

## บทคัดย่อ

วัตถุประสงค์ของรายงานฉบับนี้คือการพัฒนากระบวนการตัดสินใจแบบออนไลน์เพื่อการค้นหาเส้นทางที่รองรับคุณภาพการบริการในเครือข่ายเคลื่อนที่แบบแอตซอค โดยสามารถลดโอเวอร์เฮดในกระบวนการสื่อสาร เพิ่มประสิทธิภาพของระบบในระยะยาว และสามารถปฏิบัติการการได้อย่างดีเยี่ยมภายใต้ข้อมูลที่ไม่ชัดเจนสำหรับเครือข่ายที่มีรูปร่างเครือข่ายแบบพลวัต องค์ความรู้ที่ได้จากรายงานฉบับนี้คือการประยุกต์ใช้กระบวนการเรียนรู้แบบรีอินฟอร์สเมนต์และกลยุทธ์พาท แคชซิง สามารถลดขนาดของเมสเสจโอเวอร์เฮดในการค้นหาเส้นทางที่รองรับคุณภาพการบริการในเครือข่ายเคลื่อนที่แบบแอตซอค

รายงานฉบับนี้ทำการกำหนดปัญหาการควบคุมเมสเสจโอเวอร์เฮดสำหรับการค้นหาเส้นทางที่รองรับคุณภาพการบริการในเครือข่ายเคลื่อนที่แบบแอตซอคให้เป็นกระบวนการการตัดสินใจแบบมาร์คอฟภายใต้สภาวะการณ์ที่สังเกตได้บางส่วน (partially observable Markov decision process หรือ POMDP) ด้วยการนำวิธีการตรวจสอบด้วยตั๋ว (Ticket-based probing หรือ TBP) แบบเดิมมาผนวกกับกระบวนการเรียนรู้แบบรีอินฟอร์สเมนต์สำหรับ POMDP ที่เรียกว่า วิธีออนโพลิซี มอนติ คาร์โล (on-policy Monte Carlo หรือ ONMC) และกลยุทธ์พาท แคชซิง (path caching) เพื่อใช้หานโยบายที่เหมาะสมในการค้นหาเส้นทางที่รองรับคุณภาพการบริการสำหรับเครือข่ายเคลื่อนที่แบบแอตซอค จากผลการทดลองพบว่าภายใต้กระบวนการ POMDP สามารถเลือกนโยบายที่ดีที่สุดสำหรับการจำหน่ายตั๋ว ซึ่งแสดงให้เห็นในรูปของผลตอบแทนสะสมต่อเอพพิไซด์เมื่อเปรียบเทียบกับวิธี TBP แบบเดิมและวิธี ONMC ที่ไม่มีพาท แคชซิง นอกจากนี้กระบวนการที่นำเสนอสามารถเพิ่มประสิทธิภาพในการควบคุมเมสเสจค้นหา โดยสามารถลดขนาดของเมสเสจโอเวอร์เฮดได้ดีกว่าวิธีดั้งเดิม ซึ่งแสดงในรูปของอัตราความสำเร็จและมูลค่าเฉลี่ยตลอดเส้นทาง

## Abstract

The underlying aim of this report is to develop on-line decision-making algorithm for QoS routing in mobile ad hoc networks (MANETs) which would minimize communication overhead, maximize the overall long-term performance criterion and can perform well under the presence of uncertainty for dynamic topology networks. The contributions in this report is the experimental evidence that, RL techniques equipped with suitable path caching strategies can be employed to reduce the amount of message overhead in QoS routing in MANETs.

A novel partially observable Markov decision process (POMDP) formulation of a message overhead control problem for QoS routing in MANETs is introduced. The proposed scheme integrates the original the Ticket-Based Probing (TBP) scheme with a reinforcement learning method for POMDPs, called the on-policy first-visit Monte Carlo method with path caching (ONMCP) scheme, is applied to support QoS routing at the network level in a MANET. Results obtained from various scenarios of mobility and imprecise information, and stringent QoS requirements show that the POMDP framework can achieve good ticket-issuing policies, in terms of the accumulated reward per episode when compared to the original heuristic TBP scheme and the ONMC scheme without path caching. Furthermore, our approach can lead to more efficient control of search messages, i.e., a reduction of message overhead with marginal difference in the success ratio and average path cost.