บทที่ 3

กระบวนการเรียนรู้แบบรีอินฟอสเมนท์สำหรับการค้นพบเส้นทางในเครือข่ายเคลื่อนที่แบบแอดฮอคด้วยกลยุทธ์พาธ แคชชิ่ง

3.1 กล่าวนำ

เนื้อหาบทนี้กล่าวถึงวิธีการเลือกเส้นทางที่รองรับคุณภาพการบริการในเครือข่ายเคลื่อนที่แบบแอดฮอค โดยมีวัตถุประสงค์เพื่อให้เกิดข้อตกลงสำหรับการเกิดสมดุลแลกเปลี่ยนระหว่างการเพิ่มความสำเร็จในการค้นหาเส้นทางและการใช้เมสเสจค้นหาในปริมาณต่ำ โดยการประยุกต์ใช้กระบวนการเรียนรู้แบบรีอินฟอร์สเมนท์ด้วยวิธีออนโพลิซี มอนติ คาร์โล (ONMC) ด้วยกลยุทธ์พาธ แคชชิ่ง ซึ่งวิธีการดังกล่าวเหมาะสำหรับการตัดสินใจที่มีการแบ่งเรียนรู้เป็นเอพพิโซด

เนื้อหาสำคัญที่จะกล่าวถึงในบทนี้ ได้แก่

1. กล่าวแนะนำการจัดการคุณภาพการบริการสำหรับเส้นทางในเครือข่ายเคลื่อนที่แบบแอดฮอค ด้วยกระบวนการตัดสินใจแบบมาร์คอฟภายใต้สภาวการณ์ที่สังเกตได้บางส่วน
2. การหาเส้นทางด้วยวิธี TBP และกลยุทธ์พาธ แคชชิ่ง
3. การเปรียบเทียบประสิทธิภาพการเลือกเส้นทางด้วยวิธีออนโพลิซี มอนติ คาร์โล ด้วยกลยุทธ์พาธ แคชชิ่ง (ONMCP) เทียบกับวิธีการเลือกเส้นทางที่มีอยู่เดิมอีก 2 วิธี ได้แก่ วิธี TBP และวิธี ONMC

3.2 คุณภาพการบริการสำหรับเส้นทางในเครือข่ายเคลื่อนที่แบบแอดฮอค

องค์ประกอบสำคัญของการจัดการคุณภาพการบริการของเส้นทาง (QoS routing) ในเครือข่ายเคลื่อนที่แบบแอดฮอค (mobile ad hoc network หรือ MANET) คือ ข้อมูลเกี่ยวกับแหล่งพลังงานที่เหลืออยู่ในโหนดเคลื่อนที่ ข้อมูลดังกล่าวขึ้นอยู่กับการอัพเดทข้อมูลระหว่างโหนดเคลื่อนที่

ดังนั้นการแลกเปลี่ยนข้อมูลระหว่างโหนดเคลื่อนที่จึงเป็นสิ่งจำเป็นอย่างยิ่ง การแลกเปลี่ยนข้อมูลนั้นจะกระทำเป็นช่วงเวลา หรือ เมื่อพบว่ารูปร่างเครือข่ายมีการเปลี่ยนแปลง อย่างไรก็ตามยังมีข้อมูลที่คลุมเครือสืบเนื่องมาจากเมสเสจบอกการอัพเดทถูกส่งมาช้าหรือมีการสูญหาย ซึ่งอาจทำให้ข้อมูลถูกยับยั้ง

ข้อมูลของเครือข่ายที่ถูกต้องนั้นยากต่อการสังเกต เนื่องมาจากโหนดเคลื่อนที่แต่ละตัวจะถูกพบจากการเฝ้าสังเกต (observation) จากสิ่งแวดล้อมของตนเองเท่านั้น ซึ่งอาจทำให้ข้อมูลที่รวบรวมมาไม่สมบูรณ์และอาจเกิดความผิดพลาดได้ ข้อมูลที่ได้จากการเฝ้าสังเกตเครือข่ายของโหนดเคลื่อนที่แต่ละตัวจะถูกนำไปใช้ในการตัดสินใจในเรื่องต่างๆ เช่น ต้องใช้เมสเสจควบคุมจำนวนเท่าไรเพื่อใช้ในการค้นหาเส้นทางการส่งข้อมูลที่เป็นไปได้ เป็นต้น

ภายใต้การพิจารณาสมมติฐานเกี่ยวกับการเคลื่อนที่และข้อมูลด้านแหล่งจ่ายพลังงาน มีความเป็นไปได้ที่จะจำลองโมเดลการส่งสถานะ (state transitions) ให้เป็นกระบวนการมาร์คอฟ [11] และเนื่องจาก สถานะ (state) ที่ถูกต้องของเครือข่ายไม่สามารถระบุได้ ดังนั้นจึงมีการใช้แบบจำลองการตัดสินใจที่เรียกว่า กระบวนการตัดสินใจแบบมาร์คอฟภายใต้สภาวการณ์ที่สังเกตได้บางส่วน (partially observable Markov decision process หรือ POMDP) ภายในเครือข่ายเคลื่อนที่แบบแอดฮอค โดยมีวัตถุประสงค์เพื่อค้นหานโยบายที่ให้ประสิทธิภาพสูงสุดภายใต้ขอบเขตจำกัด โดยที่ปัญหาขอบเขตจำกัดจะถูกพิจารณาในที่นี้เนื่องมาจากการแลกเปลี่ยนข้อมูลระหว่างโหนดเคลื่อนที่จะกระทำเป็นแบบเอพพิโซด (episode) โดยเอพพิโซดจะเริ่มต้นหลังจากมีการแลกเปลี่ยนเมสเสจเกิดขึ้นและจะสิ้นสุดเมื่อมีการแลกเปลี่ยนเมสเสจถัดมา

3.3 วิธีการค้นหาเส้นทางแบบ TBP และพาธ แคชชิ่ง

วิธีออนโพลิซี เฟิร์ส-วิสิท มอนติ คาร์โล (on-policy first-visit Monte Carlo หรือ ONMC) [7] ถูกนำมาใช้เพื่อหานโยบายภายใต้การเฝ้าสังเกตสิ่งแวดล้อมในกระบวนการตัดสินใจแบบมาร์คอฟภายใต้สภาวการณ์ที่สังเกตได้บางส่วน (partially observable Markov decision process หรือ POMDP)

วิธี ONMC จากกระบวนการ POMDP ถูกรวมเข้าไปในวิธีการค้นหาเส้นทางที่เรียกว่า Ticket-Based Probing (TBP) โดยวิธี TBP เป็นอัลกอริธึมค้นหาเส้นทางจากหลายเส้นทางแบบกระจาย (multipath distributed routing algorithm) สำหรับระบบที่มีเวลาหน่วงตลอดเส้นทาง (end-to-end delay) หรือการร้องขอแบนวิธด์สำหรับข้อมูลสถานะที่มีความคลุมเครือระดับสูง [10] จุดประสงค์ของ

อัลกอริธึมนี้คือการเลือกเส้นทางการส่งข้อมูลที่เหมาะสมด้วยความน่าจะเป็นในการส่งสำเร็จสูงสุดสำหรับเครือข่ายที่มีรูปร่างเครือข่ายแบบพลวัติด้วยข้อมูลที่ไม่แม่นยำ แนวคิดพื้นฐานของอัลกอริธึมนี้คือ เมื่อโหนดต้นทาง s ต้องการเส้นทางการส่งที่เหมาะสมกับเวลาหน่วง (หรือ แบนวิดธ์) ที่ถูกร้องขอต่อโหนดปลายทาง d จำนวนของโพรบ (เมสเสจค้นหา) ถูกส่งจากโหนด s ไปยังโหนด d จำนวนโพรบทั้งหมดที่ใช้สำหรับค้นหาเส้นทางถูกควบคุมโดย จำนวนตั้งต้นของตั๋วเชิงตรรกะ (logical tickets) หรือ M_0 โดยพารามิเตอร์ M_0 จะถูกคำนวณที่โหนดต้นทาง s เมื่อโหนดข้างเคียง j ได้รับโพรบจากโหนด s โหนด j จะทำสำเนาโพรบนั้นไว้และทำการคำนวณจำนวนของตั๋วเพื่อดำเนินการทำสำเนาโพรบอีกครั้ง การคำนวณตั๋วที่โหนด j จะกระทำภายใต้ข้อมูลตลอดเส้นทางที่หามาได้ (นั่นคือ จากโหนด j ไปยังโหนด d) และจะไม่เกินกว่าจำนวนของตั๋วในโพรบที่โหนด j ได้รับมา แต่ละโพรบจะถือตั๋วไว้อย่างน้อยหนึ่งใบ และตั๋วทั้งหมดที่อยู่ในเครือข่ายจะถูกควบคุมโดยความแปรปรวนของพารามิเตอร์ M_0

3.3.1 การคำนวณตั๋วตั้งต้น: ภาพรวมของวิธี TBPดั้งเดิม

งานวิจัยนี้ศึกษาปัญหาการเลือกเส้นทางที่มีมูลค่าเวลาหน่วงน้อยที่สุด (delay-constrained least-cost routing) พิจารณาการร้องขอการเชื่อมต่อของโหนดต้นทาง โหนดปลายทาง และความต้องการของค่าเวลาหน่วงเฉลี่ยตลอดเส้นทาง (mean end-to-end delay) คือ s, d และ D_{req} ตามลำดับ กำหนดให้ D_{ij} คือค่าเวลาหน่วงเฉลี่ยของการเชื่อมต่อ (mean link delay) ระหว่างโหนด i และโหนด j โดยที่เวลาหน่วงเฉลี่ยตลอดเส้นทางของเส้นทางที่มีเวลาหน่วงน้อยที่สุดคือ r^* และ $D_n(d) = \sum_{(i,j)\in r^*} D_{ij}$ โดยพารามิเตอร์ $\Delta D_n(d)$ คือค่าความแปรปรวนของเวลาหน่วงเฉลี่ยตลอดเส้นทาง หาได้จาก

$$\Delta D_n^{new}(d) = \rho \Delta D_n^{old}(d) + (1-\rho)\beta \,|\, D_n^{new}(d) - D_n^{old}(d) \,| \qquad (3.1)$$

พารามิเตอร์ ρ คือตัวแปรเพิกเฉย ซึ่งใช้กำหนดความเร็วในการเพิกเฉยค่า $\Delta D_n^{old}(d)$ และพารามิเตอร์ $1-\rho$ ใช้กำหนดความรวดเร็วของการทำให้ค่า $\Delta D_n^{new}(d)$ ลู่เข้าสู่ $|D_n^{new}(d) - D_n^{old}(d)|$ และ β คือพารามิเตอร์สำหรับรองรับค่า $\Delta D_n^{new}(d)$ ในงานวิจัย [1] จำนวนของตั๋ว (M_0) หาได้จาก $M_0 = Y_0 + G_0$ เมื่อ Y_0 และ G_0 คือตั๋วสีเหลืองและตั๋วสีเขียวตามลำดับ โดยตั๋วสีเหลือง

สำหรับโอกาสสูงสุดในการค้นพบเส้นทางที่เหมาะสม ในขณะที่ตั๋วสีเขียวหมายถึงโอกาสสูงสุดในการพบเส้นทางที่มีมูลค่าต่ำ

พารามิเตอร์ Y_0 ถูกกำหนดจากกฎฮิวริสติก (heuristic rules) [10]

$$Y_0 = \begin{cases} 1 & , D_{reg} > D_{hi} \\ \dfrac{D_s(d) + \Delta D_s(d) - D_{req}}{2 \times \Delta D_s(d)} \times \theta_Y & , D_{lo} \le D_{req} \le D_{hi} \\ 0 & , D_{req} < D_{lo} \end{cases} \tag{3.2}$$

เมื่อ $D_{hi} = D_s(d) + \Delta D_s(d),\ D_{lo} = D_s(d) - \Delta D_s(d)$ และ θ_Y คือพารามิเตอร์สำหรับกำหนดจำนวนสูงสุดของตั๋วสีเหลือง

สำหรับพารามิเตอร์ G_0 จะแตกต่างจากกฎฮิวริสติกเล็กน้อย

$$G_0 = \begin{cases} 1 & , D_{reg} > \Theta D_{hi} \\ \dfrac{\Theta(D_s(d) + \Delta D_s(d)) - D_{req}}{\Theta(D_s(d) + \Delta D_s(t)) - D_s(d)} \times \theta_G & , D_s(d) \le D_{req} \le \Theta D_{hi} \\ \dfrac{D_{req} - D_s(d) + \Delta D_s(d)}{\Delta D_s(d)} \times \theta_G & , D_{lo} \le D_{req} \le D_s(d) \\ 0 & , D_{req} < D_{lo} \end{cases} \tag{3.3}$$

เมื่อ θ_G คือพารามิเตอร์สำหรับกำหนดจำนวนสูงสุดของตั๋วสีเขียว ค่า $\Theta > 1$ใช้กำหนดระดับตั้งต้น (threshold) ที่เกินกว่า D_{req} ซึ่งใช้สำหรับค้นหาเส้นทางที่มีเวลาหน่วงสูงมาก

3.3.2 การคำนวณตั๋วตั้งต้น: วิธี TBP ภายใต้กระบวนการ ONMC

กระบวนการ ONMC สำหรับ POMDPs ถูกนำมาใช้ทั้งในระบบจริงหรือระบบจำลองเพื่อหานโยบายการแจกจ่ายตั๋วที่ดีในแง่ของสมดุลแลกเปลี่ยนระหว่างจำนวนของตั๋วที่ถูกแจกจ่ายไปกับความน่าจะเป็นในการค้นพบเส้นทางที่เหมาะสม นอกจากนี้วิธีนี้ยังมีการแทนที่การคำนวณค่า M_0 จากกฎฮิวริสติกในสมการที่ (3.2) และ (3.3) ด้วยการเลือก M_0 จากเซตจำกัดในลำดับของกระบวนการตัดสินใจแบบมาร์คอฟ

พิจารณาเซตของโหนดเคลื่อนที่ N ในเครือข่ายเคลื่อนที่แบบแอดฮอค แต่ละโหนดจะเก็บข้อมูลของเวลาหน่วงตลอดเส้นทางของทุกโหนดปลายทางไว้ สำหรับคู่ของโหนดต้นทาง-ปลายทาง (s,d) จะมีเซตของการเฝ้าสังเกตดังนี้

$$O_{sd} = \{[q_D(m), q_{\Delta D}(l)]: \ 1 \le m \le n, \ 1 \le l \le n_\Delta\}$$

เมื่อ n (n_Δ) คือช่วงของเวลาหน่วงตลอดเส้นทางแบบเต็มหน่วย และ $q_D(m)$ $(q_{\Delta D}(l))$ คือช่วงลำดับที่ m (l) บนช่วงจำนวน $[0,\infty)$

พิจารณา $o_k \in O_{sd}$ ที่เวลา k โหนด s จะเลือกการกระทำ $a_k \in A = \{0,...,M_{\max}\}$ เมื่อ $M_{\max}$ คือจำนวนตั๋วสูงสุดที่ยอมรับได้ โดยที่ตั๋วสีเขียวจะไม่ถูกพิจารณา $(G_0 = 0)$ และจะมีเพียงตั๋วสีเหลืองเท่านั้นที่ถูกพิจารณา จึงทำให้ $M_{\max} = \theta_Y$ เพื่อเป็นการเน้นความสำคัญไปที่การหาเส้นทางที่เหมาะสมมากกว่าการหาเส้นทางที่มีมูลค่าต่ำ ถ้ามีการค้นพบเส้นทางที่เหมาะสมอย่างน้อย 1 เส้นทาง จะได้รับผลตอบแทนเป็น $g(o_k, a_k)$ ถ้าไม่เช่นนั้นแล้ว การกระทำดังกล่าวจะถูกลงโทษ ซึ่งผลตอบแทนนิยามโดย

$$g(\cdot, a_k) = \begin{cases} \varsigma_j - \log a_k & , a_k > 0, X = \aleph \\ -(\varsigma_j - \log a_k) & , a_k > 0, X = 0 \\ -\log a_k & , a_k > 0, X > \aleph \\ 0 & , a_k = 0 \end{cases} \qquad (3.4)$$

เมื่อ $\varsigma_j \in R^+$ คือผลตอบแทนที่ได้รับสำหรับการบริการโหนด j X คือจำนวนเส้นทางการส่งที่ถูกค้นพบ และ $\aleph$ คือ จำนวนเส้นทางที่ต้องการสูงสุดจากเส้นทางที่ค้นพบ

ถ้าเส้นทางที่เหมาะสมที่ถูกพบมีหลายเส้นทาง โหนดปลายทาง d จะเลือกเส้นทางที่มีมูลค่าน้อยที่สุด จากนั้นจะส่งเมสเสจตอบกลับ (acknowledge message) ซึ่งประกอบด้วยค่าใหม่ของค่าเวลาหน่วงเฉลี่ยตลอดเส้นทาง ไปยังโหนดต้นทาง s ด้วยเส้นทางที่ถูกเลือกไว้ หลังจากได้รับเมสเสจตอบกลับแล้ว โหนด s จะอัพเดทข้อมูลเครือข่ายของตน นั่นคือการอัพเดท $D_s^{new}(d)$ และ $\Delta D_s^{new}(d)$ ที่คำนวณได้จากสมการ (3.1) กระบวนการดังกล่าวจะถูกกระทำซ้ำทุกเส้นทางการเชื่อมต่อที่ถูกร้องขอที่โหนดต้นทาง s จนกระทั่งเกิดการแลกเปลี่ยนของเวกเตอร์ระยะทางที่โหนด s ดังนั้นวัตถุประสงค์ของการใช้ OPMC เพื่อใช้กำหนดนโยบายที่ใกล้เคียงนโยบายที่ดีที่สุด $\pi: O_{sd} \to A$ ภายใต้กระบวนการเฝ้าสังเกต

3.3.3 พาธ แคชชิ่ง

วิธีพาธ แคชชิ่ง (path caching) ภายใต้กระบวนการ ONMC [11] ถูกนำมาใช้ในการหาเส้นทางสำหรับทุกการเชื่อมต่อที่ถูกร้องขอ เพื่อหลีกเลี่ยงการร้องขอเส้นทางที่ถี่เกินไป แต่ละโหนดจึงมีการเก็บรักษาเส้นทางไว้ [13], [14] ดังนั้นกลยุทธ์พาธ แคชชิ่งจึงช่วยลดโอเวอร์เฮด (overhead) ในเครือข่าย MANETs ได้ ดังนั้นในงานวิจัยฉบับนี้จึงมีการใช้กลยุทธ์พาธ แคชชิ่ง ซึ่งถูกสนับสนุนด้วยวิธี TBP โดยพาธ แคช (path cache) เป็นเซตของเส้นทางซ้ำซ้อนที่ถูกค้นพบด้วยวิธี TBP ขนาดของพาธ แคชขึ้นอยู่กับระดับความซ้ำซ้อนของเส้นทางที่ต้องการ

3.4 การทดสอบและวิเคราะห์ผล

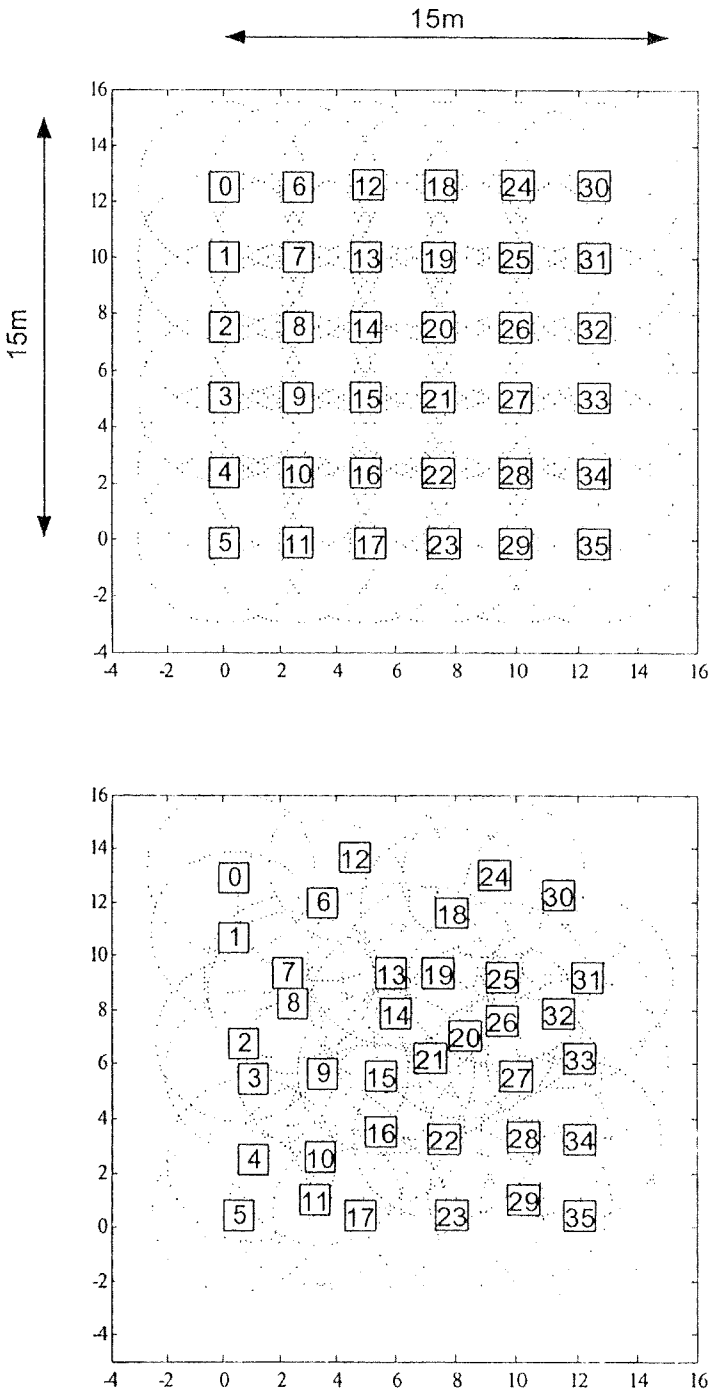
ประสิทธิภาพสำหรับการประยุกต์วิธี TBP ภายใต้กระบวนการ ONMC ในเครือข่ายเคลื่อนที่แบบแอดฮอคจะถูกประเมินด้วยแบบจำลองระบบในคอมพิวเตอร์ งานวิจัยฉบับนี้ทำการทดสอบประสิทธิภาพของระบบในหลายด้านด้วยกัน กล่าวคือ

1) ผลตอบแทนสะสม ซึ่งเป็นค่าที่ได้จากผลตอบแทนสะสมจากทุกเอพพิโซดหารด้วยจำนวนเอพพิโซดทั้งหมด
2) อัตราความสำเร็จ ซึ่งเป็นค่าที่ได้จากจำนวนการเชื่อมต่อที่ตอบรับทั้งหมดหารด้วยจำนวนการร้องขอการเชื่อมต่อทั้งหมด
3) มูลค่าเส้นทางโดยเฉลี่ย ซึ่งเป็นค่าที่ได้จากมูลค่าทั้งหมดจากการสร้างเส้นทางเชื่อมต่อหารด้วยจำนวนการเชื่อมต่อที่ถูกสร้างขึ้น
4) จำนวนเมสเสจค้นหาโดยเฉลี่ย ซึ่งเป็นค่าที่ได้จากจำนวนของเมสเสจค้นหาทั้งหมดที่ถูกส่งไปหารด้วยจำนวนการร้องขอการเชื่อมต่อทั้งหมด

พิจารณาโหนดเคลื่อนที่ 36 โหนดในเครือข่ายเคลื่อนที่แบบแอดฮอคบนพื้นที่15x15 เมตร^2 รูปร่างเครือข่ายจะถูกสุ่มอย่างสม่ำเสมอด้วยแบบจำลองการเคลื่อนที่ ความเร็วของโหนดเคลื่อนที่มีค่าอยู่ระหว่าง 0.3-0.7 เมตร/วินาที แต่ละโหนดมีรัศมีการส่ง 3 เมตร ลิงค์เชื่อมต่อ (link) จะถูกสร้างขึ้นระหว่าง 2 โหนดเคลื่อนที่ใดๆที่อยู่ภายใต้รัศมีการส่งที่กำหนด

การร้องขอการเชื่อมต่อจะถูกสร้างขึ้นที่โหนดต้นทางด้วยอัตราการเชื่อมต่อ 0.2 วินาที มูลค่าของแต่ละลิงค์จะมีค่าสม่ำเสมออยู่ในช่วง [0,1] แต่ละลิงค์เชื่อมต่อระหว่างโหนด i และโหนด j จะมีเวลาหน่วง 2 แบบเกิดขึ้น คือ ดีเวลาหน่วงเฉลี่ยที่เกิดขึ้นจริง (D_{ij}) และเวลาหน่วงเฉลี่ยสำหรับการ

ประกาศ ($\tilde{D}_{ij}$) เวลาหน่วงในแบบหลังจะถูกประกาศไปที่เครือข่ายและนำมาใช้เพื่อคำนวณค่าเวลาหน่วงเฉลี่ยตลอดเส้นทาง $D_j(d)$ สำหรับทุกโหนด j และ d ในเครือข่ายเคลื่อนที่แบบแอดฮอค เวลาหน่วงเฉลี่ยที่เกิดขึ้นจริงในแต่ละลิงค์จะถูกสร้างขึ้นอย่างสม่ำเสมอในช่วง [0,50] มิลลิวินาที

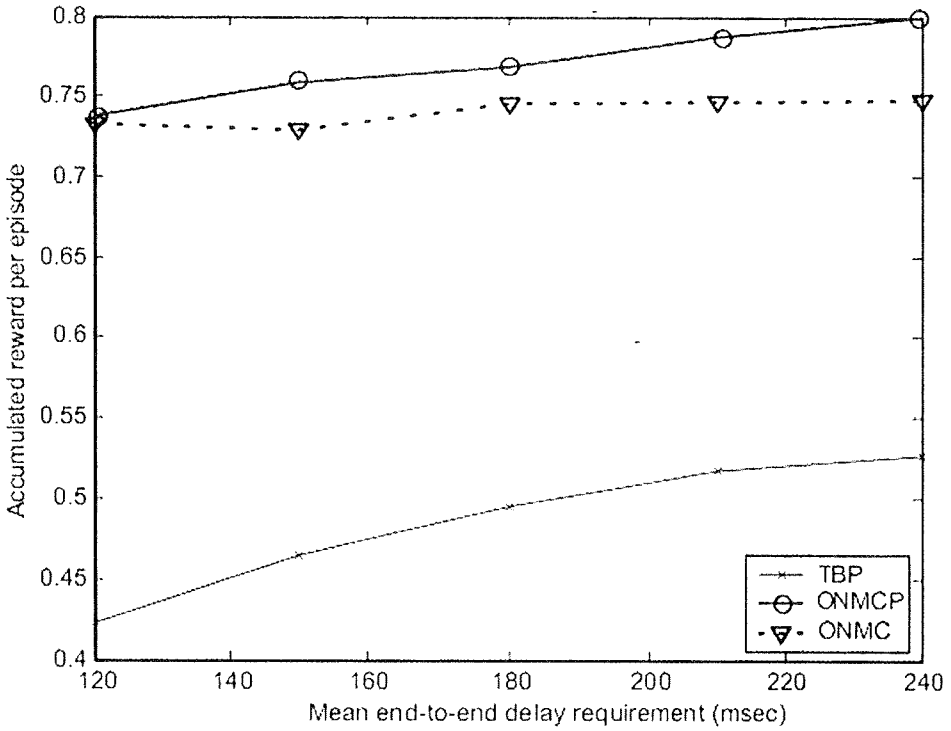


รูปที่ 3.1 แบบจำลองเครือข่ายโหนดเคลื่อนที่ 36 โหนดในเครือข่ายเคลื่อนที่แบบแอดฮอคบนพื้นที่ 15x15 เมตร^2 แต่ละโหนดมีรัศมีการส่ง 3 เมตรแสดงโดยวงกลมเส้นประ รูปบนแสดงความสัมพันธ์เริ่มต้นของโหนดเคลื่อนที่ รูปล่างแสดงถึงความสัมพันธ์โหนดเคลื่อนที่ที่เปลี่ยนไปหลังจากเริ่มการทดสอบ

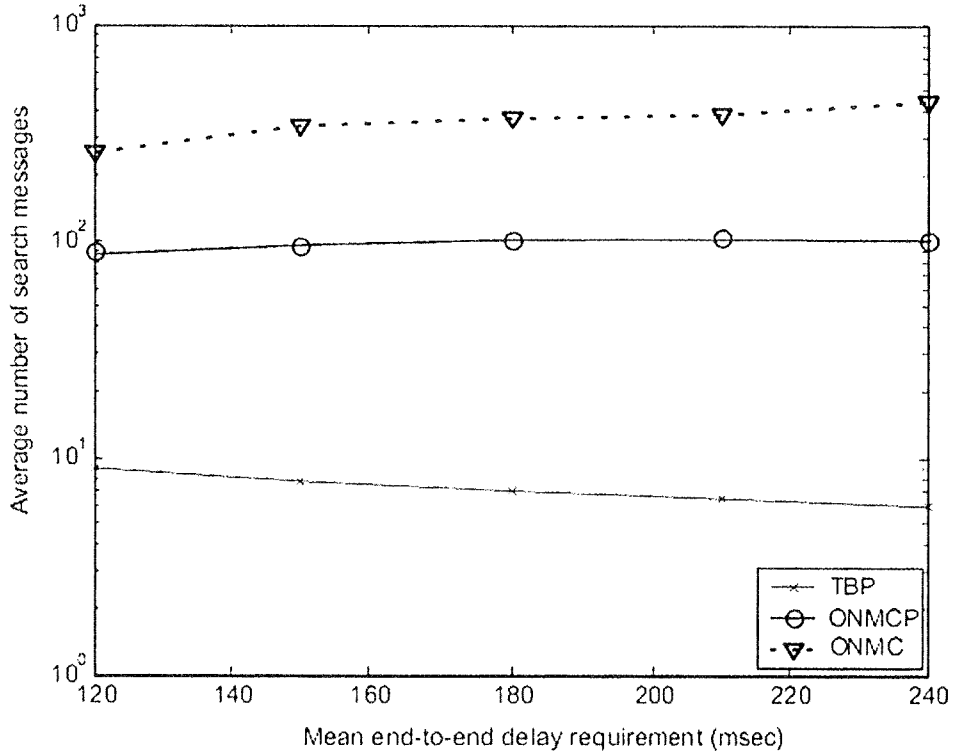
กระบวนการเลือกเส้นทาง 3 วิธีถูกนำมาทดสอบประสิทธิภาพ คือ วิธีTBP วิธี TBP ภายใต้กระบวนการ ONMC และวิธี TBP ภายใต้กระบวนการ ONMC ด้วยพาธ แคชชิ่ง (ONMCP) ทั้ง 3 กระบวนการนี้จะพิจารณาเพียงตั๋วสีเหลืองโดยมี $M_{\max} = \theta_Y = 100$ ใบ เพื่อเน้นความสำคัญในการหาเส้นทางที่เป็นไปได้มากว่าการหาเส้นทางที่มีมูลค่าต่ำ นอกจากนี้การร้องขอการเชื่อมต่อจะถูกปฏิเสธทันทีถ้าเวลาหน่วงเฉลี่ยตลอดเส้นทางที่ร้องขอมีค่ามากกว่าเวลาหน่วงตลอดเส้นทางสูงสุดที่เป็นไปได้

กำหนดให้เซตของการกระทำ (เมื่อการร้องขอการเชื่อมต่อไม่ถูกปฏิเสธ) สำหรับทุกอัลกอริธึมกำหนดโดย $M_0 \in A = \{1,10,20,...,100\}$ เวลาหน่วงเฉลี่ยตลอดเส้นทางและการเปลี่ยนแปลงของเวลาหน่วง (ในหน่วย มิลลิวินาที) $q_D(m) \in \{[0,10),\ [10,20),\ ...\ ,[250,\infty)\}$ และ $q_{\Delta D}(l) \in \{[0,10),[10,\infty)\}$ เมื่อ m = 1,...,26 l =1,2 และ $q_D(m)$ คือช่วงการแบ่งที่ m ของเวลาหน่วงเฉลี่ยตลอดเส้นทางระหว่างโหนด s และโหนด d และ $q_{\Delta D}(l)$ คือช่วงการแบ่งที่ l ของการเปลี่ยนแปลงของเวลาหน่วงเฉลี่ยตลอดเส้นทางระหว่างสองโหนดเคลื่อนที่ วิธี ONMC และ ONMCP จะถูกฝึกทั้งหมด 4×10^6 ครั้งของการร้องขอการเชื่อมต่อ หลังจากนั้นจึงจะนำมาทดสอบประสิทธิภาพและเปรียบเทียบกับวิธี TBP โดยทุกอัลกอริธึมจะถูกทดสอบด้วยการรันทั้งหมด 1×10^6 ครั้งของการร้องขอการเชื่อมต่อ

รูปที่ 3.2 แสดงให้เห็นว่าผลตอบแทนเฉลี่ยสะสมต่อเอพพิโซดจะเพิ่มขึ้นเมื่อความต้องการของเวลาหน่วงเฉลี่ยตลอดเส้นทางเพิ่มขึ้น ที่เป็นเช่นนี้เพราะ เมื่อความต้องการของเวลาหน่วงเฉลี่ยตลอดเส้นทางเพิ่มขึ้น ทำให้ง่ายขึ้น ดังนั้นผลตอบแทนเฉลี่ยสะสมจึงมีค่าเพิ่มขึ้น นอกจากนี้เมื่อความต้องการของเวลาหน่วงเฉลี่ยตลอดเส้นทางเพิ่มขึ้น จะเห็นว่าวิธี ONMCP มีประสิทธิภาพดีกว่าวิธี ONMC เนื่องจากมีตั๋วจำนวนน้อยมากถูกแจกจ่ายไปเมื่อพาท แคชชิ่งถูกนำมาใช้ รูปที่ 3.3 ชี้ให้เห็นว่าวิธี TBP ให้จำนวนเมสเสจค้นหาเฉลี่ยน้อยที่สุดแม้แต่บริเวณที่มีค่าผลตอบแทนสะสมต่ำสุดต่อเอพพิโซด วิธี ONMCP มีจำนวนเมสเสจค้นหาเฉลี่ยน้อยกว่าวิธี ONMC เนื่องมาจากการใช้งานพาท แคชชิ่ง

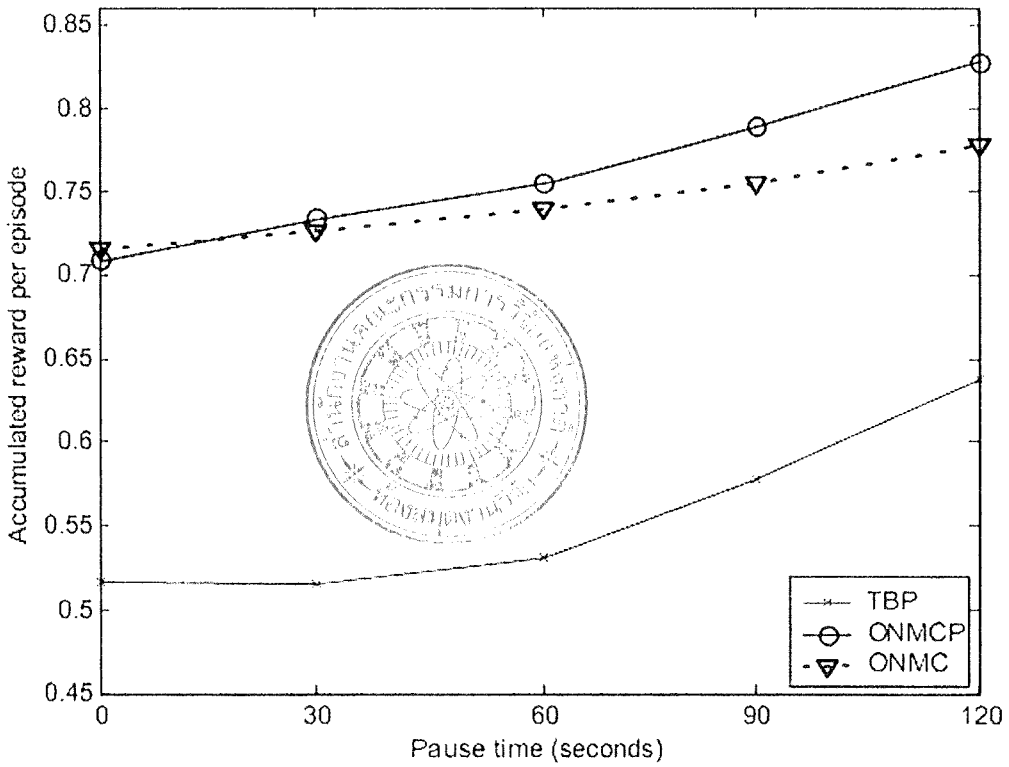


รูปที่ 3.2 ผลตอบแทนเฉลี่ยสะสมต่อเอพพิโซดที่อัตราความคลุมเครือ 0.5

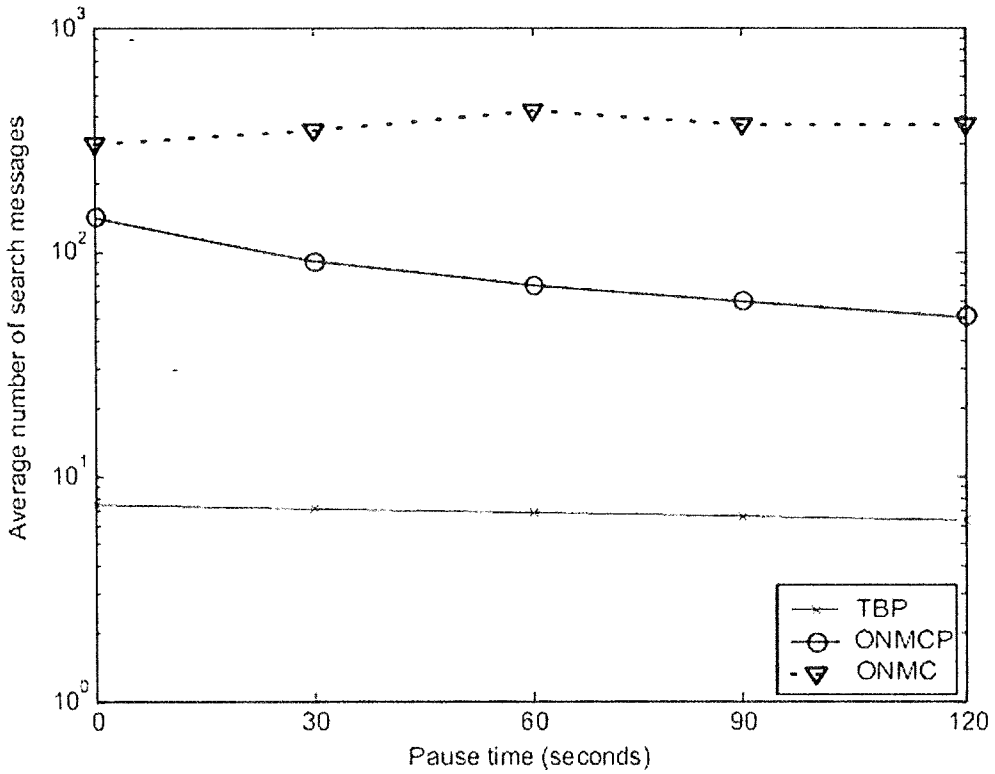


รูปที่ 3.3 จำนวนของเมสเสจค้นหาเฉลี่ยที่อัตราความคลุมเครือ 0.5

การทดลองสุดท้ายแสดงถึงผลกระทบจากการเคลื่อนที่ของทุกอัลกอริธึม (โดยการเพิ่มเวลาเพื่อให้โหนดอยู่ในสภาวะคงที่ซึ่งเรียกว่า เวลาหยุดพักชั่วขณะ หรือ pause time) ภายใต้ความต้องการของเวลาหน่วงเฉลี่ยตลอดเส้นทางคงที่ รูปที่ 3.4 แสดงให้เห็นว่าวิธี ONMCP และวิธี ONMC มีความแตกต่างของผลตอบแทนเฉลี่ยสะสมต่อเอพพิโซดเพียงเล็กน้อยเท่านั้น ส่วนวิธี TBP ให้ผลตอบแทนเฉลี่ยสะสมต่อเอพพิโซดน้อยที่สุด ทั้งนี้เนื่องมาจากเมื่อโหนดอยู่ในสภาวะคงที่นานขึ้นทำให้ง่ายต่อการค้นหาเส้นทาง ดังนั้นผลตอบแทนเฉลี่ยสะสมต่อเอพพิโซดของทุกอัลกอริธึมจึงเพิ่มขึ้นนั่นเอง รูปที่ 3.5 ชี้ให้เห็นว่าวิธี ONMCP และ ONMC มีจำนวนเมสเสจค้นหาเฉลี่ยลดลงเนื่องจากเส้นทางการส่งที่เป็นไปได้ถูกค้นพบง่ายขึ้นอันเนื่องมาจากโหนดอยู่ในสภาวะคงที่นานขึ้น ดังนั้นทั้งสองวิธีจึงเรียนรู้การแจกจ่ายตั๋วให้น้อยลงเพื่อให้มีการใช้จำนวนของเมสเสจค้นหาน้อยที่สุด อย่างไรก็ตามจะเห็นว่าวิธี ONMCP มีการสร้างจำนวนเมสเสจค้นหาเฉลี่ยน้อยที่สุดเนื่องจากการใช้พาท แคชซึ่งสามารถหลีกเลี่ยงความถี่ในการร้องขอการค้นหาเส้นทางได้



รูปที่ 3.4 ผลตอบแทนเฉลี่ยสะสมต่อเอพพิโซดที่ช่วงเวลาหยุดพักชั่วขณะที่แตกต่างกัน



รูปที่ 3.5 จำนวนของเมสเสจค้นหาเฉลี่ยที่ช่วงเวลาหยุดพักชั่วขณะที่แตกต่างกัน

3.5 สรุป

งานวิจัยฉบับนี้นำเสนอวิธี TBP ภายใต้กระบวนการเรียนรู้แบบรีอินฟอร์สเมนท์และพาธ แคชชิ่ง ซึ่งเรียกกระบวนการนี้ว่า กระบวนการ ONMCP โดยนำไปประยุกต์ใช้ในการค้นหาเส้นทางที่มีการรับรองคุณภาพการบริการสำหรับเครือข่ายเคลื่อนที่แบบแอดฮอค ผลการจำลองระบบแสดงให้เห็นว่ากระบวนการที่ได้นำเสนอสามารถได้มาซึ่งนโยบายในการแจกจ่ายตั๋วที่ดีในด้านของผลตอบแทนสะสมต่อเอพพิโซด เมื่อเทียบกับวิธี TBP ดั้งเดิมและวิธี ONMC นอกจากนี้วิธี ONMCP ยังสามารถลดโอเวอร์เฮดจากการค้นหาเส้นทางอันเนื่องมาจากการใช้กลยุทธ์พาธ แคชชิ่งอีกด้วย