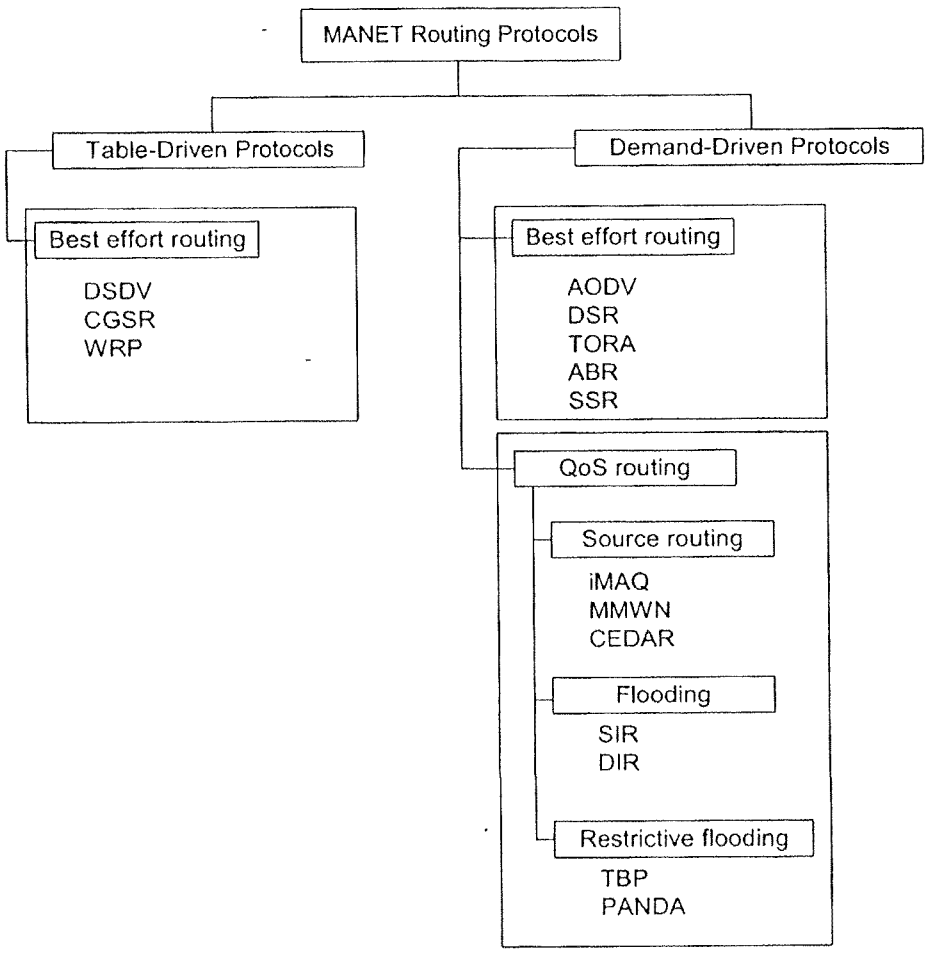
บทที่1
ที่มาและความสำคัญ

บทนี้กล่าวถึงพื้นฐานเครือข่ายเคลื่อนที่แบบแอดฮอค มุ่งเน้นถึงความสำคัญของปัญหาการหาเส้นทางที่รองรับกับคุณภาพการบริการในเครือข่ายเคลื่อนที่แบบแอดฮอค อีกทั้งยังกล่าวถึงแรงจูงใจในการนำกระบวนการเรียนรู้แบบรีอินฟอสเม้นท์มาใช้เพื่อปรับปรุงโปรโตคอลค้นหาเส้นทางที่มีอยู่เดิมให้เกิดประสิทธิภาพดียิ่งขึ้น ซึ่งเป็นเป้าหมายหลักของงานวิจัยนี้

1.1 ความสำคัญของปัญหา

เครือข่ายเคลื่อนที่แบบแอดฮอค (Mobile ad hoc network หรือ MANET) เป็นเครือข่ายเพื่อการติดต่อสื่อสาร เมื่อทุกโหนดภายในเครือข่ายต่างร่วมมือกันเพื่อเชื่อมต่อเส้นทางในเครือข่ายโดยปราศจากควบคุมจากศูนย์กลาง คุณลักษณะทั่วไปของเครือข่ายเคลื่อนที่แบบแอดฮอค คือ ทุกโหนดภายในเครือข่ายสามารถเคลื่อนที่ได้อย่างอิสระภายใต้ขอบเขตของแบนวิธ, ความจุของเส้นทางเชื่อมต่อ และรูปแบบของเครือข่ายที่ไม่สามารถทำนายได้ แต่ละโหนดจะมีระยะการส่งข้อมูลที่จำกัด โหนดต้นทางจะติดต่อสื่อสารกับโหนดปลายทางที่อยู่นอกเหนือรัศมีการส่งด้วยการใช้โหนดข้างเคียง ดังนั้นทุกๆโหนดในเครือข่ายจึงมีความสามารถเช่นเดียวกับตัวค้นหาเส้นทางเคลื่อนที่ที่สามารถส่งต่อข้อมูลและเป็นเสมือนโฮสได้ในเวลาเดียวกัน จากคุณลักษณะของเครือข่ายเคลื่อนที่แบบแอดฮอคจะเห็นได้ว่า เมื่อแต่ละโหนดในเครือข่ายสามารถเคลื่อนที่ได้อย่างเป็นอิสระต่อกันส่งผลให้การค้นหาเส้นทางจากต้นทางไปยังปลายทางนั้นซับซ้อนมากกว่าระบบเครือข่ายไร้สายทั่วไป เนื่องจากข้อมูลเส้นทางการสื่อสารนั้นจะมีการเปลี่ยนแปลงอยู่ตลอดเวลาตามสภาพของเส้นทางและรูปร่างเครือข่าย

ในเครือข่าย MANETs งานวิจัยส่วนใหญ่มุ่งเน้นไปที่การพัฒนาโปรโตคอลเลือกเส้นทาง (Routing protocol) เพื่อค้นหา เลือก และบำรุงรักษาเส้นทางการส่งข้อมูลที่สั้นที่สุด เพื่อให้เกิดการถ่ายโอนข้อมูลที่ดีที่สุด โปรโตคอลเลือกเส้นทางสามารถจำแนกได้เป็น 2 ประเภทคือ โปรโตคอลเลือกเส้นทางด้วยตารางเส้นทาง (Table-driven routing protocol) และโปรโตคอลเลือกเส้นทางตามอุปสงค์ (demand-driven routing protocol) ซึ่งแสดงในรูปที่ 1.1



รูปที่ 1.1 โปรโตคอลค้นหาเส้นทางในเครือข่ายเคลื่อนที่แบบแอดฮอค

โดยโปรโตคอลเลือกเส้นทางที่ขับเคลื่อนด้วยตารางเส้นทางจะ จัดเตรียมตารางเส้นทางเพื่อเก็บข้อมูลเกี่ยวกับเส้นทางจากโหนดหนึ่งไปยังทุกๆโหนดในเครือข่ายไว้ล่วงหน้า ทำให้แต่ละโหนดไม่เสียเวลาในการประมวลผลเลือกเส้นทาง อย่างไรก็ตามด้วยการกระทำเช่นนี้ทำให้สิ้นเปลืองพื้นที่จัดเก็บข้อมูลตารางเส้นทาง ตัวอย่างโปรโตคอลประเภทนี้ได้แก่ โปรโตคอลเลือกเส้นทางจากลำดับของระยะทางจากเวกเตอร์ปลายทาง (Destination-sequenced distance-vector protocol หรือ DSDV) [1] โปรโตคอลเลือกเส้นทางด้วยการสับเปลี่ยนหัวหน้ากลุ่มเกตเวย์ (Clusterhead gateway switch routing หรือ CGSR) [2] โปรโตคอลเลือกเส้นทางแบบไร้สาย (Wireless routing protocol หรือ WRP) [3]

สำหรับโปรโตคอลเลือกเส้นทางตามฟังก์ชันอุปสงค์นั้น เส้นทางจะถูกค้นหาและถูกบำรุงรักษาตามการร้องขอ โหนดต้นทางจะเริ่มกระบวนการค้นหาเส้นทางก็ต่อเมื่อต้องการเส้นทางเพื่อส่งข้อมูลไปยังโหนดปลายทางเท่านั้น กระบวนการนี้จึงสามารถหลีกเลี่ยงโอเวอร์เฮด (overhead) ขนาด

ใหญ่ในการเก็บรักษาตารางเส้นทางจากโปรโตคอลเลือกเส้นทางที่ขับเคลื่อนด้วยตารางเส้นทางลงได้ ตัวอย่างโปรโตคอลประเภทนี้ได้แก่ โปรโตคอลเลือกเส้นทางด้วยการใช้เวกเตอร์ระยะทางตามฟังก์ชั่นอุปสงค์ภายในเครือข่ายแอดฮอค (Ad hoc on-demand distance vector routing หรือ AODV) [4] โปรโตคอลเลือกเส้นทางจากต้นทางแบบพลวัติ (Dynamic source routing หรือ DSR) [5] โปรโตคอลเลือกเส้นทางโดยใช้ลำดับเวลา (Temporally ordered routing algorithm หรือ TORA) [6] โปรโตคอลเลือกเส้นทางด้วยการเปลี่ยนหมู่ (Associativity-based routing protocol หรือ ABR) [7] และโปรโตคอลเลือกเส้นทางด้วยการปรับเสถียรภาพของสัญญาณ (Signal stability-based routing protocol หรือ SSR) [8]

จากโปรโตคอลเลือกเส้นทางที่กล่าวมาข้างต้นไม่ได้มีการรองรับคุณภาพการบริการ (QoS) ของเส้นทาง เช่น แบนวิธและเวลาหน่วงตลอดเส้นทาง (end-to-end bandwidth and delay) และเงื่อนไขของเวลาหน่วง อย่างไรก็ตามการรับรอง QoS ของเส้นทางในเครือข่าย MANETs ที่มีรูปร่างเครือข่ายแบบพลวัติเป็นเรื่องยากเนื่องมาจาก ประการแรก QoSของเส้นทางต้องการการจองทรัพยากรตลอดเส้นทางการส่งข้อมูลระหว่างคู่โหนดต้นทางไปยังโหนดปลายทาง การจองทรัพยากรดังกล่าวขึ้นอยู่กับอัลกอริธึมค้นหาเส้นทางซึ่งต้องอาศัยข้อมูลสถานะของพลังงานและรูปร่างเครือข่ายที่แม่นยำ อีกทั้งข้อมูลดังกล่าวยังเป็นข้อมูลที่ไม่สามารถระบุได้อย่างชัดเจนสำหรับเครือข่ายMANETs ดังนั้นการตัดสินใจเลือกเส้นทางด้วยข้อมูลที่ไม่แน่ชัดหรือข้อมูลที่ไม่มีการอัพเดทอาจทำให้ได้เส้นทางที่ไม่เหมาะสม ประการที่สองทุกโหนดในเครือข่ายสามารถเคลื่อนที่ได้ ดังนั้นเส้นทางที่เชื่อมต่ออยู่อาจขาดหายได้ตลอดเวลา ซึ่งอาจทำให้เกิดปัญหาการรักษาเส้นทางเชื่อมต่อตามมา ดังนั้นการประกันคุณภาพของเส้นทางในเครือข่ายประเภทนี้จึงแทบเป็นไปไม่ได้เลยถ้าโหนดในเครือข่ายมีการเคลื่อนที่มากเกินไป ทำให้งานวิจัยส่วนใหญ่มักพิจารณาเครือข่ายMANETsที่มีการเปลี่ยนแปลงรูปร่างเครือข่ายที่ไม่รวดเร็วมากนัก

เพื่อบรรเทาการพิจารณาการเลือกเส้นทางจากข้อมูลที่ไม่แน่ชัด งานวิจัยส่วนใหญ่จึงนำเสนอโปรโตคอลเลือกเส้นทางที่รองรับQos ในเครือข่ายเคลื่อนที่แบบแอดฮอคซึ่งอาศัยการฟลัดดิ้ง (flooding) เพื่อหาเส้นทาง อาทิ อัลกอริธึมเลือกเส้นทางด้วยโหนดต้นทาง (Source-initiated routing algorithm หรือ SIR) [9] อัลกอริธึมเลือกเส้นทางด้วยโหนดปลายทาง (Destination-initiated routing algorithm หรือ DIR) [9] อย่างไรก็ตามการฟลัดดิ้งมีขีดจำกัดเมื่อเครือข่ายมีขนาดใหญ่ขึ้น

ในทางตรงกันข้ามการฟลัดดิ้งแบบวงจำกัด (restrictive flooding) ถูกนำเสนอขึ้นซึ่งเป็นวิธีที่ก้ำกึ่งระหว่างการฟลัดดิ้ง (แบบไม่จำกัด) และการเลือกเส้นทางโดยโหนดต้นทาง (source

routing) โดยการฟลัดดิ้ง แบบวงจำกัดยังคงมีการหาเส้นทางด้วยวิธีการเดิมอยู่ แต่เมสเสจฟลัดดิ้งจะถูกควบคุมด้วยหน่วยวัดบางประการที่โหนดต้นทางด้วยการใช้ข้อมูลที่ครอบคลุม (global information) เช่น วิธีการตรวจสอบด้วยตั๋ว (Ticket-based probing หรือ TBP) [10] ซึ่งวิธีนี้สามารถควบคุมการฟลัดดิ้งที่ถี่เกินไปได้ด้วยการจำกัดจำนวนตั๋วเชิงตรรกะ (logical ticket) ที่โหนดต้นทาง อย่างไรก็ตาม วิธีการนี้ยังมีประเด็นเปิดที่ควรพิจารณาอยู่ นั่นคือ การคำนวณจำนวนตั๋วเชิงตรรกะที่เหมาะสมสำหรับการแจกจ่ายที่โหนดต้นทาง นอกจากนี้วิธีนี้ยังอาศัยกฎฮิวริสติก (heuristic rule) สำหรับการคำนวณตั๋วอีกด้วย งานวิจัย [11] ได้นำวิธี TBP แบบเดิมมาผนวกกับกระบวนการเรียนรู้แบบรีอินฟอร์สเมนท์ (reinforcement learning หรือ RL) ซึ่งเป็นวิธีที่สามารถเรียนรู้ นโยบายที่ดีสำหรับการจำหน่ายตั๋วด้วยการเลือกการกระทำโต้ตอบกับสิ่งแวดล้อมโดยตรงโดยใช้กระบวนการตัดสินใจอย่างมีเหตุผลในรูปแบบออนไลน์ ซึ่งสามารถหลีกเลี่ยงการคำนวณตั๋วที่ผูกกับกฎฮิวริสติกจากวิธี TBPเดิมได้อีกด้วย อย่างไรก็ตาม วิธีการนี้ยังคงมีโอเวอร์เฮดขนาดใหญ่อันเนื่องมาจากความถี่ในการถูกร้องขอให้ค้นหาเส้นทางใหม่ทุกครั้งที่มีการร้องขอเส้นทาง

รายงานวิจัยฉบับนี้ทำการพัฒนาโปรโตคอลค้นหาเส้นทางที่รองรับคุณภาพการบริการในเครือข่ายเคลื่อนที่แบบแอดฮอค โดยนำวิธีค้นหาเส้นทางแบบ TBP ที่มีอยู่เดิมมาผนวกกับกระบวนการเรียนรู้แบบรีอินฟอร์สเมนท์ภายใต้กระบวนการตัดสินใจแบบมาร์คอฟภายใต้สภาวการณ์ที่สังเกตได้บางส่วน (partially observable Markov decision process หรือ POMDP) ที่เรียกว่า วิธีออนโพลิซีมอนติ คาร์โล (on-policy Monte Carlo หรือ ONMC) [11] และกลยุทธ์พาธ แคชชิ่ง (path caching) เพื่อใช้หานโยบายที่เหมาะสมสำหรับการค้นหาเส้นทางที่รับรองคุณภาพการบริการ (QoS routing) อีกทั้งยังสามารถลดโอเวอร์เฮดในการค้นหาเส้นทาง (routing overhead) สำหรับเครือข่ายเคลื่อนที่แบบแอดฮอคที่มีรูปร่างเครือข่ายแบบพลวัติ

1.2 วัตถุประสงค์

วัตถุประสงค์ของรายงานฉบับนี้คือ

1.2.1 เพื่อพัฒนากระบวนการตัดสินใจสำหรับการหาเส้นทางในเครือข่ายเคลื่อนที่แบบแอดฮอคที่สามารถลดโอเวอร์เฮดในการค้นหาเส้นทาง และเพิ่มประสิทธิภาพของเครือข่ายในระยะยาว

1.2.2 เพื่อประยุกต์ใช้กระบวนการเรียนรู้แบบรีอินฟอร์สเม้นท์กำหนดปัญหาการค้นหาเส้นทางที่รอบรับคุณภาพการบริการสำหรับเครือข่ายเคลื่อนที่แบบแอดฮอคที่มีรูปแบบเครือข่ายแบบพลวัติ

1.2.3 เพื่อประเมินประสิทธิภาพของการผนวกวิธี POMDP RL เข้ากับวิธี TBP ด้วยการเปรียบเทียบผลกับโปรโตคอลหาเส้นทางที่มีอยู่ โดยพิจารณาจากผลตอบแทนสะสมต่อเอพพิโซด และจำนวนเมสเสจค้นหาโดยเฉลี่ย

1.2.4 เพื่อเปรียบเทียบข้อแลกเปลี่ยนในการพบเส้นทางที่รอบรับคุณภาพการบริการสำหรับกระบวนการเรียนรู้แบบรีอินฟอร์สเม้นท์

1.3 ส่วนประกอบของรายงานวิจัย

ส่วนที่เหลือของรายงานวิจัยฉบับนี้ประกอบด้วย บทที่2 กล่าวถึงทฤษฎีพื้นฐานของกระบวนการเรียนรู้แบบรีอินฟอร์สเม้นท์ซึ่งเป็นส่วนจำเป็นที่ทำให้เกิดองค์ความรู้ในรายงานวิจัยฉบับนี้ ประการแรกอธิบายแนวคิดของกระบวนการตัดสินใจแบบมาร์คอฟ (Markov Decision Process หรือ MDP) และแนะนำกระบวนการเรียนรู้แบบรีอินฟอร์สเมนท์เพื่อหาผลเฉลยของกระบวนการการตัดสินใจแบบมาร์คอฟที่กำหนดขึ้น โดยใช้กระบวนการเรียนรู้แบบรีอินฟอร์สเมนท์ที่เรียกว่าวิธี ออนโพลิซี มอนติ คาร์โล (On-policy Monte Carlo หรือ ONMC) ซึ่งเรียนรู้จากประสบการณ์ที่เกิดจากผลของการกระทำที่ส่งผลต่อสิ่งแวดล้อมในแต่ละเอพพิโซด (episode)

บทที่3 กล่าวถึงการศึกษาโปรโตคอลเลือกเส้นทางที่รองรับคุณภาพการบริการสำหรับเครือข่ายเคลื่อนที่แบบแอดฮอค การประยุกต์วิธี TBP ผนวกกับกระบวนการเรียนรู้แบบรีอินฟอร์สเมนท์แบบ POMDP ที่เรียกว่า วิธีออนโพลิซี มอนติ คาร์โล (ONMC) ด้วยการใช้กลยุทธ์พาธ แคชชิ่ง ซึ่งเกิดสมดุลแลกเปลี่ยนระหว่างการเพิ่มความสำเร็จในการค้นพบเส้นทางและการใช้เมสเสจค้นหาในปริมาณต่ำ พร้อมทั้งเปรียบเทียบประสิทธิภาพของเส้นทางในแง่ของ ผลตอบแทนสะสมต่อเอพพิโซดและจำนวนเมสเสจค้นหาโดยเฉลี่ย ซึ่งวิธี ONMC ด้วยการใช้กลยุทธ์พาธ แคชชิ่ง จะถูกเปรียบเทียบประสิทธิภาพกับโปรโตคอลเลือกเส้นทางที่ใช้อยู่สำหรับเครือข่ายเคลื่อนที่แบบแอดฮอค

บทที่4 กล่าวถึงการสรุปผล และแนวทางการพัฒนาในอนาคต