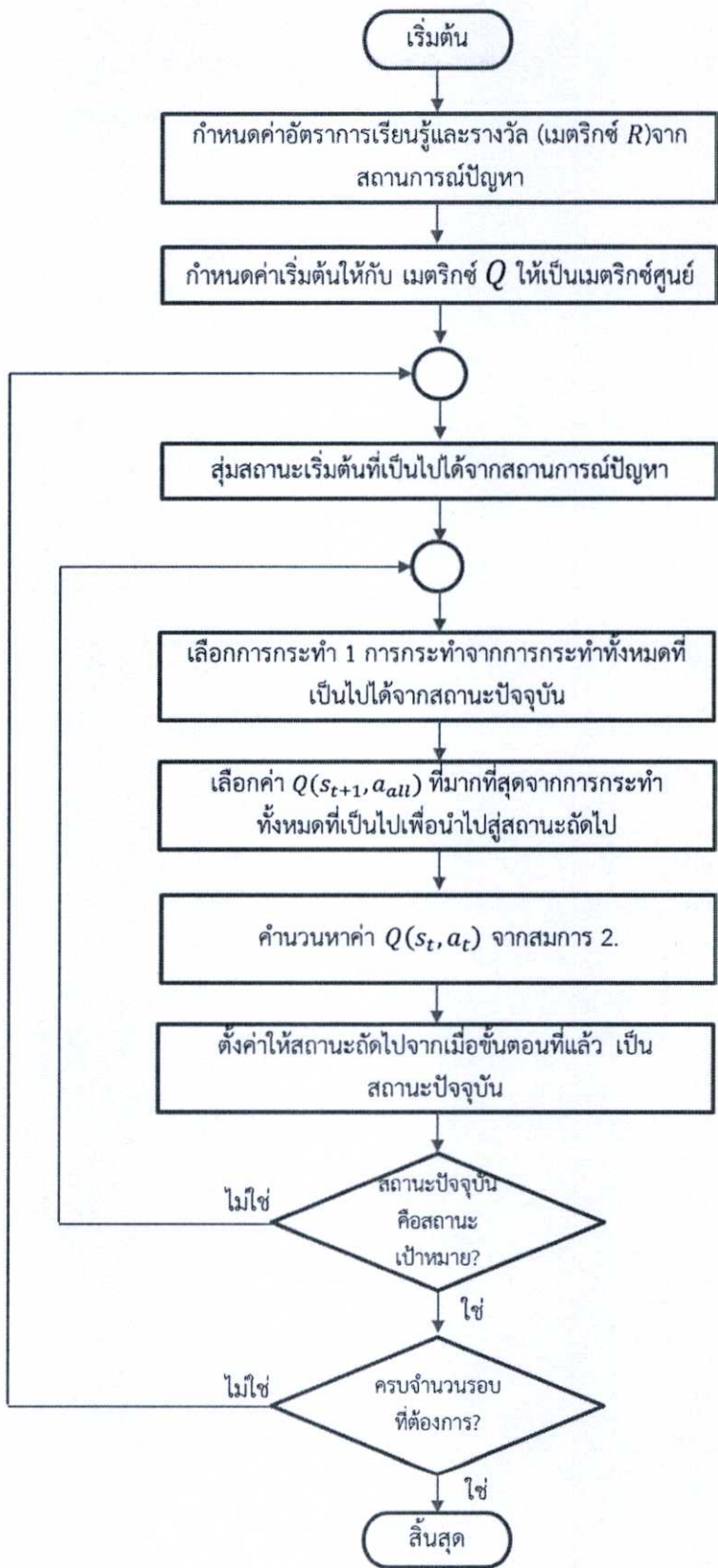
s_{t+1} คือ สถานะถัดไป

a_{all} คือ การกระทำทั้งหมดในเมตริกซ์ที่มีให้เลือกในสถานะถัดไป

จากสมการที่ 2. อธิบายได้ว่า เมตริกซ์ Q หาได้จาก เมตริกซ์ R บวกด้วย ผลคูณของอัตราการเรียนรู้กับค่าในเมตริกซ์ Q ที่มีค่าการกระทำมากที่สุดของสถานะถัดไป

การเลือกสถานะเริ่มต้นของการเรียนรู้ จะเลือกจากสถานะที่เป็นรางวัลที่เป็นไปได้ คือ มากกว่า ศูนย์ โดยถ้ามีมากกว่า 1 สถานะที่เป็นไปได้ที่ให้เลือก จะเลือกด้วยการสุ่ม

ขั้นตอนการทำงานของการเรียนรู้แบบควิจะสำรวจ (explore) สถานะทุกสถานะจนกว่าการแก้ปัญหาจะถึงเป้าหมาย การสำรวจ 1 รอบ (episode) หมายถึงการสำรวจสถานะเริ่มต้นไปจนถึงสถานะเป้าหมาย ซึ่งเมื่อถึงเป้าหมายในการสำรวจ 1 รอบ agent จะทำการสำรวจรอบใหม่ โดยเริ่มกำหนดสถานะเริ่มต้นอีกและทำวนจนถึงสถานะเป้าหมายไปเรื่อยๆ จนกว่าจะถึงกำหนดการสำรวจที่เรากำหนดไว้ตั้งแต่แรก ดังที่แสดงในรูปที่ 2.14



รูปที่ 2.14 Flow chart ของการเรียนรู้แบบคิว

2.5 การประเมินประสิทธิภาพการตรวจจับสเปกตรัม

2.5.1. กราฟแสดงความสัมพันธ์ระหว่างผลบวกจริงและผลบวกปลอม (Receiver Operating Characteristic curve – ROC curve)

กราฟ ROC ย่อมาจาก Receiver Operating Characteristic เป็น กราฟ ROC ที่ได้รับความนิยมใช้เพื่อวัดประสิทธิภาพของหลากหลายวัตถุประสงค์ เช่น ใช้กับการตรวจจับสัญญาณเพื่อวัดประสิทธิภาพของอัตราการตรวจจับถูกต้อง (hit rate) และอัตราการตรวจจับผิด (false alarm), ใช้ในทางการแพทย์ การรักษา การวินิจฉัยโรค เป็นต้น

ROC เริ่มจากปัญหาการจัดหมวดหมู่ที่มีเพียง 2 หมวดหมู่ โดยที่แต่ละตัวอย่าง (instance) แทนให้เป็นสัญลักษณ์ I ประกอบไปด้วยเซต $\{p, n\}$ ซึ่งหมายถึงหมวดหมู่ที่เป็นบวก (positive class) และหมวดหมู่ที่เป็นลบ (negative class) เพื่อเป็นการแยกระหว่างคลาสที่แท้จริงกับคลาสที่ถูกทำลาย จึงแทนสัญลักษณ์การจัดหมวดหมู่บวก, ลบ ของคลาสที่แท้จริงเป็น $\{p, n\}$ และคลาสที่ถูกทำนายเป็น $\{Y, N\}$ ซึ่งมีความหมายแทนคลาสที่เป็นบวกลบเหมือนกันเพียงแต่ใช้สัญลักษณ์ต่างกัน

จากการจัดหมวดหมู่ (classifier) และ ตัวอย่าง (instance) จะเกิดผลลัพธ์ที่เป็นไปได้ 4 ผลลัพธ์นั้นคือ

- ถ้าตัวอย่างที่จริงเป็นบวกและถูกจัดว่าเป็นบวก จะเรียกว่าเป็น บวกจริง (true positive)
- ถ้าตัวอย่างที่จริงเป็นบวกและถูกจัดว่าเป็นลบ จะเรียกว่าเป็น ลบเท็จ (false negative)
- ถ้าตัวอย่างที่จริงเป็นลบและถูกจัดว่าเป็นลบ จะเรียกว่าเป็น ลบจริง (true positive)
- ถ้าตัวอย่างที่จริงเป็นลบและถูกจัดว่าเป็นบวก จะเรียกว่า ลบเท็จ (false negative)

ในรูปแบบทั่วไป จะแทนเซตของตัวอย่างและเซตของตัวจัดหมวดหมู่เป็นเมตริกซ์ 2×2 เรียกว่า เมตริกซ์ความสับสน (confusion matrix)

ตารางที่ 2.1 เมตริกซ์ความสับสน

ผลที่เกิดการจัดกลุ่ม	ผลที่เป็นจริง	
	p	n
Y	บวกจริง (TP)	บวกเท็จ (FP)
N	ลบเท็จ (FN)	ลบจริง (TN)
ผลรวมสดมภ์	P	N

$$\text{True positive rate} \approx \frac{TP}{P} \quad (2.10)$$

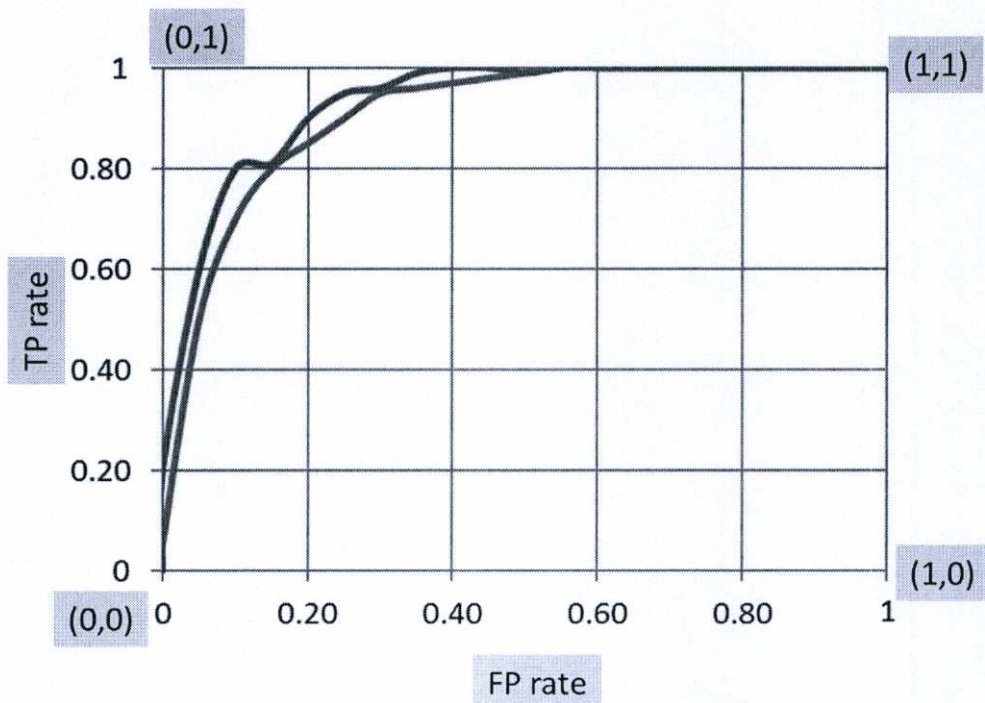
$$\text{False Positive rate} \approx \frac{FP}{N} \quad (2.11)$$

จากตารางที่ 2. แสดงถึงเมตริกซ์ความสับสน สมการที่ 2.9 True positive rate คืออัตราการจับผิดที่ถูกต้อง (ทำนายถูก) สมการที่ 2.10 False positive rate คืออัตราการจับผิดที่ผิด (ทำนายผิด)

คำศัพท์อื่นๆ ที่เกี่ยวข้องกับการเขียนกราฟ ROC มีดังต่อไปนี้

- ความไว (Sensitivity) เท่ากับ อัตราการจับผิดที่ถูกต้อง
- ความจำเพาะ (Specificity) เท่ากับ 1-อัตราการจับผิดที่ผิด

กราฟ ROC เป็นกราฟสองมิติซึ่งอัตราการจับผิดที่ถูกต้อง (TP rate) อยู่บนแกน Y และอัตราการจับผิดที่ผิด (FP rate) อยู่บนแกน X กราฟ ROC สามารถอธิบายความสัมพันธ์ระหว่างผลประโยชน์ (benefit) และ มูลค่า (cost) จากรูปที่ 2.15



รูปที่ 2.15 ตัวอย่างการเขียนกราฟ ROC

บทที่ 3

วรรณกรรมที่เกี่ยวข้อง

3.1 การตรวจจับสเปกตรัม (Spectrum sensing)

Cabric และคณะ [4] ได้ทำการทดลองเปรียบเทียบประสิทธิภาพของการตรวจจับสเปกตรัมหลักๆ ที่เป็นต้นแบบในการศึกษาต่อในปัจจุบัน คือ การทำแมทซ์ฟิลเตอร์ (matched filtering), การตรวจจับพลังงาน (energy detection), และการตรวจจับรูปแบบของสัญญาณที่ส่งพร้อมกับการกล้ำสัญญาณ (cyclostationary feature detection) และได้จำแนกข้อดีข้อเสียของแต่ละหลักการ โดยสรุปได้ดังนี้

1. การทำแมทซ์ฟิลเตอร์ ข้อดีคือรองรับอัตราส่วนสัญญาณต่อเสียงรบกวนได้สูงสุด ข้อเสียคือก่อนที่จะทำการแมทซ์ฟิลเตอร์ต้องมีความรู้ก่อนหน้า (Priori Knowledge) ทั้งใน PHY เลเยอร์และ MAC เลเยอร์ ยกตัวอย่างเช่น รูปแบบในการกล้ำสัญญาณ, รูปแบบลำดับการส่ง, รูปร่างของคลื่นสัญญาณ, รูปแบบการจัดแพ็คเกจ
2. การตรวจจับพลังงาน ข้อดีคือเป็นการตรวจจับที่ง่าย ไม่ยุ่งยาก ไม่ต้องมีข้อมูลก่อนหน้า ไม่ว่าจะรูปแบบสัญญาณจะเป็นอย่างไร ก็สามารถใช้การตรวจจับแบบนี้ได้ แต่จะมีประสิทธิภาพไม่ดีเมื่อมีอัตราส่วนสัญญาณต่อเสียงรบกวนน้อย
3. การตรวจจับรูปแบบของสัญญาณที่ส่งพร้อมกับการกล้ำสัญญาณ ข้อดีคือมีประสิทธิภาพในการตรวจจับที่ดี ไม่ว่าจะมีความถี่สัญญาณรบกวนมาก หรือสัญญาณถูกกล้ำมาซับซ้อนเพียงใด เพียงแต่เป็นการตรวจจับที่ซับซ้อน มีกรรมวิธียุ่งยาก

การทดลองประสิทธิภาพของ [4] ได้ทำการทดลองโดยใช้สัญญาณที่ได้รับการกล้ำสัญญาณด้วย 4-FSK การตรวจจับแบบ cyclostationary สามารถทน SNR ได้ต่ำสุดถึง -20dB ในขณะที่ energy detection ทนได้แค่ 10dB

3.1.1 การตรวจจับสเปกตรัมโดยใช้ความรู้เรื่องการเรียนรู้แบบเสริมแรง

การตรวจจับสเปกตรัมโดยใช้การเรียนรู้แบบเสริมแรงและการเข้าใช้ประโยชน์ของสเปกตรัมของระบบวิทยุรู้คิดได้รับการเสนอใน [7][8][9] โดยที่ใน [7] และ [8] ได้นำเสนอการตรวจจับสเปกตรัมโดยที่กระจายความสำคัญให้ทั้งการตรวจจับที่ดี, การรับส่งที่ดี, และการสับเปลี่ยนช่องสัญญาณให้กับผู้ใช้ CR ในระบบให้ใช้ได้อย่างทั่วถึง การวิจัยใน [9] ใช้การเรียนรู้แบบเสริมแรงแบบการเรียนรู้ Q โดยใช้การทำงานของตัวควบคุมจากส่วนกลาง (fusion center) ร่วมด้วย และการเลือกแบบ ϵ -greedy โดยแยกการกระทำออกเป็น 2 ระยะ ในระยะแรก ตัวควบคุมจากส่วนกลางจะให้ CR หาแถบความถี่ที่ดีที่สุดด้วยการตรวจจับสเปกตรัมเพื่อที่ CR จะได้ใช้ประสิทธิภาพของการ

รับส่งได้มากที่สุด ในระยะที่สอง CR ที่ได้รับเลือกให้เป็นผู้ตรวจจับสเปกตรัมในระยะแรกจะต้องมีความน่าจะเป็นในการตรวจจับผิดพลาดให้น้อยที่สุดเท่าที่จะเป็นไปได้

3.1.1.1 การตรวจจับสเปกตรัมโดยใช้ความรู้เรื่องการเรียนรู้แบบคิว

Zhe และ Robert [10] ได้นำเสนอวิจัยและการทดลองของการตรวจจับสเปกตรัม โดยการใช้หลักการการเรียนรู้แบบคิว ในปี ค.ศ.2011 งานวิจัยนี้ออกแบบโมเดลในการทดลองว่า ทำการทดลองกับสัญญาณ Wi-Fi แบบสถานการณ์จริง ในห้องปิด โดยใช้อุปกรณ์ปล่อยสัญญาณ Wi-Fi เชื่อมต่อกับอินเทอร์เน็ต ทำการดาวน์โหลดด้วยแบนวิธ 2.3 เมกกะไบต์ต่อวินาที และใช้เสาอากาศ 4 ต้น ในการเป็นตัวแทนผู้ใช้งานหลัก 1 ต้น และผู้ใช้สำรอง 3 ต้น ใช้ข้อสลิโอสโคปดิจิทัลใช้สารเรืองแสง (Digital Phosphor Oscilloscope – DPO) ใช้จับสัญญาณในการทดลอง โดยจับสัญญาณในระยะเวลา 40 มิลลิวินาที และแสดงผลด้วยกราฟแสดงเวลา

งานวิจัยของ Zhe และ Robert นี้มีหลักการและตัวแปรของการตรวจจับสเปกตรัม คู่กับการเรียนรู้แบบคิว ดังนี้

- กำหนดจำนวนสถานะของการเรียนรู้เป็นการมีอยู่หรือไม่ของสเปกตรัม คือ สถานะเป็น 1 หมายถึง สเปกตรัมมีผู้ใช้งานอยู่, สถานะเป็น 0 คือ สเปกตรัมพร้อมใช้งาน สมการที่ 3. เป็นการหาจำนวนสถานะเพื่อนำไปใช้ในการเรียนรู้แบบคิว

$$s_{t,f} = \sum_{i=1}^N d_{i,t,f} \cdot 2^{i-1} \quad (3.1)$$

- กำหนดรางวัลของการเรียนรู้เป็น ค่าคงที่ โดยแบ่งเป็น 2 กลุ่มดังสมการที่ 3.

$$r = \begin{cases} R_p & a = c \\ R_n & a \neq c \end{cases} \quad (3.2)$$

- ให้อัตราการเรียนรู้เป็น 0.1 และอัตราส่วนลดเท่ากับ 0.1
- ใช้ความถี่ Wi-Fi ที่ 2.418 กิกะเฮิรตซ์

เมื่อ

a คือ การกระทำที่เลือกกระทำ

c คือ สถานะของช่องสัญญาณที่แท้จริง

ขั้นตอนการทำงานของ การตรวจจับสเปกตรัมโดยใช้ความรู้เรื่องการเรียนรู้แบบคิว

- ขั้นตอนที่ 1. คำนวณหาจำนวนสถานะ S จากสมการที่ 3.1
- ขั้นตอนที่ 2. เลือกการกระทำที่มีค่าการกระทำสูงสุด
- ขั้นตอนที่ 3. รับรางวัลจากการเลือกการกระทำจากขั้นตอนที่ 2.
- ขั้นตอนที่ 4. ปรับปรุงค่า $Q(s, a)$ จากสมการการเรียนรู้แบบคิวในบทที่ 2 สมการที่ 2.7
- ขั้นตอนที่ 5. ทำซ้ำขั้นตอนที่ 1.

การทดลองของการทดลองนี้ อ่านค่าได้เพียงในรูปแบบของโดเมนเวลา (time domain) เนื่องจากใช้เพียงออสซิลโคปในการวัดเท่านั้น ซึ่งการทดลองที่เกี่ยวกับสัญญาณคลื่นความถี่วิทยุ ควรวิเคราะห์ที่ได้หลากหลายกว่านี้

3.1.1.2 การตรวจจับสเปกตรัมที่ใช้การเรียนรู้แบบซาร์ซา (Sarsa learning)

[6] ได้นำเสนอการกระจายของ CR ที่มีหลายผู้ใช้ (multiuser) หลายช่วงความถี่ (Multiband) ด้วยวิธีการเรียนรู้แบบเสริมแรงที่มีหลายตัวเรียนรู้ (multiagent) โดยกล่าวว่า เพื่อเพิ่มประสิทธิภาพในการตรวจจับการใช้งานของสเปกตรัม ในเครือข่ายของผู้ใช้ CR ต้องมีการกระจายการผลของการตรวจจับให้แก่ผู้ใช้ CR อื่นๆ ในเครือข่ายด้วย ทั้งเป็นการประหยัดเวลาและทรัพยากรของผู้ใช้ CR เอง ผู้วิจัยนี้ต้องการนโยบายการตรวจจับสเปกตรัมที่มีความน่าจะเป็นในการตรวจจับมากที่สุดและลดความน่าจะเป็นของการตรวจจับผิดพลาดให้น้อยที่สุดที่เป็นไปได้ โดยได้ทำจำลองผลจากการตรวจจับหาช่องสัญญาณบนความถี่ที่มีช่องสัญญาณมาก ที่ใช้งานได้เครือข่าย CR

หลักการที่ [6] นำเสนอ

- ใช้การเรียนรู้แบบเสริมแรงที่มีชื่อว่า Sarsa ด้วยขั้นตอนแบบฟังก์ชันเชิงเส้นเพื่อลดมิติและการคำนวณที่ซับซ้อนในคู่ของสถานะและการกระทำของเอเจนต์ (state-action pair) ในระบบวิทยุรู้คิด
- ผู้วิจัยได้จำลองสร้างสถานการณ์โครงข่าย CR แบบที่ไม่มี ตัวควบคุมจากส่วนกลางหรือสถานีฐาน เนื่องจากสถานการณ์ก็อาจเกิดขึ้นในชีวิตประจำวัน

3.2 การใช้คลื่นความถี่ร่วมกัน (Spectrum sharing)

Tao และคณะ [5] ได้เสนอการแจกจ่ายคลื่นความถี่วิทยุสำหรับวิทยุรู้คิดซึ่งช่วยลดความจำเป็นในการทำการตรวจจับสเปกตรัม โดยใช้การเรียนรู้แบบเสริมแรงเข้ามาช่วยแทนที่จะตรวจจับคลื่นความถี่วิทยุทั้งหมดที่ใช้งานได้ และได้เปรียบเทียบประสิทธิภาพกับการที่ไม่ได้ใช้การเรียนรู้เข้ามาช่วยในการแบ่งคลื่นความถี่วิทยุกันใช้อีกด้วย

วิทยุรู้คิดใน [5] ได้กำหนดค่านิยามไว้จาก 6 คือ “วิทยุรู้คิดเป็นวิทยุที่ตื่นตัวตลอดเวลาและสามารถรับรู้ถึงสิ่งแวดล้อมที่มันอยู่, เรียนรู้จากสิ่งแวดล้อมของมันและปรับตัวเพื่อการใช้งานของ

ตัวเองได้” หนึ่งในวิธีการที่จะทำให้วิทยุคิดค้นตัวตลอดเวลาคือการตรวจจับการใช้งานคลื่นความถี่สเปกตรัม (Spectrum Sensing) แต่เป็นกระบวนการที่สิ้นเปลืองพลังงานและเวลาเป็นอย่างมาก จึงคิดค้นหาวิธีการแบ่งการใช้สเปกตรัมร่วมกัน (Spectrum sharing) โดยการเรียนรู้สิ่งแวดล้อมของการติดต่อสื่อสารแบบไร้สายโดยการเรียนรู้แบบเสริมแรง (Reinforcement Learning)

สิ่งที่วิจัยนี้ทำและเปรียบเทียบประสิทธิภาพของระบบ แบ่งออกเป็น 3 หัวข้อได้แก่

1. ผู้ใช้สำรองทุกตัวทำการตรวจจับสเปกตรัมเองทุกตัว ทุกช่องความถี่ที่มี
2. แบ่งกลุ่มผู้ใช้สำรองทำการตรวจจับสเปกตรัมในกลุ่มช่องความถี่ที่กำหนดไว้ให้
3. ผู้ใช้สำรองไม่ต้องทำการตรวจจับสเปกตรัม ผู้ใช้สำรองใช้ช่องความถี่ที่กำหนดไว้ให้

การเรียนรู้แบบเสริมแรงเป็นขั้นตอนการคำนวณทางคณิตศาสตร์เพื่อที่จะเรียนรู้จากพฤติกรรมที่โต้ตอบและเพิ่มประสิทธิภาพในการตรวจจับหรือการแบ่งกันใช้ของช่องความถี่วิทยุที่ไม่มีใช้งาน โดยการใช้รางวัลเป็นตัวอ้างอิงว่าควรปรับน้ำหนัก (Weight) ไปในทิศทางใด และน้ำหนักเป็นตัวเลขบ่งบอกถึงความสำคัญและลำดับความสำคัญของช่องสัญญาณที่วิทยุคิดจะเข้าไปใช้ [5] ได้เสนอคำว่า pre-play ซึ่งหมายถึง สถานะของการกระจายตัวของผู้ใช้สำรองที่กำลังหาจุดที่ดีที่และเรียนรู้จากประสบการณ์ของการหาจุดที่ดีที่สุดนั้น [5] แทนผู้ใช้สำรอง หรือ CR นั้นว่าเป็นผู้เล่น (player) กล่าวว่า ในสถานะ pre-play นั้น ผู้เล่นจะสำรวจคลื่นความถี่ที่สามารถใช้งานได้โดยการเข้าถึงแบบกายภาพ ซึ่งผู้เล่นแต่ละคนจะมีค่าความน่าจะเป็นในการสำรวจคลื่นความถี่วิทยุที่เท่ากันทั้งหมด และจะปรับค่าน้ำหนักของทรัพยากรที่ถูกใช้แล้วนั้นเพื่อผู้เล่นเหล่านี้จะต้องถูกปรับปรุงตัวเองหลังจากที่ผ่านการเรียนรู้แต่ละครั้งไปในทันที หลักการนี้คือการเปลี่ยนค่าน้ำหนักในทันที (temporal difference) ตามหลักการของการให้คะแนน น้ำหนักของการใช้สเปกตรัมที่ประสบความสำเร็จหรือการใช้สเปกตรัมที่เป็นสเปกตรัมที่ใช้งานได้จะต้องทำการปรับเพิ่มค่าน้ำหนัก และในทางตรงกันข้ามถ้าการใช้สเปกตรัมไม่ประสบความสำเร็จเช่น สเปกตรัมนั้นว่างแล้วไม่ได้ใช้ หรือ ไปใช้สเปกตรัมที่มีผู้ใช้งานอยู่แล้ว จะต้องได้รับการปรับลดค่าน้ำหนัก โดยในเกมนี้ ผู้เล่นหรือผู้ใช้ CR จะต้องเรียนรู้การเลือกช่องสัญญาณในการติดต่อสื่อสารที่ดี

ขั้นตอนในการทดลองอัลกอริทึม ของหลักการแบ่งสเปกตรัมใช้ร่วมกันของ [5] ได้ทำการจับคู่การรับส่งของแต่ละโหนด กระจายกันอยู่ในพื้นที่สี่เหลี่ยม และเรียงตามขั้นตอนดังนี้



รูปที่ 3.1 ขั้นตอนการแบ่งสเปกตรัมใช้ร่วมกัน

ขั้นตอนที่ 1 ประเมินสถานะ

ผู้ใช้ CR ถูกแทนด้วยคู่ของการรับการส่งแทนด้วย U_i โดยที่ i คือตัวเลขแทนผู้ใช้ CR ได้รับการเริ่มต้นค่าของการปรับค่าน้ำหนักอยู่ที่ 5 ถ้าเป็นการติดต่อสื่อสารที่ดี จะได้รับปรับค่าน้ำหนักให้เพิ่ม แต่ถ้าเป็นการติดต่อสื่อสารที่แย่ จะได้รับการปรับค่าน้ำหนักในลดลง

ขั้นตอนที่ 2 ตรวจสอบสเปกตรัม

U_i จะทำการตรวจสอบระดับสัญญาณรบกวนบนช่องสัญญาณด้วยการตรวจสอบแบบตรวจสอบพลังงานคือถ้ามีสัญญาณรบกวนน้อยกว่าเกณฑ์ก็สามารถใช้งานช่องสัญญาณนั้นได้ แต่ถ้าตรวจพบว่ามีสัญญาณรบกวนมากกว่าเกณฑ์ก็ไม่สามารถใช้งานช่องสัญญาณนั้นได้ แล้วเริ่มทำแบบเดิมกับช่องสัญญาณใหม่

ขั้นตอนที่ 3 คำนวณ SINR

ในฝั่งรับจะคำนวณ SINR เพื่อปรับปรุงคุณภาพของการติดต่อสื่อสารของช่องสัญญาณ [5] จะคำนวณ SINR ที่เป็นเกณฑ์ไว้ แล้วคำนวณ SINR ของแต่ละ U_i ถ้า SINR ของ U_i มากกว่า SINR ที่เป็นเกณฑ์ แปลว่าการส่งประสบผลสำเร็จดี และจะเพิ่มค่าให้กับการปรับน้ำหนักให้มีค่าเพิ่มขึ้น แต่ถ้า SINR ของ U_i น้อยกว่าค่าเกณฑ์ U_i จะทำการทิ้งการติดต่อช่องสัญญาณนั้นทิ้งและปรับค่าน้ำหนักเป็นการลงโทษ หรือทำให้ค่าปรับน้ำหนักน้อยลง

จากการทดลองของ [5] ผลการทดลองพบว่า การตรวจสอบสเปกตรัมเพียงโดยจัดกลุ่มผู้ใช้สำรองแบ่งใช้สเปกตรัมที่กำหนดเท่านั้น ใช้เวลาและพลังงานน้อยกว่า การตรวจสอบสเปกตรัมของผู้ใช้สำรองกับทุกสเปกตรัมถึง 23 เปอร์เซ็นต์ แต่เมื่อเปรียบเทียบกับกรณีไม่ทำการตรวจสอบสเปกตรัม การไม่ทำการตรวจสอบสเปกตรัมทำให้การใช้งานสเปกตรัมไม่มีประสิทธิภาพที่ดี

3.3 การดำเนินการจำลองระบบวิทยุรู้คิด (Implementation)

เนื่องจากระบบวิทยุรู้คิดไม่ได้เป็นเพียงวิทยุธรรมดาทั่วไป แต่เป็นการใช้หลักการ Software Defined Radio (SDR) เข้ามาช่วยให้ระบบวิทยุรู้คิดนี้ให้เป็นระบบวิทยุที่ชาญฉลาดทันสมัย ในปัจจุบันมีการศึกษาและการทดลองระบบวิทยุรู้คิดมากมาย เช่น จำลองระบบวิทยุจาก GNU [11] ซึ่งเป็นระบบจำลองวิทยุที่ให้ดาวน์โหลดได้ฟรี, การทดลองในสภาพจริง มีการรับ-ส่งข้อมูลจริง ในอากาศ โดยมีเครื่องมือช่วย เช่น USRP ย่อมาจาก Universal Software Radio Peripheral System Model and test bed description

งานวิจัยของ Gulnur และคณะ [11] ได้ทำการทดลองการตรวจสอบสเปกตรัมแบบการตรวจสอบพลังงาน และการเข้าถึงสเปกตรัมแบบไดนามิก โดยใช้อุปกรณ์ SDR ที่ชื่อว่า USRP เชื่อมต่อกับคอมพิวเตอร์ ทำงานเป็นวิทยุ ควบคู่กับโปรแกรม LabVIEW® บนสถานการณ์สิ่งแวดล้อมจริง และมีการสร้างสัญญาณรบกวนที่สามารถปรับเพื่อวัดประสิทธิภาพในการจำลองได้ตามต้องการและครอบคลุม โดยโมเดลที่ Gulnur สร้างขึ้นมีหลักการดังต่อไปนี้

- ใช้ USRP ต่อกับคอมพิวเตอร์ เป็น ผู้ใช้งานหลัก 1 ตัว
- ใช้ USRP ต่อกับคอมพิวเตอร์ เป็น ผู้ใช้สำรอง 1 ตัว

- ใช้ USRP ต่อกับคอมพิวเตอร์ เป็น ตัวตรวจจับสเปกตรัม 1 ตัว
 - ใช้ USRP ต่อกับคอมพิวเตอร์ เป็น ตัวสร้างสัญญาณรบกวน 1 ตัว
- สัญญาณที่ได้รับที่ตัวตรวจจับสเปกตรัมเป็นดังนี้

$$r_{SU}(k) = h(k)x_{PU}(k) + f(k)x_{SU}(k) + n_n(k) + n_g(k) \quad (3.3)$$

โดยที่

$x_{PU}(k)$ คือสัญญาณที่ส่งมาจากผู้ใช้งานหลัก

$x_{SU}(k)$ คือสัญญาณที่ส่งมาจากผู้ใช้สำรอง

$h(k)$ คือค่าสัมประสิทธิ์ช่องสัญญาณของการส่งของผู้ใช้งานหลัก

$f(k)$ คือค่าสัมประสิทธิ์ช่องสัญญาณของสัญญาณรบกวน

$n_n(k)$ คือสัญญาณรบกวนจากธรรมชาติ

$n_g(k)$ คือสัญญาณรบกวนจากตัวสร้างสัญญาณรบกวนเป็นแบบเกาส์เซียน σ^2

k คือ ช่องเวลา

USRP ที่ทำหน้าที่เป็น ผู้ใช้งานสำรอง จะรับข้อมูลจากตัวตรวจจับสเปกตรัมว่า ช่องสัญญาณนั้นใช้ได้หรือไม่ แล้วจึงดำเนินการต่อ ตัวตรวจจับเมื่อได้รับสัญญาณจากสมการที่ 3. จะทำการตรวจจับสเปกตรัมแบบตรวจจับพลังงาน ที่ผู้วิจัยนี้เลือกใช้การตรวจจับดังกล่าวนี้เป็นเพราะใช้ทรัพยากรน้อย และมีความยุ่งยากน้อย โดยมีสมมติฐาน 2 แบบ คือ H_0 ไม่พบการใช้งานของผู้ใช้หลักอยู่ สามารถใช้งานได้ และ H_1 มีการใช้งานหลักอยู่ ไม่สามารถใช้งานได้ อธิบายได้เป็นสมการดังนี้

$$H_0: r_{SU}(k) = f(k)x_{SU}(k) + n_n(k) + n_g(k) \quad (3.4)$$

$$H_1: r_{SU}(k) = h(k)x_{PU}(k) + n_n(k) + n_g(k) \quad (3.5)$$

คำนวณค่าที่จะนำไปเทียบกับค่าเกณฑ์โดยใช้ผลรวมของสัญญาณที่รับเข้ามากำลังสอง

$$T(k) = \sum_{i=k-w}^k |r_{SU}(i)|^2 \quad (3.6)$$

เมื่อ w คือขนาดหน้าต่างของการตรวจจับสเปกตรัม เมื่อเข้าฟังก์ชันสร้างตัวทดสอบแล้ว ก็นำมาเทียบกับเกณฑ์ และเพื่อวัดประสิทธิภาพของโมเดล วิจัยนี้ได้วัดความน่าจะเป็นในการเกิดการตรวจจับผิดพลาด และความน่าจะเป็นของการตรวจจับถูกต้อง

การทดลองของวิจัยนี้ [11] ได้ทำการวัดจำนวนครั้งของผู้ใช้สำรองใช้ประโยชน์ของสเปกตรัม และวัดจำนวนครั้งของผู้ใช้สำรองไม่เข้าใช้สเปกตรัมที่ว่างอยู่

บทที่ 4

การตรวจจับสเปกตรัมของระบบวิทยุรู้คิดบนเครื่อง USRP

ในบทนี้จะเป็นการอธิบายการตั้งค่าการทดลองตรวจจับสเปกตรัมด้วยการตรวจจับพลังงาน โดยใช้หลักการการเรียนรู้แบบควิรร่วมด้วย การทดลองใช้สภาพแวดล้อมจริง ด้วยอุปกรณ์ USRP และใช้โปรแกรม LabVIEW® ในการสั่งการดำเนินการเครื่อง USRP ทำการตรวจวัดประสิทธิภาพของระบบด้วยกราฟ ROC ระหว่างความน่าจะเป็นในการตรวจจับถูกต้องและความน่าจะเป็นในการเกิดสัญญาณหลอก(ในบทที่ 2) โดยจะแบ่งระบบออกเป็น 2 ส่วนคือ กระบวนการที่นำเสนอในการตรวจจับพลังงานโดยใช้การเรียนรู้ของควิรช่วย และการติดตั้งฮาร์ดแวร์ระบบวิทยุรู้คิด

4.1 กระบวนการตรวจจับสเปกตรัมที่นำเสนอ

ในหัวข้อนี้นำเสนอกระบวนการที่นำเสนอการตรวจจับสเปกตรัมแบบการตรวจจับพลังงาน กระบวนการเรียนรู้ของควิร กระบวนการปรับปรุงพารามิเตอร์ของวิทยุระหว่างที่ระบบกำลังดำเนินการติดตั้งฮาร์ดแวร์ของระบบวิทยุรู้คิดที่นำเสนอเบื้องต้น

4.1.1 กระบวนการตรวจจับแบบตรวจจับพลังงาน

การตรวจจับสเปกตรัมแบบตรวจจับพลังงานเป็นรูปแบบการตรวจจับสเปกตรัมที่งานวิจัยนี้เลือกใช้ เนื่องจากมีข้อดีคือ ไม่ว่าสัญญาณที่ได้รับมาจะเป็นรูปแบบการกล้ำสัญญาณอย่างไร ก็สามารถตรวจจับได้ เพียงแต่ประสิทธิภาพในการตรวจจับจะลดลงเมื่อมีอัตราความเข้มของสัญญาณต่ำเมื่อเทียบกับชั้นสัญญาณรบกวน (SNR ต่ำ) วัตถุประสงค์ของการทำวิจัยนี้คือ ปรับปรุงให้การตรวจจับสเปกตรัมแบบตรวจจับพลังงานนี้มีประสิทธิภาพที่ดีมากยิ่งขึ้น

โดยกระบวนการตรวจจับสเปกตรัมอธิบายเป็นไดอะแกรมได้ดังนี้



รูปที่ 4.1 กระบวนการตรวจจับสเปกตรัมแบบตรวจจับพลังงาน

4.1.2 กระบวนการเรียนรู้แบบควิร

4.1.2.1 พารามิเตอร์ของการเรียนรู้แบบควิร

พารามิเตอร์ที่ใช้งานในการเรียนรู้ของควิร ตามหลักการเรียนรู้ของควิรที่กล่าวในบทที่ 2 สามารถแจกแจงรายละเอียดได้ตามตารางที่ 4.

ตารางที่ 4.1 พารามิเตอร์ในการใช้งานการเรียนรู้ของคิว

พารามิเตอร์	ค่าของพารามิเตอร์
Agent	ผู้ใช้งานสำรอง
State	ช่องสัญญาณทั้งหมดในการตรวจจับ
Action	ช่องสัญญาณที่เป็นไปได้สำหรับ state นั้นๆ
Reward	เท่ากับ 1 เมื่อพบผู้ใช้งานหลัก เท่ากับ 0 เมื่อไม่พบผู้ใช้งานหลัก
Learning rate	0.25
Discount factor	0.2

4.1.2.2 ขั้นตอนการเรียนรู้แบบคิวที่นำเสนอ

ขั้นตอนการทำงานของระบบการเรียนรู้แบบคิวที่กระทำโดย agent คือผู้ใช้งานสำรองสามารถอธิบายได้ดังนี้

ขั้นตอนที่ 1 กำหนด Q-table ให้เป็นเมตริกซ์ศูนย์

ขั้นตอนที่ 2 เลือก state (รอบแรกเป็นการสุ่ม)

ขั้นตอนที่ 3 เลือก action จากค่า Q-table ที่มีค่ามากที่สุดที่สามารถเป็นไปได้ที่จะเป็นสถานะถัดไป

ขั้นตอนที่ 4 ทำการตรวจจับสัญญาณบนช่องสัญญาณที่เลือกจาก action ที่ได้รับ แล้วนำผลไปคำนวณ reward ตามตารางที่ 4.

ขั้นตอนที่ 5 อัปเดตค่า Q ใน Q-table ตามสมการที่ 2.9

ขั้นตอนที่ 6 ตรวจสอบค่า Q ที่ได้อัปเดตไปในขั้นตอนที่ 5 กับค่า Q ที่อัปเดตในขั้นตอนที่ 5 ในรอบการเรียนรู้ก่อนหน้านี้ ว่ามีค่าที่ดีขึ้นหรือไม่ ถ้าดีขึ้น ให้ข้ามไปทำขั้นตอนที่ 8 แต่ถ้าแย่ลงให้ทำขั้นตอนที่ 7 ต่อ (ในรอบแรก ให้ข้ามไปทำขั้นตอนที่ 8)

ขั้นตอนที่ 7 ทำการปรับค่าอัตราขยาย (gain) ของวิทย์ ซึ่งจะกล่าวในหัวข้อที่ 4.1.3

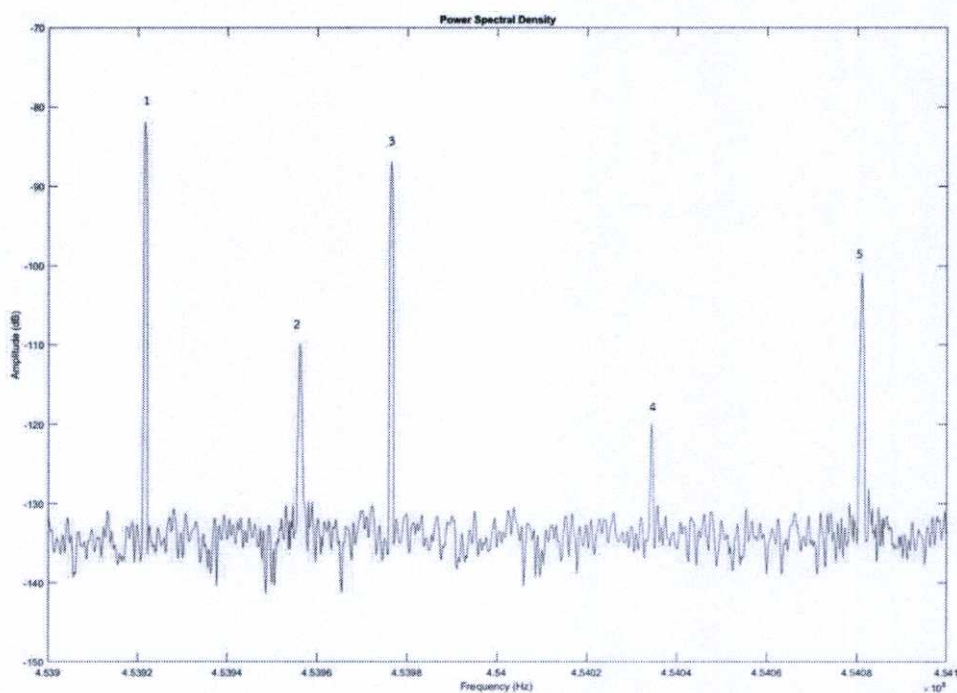
ขั้นตอนที่ 8 วนซ้ำขั้นตอนที่ 2 จนครบจำนวนรอบที่ต้องการ

4.1.3 กระบวนการปรับปรุงพารามิเตอร์ของวิทย์ในวิทย์รู้คิด

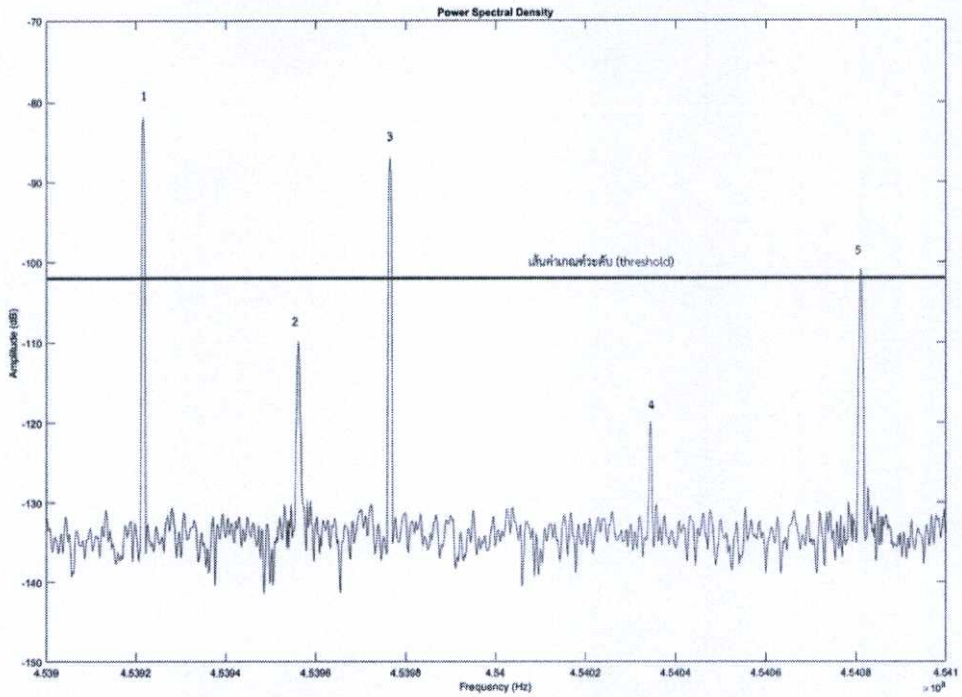
จากขั้นตอนการเรียนรู้แบบคิวที่กล่าวในหัวข้อ 4.1.2 ไปแล้วนั้น มีขั้นตอนย่อยอีกขั้นตอนหนึ่งซึ่งกระทำเพื่อปรับปรุงประสิทธิภาพในการตรวจจับให้ดียิ่งขึ้น โดยนำผลจากการเรียนรู้มาการปรับค่าอัตราขยาย (gain) มีกระบวนการดังนี้

เมื่อระบบการเรียนรู้ของคิว คำนวณค่า Q ได้ต่ำลงจากรอบการประมวลผลก่อนหน้านี้ เนื่องจากเรียนรู้ช่องสัญญาณแล้วไม่พบว่ามีผู้ใช้หลักอยู่ การเพิ่มโอกาสที่จะตรวจจับถูกต้องคือเมื่อเปรียบเทียบค่า Q ปัจจุบัน มีค่าต่ำกว่าค่า Q ของการประมวลผลก่อนหน้านี้ ระบบจะทำการปรับปรุงอัตราการขยาย เพื่อให้ยกสัญญาณที่รับเข้ามาในรอบถัดไป ให้มีค่าสูงขึ้น เพื่อให้สัญญาณที่อาจจะต่ำ

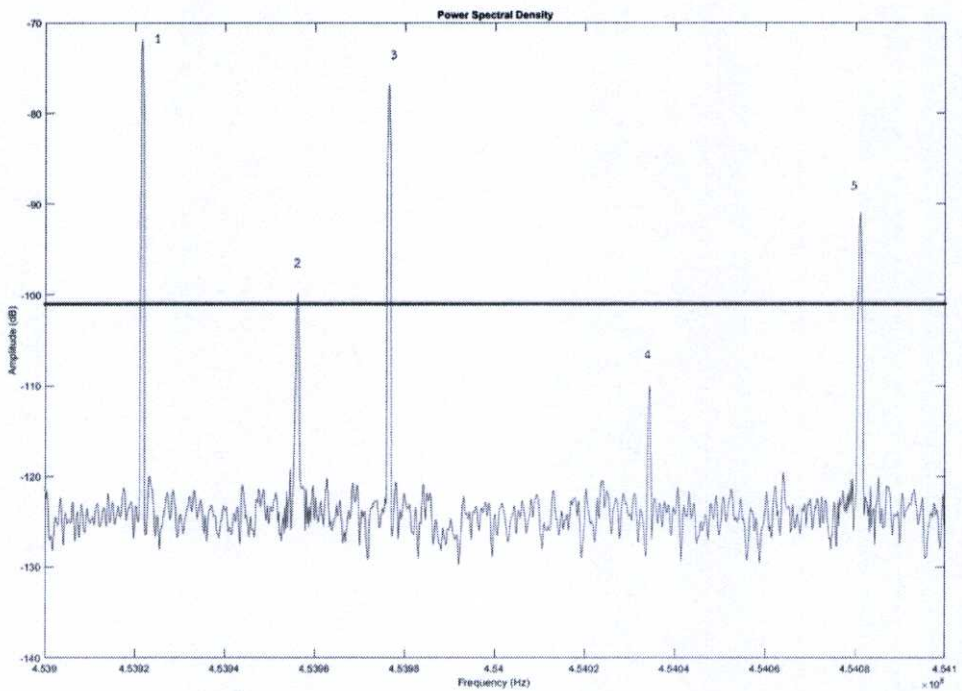
กว่าระดับค่าเกณฑ์เพียงเล็กน้อย ซึ่งวิทยุรู้คิดอาจตัดสินใจว่า ช่องสัญญาณนั้น ไม่มีผู้ใช้งาน ทำให้ใช้งานในช่องสัญญาณนั้น แล้วจะทำให้ส่งผลสัญญาณรบกวนไปสู่ผู้ใช้งานหลักได้ โมเดลนี้จึงทำการออกแบบให้วิทยุรู้คิดปรับปรุงอัตราขยายที่รับสัญญาณเข้ามาโดยยกทั้งสัญญาณและชั้นสัญญาณรบกวนขึ้น เพื่อให้ช่องสัญญาณที่อาจมีสัญญาณของผู้ใช้งานหลักปรากฏอยู่ และวิทยุรู้คิดไม่สามารถใช้งานช่องสัญญาณนั้นได้ ทำให้ประสิทธิภาพของการตรวจจับแบบตรวจจับพลังงานมีประสิทธิภาพในการตรวจจับมากขึ้น หรือน่าจะเป็นในการตรวจจับผิดน้อยลง (probability of miss detection) อธิบายได้จากรูปที่ 4.2 ถึง 4.4 ในรูปที่ 4.2 คือภาพที่ได้จากผู้ใช้วิทยุรู้คิด ได้สัญญาณมา และเมื่อสมมติว่าทำการหาค่าเกณฑ์ในการตรวจจับสเปกตรัม ได้เส้นสีแดง ตามรูปที่ 4.3 จะพบว่า ช่องสัญญาณที่หมายเลข 2 ไม่ผ่านเกณฑ์ในการตรวจจับ วิทยุรู้คิดอาจเข้าใจว่าช่องสัญญาณนั้นไม่มีผู้ใช้งานอยู่ จึงเข้าใช้งาน อาจทำให้เกิดสัญญาณรบกวนสู่ผู้ใช้งานหลักได้ โมเดลการแก้ปัญหาในวิจัยนี้ ได้แก้ไขปัญหามาโดยการเพิ่มอัตราขยายสัญญาณในภาครับ เพื่อยกทั้ง noise floor และสัญญาณขึ้น เพื่อให้ช่องสัญญาณที่อาจจะผ่านเกณฑ์ได้ดังรูปที่ 4.4



รูปที่ 4.2 ตัวอย่างสเปกตรัม

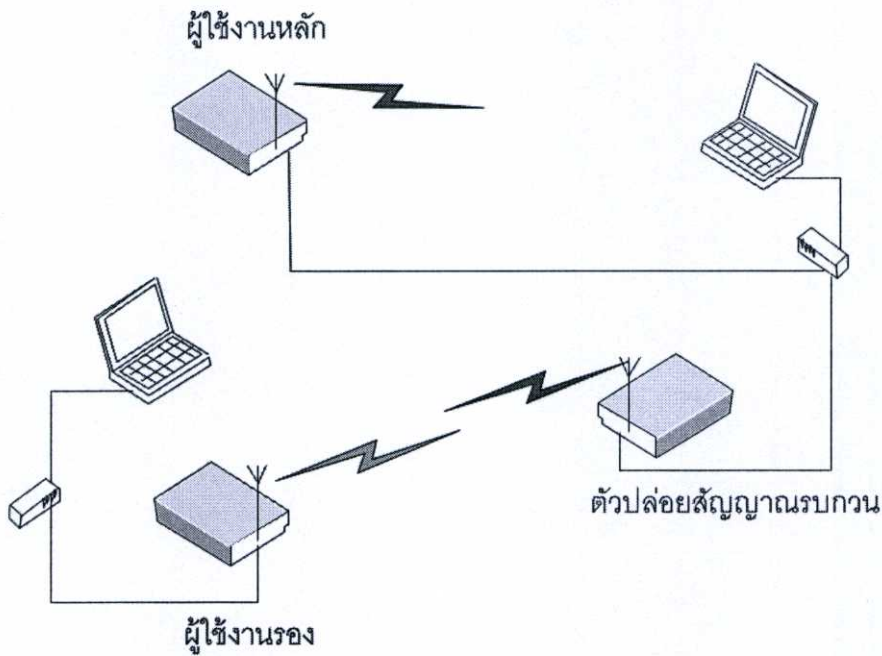


รูปที่ 4.3 ตัวอย่างค่าเกณฑ์ระดับ



รูปที่ 4.4 ตัวอย่างการยกสัญญาณและ noise floor

4.2 การติดตั้งในส่วนของฮาร์ดแวร์



รูปที่ 4.5 แบบจำลองโมเดลในการติดตั้งฮาร์ดแวร์ของระบบวิทยุรู้คิด

โมเดลที่ผู้วิจัยออกแบบระบบวิทยุรู้คิดในงานวิจัยนี้ แสดงดังรูปที่ 4.5 ใช้ USRP เชื่อมต่อกับคอมพิวเตอร์ด้วยสายแลนระดับกิกะบิต จำนวน 3 คู่ คือ ผู้ใช้งานหลัก, ผู้ใช้สำรอง, ตัวปล่อยสัญญาณรบกวน โดยแต่ละตัวมีหน้าที่ดังต่อไปนี้

ผู้ใช้งานหลัก คือ ผู้เป็นเจ้าของสเปกตรัม จะใช้งานสเปกตรัมได้ไม่จำกัด

ผู้สำรอง คือ ผู้ต้องฉกฉวยโอกาสในการเข้าใช้สเปกตรัมโดยไม่ส่งผลกระทบต่อผู้ใช้งานหลัก ด้วยการตรวจจับสเปกตรัมก่อนที่จะเข้าใช้ช่องสัญญาณนั้นๆ

ตัวปล่อยสัญญาณรบกวน เพื่อที่จะวัดประสิทธิภาพของระบบ ในสิ่งแวดล้อมที่มีสัญญาณรบกวนแตกต่างกันไป จึงมีตัวปล่อยสัญญาณรบกวนในระบบ

การติดต่อสื่อสารในรูปแบบคลื่นความถี่วิทยุระหว่าง USRP จะติดต่อกันแบบไร้สาย ผ่านเสาอากาศรอบทิศทาง (omnidirectional antenna) แต่การติดต่อระหว่าง USRP กับคอมพิวเตอร์จะเป็นการรับส่งพารามิเตอร์ การตั้งค่า และสตรีมในการประมวลผลผ่านทางสายกิกะบิตแลนเท่านั้น โดยมีรายละเอียดในการตั้งค่าเพื่อการติดต่อดังนี้

1. การติดต่อระหว่าง USRP และคอมพิวเตอร์ ใช้การติดต่อผ่านระบบ Ethernet

ตารางที่ 4.2 เลขที่อยู่ไอพีของ USRP และ คอมพิวเตอร์

อุปกรณ์	เลขที่อยู่ไอพี
USRP ที่เป็นผู้ใช้งานหลัก	192.168.1.1/24
USRP ที่เป็นผู้ใช้สำรอง	192.168.1.2/24
USRP ที่เป็นตัวสร้างสัญญาณรบกวน	192.168.1.3/24
เครื่องคอมพิวเตอร์ที่ 1	192.168.1.4/24
เครื่องคอมพิวเตอร์ที่ 2	192.168.1.5/24

2. การติดต่อระหว่าง USRP และ USRP

ตารางที่ 4.3 พารามิเตอร์ในการส่งของ USRP

พารามิเตอร์	ค่าที่ใช้
ความถี่	2.4GHz
เสาอากาศ	เสาอากาศรับรอบตัว (รุ่น VERT400)
ความแรงสัญญาณ	-90 ถึง -95 dB
รูปแบบในการกล้ำสัญญาณ	16QAM, QPSK, BPSK
Sampling rate	25 Ms/s ที่ 16 บิต I/Q

บทที่ 5

การทดลองและการเปรียบเทียบผลการทดลอง

ในบทนี้จะกล่าวถึงการทดลองและแสดงผลการเปรียบเทียบประสิทธิภาพของการตรวจจับสเปกตรัมแบบตรวจจับพลังงาน และมีการเรียนรู้แบบคิวนเป็นตัวช่วย ซึ่งทำการทดลองในสภาพแวดล้อมจริงด้วยเครื่องมือ USRP ใช้โมเดลที่กล่าวไปในบทที่ 4

การทดลองผู้ใช้งานหลักส่งสัญญาณโดยการกล้ำสัญญาณแบบ 16QAM, QPSK, BPSK บนความถี่ 2.4 GHz และมีสัญญาณรบกวนเกาส์เซียนสีขาวแบบบวกด้วยค่าเฉลี่ย (mean) เท่ากับ 0 และมีการจำลองการลดทอนของสัญญาณขึ้นอยู่กับสภาพแวดล้อมตามตาราง และเพื่อตรวจสอบประสิทธิภาพที่ดีของการตรวจจับสเปกตรัมที่นำเสนอ ได้ทดลองบนความน่าจะเป็นในการเกิด false alarm ด้วยความผิดพลาดตั้งแต่ 1 ใน 100 (10^{-2}), 1 ใน 10,000 (10^{-4}) และ 1 ใน 1,000,000 (10^{-6})

ตารางที่ 5.1 เลขชี้กำลังการสูญเสียในเส้นทางสำหรับสภาพแวดล้อมต่างๆ

สภาพแวดล้อม	เลขชี้กำลังการสูญเสียในเส้นทาง (σ)
เขตเมืองหนาแน่น	5
เขตชานเมือง	3
เขตชนบท	2.5

ในการทดลองผู้ใช้สำรองเปิดแบนด์วิดท์สำหรับตรวจจับทั้งหมด ประมาณ 200kHz แบ่งช่องสัญญาณในการตรวจจับเป็นช่องเล็กๆ จำนวน 64 ช่องตามขนาดของ FFT ซึ่งแต่ละช่องสัญญาณในการตรวจจับจะมีแบนด์วิดท์เท่ากับ $200\text{kHz}/64$ ประมาณ 3 kHz

การแสดงผลในรูปของกราฟ ROC คือค่าระหว่างความน่าจะเป็นในการตรวจจับถูกต้อง (Probability of detection - P_d) และความน่าจะเป็นในตรวจจับผิดแบบ false alarm (Probability of false alarm - P_{fa}) ซึ่ง P_d คือความน่าจะเป็นที่เป็นบวกจริง หมายถึง ตรวจจับได้ว่ามีสัญญาณของผู้ใช้งานหลักและมีการใช้งานของผู้ใช้งานหลักจริง ส่วน P_{fa} คือความน่าจะเป็นบวกเท็จ หมายความว่า ตรวจจับได้ว่ามีสัญญาณของผู้ใช้งานหลักแต่แท้ที่จริงไม่มีการใช้งานของผู้ใช้งานหลัก มีเพียงแต่สัญญาณรบกวนเท่านั้น

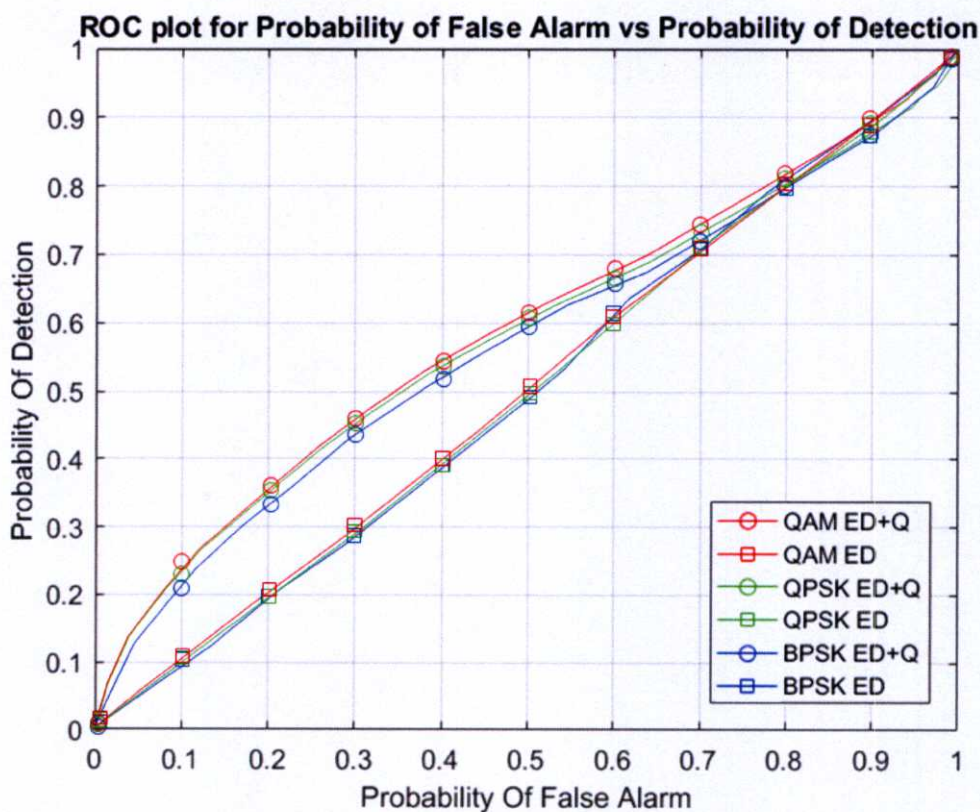
ในการทดลองได้ใช้พารามิเตอร์ดังตารางที่ 5.2 และได้แยกการทดลองตามสภาพแวดล้อมที่แตกต่างกันออกไปในหัวข้อที่ 5.2 ถึง 5.4

ตารางที่ 5.2 สรุปพารามิเตอร์ทั้งหมดที่ใช้ในการทดลอง

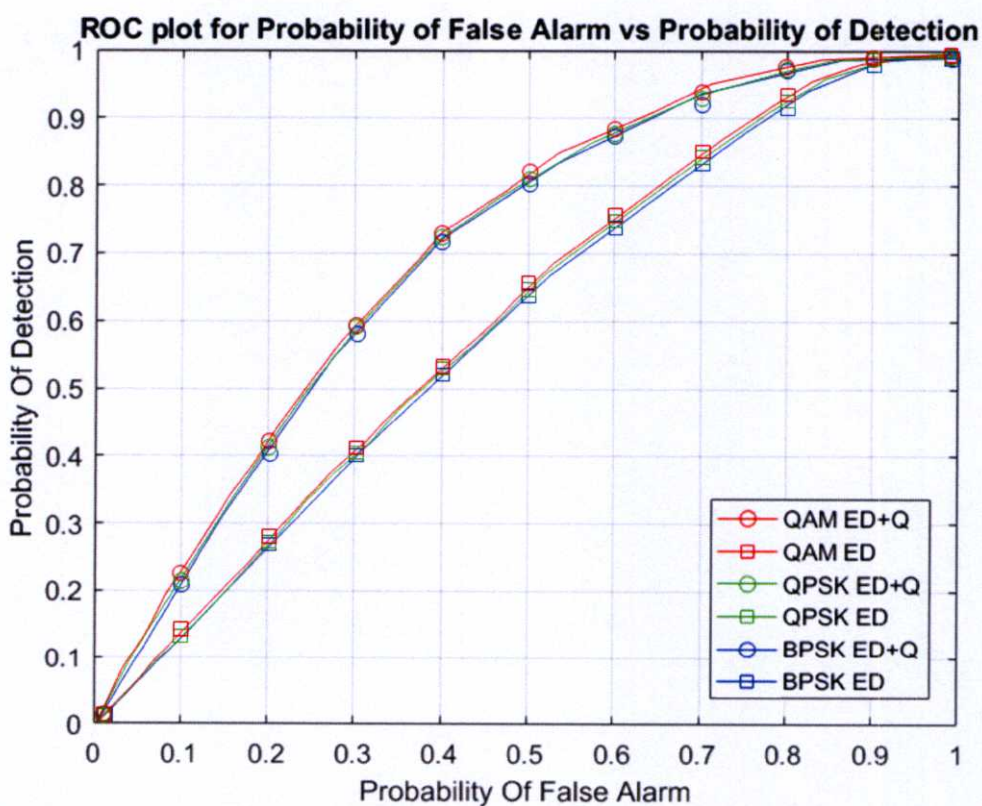
พารามิเตอร์	ค่าของพารามิเตอร์
Iteration ในการทำการเรียนรู้แบบคิว	50,000 รอบ
Agent	ผู้ใช้งานสำรอง
State	ช่องสัญญาณทั้งหมดในการตรวจจับ
Action	ช่องสัญญาณที่เป็นไปได้สำหรับ state นั้นๆ
Reward	1 = พบผู้ใช้งานหลัก 0 = ไม่พบผู้ใช้งานหลัก
Learning rate	0.25
Discount factor	0.2
รูปแบบการกล้ำสัญญาณ	16QAM, QPSK, BPSK
เลขชี้กำลังในการสูญเสียเส้นทาง	5 = เขตเมืองหนาแน่น 3 = เขตชานเมือง 2.5 = เขตชนบท
ความน่าจะเป็นในการเกิด false alarm	10^{-2} , 10^{-4} , 10^{-6}
กระบวนการในการทดลอง	ED+Q หมายถึงกระบวนการที่นำเสนอในบทที่ 4 ED หมายถึง Energy detection
ความถี่	2.4GHz
ความแรงสัญญาณ	-90 ถึง -95 dB
Sampling rate	25 Ms/s(mega sample per second) ที่ 16 บิต I/Q

5.1 ผลการทดลองเมื่อจำลองสภาพแวดล้อมเป็นเขตเมืองหนาแน่น

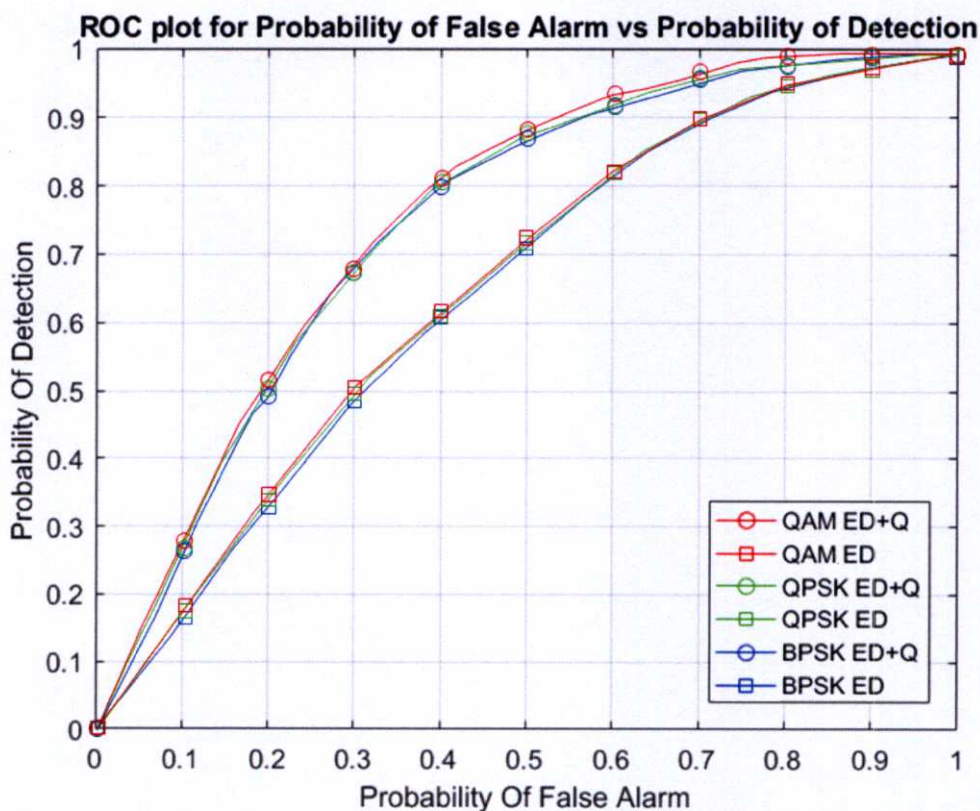
การทดลองที่จำลองสภาพแวดล้อมเป็นเขตเมืองหนาแน่น โดยใช้เลขชี้กำลังในการสูญเสียเส้นทางเป็น 5 ได้ผลการทดลองแสดงเป็น ROC curve ตามรูปที่ 5.1 ถึง 5.3 แสดงการเปรียบเทียบคุณภาพของการตรวจจับสเปกตรัมระหว่างกระบวนการที่นำเสนอ (ED+Q) กับกระบวนการตรวจจับสเปกตรัมแบบตรวจจับพลังงานแบบดั้งเดิม ดังจะเห็นได้ว่า ค่าความน่าจะเป็นในการตรวจจับถูกต้องของกระบวนการที่นำเสนออยู่ใกล้ 100% มากกว่า กระบวนการตรวจจับพลังงานแบบดั้งเดิม ซึ่งหมายถึงดีกว่า ถึงแม้ว่ากระบวนการที่นำเสนอยังมีค่าต่ำอยู่ เนื่องจากสัญญาณมีค่า SNR ต่ำมาก แต่ก็ให้ผลที่ดีกว่ากระบวนการดั้งเดิม



รูปที่ 5.1 ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตเมืองหนาแน่น, $SNR = 13.7, P_{fa} = 10^{-2}$



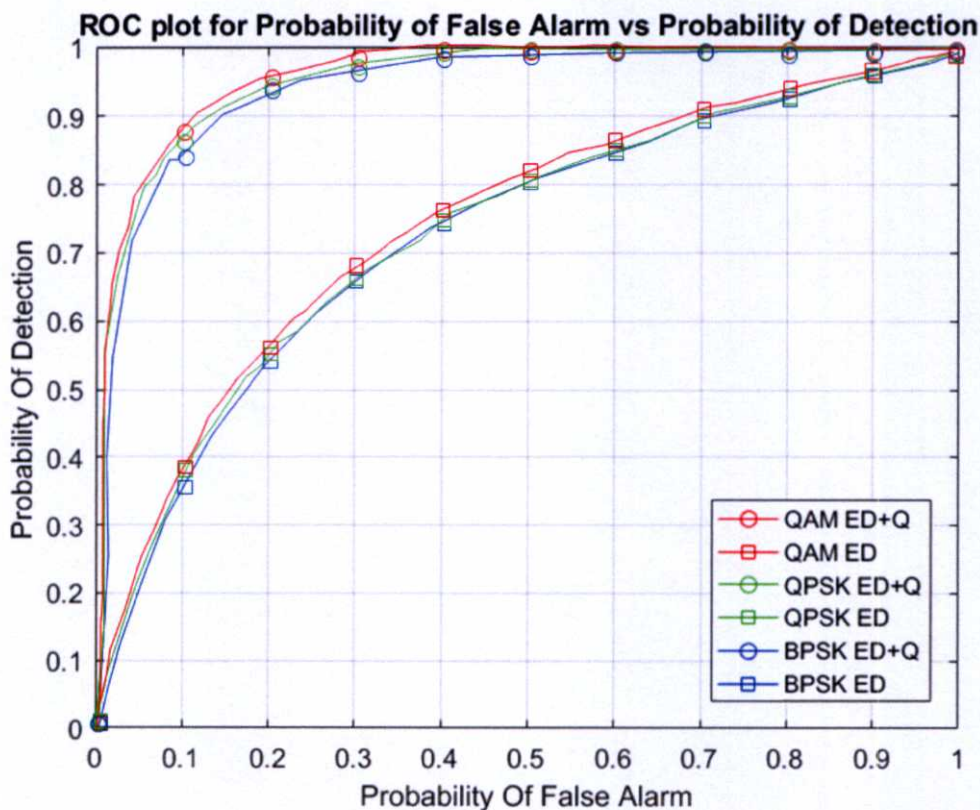
รูปที่ 5.2 ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตเมืองหนาแน่น, $SNR = 13.7, P_{fa} = 10^{-4}$



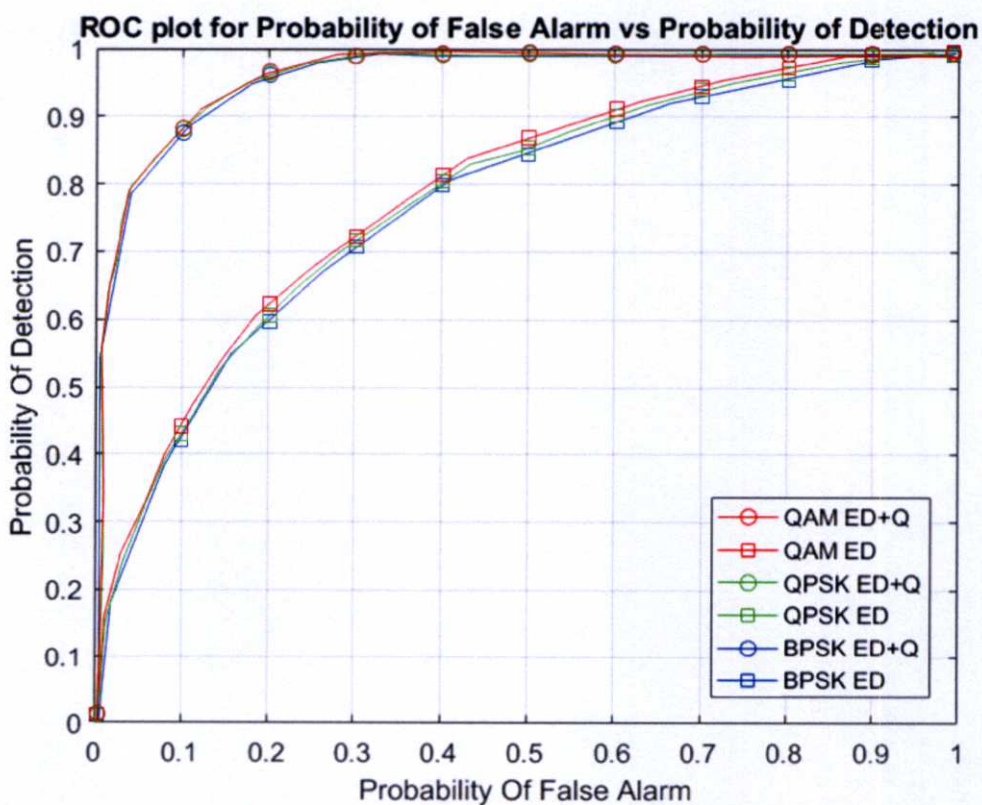
รูปที่ 5.3 ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตเมืองหนาแน่น, SNR = 13.7, $P_{fa} = 10^{-6}$

5.2 ผลการทดลองเมื่อจำลองสภาพแวดล้อมเป็นเขตชานเมือง

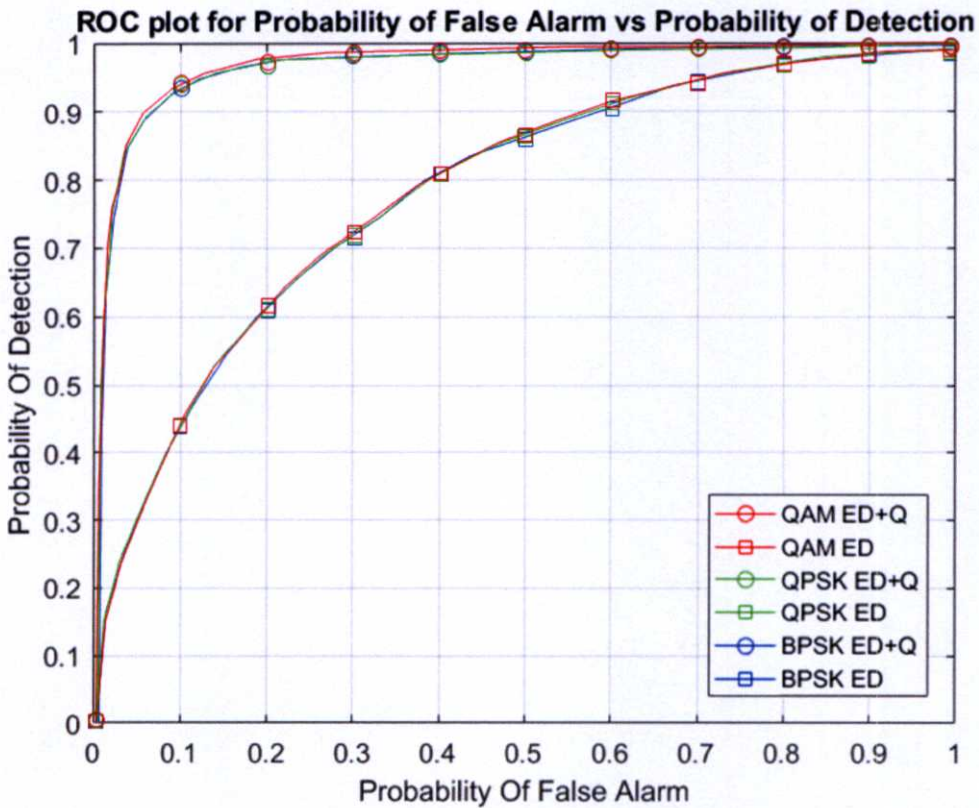
การทดลองที่จำลองสภาพแวดล้อมเป็นเขตชานเมือง โดยการใช้เลขชี้กำลังในการสูญเสียเส้นทางเป็น 3 ได้ผลการทดลองแสดงเป็น ROC curve ตามรูปที่ 5.4 ถึง 5.6 ในการจำลองในเขตชานเมือง ซึ่งทำให้สัญญาณที่ผ่านเข้ามาในระบบมีค่า SNR ที่สูงขึ้น ทำให้ค่าความน่าจะเป็นในการตรวจจับถูกต้องของทั้งสองกระบวนการเข้าใกล้ 100% มากขึ้น แต่กลุ่มเส้นที่แสดงถึง ROC curve ของกระบวนการที่นำเสนอ ยังคงเข้าใกล้ 100% มากกว่าของกระบวนการแบบดั้งเดิม ทำให้สรุปได้ว่ากระบวนการที่นำเสนอมีคุณภาพที่ดีกว่ากระบวนการแบบดั้งเดิม



รูปที่ 5.4 ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตชานเมือง, $SNR = 24.5, P_{fa} = 10^{-2}$



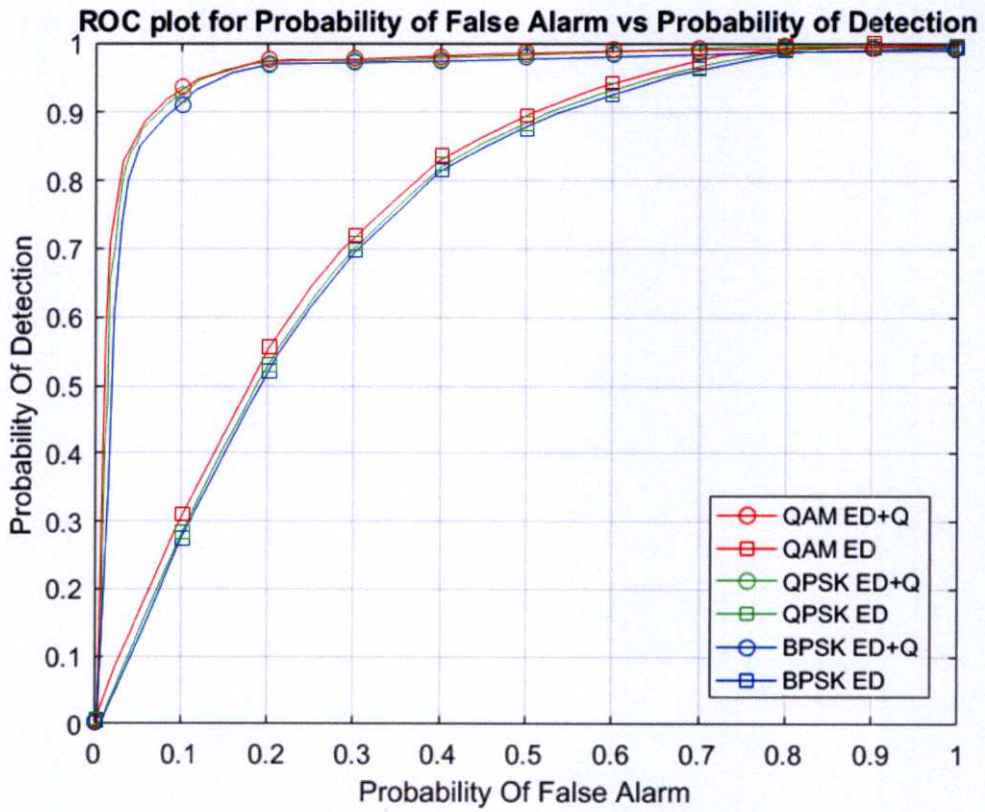
รูปที่ 5.5 ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตชานเมือง, $SNR = 24.5, P_{fa} = 10^{-4}$



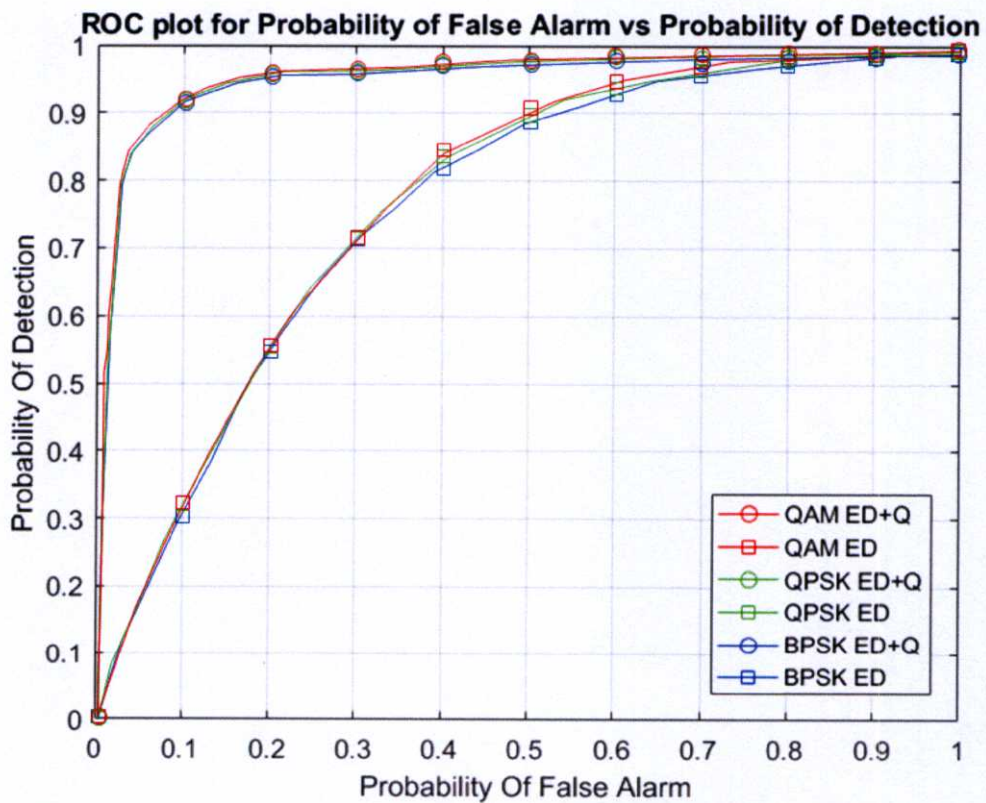
รูปที่ 5.6 ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตชานเมือง, SNR = 24.5, $P_{fa} = 10^{-6}$

5.3 ผลการทดลองเมื่อจำลองสภาพแวดล้อมเป็นเขตชนบท

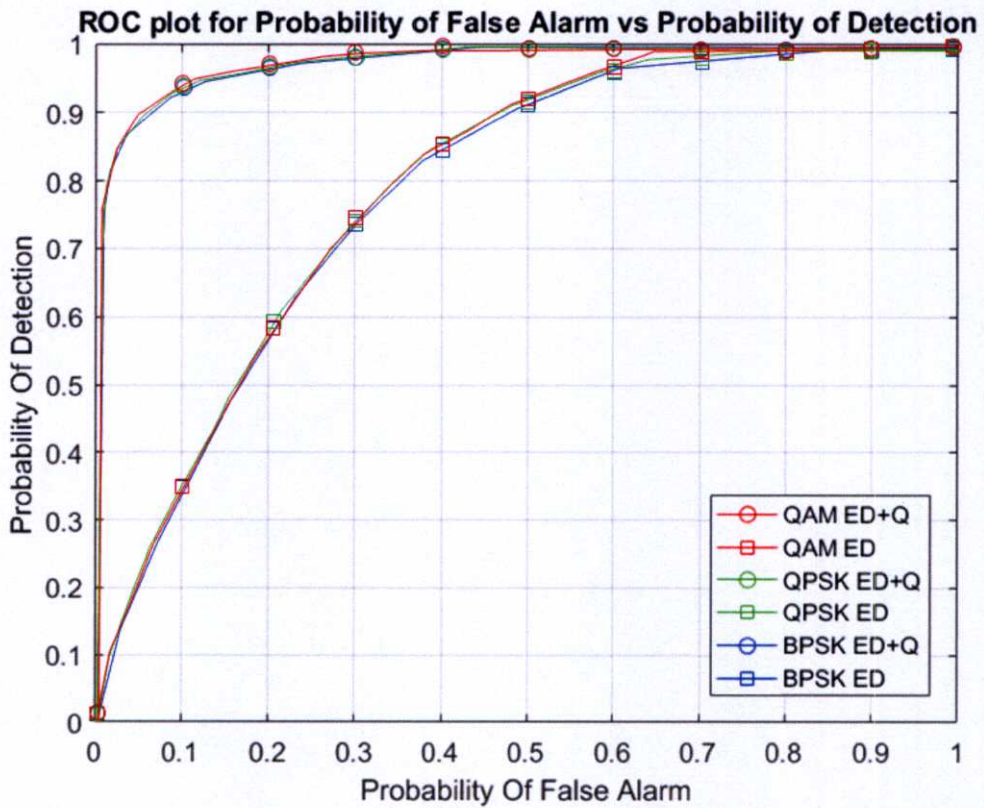
การทดลองที่จำลองสภาพแวดล้อมเป็นเขตชนบท โดยการใช้เลขชี้กำลังในการสูญเสียเส้นทางเป็น 2.5 ได้ผลการทดลองแสดงเป็น ROC curve ตามรูปที่ 5.7 ถึง 5.9 ในการจำลองนี้ เป็นการจำลองที่ทำให้เห็นคุณภาพที่ดีที่สุดของการกระบวนกรที่นำเสนอ เนื่องจากค่าความน่าจะเป็นในการตรวจจับถูกเข้าใกล้ 100% มากที่สุด ทำให้สรุปได้ว่า SNR มีผลต่อการตรวจจับสเปกตรัมแบบตรวจจับพลังงานเป็นอย่างมาก แต่ถ้ามีกระบวนการในการช่วยในการตรวจจับที่ดี ก็ทำให้การตรวจจับสเปกตรัมมีความน่าจะเป็นในการตรวจจับถูกต้องมากยิ่งขึ้น



รูปที่ 5.7 ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตชนบท, SNR = 30.7, $P_{fa} = 10^{-2}$



รูปที่ 5.8 ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตชนบท, SNR = 30.7, $P_{fa} = 10^{-4}$



รูปที่ 5.9 ROC curve ในการจำลองสภาพแวดล้อมเป็นเขตชนบท, SNR = 30.7, $P_{fa} = 10^{-6}$

5.4 เปรียบเทียบผลการทดลองจากกราฟ ROC curve

จากกราฟ ROC curve ที่ได้แสดงผลการทดลองในสภาพสิ่งแวดล้อมต่างๆ ในหัวข้อที่ 5.1 ถึง 5.3 จะเห็นได้ว่า การจำลองในสภาพแวดล้อม ที่มีประสิทธิภาพสูง จะมีจุดตัด (cut-off point) ที่ค่า probability of false alarm (ในแกน x) ต่ำ และค่า probability of detection (ในแกน y) สูง ส่งผลให้ ROC curve นั้น เข้าชิดมุมซ้ายบนมากที่สุด นั่นคือ ผลของการตรวจจับสเปกตรัมแบบที่วิทยานิพนธ์นี้นำเสนอ (ED+Q) บนการจำลองในสภาพแวดล้อมในเขตชนบทที่ดีที่สุด รองลงมาคือสภาพแวดล้อมในเขตชานเมือง และสุดท้ายคือสภาพแวดล้อมในเขตเมืองหนาแน่น

สรุปผลการทดลองกระบวนการตรวจจับสเปกตรัมที่วิทยานิพนธ์นี้นำเสนอ เมื่อเปรียบเทียบกับ การตรวจจับสเปกตรัมแบบการตรวจพลังงานแบบดั้งเดิม ให้ผลว่าประสิทธิภาพในการตรวจจับได้ ดีกว่ามากในสภาพแวดล้อมเดียวกัน และถึงแม้สัญญาณจากผู้ใช้งานหลักเป็นแบบใด ก็ไม่ได้ทำให้ผลของการตรวจจับแตกต่างกันเพียงใด

บทที่ 6

สรุปผลการวิจัยและข้อเสนอแนะ

ระบบวิทยุรู้คิดเป็นการปรับปรุงระบบวิทยุธรรมดา ให้มีประสิทธิภาพในการจัดสรรทรัพยากร อันมีจำกัดอย่างเช่นคลื่นความถี่วิทยุ ซึ่งระบบวิทยุรู้คิดมีกระบวนการมากมาย ที่จะทำให้วิทยุธรรมดา เป็นวิทยุที่ชาญฉลาด เช่น การตรวจจับสเปกตรัม ว่ามีการใช้งานของผู้ใช้งานหลักหรือไม่ เพื่อที่จะได้ ใช้งานคลื่นความถี่นั้น โดยไม่ส่งผลกระทบต่อผู้ใช้งานหลัก ซึ่งในวิทยานิพนธ์นี้ได้ทำการศึกษาเรื่องการ ตรวจจับสเปกตรัมด้วยการตรวจจับพลังงาน ซึ่งก่อนหน้านี้ก็มีการศึกษาพัฒนาเรื่องนี้มาก่อนหน้าแล้ว แต่ที่วิทยานิพนธ์นี้เพิ่มขึ้นมาคือการนำการเรียนรู้แบบเสริมแรงเข้ามาช่วยเพื่อให้ระบบวิทยุรู้คิด มีความจำเปรียบเสมือนสมองของมนุษย์ โดยที่ไม่ได้ตรวจจับเพียงอย่างเดียว แต่ทำการเก็บสถานะสิ่ง ต่างๆ ที่เป็นพารามิเตอร์ในการรับส่งวิทยุ เพื่อทำการเรียนรู้ในสิ่งแวดล้อมที่ไม่อาจคาดเดาได้อย่างเช่น สิ่งแวดล้อมในการส่งคลื่นความถี่วิทยุ

อีกทั้งวิทยานิพนธ์นี้ ได้ทำการทดลองระบบวิทยุรู้คิดบนสภาพสิ่งแวดล้อมจริง ด้วยเครื่อง อุปกรณ์ที่เรียกว่า SDR เป็นวิทยุรุ่นใหม่ที่เพิ่มความสามารถให้กับการใช้งานที่มีความต้องการ หลากหลายในปัจจุบัน ในการทดลองได้ทำการเปรียบเทียบบนการจำลองที่แสดงถึงสถานะ สภาพแวดล้อมต่างๆ กันในประเทศไทย คือ เขตเมืองหนาแน่น เขตชานเมืองและเขตชนบท และด้วย การส่งสัญญาณจากผู้ใช้หลักที่หลากหลายการกล้าสัญญาณ ได้แก่ 16QAM, QPSK และ BPSK ซึ่ง เป็นการกล้าสัญญาณของสัญญาณดิจิทัลที่นิยมใช้ในปัจจุบัน และยังทำการทดลองบนย่านความถี่ โทรศัพท เพื่อให้เห็นว่า โมเดลนี้สามารถใช้งานได้กับสัญญาณรูปแบบต่างๆ ได้หลากหลาย แต่ ประสิทธิภาพการตรวจจับจะแยกลงเมื่อ SNR มีค่าน้อยลง

เอกสารอ้างอิง

- [1] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, 1999.
- [2] D. Ariananda, M. Lakshmanan, and H. Nikookar, "A survey on spectrum sensing techniques for cognitive radio," *2009 Second International Workshop on Cognitive Radio and Advanced Spectrum Management*, 2009.
- [3] "USRP," - *National Instruments*. [Online]. Available: <http://www.ni.com/sdr/usrp>. [Accessed: 13-Sep-2016].
- [4] D. Cabric, S. Mishra, and R. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004*.
- [5] T. Jiang, D. Grace, and Y. Liu, "Cognitive radio spectrum sharing schemes with reduced spectrum sensing requirements," *IET Seminar on Cognitive Radio and Software Defined Radio: Technologies and Techniques*, 2008.
- [6] J. Lunden, S. R. Kulkarni, V. Koivunen, and H. V. Poor, "Multiagent Reinforcement Learning Based Spectrum Sensing Policies for Cognitive Radio Networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 5, pp. 858–868, 2013.
- [7] U. Berthold, F. Fu, M. van der Schaar, and F. K. Jondral, "Detection of spectral resources in cognitive radios using reinforcement learning," in *Proc. 3rd IEEE Symp. New Frontiers Dynamic Spectrum Access Netw. (DySPAN '08)*, Chicago, IL, USA, Oct. 14–17, 2008, pp. 1–5.
- [8] M. Di Felice, K. R. Chowdhury, W. Meleis, and L. Bononi, "To sense or to transmit: A learning-based spectrum management scheme for cognitive radio mesh networks," in *Proc. 5th IEEE Workshop Wireless Mesh Netw. (WIMESH 2010)*, Boston, MA, USA, Jun. 21, 2010.
- [9] J. Oksanen, J. Lundén, and V. Koivunen, "Reinforcement learning based sensing policy optimization for energy efficient cognitive radio networks," *Neurocomput.—Spec. Iss. Mach. Learn. for Signal Process.*, vol. 80, pp. 102–110, Mar. 2012.

- [10] Z. Chen and R. C. Qiu, "Cooperative spectrum sensing using Q-learning with experimental validation," 2011 Proceedings of IEEE Southeastcon, 2011.
- [11] J. Lunden, S. R. Kulkarni, V. Koivunen, and H. V. Poor, "Multiagent Reinforcement Learning Based Spectrum Sensing Policies for Cognitive Radio Networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 5, pp. 858–868, 2013.
- [12] Uyanik, Gulnur Selda, Ozge Cepheli, Gunes Karabulut Kurt, and Sema Oktug. "Implementation and performance evaluation of dynamic spectrum access using software defined radios." *2013 First International Black Sea Conference on Communications and Networking (BlackSeaCom) (2013)*.
- [13] L. R. Faganello, R. Kunst, C. B. Both, L. Z. Granville, and J. Rochol, "Improving reinforcement learning algorithms for dynamic spectrum allocation in cognitive sensor networks," 2013 IEEE Wireless Communications and Networking Conference (WCNC), 2013.
- [14] M. Li, Y. Xu, and J. Hu, "A Q-Learning based sensing task selection scheme for cognitive radio networks," 2009 International Conference on Wireless Communications & Signal Processing, 2009.

ภาคผนวก
ผลงานที่ได้รับการตีพิมพ์



The 19th International Conference on Advanced Communications Technology

“Opening New Era of Smart Society”

<http://www.icact.org>

Phoenix Park, Pyeongchang
Korea (South)
Feb. 19 ~ 22, 2017



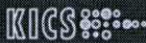
PROCEEDING & JOURNAL

IEEE Catalog Number: CFP17561-ART
ISBN: 978-89-968650-9-4
ISSN: 1738-9445

Organizers



Sponsors



Open Standards and
Internet Association



Korea
Institute of
Information
Security &
Cryptography

The Implementation of Spectrum Sensing and Spectrum Allocation on Cognitive Radio

Pitcha Rungsawang, Amnach Khawne

Department of Computer Engineering, Faculty of Engineering,
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
pitcha.rsw@gmail.com, amnach.kh@kmitl.ac.th

Abstract— Cognitive Radio is a system designed to use radio spectrum in the most efficient way possible—as demands for radio spectrum have been heavily increased nowadays. The system works by allowing unlicensed users to use radio channel on the licensed users' spectrum without interfering with the licensed users. The system needs spectrum sensing to verify the usage of each specific spectrum and in ever-shifting environment, Reinforcement Learning is also used for establishing spectrum allocation as well.

Keywords— Cognitive Radio, Spectrum Sensing, Spectrum Allocation, Reinforcement Learning, USRP, LabView

I. Introduction

Radio Spectrum is a primary form of wireless communication. In order to reserve radio frequency, the formal request need to be accepted by a national organization which oversee the management of radio frequency. However, since there are massive usage from countless users and active equipment, the management and distribution of radio frequency are usually insufficient. So the new method of radio frequency management is needed to offer more efficiency and flexibility: Cognitive Radio (CR) [1].

Cognitive Radio allows the unlicensed secondary users (SU) to use the frequency without affecting the licensed Primary Users (PU) by granting usage to the SU while the frequency is not in use by the PU—using the spectrum sensing [2] method to detect usage and look for spectrum hole in the environment. Nevertheless, the difficulty of this method is detecting the user's signal within noise-heavy environment. Speed and accuracy are also additional concerns.

Reinforcement Learning is the most suitable method for the ever-shifting environment; using mathematical calculation to find the best possible spot. Even after arriving at the answer, the system will still take surrounding environment into calculation to find the best long-term answer.

The goal of this experiment is to learn and develop the capability of Cognitive Radio by using Reinforcement Learning. The experiment is conducted on the vacant radio frequency slot, using Universal Software Radio Peripheral (USRP) [3] equipment to study radio frequency communication, along with LabView as a program to generate

environment simulation and assemble .m files to use in MATLAB.

II. Literature Review

Cognitive Radio[1] was introduced by Joseph Mitola and Gerald Q. Maguire in 1999, and keep on developing since then: using Machine-Learning[5] and artificial intelligence[4] to improve the system's performance. Reinforcement Learning is also the subsection of Machine-Learning, suitable for learning about environment surrounding the agent. The process is based on trial and error to finding the best possible state; giving reward when choosing good states and punishment when choosing bad ones[6]. All of those rewards and punishments will then be send back to the agent in order to estimate a pattern which can score the highest—finding the balance between exploration and exploitation. The advantage of this algorithm is even if the best answer is found; the system will continue to learn from available data to try to generate even better answer—causing the algorithm to improve and balance itself out[7].

The aforementioned reward-punishment process is called Weighting Factors[8]. The values of the whole factors are shown in TABLE I. Based on the degree of success, either a reward or a punishment is assigned to the weight of the used spectrum.

TABLE I. Weighting Factors [8]

Patterns	Before		After	
	Reward	Punishment	Reward	Punishment
Light	1	1	1	-1
Mild	1	0	1	0
Heavy	1	0.5	1	0

The main difference between these patterns is the values assigned to punishment factors. Under normal circumstances, the absolute values of the reward value and the punishment value are equal—hence the 'Light Penalty' pattern. If the attempt for communication fails, however, the weight is directly reduced to zero. Therefore we call it the 'Mild Penalty'. If the user picks up a channel randomly and keeps

using it as long as the quality of communication in that channel is above the requirement, weights will be reduced by 50% immediately. We refer to this pattern as the 'Heavy Penalty' pattern.

There are many past studies dedicated to improving the efficiency of Cognitive Radio, including communication simulation on MATLAB and communication experiment in real time. M. A. K. Abdulsattar and Z. A. Hussein had conducted Energy Detector experiment in real time[9], using TMS320C6713 as digital processor. N. Sharma, D. B. Rawat[10]. also conducted Dynamic Spectrum Access in Cognitive Radio using USRP and LabView.

III. Spectrum Sensing

Spectrum Sensing is mainly used for detecting the usage of primary user on specific frequency. There are many different methods available for the process: Energy Detection, Cyclostationary Feature Detection, Matched Filter Detection, etc. This experiment, however, focuses on Energy Detection. The main drawback of this process is its low efficiency, but it compensates in simplicity; it can be conducted even without prior knowledge of incoming signals. This is the reason for many studies conducted to find a way to improve its efficiency.

The process of detection is as followed:

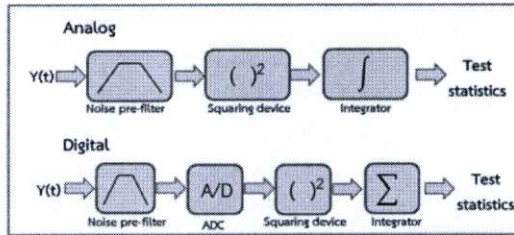


Figure 1. Process of Energy Detection

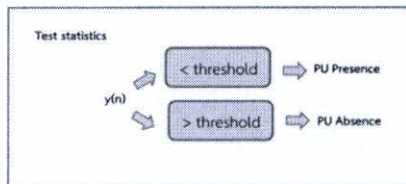


Figure 2. Test statistics

When the system received the input signal $y(t)$, it will generate the sampling by randomize all signals in and send it along as $y(n)$. Then finding the energy by times the amplitude by two and divide by the summary of all amplitudes. The resulting energy will then be compared to the threshold from [11]. If the energy pass the threshold, it means there is a user currently occupied the frequency. If not, then the frequency is vacant.

Threshold used for this experiment is as followed:

$$\lambda_k^* = \sqrt{2} \operatorname{erfc}^{-1}(2P_{fk}), k \neq i$$

There are two hypotheses following the experiment:

$$T_{ED} = \begin{cases} \text{Normal}(\sigma_w^2, \frac{2\sigma_w^4}{N_{ED}}) & ; H_0 \\ \text{Normal}(\sigma_w^2 + \sigma_s^2, \frac{2(\sigma_w^2 + \sigma_s^2)^2}{N_{ED}}) & ; H_1 \end{cases}$$

Normal is Normal Distribution, σ_s^2 is Signal Variation and σ_w^2 is Noise Variation

The efficiency of Spectrum Sensing depends on Probability of Detection (P_d) and Probability of False Alarm (P_f) which can be estimated using these equations:

$$P_d = Q\left(\frac{\lambda_{ED} - (\sigma_w^2 + \sigma_s^2)}{\sqrt{\frac{2}{N_{ED}}(\sigma_w^2 + \sigma_s^2)}}\right)$$

$$P_f = Q\left(\frac{\lambda_{ED} - \sigma_w^2}{\sqrt{\frac{2}{N_{ED}}\sigma_w^2}}\right)$$

λ_{ED} is the Threshold of ED and $Q(\cdot)$ is the Gaussian distribution

IV. Experiment on USRP and LabView

We conducted this experiment by using NI USRP-2920, which is hardware to be used with computer to imitate the signal process of radio. Alongside computers with MATLAB and Labview installed.

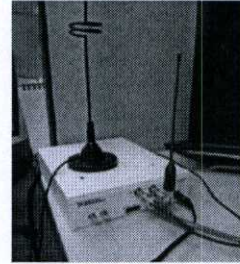


Figure 3. USRP 2920

Labview is a program which includes many virtual instruments for research and development alongside USRP or creates simulation by itself. Like programming, users can create and design their own simulation with plenty of tools at their disposal: operator, gate, if condition, while loop, for loop, and virtual instruments. It can also be used to analyze incoming signal, simulate signal from different modulations, and visualize embedded control systems.

We used LabView with USRP to send both signal and noise in real time for USRP to do spectrum sensing with energy detection. The result was unsatisfactory, since energy detection is inefficient within a noise-heavy environment with

low SNR—according to [12]. So we used Q-learning to analyze incoming signal and adjust gain on the radio by using LabView to control USRP.

Theoretically, Q-learning is a trial and error process. When SNR is low, we would adjust gain until we receive answer or acceptable state, but Q-learning will not stop there. It will continue to explore; using randomization to find better answer—making it suitable for spectrum sensing in ever shifting and unpredictable environment, as in the following formula:

$$Q(s, a) = r_{(s,a)} + \gamma \arg \max(Q(s', a'))$$

$Q(s, a)$ is a matrix function following:

$$Q(s, a) = \text{zeros}(\text{length}(\text{state}), \text{length}(\text{action}))$$

r is intermediate reward and γ is discount factor of learning and $Q(s', a')$ is a matrix function following:

$$Q(s', a') = \text{zeros}(\text{len}(\text{new state}), \text{len}(\text{new action}))$$

We used weight adjustment setting from table 1 in chapter 2. Learning rate, discount factor, and exploration probability were used with the default setting of 0.5.

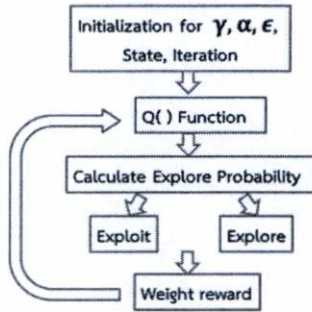


Figure 4. Q-learning Flowchart

The communication between USRP and experimental computer was conducted via Ethernet. We set the IP address of USRP as 192.168.1.2 255.255.255.0 and IP address of computer as 192.168.1.3 255.255.255.0.

We inputted the following parameter into USRP: specific frequency, USRP's IP address, Antenna, Signal Amplification Rate, and Sampling Rate. From there, we generated the Signal to Noise Ratio (SNR) and calculated with the Sampling signal using energy detection, estimated the result with P_d and P_f , and put the findings into graphs.

Below are block diagrams from LabView:

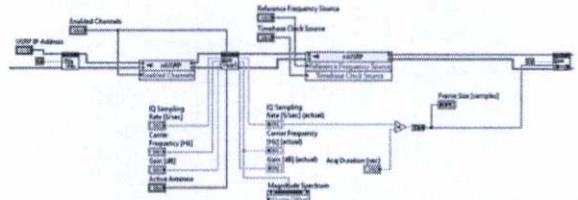


Figure 5. USRP Open Rx Session on Configure Functions

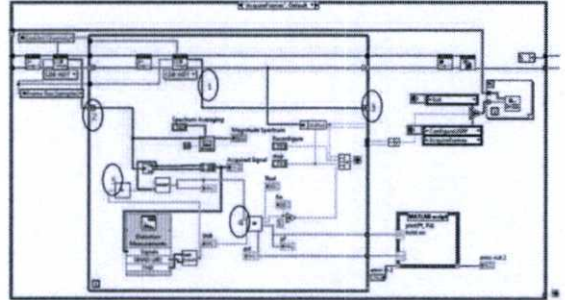


Figure 6. USRP read, write, process signal, energy detector and using SNR to weight for Q-learning

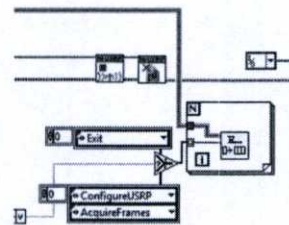


Figure 7. USRP Close Session and abort

Opening Rx session of USRP is showed in figure 5. We configured IP address of USRP, antenna number, sampling rate, etc. It's initializing for radio first. Then we obtain discrete signal. We calculated SNR to energy detection before use it into reward function of Q-learning that we designed in figure 6. Figure 7 we terminated session of USRP at end of program.

From figure 6, point 1, 2, and 3 are complex waveform ($y(n)$) data lines, which we would use in a formula. The process is meant to create waveform of time domain and frequency domain, as well as using $y(n)$ to find SNR in energy detection. Point 4 is a MATLAB script which used SNR and $y(n)$ to calculate P_d and P_f from the theory in chapter 2. Point 5 is the result of using SNR from each iteration to calculate weight and adjust gain of the radio. After that, we plot a ROC graph between P_f and P_d .

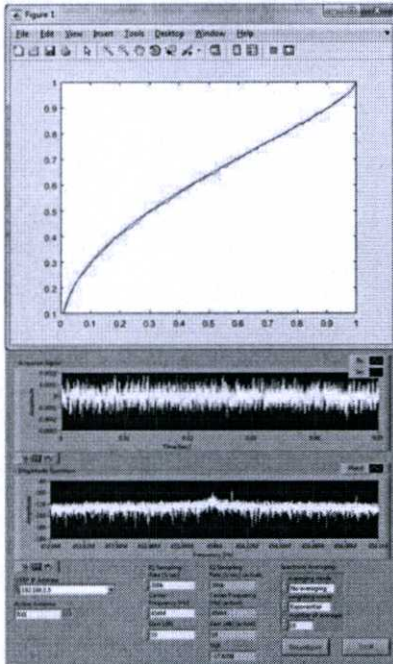


Figure 8. ROC curve between Pd, Pf at SNR = 17.8208 dB

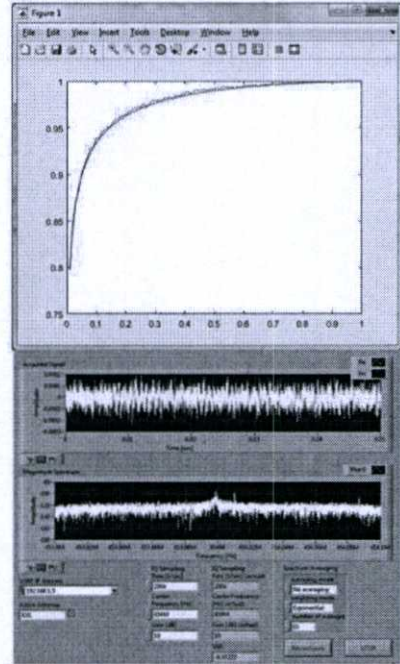


Figure 10. ROC curve between Pd, Pf at SNR = 9.37222 dB

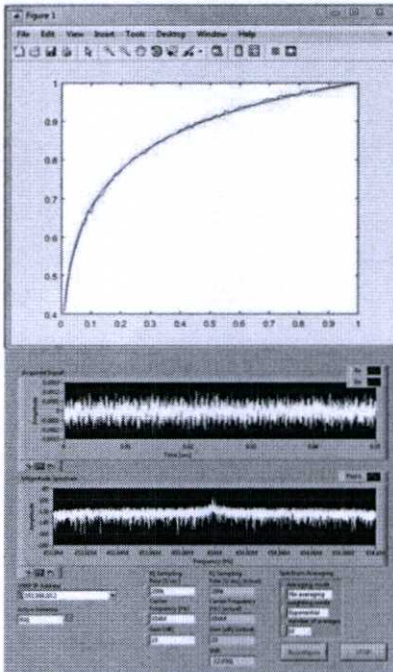


Figure 9. ROC curve between Pd, Pf at SNR = 12.0501 dB

Figure 8 is a result prior to q-learning. Figure 9 and 10 are results after q-learning.

From the result, we can concluded that the outcome of energy detection depend heavily on SNR. When Q-learning is included in the radio, signal detection will become more efficient.

V. Experimental result

The efficiency of spectrum sensing can be compared to Receiver Operator Characteristic (ROC) Curve, generating relation graphs between X as a true positive rate or Sensitivity and Y as false positive rate or 1- Specificity. Therefore, true positive rate of this experiment is probability of detection and false positive rate is probability of false alarm. The best result is the spot on curve farthest from diagonal line. Figure 8 and 9 showed the best result that we had experiment and limit environment.

VI. Conclusion

In this experiment, we conduct spectrum sensing using energy detection method and reward feedback with SNR to weight parameter on the USRP in real time. The efficiency of detection depends largely on the level of noise that real environment. In the future, reinforcement learning will be used alongside energy detection to improve efficiency of Cognitive Radio system for further experiment on USRP.

References

- [1] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, 1999.
- [2] D. Ariananda, M. Lakshmanan, and H. Nikoogar, "A survey on spectrum sensing techniques for cognitive radio," *2009 Second International Workshop on Cognitive Radio and Advanced Spectrum Management*, 2009.
- [3] "USRP," - *National Instruments*. [Online]. Available: <http://www.ni.com/sdr/usrp>. [Accessed: 13-Sep-2016].
- [4] N. Abbas, Y. Nasser, and K. E. Ahmad, "Recent advances on artificial intelligence and learning techniques in cognitive radio networks," *J Wireless Com Network EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, 2015.
- [5] M. Bkassiny, Y. Li, and S. K. Jayaweera, "A Survey on Machine-Learning Techniques in Cognitive Radios," *IEEE Communications Surveys & Tutorials IEEE Commun. Surv. Tutorials*, vol. 15, no. 3, pp. 1136–1159, 2013.
- [6] R. S. Sutton and A. G. Barto, *Reinforcement learning: an introduction*. Cambridge, MA: MIT Press, 1998.
- [7] T. Jiang, D. Grace, and Y. Liu, "Two-stage reinforcement-learning-based cognitive radio with exploration control," *IET Communications*, vol. 5, no. 5, pp. 644–651, 2011.
- [8] T. Jiang, D. Grace, and Y. Liu, "Performance of cognitive radio reinforcement spectrum sharing using different weighting factors," *2008 Third International Conference on Communications and Networking in China*, 2008.
- [9] M. A. K. Abdulsattar and Z. A. Hussein, "Energy Detector with Baseband Sampling for Cognitive Radio: Real-Time Implementation," *Wireless Engineering and Technology WET*, vol. 03, no. 04, pp. 229–239, 2012.
- [10] N. Sharma, D. B. Rawat, B. B. Bista, and S. Shetty, "A Testbed Using USRP(TM) and LabView(R) for Dynamic Spectrum Access in Cognitive Radio Networks," *2015 IEEE 29th International Conference on Advanced Information Networking and Applications*, 2015.
- [11] Y.-C. Liang, Y. Zeng, E. Peh, and A. T. Hoang, "Sensing-Throughput Tradeoff for Cognitive Radio Networks," *2007 IEEE International Conference on Communications*, 2007.
- [12] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.



Pitcha RUNGSAWANG was born in Petchaburi, Thailand on May 30, 1992. I graduated with a B.E. degree in computer engineering from King Mongkut's Institute of Technology Ladkrabang Bangkok, Thailand, in 2014. I'm interested in cognitive radio based on machine learning, signal intelligence, and wireless communication.



Annach KHAWNE was born in Suphumburi, Thailand, on August 6, 1974. He received the B. E. degree in telecommunication engineering and the M. E. degree in electrical engineering both from King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand, in 1998 and 2002, respectively. He received the Ph.D degree in Interdisciplinary Studies from Tokai University, Japan, in 2011. He is currently a lecturer in faculty of engineering, King Mongkut's Institute of Technology Ladkrabang. His research interests include image watermarking, image compression, and network security.

ประวัติผู้เขียน

ชื่อ-นามสกุล

นางสาวพิชชา รุ่งสว่าง

วัน เดือน ปีเกิด

30 พฤษภาคม 2535 ที่จังหวัดเพชรบุรี

ที่อยู่

109/69 ถนนพหลโยธิน แขวงอนุสาวรีย์ เขตบางเขน จังหวัด
กรุงเทพฯ 10220

ประวัติการศึกษา

พ.ศ. 2557

วิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง