บทที่ 4
บทสรุป

4.1 บทสรุป

ในรายงานวิจัยฉบับนี้นำเสนอกระบวนการเรียนรู้แบบรีอินฟอร์สเมนท์ภายใต้การเฝ้าสังเกตสิ่งแวดล้อมในกระบวนการตัดสินใจแบบมาร์คอฟภายใต้สภาวการณ์ที่สังเกตได้บางส่วน (partially observable Markov decision process หรือ POMDP) ด้วยวิธีการที่เรียกว่า ออนโพลิซี มอนติ คาร์โล (On-policy Monte Carlo หรือ ONMC) เพื่อปรับปรุงวิธีหาเส้นทางที่รองรับคุณภาพในการบริการสำหรับเครือข่ายเคลื่อนที่แบบแอดฮอคที่มีอยู่เดิมนั่นคือ วิธีตรวจสอบด้วยตั๋ว (Ticket-based probing หรือ TBP) โดยมีวัตถุประสงค์เพื่อหาสมดุลแลกเปลี่ยนระหว่างการใช้จำนวนเมสเสจค้นหาและความสำเร็จในการค้นพบเส้นทาง โดยองค์ความรู้ที่ได้รับในรายงานวิจัยนี้ได้แก่

4.1.1 การกำหนดปัญหา

กำหนดปัญหาการค้นหาเส้นทางให้เป็นกระบวนการตัดสินใจแบบมาร์คอฟภายใต้สภาวการณ์ที่สังเกตได้บางส่วน (partially observable Markov Decision Process หรือ MDP) เพื่อหาสมดุลแลกเปลี่ยนระหว่างการใช้จำนวนเมสเสจค้นหาและความสำเร็จในการค้นพบเส้นทางที่รองรับคุณภาพการบริการ (QoS routing) ในเครือข่ายเคลื่อนที่แบบแอดฮอค

4.1.2 การค้นหาเส้นทางที่รองรับคุณภาพการบริการในเครือข่ายเคลื่อนที่แบบแอดฮอคด้วยกระบวนการเรียนรู้แบบรีอินฟอร์สเมนท์

ในบทที่ 3 นั้นได้ประยุกต์ใช้กระบวนการ POMDP RL เพื่อค้นหาเส้นทางที่รองรับคุณภาพการบริการ (QoS routing) ในเครือข่ายเคลื่อนที่แบบแอดฮอคด้วยการใช้วิธีการ ONMC ผนวกเข้ากับวิธี TBP และกลยุทธ์พาธ แคชชิ่งเพื่อเรียนรู้นโยบายที่ดีในการแจกจ่ายตั๋วที่โหนดต้นทาง และสามารถลดจำนวนโอเวอร์เฮดในการค้นหาเส้นทาง

ผลการทดลองที่ได้จากการจำลองระบบบนคอมพิวเตอร์แสดงให้เห็นว่าวิธี TBP ภายใต้กระบวนการ ONMC ด้วยกลยุทธ์พาธ แคชชิ่ง (ONMCP) สามารถเรียนรู้นโยบายการจำหน่ายตั๋วที่ดี ในแง่ของการให้ผลตอบแทนสะสมต่อเอพพิโซดสูงสุดเมื่อเปรียบเทียบกับวิธี TBP ดั้งเดิมและวิธี TBP ภายใต้กระบวนการ ONMC ในขณะที่มีการใช้เมสเสจค้นหาน้อยกว่า

วิธี TBP ภายใต้กระบวนการ ONMC มีความต้องการในการคำนวณเพื่อให้ได้การตัดสินใจในหนึ่งครั้งต้องใช้การโต้ตอบ $O(|S||A|)$ ครั้ง เมื่อ $|S|$ และ $|A|$ แทนขนาดของเซตของสถานะและการทำงานที่เป็นไปได้ทั้งหมดตามลำดับ โดยเมื่อขนาดของเครือข่ายใหญ่ขึ้นก็ไม่ได้ส่งผลต่อความต้องการในการคำนวณนี้ อย่างไรก็ตามความต้องการด้วยหน่วยความจำจะแปรผันตรงกับจำนวนโหนดที่อยู่ในเครือข่าย ในทางปฏิบัติ การใช้โครงสร้างลำดับชั้นถูกนำมาใช้แก้ไขปัญหาความต้องการด้านหน่วยความจำที่เพิ่มขึ้นเพื่อนำไปปฏิบัติจริงกับเครือข่ายเคลื่อนที่แบบแอดฮอคขนาดใหญ่ได้

4.2 งานวิจัยในอนาคต

4.2.1 การรักษาเส้นทาง

รายงานวิจัยฉบับนี้กล่าวครอบคลุมเนื้อหาของการค้นหาเส้นทางเมื่อมีการร้องขอการเชื่อมต่อเท่านั้น อย่างไรก็ตามกระบวนการเรียนรู้แบบรีอินฟอร์สเมนท์สามารถขยายองค์ความรู้เพื่อการพัฒนาเป็นอัลกอริธึมรักษาเส้นทางเพื่อคงเส้นทางสื่อสารไว้ ในเครือข่าย MANETs เมื่อมีความเสียหายของเส้นทาง เส้นทางใหม่จะถูกสร้างขึ้นมาอีกครั้ง [5] หรือ มีการซ่อมบำรุงเส้นทางเก่า [7], [9] โดยที่การสร้างเส้นทางใหม่จะไปเพิ่มโอเวอร์เฮดในการสร้างเส้นทางได้ แต่ก็ยังมีการใช้ต้นทุนต่ำกว่าการซ่อมบำรุงเส้นทาง ดังนั้นกระบวนการตัดสินใจแบบ MDP สามารถนำไปใช้แก้ปัญหาการรักษาเส้นทางเพื่อหาสมดุลแลกเปลี่ยนระหว่างประสิทธิภาพของเส้นทางและโอเวอร์เฮดที่เพิ่มขึ้น

4.2.2 การพิจารณาพลังงานจากแบตเตอร์รี่

เครือข่าย MANETs ที่ถูกพิจารณาในรายงานวิจัยฉบับนี้สมมุติว่าแหล่งพลังงานจากแบตเตอร์รี่ของแต่ละโหนดเคลื่อนที่มีระดับพลังงานคงที่ตลอดการทดลอง ดังนั้นการเชื่อมต่อระหว่างโหนดจะไม่เกิดขึ้นเมื่อโหนดอยู่ห่างกันมากเกินไปเท่านั้น ดังนั้นประเด็นที่น่าสนใจคือการขยายกรอบงานไปสู่การพิจารณาให้รัศมีการส่งข้อมูลของโหนดแปรผันตามระดับพลังงานจากแบตเตอร์รี่

4.2.3 การประสานงานข้ามชั้น

รายงานวิจัยฉบับนี้มีการพิจารณาระดับชั้นโปรโตคอลในเครือข่ายเคลื่อนที่แบบแอดฮอค ในชั้นกายภาพ (physical layer) การเข้าใช้ช่องสัญญาณในชั้นเชื่อมต่อข้อมูล (media-access control layer) และชั้นเครือข่าย (network layer) ซึ่งถูกพิจารณาแบบแยกส่วนกัน อย่างไรก็ตามการปรับปรุงประสิทธิภาพของเครือข่ายและและการลดการใช้พลังงานในเครือข่ายจำเป็นต้องทำการพิจารณาร่วมกันทั้งสามชั้น ดังนั้นสิ่งที่น่าสนใจคือการขยายกรอบงานไปสู่การพิจารณาการประสานงานข้ามชั้นกัน