การประยุกต์ใช้การเรียนรู้แบบการเสริมกำลัง ในงานปฏิบัติการทางการบินในสภาวะไม่ปกติ เพื่อแก้ปัญหาการปรับตารางเวลาเที่ยวบิน

นาย โอภาส เพลินประภาพร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2552

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

APPLYING REINFORCEMENT LEARNING TO IRREGULAR FLIGHT OPERATION TO SOLVE FLIGHT RESCHEDULING PROBLEM

Mr. Opat Plernprapaporn

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science Program in Computer Science

Department of Computer Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2009

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การประยุกต์ใช้การเรียนรู้แบบการเสริมกำลัง ในงานปฏิบัติการทางการบินในสภาวะไม่ปกติ เพื่อแก้ปัญหาการปรับตารางเวลาเที่ยวบิน
โดย	นาย โอภาส เพลินประภาพร
สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร.โปรดปราน บุณยพุกกณะ

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

..... คณบดีคณะวิศวกรรมศาสตร์

(รองศาสตราจารย์ ดร.บุญสม เลิศหิรัญวงศ์)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ

(ศาสตราจารย์ ดร.ประภาส จงสถิตย์วัฒนา)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(ผู้ช่วยศาสตราจารย์ ดร.โปรดปราน บุณยพุกกณะ)

..... กรรมการ

(ศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล)

..... กรรมการภายนอกมหาวิทยาลัย

(รองศาสตราจารย์ ดร.วีระ บุญจริง)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

โอภาส เพลินประภาพร : การประยุกต์ใช้การเรียนรู้แบบการเสริมกำลัง ในงานปฏิบัติการทางการบินในสภาวะไม่ปกติ เพื่อแก้ปัญหาการปรับตารางเวลาเที่ยวบิน. (APPLYING REINFORCEMENT LEARNING TO IRREGULAR FLIGHT OPERATION TO SOLVE FLIGHT RESCHEDULING PROBLEM) อ. ที่ปรึกษาวิทยานิพนธ์หลัก: ผู้ช่วยศาสตราจารย์ ดร. โปรดปราน บุณยพุกกณะ, 85 หน้า.

ในขณะที่งานวิจัยส่วนใหญ่พยายามที่จะแก้ปัญหาความผิดปกติที่เกิดขึ้นในงานปฏิบัติการทางการบินด้วยต้นแบบทางคณิตศาสตร์ หรือ ทฤษฎีการไหลในเครือข่าย แต่ยังมีงานวิจัยเพียงส่วนน้อยที่ได้ประยุกต์วิธีการเรียนรู้ที่เป็นประโยชน์ในกรณีดังกล่าวนี้ เนื่องจากจะทำให้ได้ผลลัพธ์แบบทันกาล วิทยานิพนธ์นี้ได้ประยุกต์ใช้การเรียนรู้ที่เพื่อแก้ปัญหาการปรับตารางเวลาเที่ยวบินสำหรับงานปฏิบัติการภายในประเทศ โดยผลการทดลองสรุปได้ว่าการเรียนรู้ที่ให้ผลลัพธ์ใกล้เคียงกับวิธีขยายและจำกัดเขตภายใต้สภาพแวดล้อมเดียวกันและเสนอผลเฉลยที่แบบทันกาลอีกด้วย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชาวิศวกรรมคอมพิวเตอร์ ลายมือชื่อนิสิต.....
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....
ปีการศึกษา 2552

4970721221 : MAJOR COMPUTER SCIENCE

KEYWORDS: Q-LEARNING / ACTION / REWARD / FLIGHT SCHEDULE / IRREGULAR SITUATION

OPAT PLERNPRAPAPORN: APPLYING REINFORCEMENT LEARNING TO IRREGULAR FLIGHT OPERATION TO SOLVE FLIGHT RESCHEDULING PROBLEM. THESIS ADVISOR: ASSISTANT PROFESSOR PROADPARN PUNYABUKKANA, Ph.D., 85 pp.

While most researches attempt to solve flight operation irregularity with mathematieal model or network flow theory, but only few researches have applied Q-Learning method which can be beneficial to this particular case since it offers solution at real-time. This thesis aims at applying Q-Learning to solve flight rescheduling problem for domestic operation. The experiment suggested that the performance of Q-Learning is comparable to that of branch-and-bound in term of results. Furthermore, it was able to offer the solution at real-time.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Department: Computer Engineering Student's Signature โอภาส เพ็ญประภา
Field of Study: Computer Science Advisor's Signature
Academic Year: 2009

กิตติกรรมประกาศ

ขอขอบพระคุณ บิดา มารดา ที่เป็นกำลังใจสำคัญในการทำวิทยานิพนธ์ พร้อมทั้งให้ความช่วยเหลือ และสนับสนุนในทุกๆ ด้านแม้ยามเจ็บป่วยท่านก็เป็นกำลังใจให้เสมอ จนทำให้วิทยานิพนธ์ฉบับนี้สำเร็จลงได้

ขอขอบพระคุณอาจารย์ที่ปรึกษา ผศ.ดร.โปรดปราน บุณยพุกกณะ พิตรสาธร ที่เป็นผู้ให้แนวคิด และให้คำปรึกษา จนทำให้วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วง ขอขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ ศ.ดร.ประภาส จงสถิตย์วัฒนา ศ.ดร.บุญเสริม กิจศิริกุล และ รศ.ดร.วีระ บุญจริง ที่ได้ช่วยให้ความคิดเห็น และเสนอแนะข้อแก้ไขต่างๆ รวมถึง ผศ.ดร.มาโนช โลหเตปานนท์ ที่มาสอนปัญหาต่างๆ ในสายการบิน และการแก้ไขปัญหาเหล่านั้นที่เกิดขึ้น เพื่อทำให้วิทยานิพนธ์ฉบับนี้สมบูรณ์ ถ้าหากปราศจากความร่วมมือจากพวกท่านเหล่านี้วิทยานิพนธ์ฉบับนี้คงไม่อาจสำเร็จได้ และขอขอบคุณบุคลากร รวมถึงเจ้าหน้าที่ที่การบินไทยที่ให้การต้อนรับนำเสนอปัญหาและข้อมูลที่เป็นประโยชน์ต่อการวิจัย ขอขอบคุณเพื่อนๆ ที่ให้กำลังใจเสมอมาทั้งในยากทุกข์ และยามสุข

และท้ายที่สุดนี้ขอขอบพระคุณนักวิจัยทั้งหลายที่พยายามคิดค้น ค้นหาสิ่งใหม่ๆ การแก้ไขปรับปรุง และการพัฒนาที่ไม่เคยหยุด เพื่อเป็นแนวทางในการดำเนินงานวิจัยของคนรุ่นหลังต่อไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญ ..... ช
สารบัญตาราง ..... ฌ
สารบัญภาพ ..... ญ
บทที่ 1 1
1.1 ความเป็นมาและความสำคัญของปัญหา 1
1.2 วัตถุประสงค์ของการวิจัย 2
1.3 ขอบเขตการวิจัย 2
1.4 ขั้นตอนและวิธีการดำเนินงานวิจัย 3
1.5 ประโยชน์ที่คาดว่าจะได้รับ 3
บทที่ 2 4
2.1 ทฤษฎีที่เกี่ยวข้อง 4
2.1.1 การจัดตารางเวลาเที่ยวบินภายในประเทศของสายการบิน 4
2.1.2 ปัญหาและการแก้ไขความผิดปกติในตารางเวลาเที่ยวบิน 7
2.1.3 การเรียนรู้แบบการเสริมกำลังและการเรียนรู้คิว 8
2.2 การทบทวนวรรณกรรม และงานวิจัยที่เกี่ยวข้อง 9
บทที่ 3 12
3.1 การประยุกต์การเรียนรู้คิว 12
3.1.1 สถานะ 12
3.1.1.1 การนิยามกลุ่มของเครื่องบินที่ถูกใส่ให้แต่ละเที่ยวบิน 14
3.1.1.2 การนิยามความผิดปกติ 17
3.1.2 การกระทำ 20
3.1.3 รางวัล 24
3.1.4 การลู่เข้า 25
3.1.5 ฟังก์ชันการประมาณ 26
บทที่ 4 29
4.1 การทดลองการแสดงสถานะในตารางเวลาเที่ยวบิน 29
4.2 การทดลองการปรับตารางเวลาเที่ยวบินใหม่และคำนวณค่ารางวัล 31
4.3 การทดลองการลู่เข้า 36

4.4 การทดลองประยุกต์การเรียนรู้คิวในตารางเวลาเที่ยวบิน 38
4.5 การทดลองเปรียบเทียบผลของการเรียนรู้คิว 41
4.6 การทดลองหาเวลาที่ใช้ในการปรับตารางเวลาเที่ยวบิน 49
4.7 การวิเคราะห์ผล 50
บทที่ 5 53
5.1 สรุปผลการวิจัย 53
5.2 ข้อเสนอแนะ 54
รายการอ้างอิง 56
ภาคผนวก 58
ภาคผนวก ก ผลงานการตีพิมพ์ 59
ภาคผนวก ข ข้อมูลตารางเวลาเที่ยวบินที่ใช้ในการเรียนรู้คิว 66
ภาคผนวก ค ตารางค่า TOTAL_DELAY_TIME จากการเรียนรู้คิว 82
ประวัติผู้เขียนวิทยานิพนธ์ 85

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

ตาราง หน้า

ตารางที่ 1 ตัวอย่างตารางเวลาเที่ยวบินภายในประเทศ..... 5
ตารางที่ 2 ตัวอย่างตารางเวลาเที่ยวบินภายหลังจากทำการจับคู่และใช้สัญลักษณ์แทนเครื่องบิน.....16
ตารางที่ 3 การปรับตารางเวลาตามการกระทำที่กำหนดเพื่อคำนวณหาค่า *TOTAL_DELAY_TIME* 32
ตารางที่ 4 $\hat{Q}(s,a)$ ที่ได้จากการปรับปรุงค่าในแต่ละรอบ.....35
ตารางที่ 5 เปรียบเทียบค่า *TOTAL_DEAY_TIME* จากครั้งที่ให้ผลการทดลองดีที่สุดเทียบกับค่าเฉลี่ยของการทดลองทั้ง 10 ครั้ง.....39
ตารางที่ 6 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับ เมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 1 44
ตารางที่ 7 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับ เมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 2 44
ตารางที่ 8 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับ เมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 3 45
ตารางที่ 9 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับ เมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 4 45
ตารางที่ 10 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับ เมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 5 46
ตารางที่ 11 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับ เมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 6 46
ตารางที่ 12 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับ เมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 7 47
ตารางที่ 13 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับ เมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 8 47
ตารางที่ 14 เปรียบเทียบผลที่ได้จากวิธีการเรียนรู้คิวและวิธีขยายและจำกัดเขต.....48
ตารางที่ 15 เวลาที่ใช้ในการปรับตารางเวลาเที่ยวบินใหม่ด้วยการใช้ข่ายงานประสาทเป็นฟังก์ชันการประมาณ ..49

จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

ภาพประกอบ หน้า

รูปที่ 1 ตัวอย่างการแสดงเที่ยวบินในรูปของกำหนดการการหมุนเวียนเครื่องบิน 6
รูปที่ 2 ขั้นตอนวิธีการเรียนรู้คิว 9
รูปที่ 3 การกำหนดสถานะให้แต่ละเที่ยวบินในตารางเวลา 13
รูปที่ 4 แถบสีดำแสดงความผิดปกติที่เกิดขึ้นกับเครื่องบินมีผลทำให้เครื่องบินหมายเลขเครื่องดังกล่าวไม่สามารถใช้งานได้ 13
รูปที่ 5 ตัวอย่างการแสดงตารางเวลาเที่ยวบินในรูปของกำหนดการภายหลังจากการใช้วิธีการจับคู่เที่ยวบินและใช้สัญลักษณ์แทนเครื่องบิน 16
รูปที่ 6 การแทรกความผิดปกติเข้าไปในตารางเวลาเที่ยวบินเพื่อใช้เป็นข้อมูลฝึกและข้อมูลทดสอบ 19
รูปที่ 7 ความผิดปกติที่เกิดขึ้นเป็นผลให้ต้องมีการปรับตารางเวลาใหม่ โดยพิจารณาจากเงื่อนไขบังคับของตารางเวลาเที่ยวบิน 21
รูปที่ 8 กราวนด์ไทม์ของเครื่องบินในแต่ละแบบเครื่อง 22
รูปที่ 9 ตัวอย่างการปรับตารางเวลาใหม่โดยพิจารณาการเปลี่ยนเครื่องบินในแบบเครื่องเดียวกัน 23
รูปที่ 10 ตัวอย่างการปรับตารางเวลาใหม่โดยพิจารณาการเปลี่ยนเครื่องบินในแบบเครื่องที่ต่างกัน 23
รูปที่ 11 แผนภูมิสายงานแสดงการทำงานของการประยุกต์การเรียนรู้คิวเพื่อแก้ปัญหาการปรับตารางเวลาเที่ยวบิน 28
รูปที่ 12 ตารางเวลาเที่ยวบินของเครื่องบินสัญลักษณ์ aircraft_1 และ aircraft_2 พร้อมความผิดปกติในเครื่องบินสัญลักษณ์ aircraft_2 ที่เวลา 16:00 น. 29
รูปที่ 13 ตารางเวลาเที่ยวบินของเครื่องบินสัญลักษณ์ aircraft_3 และ aircraft_4 พร้อมความผิดปกติในเครื่องบินสัญลักษณ์ aircraft_3 ที่เวลา 10:00 น. 32
รูปที่ 14 รากกำลังสองเฉลี่ยจากการทดลองที่ 4.2 36
รูปที่ 15 รากกำลังสองเฉลี่ยจากการทดลองที่ 4.3 38
รูปที่ 16 เปรียบเทียบค่าเฉลี่ยของ *TOTAL_DELAY_TIME* จากการทดลองทั้ง 10 ครั้ง 41
รูปที่ 17 การเข้ารหัสชั้นอินพุต 42
รูปที่ 18 การเข้ารหัสชั้นเอาต์พุต 42
รูปที่ 19 การใช้ข่ายงานประสาทแบบแพร่กระจายย้อนกลับในการประมาณค่าคิว 43
รูปที่ 20 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่ได้จากการทดลอง 48

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การเรียนรู้แบบการเสริมกำลัง (Reinforcement learning) คือ ขั้นตอนวิธีการเรียนรู้ของเครื่องที่พบว่ามีประสิทธิภาพในการประยุกต์ใช้งานด้านต่างๆมากมาย ตัวอย่างการประยุกต์ใช้งาน เช่น งานของ Tesauro เป็นงานที่เขานำเสนอการใช้การเรียนรู้แบบเสริมกำลังในเกมแบคแกมมอน (backgammon game) [1] และงานของ Zhang และ T. G. Dietterich ซึ่งทำให้เกิดผลในการจัดกำหนดการงาน (job-shop scheduling) [2]

โดยทั่วไปการเรียนรู้แบบการเสริมกำลังเรียนรู้ที่จะเลือกการกระทำ (Action) ที่เหมาะสมที่จะกระทำลงไปเพื่อให้ถึงจุดหมาย การเรียนรู้แบบนี้จะจัดปัญหาให้อยู่ในรูปของสถานะ (State) และการกระทำ เมื่อทำการประยุกต์การกระทำใดๆ ลงไปในแต่ละสถานะ ก็จะได้ค่ารางวัล (Reward) คืนมา เป้าหมายของระบบการเรียนรู้ระบบนี้คือ การหานโยบายควบคุม (Control policy) ที่ทำให้ได้ค่ารางวัลมากที่สุด ในงานวิจัยนี้จึงขอแนะนำการประยุกต์ใช้การเรียนรู้คิว (Q-learning) อีกรูปแบบหนึ่งของการเรียนรู้แบบการเสริมกำลัง เพื่อใช้ในการแก้ปัญหาการปรับตารางเวลาเที่ยวบินใหม่เมื่อมีความผิดปกติเกิดขึ้น

ในการปฏิบัติงานทางการบินในแต่ละวันบ่อยครั้งที่สายการบินต้องเผชิญกับสถานการณ์ต่างๆ ที่ส่งผลต่อการวางแผนในการปฏิบัติงาน สถานการณ์เหล่านี้เรียกกันว่าความผิดปกติ (irregularity) ซึ่งรวมถึงเหตุการณ์ไม่ได้คาดคิดต่างๆ เช่น สภาพอากาศเลวร้าย, เครื่องบินขัดข้อง หรือเกิดเหตุผิดพลาดเกี่ยวกับบุคลากร การปฏิบัติงานเมื่อมีความผิดปกติเกิดขึ้นในตารางเวลาเที่ยวบิน ทำให้มีความจำเป็นที่จะต้องเปลี่ยนแผนการบินจากที่ได้วางแผนไว้ และยังเป็นผลให้มีการปรับตารางเวลาเที่ยวบินใหม่ และ/หรือ เปลี่ยนเส้นทางเที่ยวบิน โดยในงานนี้ให้ความสนใจเฉพาะการปรับตารางเวลาเที่ยวบินที่เครื่องบินจะทำการเปลี่ยนเครื่องบินภายในแบบเครื่องเดียวกัน และการเปลี่ยนเครื่องบินในแบบเครื่องที่ต่างกัน ส่วนความผิดปกติจะพิจารณาว่าเป็นความผิดปกติใดๆ ที่ทำให้เครื่องบินไม่สามารถใช้งานได้ในเวลาดังกล่าว จากการปรับเปลี่ยนอย่างที่ได้กล่าวย่อมทำให้เกิดเวลาล่าช้า (delay time) เป้าหมายของงานวิจัยนี้คือ ปรับตารางเวลาเที่ยวบินให้ได้เวลาล่าช้าน้อยที่สุดในเวลาที่สั้นที่สุด เพื่อให้ระบบปฏิบัติการทางการบินสามารถดำเนินต่อไปได้

งานวิจัยส่วนใหญ่ที่เกี่ยวกับงานปฏิบัติการทางการบินจะใช้งานต้นแบบทางคณิตศาสตร์ (mathematical model) หรือทฤษฎีการไหลในเครือข่าย (network flow theory) ในการแก้ปัญหา ตัวอย่างเช่น งานของ Clarke ในปี ค.ศ.1995 เขาได้แนะนำปัญหาปัญหานี้และผลของความผิดปกติที่

ส่งผลต่อการปฏิบัติการ [3] ต่อมาในปี ค.ศ.1996 เขาได้แสดงต้นแบบทางคณิตศาสตร์ที่ใช้การแก้ปัญหาแบบศึกษาสำนึก (heuristic) เข้าช่วย [4] ยังมีงานที่เกี่ยวข้องอื่นๆ เช่น งานของ Jarrah และคณะ ในปี ค.ศ.1993 [5] และงานของ Mathaisel ในปี ค.ศ.1996 [6]

ถึงอย่างไรก็ดี มีงานวิจัยเพียงส่วนน้อยที่ทราบถึงการประยุกต์ใช้งานการเรียนรู้แบบการเสริมกำลังในงานปฏิบัติการทางการบิน ซึ่งก็อาจจะเป็นผลมาจากข้อเท็จจริงที่ว่า การแก้ปัญหาด้วยการเรียนรู้ของเครื่องนั้นไม่เหมาะสมเมื่อต้องการผลในแบบทันกาล (real time) อย่างไรก็ตามการเรียนรู้คิวอีกรูปแบบของการเรียนรู้แบบการเสริมกำลังไม่จำเป็นต้องรู้รูปแบบของสิ่งแวดล้อมของมัน และในงานนี้ก็พบว่า สามารถประยุกต์การเรียนรู้คิวได้โดยการเปลี่ยนรูปแบบของการปรับตารางเวลาเที่ยวบินให้อยู่ในรูปแบบของการเปลี่ยนสถานะ (state transition) ที่มีการกระทำที่เหมาะสมในแต่ละสถานะ

1.2 วัตถุประสงค์ของการวิจัย

เพื่อนำเสนอการประยุกต์ใช้งานการเรียนรู้เรียนรู้แบบการเสริมกำลังโดยเฉพาะอย่างยิ่งการเรียนรู้คิว เพื่อแก้ปัญหาการปรับตารางเวลาเที่ยวบินในงานปฏิบัติการทางการบินในสภาวะไม่ปกติ ซึ่งคลอบคุลมการเปลี่ยนเครื่องบินภายในแบบเครื่องเดียวกัน และการเปลี่ยนเครื่องบินในแบบเครื่องที่ต่างกัน

1.3 ขอบเขตการวิจัย

1. นำเสนอการประยุกต์ในงานการเรียนรู้คิว
2. ทดสอบการประยุกต์การเรียนรู้คิวในงานปฏิบัติการทางการบินในสภาวะไม่ปกติ ข้อมูลที่ใช้ในการเรียนรู้นำมาจากตารางเวลาเที่ยวบินภายในประเทศ โดยใส่ความผิดปกติเข้าไปในตารางเวลาเป็นจำนวน 56 ตัวอย่าง เพื่อใช้ในการสอนการเรียนรู้และการทดสอบการเรียนรู้ต่อไป
3. การปรับตารางเวลาเที่ยวบิน
   - 3.1 การเปลี่ยนเครื่องบินนั้น จะคำนึงถึงการเปลี่ยนเครื่องบินภายในแบบเครื่องเดียวกัน และการเปลี่ยนเครื่องบินในแบบเครื่องที่ต่างกัน
   - 3.2 ในการปรับตารางเวลาเที่ยวบินแต่ละครั้งจะต้องคำนึงถึงข้อจำกัดของตารางเวลาเที่ยวบิน และกราวนด์ไทม์ของแต่ละแบบเครื่อง
4. ทำการสอนการเรียนรู้คิวโดยการนำตัวอย่างความผิดปกติจำนวน 50 ตัวอย่างมาทำการสอนและทดสอบจำนวน 10 ครั้ง และนำผลของการสอนการเรียนรู้คิวในครั้งที่ดีที่สุดมาทำการสอนข่ายงานประสาท และใช้ในการประเมินผลการเรียนรู้คิวต่อไป

5. แสดงการลู่เข้าของการเรียนรู้คิวภายใต้ข้อกำหนดข้างต้น

6. ประเมินผลการเรียนรู้คิว โดยการเปรียบเทียบผลที่ได้จากวิธีขยายและจำกัดเขต (branch-and-bound) จากการแบ่งตัวอย่างเป็นชุดๆ จากการเรียนรู้ชุดละ 7 ตัวอย่างมาทำการทดสอบ โดยใช้ทำการสอน 48 ตัวอย่างและทดสอบ 7 ตัวอย่างทำซ้ำกัน 8 ครั้ง

7. ใช้ข้อมูลตารางเวลาเที่ยวบินของเที่ยวบินภายในประเทศ

1.4 ขั้นตอนและวิธีการดำเนินงานวิจัย

1. ทำความเข้าใจการปฏิบัติการทางการบินและปัญหาการปรับตารางเวลาเที่ยวบิน
2. ศึกษาการเรียนรู้แบบการเสริมกำลังและการเรียนรู้คิวพร้อมการประยุกต์ใช้งาน
3. ประยุกต์การเรียนรู้คิวในปัญหาดังกล่าว
4. เตรียมข้อมูลในการทดลองจากตารางเวลาเที่ยวบินภายในประเทศ
5. ทดลองและตรวจสอบการประยุกต์ใช้งาน
6. สรุปผลการวิจัยและจัดทำเอกสาร

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ทราบถึงการประยุกต์ใช้งานการเรียนรู้เรียนรู้คิวเพื่อแก้ปัญหาการปรับตารางเวลาเที่ยวบินในสภาวะไม่ปกติ

2. ได้ทราบถึงประสิทธิภาพของการเรียนรู้คิวในการแก้ปัญหาการปรับตารางเวลาเที่ยวบินในสภาวะไม่ปกติเมื่อเปรียบเทียบกับผลที่ได้จากวิธีขยายและจำกัดเขต

3. ได้ทราบถึงแนวทางในการแก้ปัญหาการปรับตารางเวลาเที่ยวบินในสภาวะไม่ปกติในแนวทางอื่น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการปรับตารางเวลาเที่ยวบิน และการเรียนรู้แบบเสริมกำลังและการเรียนรู้คิว โดยในส่วนแรกจะกล่าวถึงตารางเวลาเที่ยวบิน, ปัญหาและแนวทางการแก้ไขเมื่อเกิดความผิดปกติขึ้นในตารางเวลาเที่ยวบิน ต่อมาจะกล่าวถึงทฤษฎีการเรียนรู้แบบเสริมกำลังและการเรียนรู้คิวที่จะนำมาใช้แก้ปัญหาการปรับตารางเวลาเที่ยวบินในสภาวะไม่ปกติในงานวิจัยนี้ ในส่วนท้ายของบทจะทำการทบทวนวรรณกรรมและงานวิจัยที่เกี่ยวข้องกับงานวิจัยชิ้นนี้

2.1 ทฤษฎีที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้องกับงานวิจัยนี้แบ่งได้เป็นสองส่วนหลักคือ ส่วนของตารางเวลาเที่ยวบิน และส่วนของการเรียนรู้แบบเสริมกำลังและการเรียนรู้คิว ในส่วนของตารางเวลาเที่ยวบินจะบอกถึงการจัดตารางเวลาเที่ยวบินภายในประเทศ และปัญหาที่อาจจะเกิดขึ้นในตารางเวลาเที่ยวบินรวมถึงผลกระทบต่อตารางเวลาเที่ยวบินและแนวทางการแก้ไขปัญหาการปรับตารางเวลาเที่ยวบินเมื่อเกิดความผิดปกติภายในตารางเวลาเที่ยวบิน ในส่วนของการเรียนรู้แบบเสริมกำลังและการเรียนรู้คิวจะกล่าวถึงหลักการของการเรียนรู้แบบเสริมกำลัง, ความเกี่ยวข้องของการเรียนรู้แบบเสริมกำลังและการเรียนรู้คิว แต่ในส่วนของความเกี่ยวข้องระหว่างตารางเวลาเที่ยวบินและการเรียนรู้แบบเสริมกำลังและการเรียนรู้คิวนั้นจะกล่าวถึงในบทวิธีดำเนินงานวิจัย ซึ่งจะเกี่ยวข้องกับการประยุกต์การใช้งานการเรียนรู้คิวในการปรับตารางเวลาเที่ยวบินในสภาวะไม่ปกติ

2.1.1 การจัดตารางเวลาเที่ยวบินภายในประเทศของสายการบิน

จากงานวิจัยของ N. Kohl และคณะ [7] และงานวิจัยของ Jens Clausen และคณะ [8] ได้กล่าวไว้เกี่ยวกับเรื่องของการจัดตารางเวลาเที่ยวบินที่มีขั้นตอนต่างๆ เริ่มตั้งแต่ขั้นตอนการวางแผนตารางเวลาเที่ยวบิน, การกำหนดแบบของเครื่องบินที่ใช้ (Fleet assignment), การจัดตารางเวลาเจ้าหน้าที่ (crew scheduling), การหมุนเวียนเครื่องบิน (aircraft rotation) และอื่นๆ ซึ่งขั้นตอนแต่ละขั้นตอนที่ได้กล่าวมานี้จำต้องใช้เวลาในการวางแผนไว้ก่อนล่วงหน้าเป็นเวลานานก่อนที่จะมาเป็นตารางเวลาเที่ยวบินที่ใช้ในการปฏิบัติการทางการบินจริง และในแต่ละขั้นตอนก็มีวิธีในการปรับปรุงและการแก้ปัญหาแบบเฉพาะส่วนเช่น ในขั้นตอนการวางแผนตารางเวลาเที่ยวบินอาจจะวางแผนตามความต้องการของผู้โดยสารหรือตามแต่ฤดูกาลเป็นต้น การกำหนดแบบเครื่องก็เป็นอีกปัญหาหนึ่งที่มีความสำคัญและใช้เวลาในการแก้ปัญหา เนื่องจากจากเกี่ยวข้องกับประโยชน์ของทาง

สายการบินที่จะได้คำนวณผลของกำไรที่จะได้รับ ในส่วนอื่นๆ ที่ไม่ได้กล่าวถึงในที่นี้ของการวางแผนการจัดตารางเวลาเที่ยวบินก็จำต้องให้ความสำคัญเช่น ส่วนภาคพื้นดินหรือส่วนจัดหาอาหารก็ตาม ขั้นตอนการจัดตารางเวลาเที่ยวบินจนเสร็จสมบูรณ์พร้อมปฏิบัติการณ์นั้นจำต้องใช้ข้อมูล, เวลา และงบประมาณเป็นอย่างมาก แต่อย่างไรก็ดีเมื่อถึงเวลาที่จะต้องทำการปฏิบัติการทางการบินจริงก็อาจมีเหตุการณ์หรือความผิดปกติที่คาดไม่ถึง มาทำให้ไม่สามารถปฏิบัติการณ์ได้ตามแผนหรือตามตารางเวลาที่ได้วางแผนกันไว้ ซึ่งเหตุการณ์หรือความผิดปกติที่เกิดขึ้นนั้นส่งผลให้ต้องมีการปรับตารางเวลาเที่ยวบินตามมา

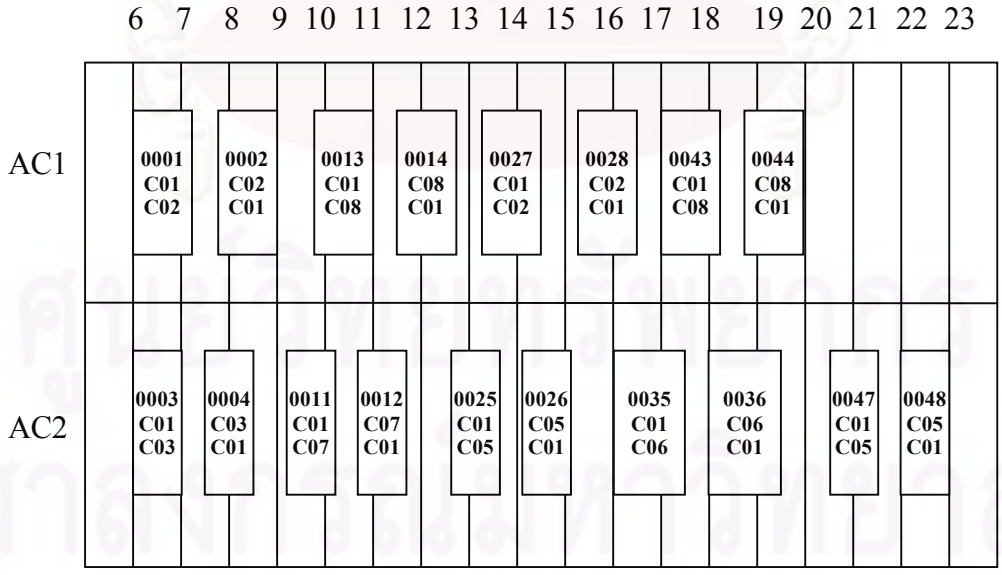
จากที่ได้กล่าวมาการจัดตารางเวลาเที่ยวบินของสายการบินนั้นได้มีการวางแผนการบินไว้ล่วงหน้าก่อนแล้ว โดยการวางแผนอาจจะแบ่งตามฤดูกาลเช่น ฤดูร้อน (summer) ฤดูหนาว (winter) หรือวางแผนการบินตามความต้องการของตลาดผู้โดยสาร อย่างไรก็ตามการวางแผนการบินจะทำก่อนการบินจริงเป็นเวลาหลายเดือน [7, 9] ดังนั้นเมื่อถึงเวลาที่จะทำการบินตามแผนที่ได้วางไว้ก็อาจจะเกิดความผิดปกติหรือสิ่งที่ไม่ได้คาดคิดไว้ได้เสมอ ในที่นี้จะขอกล่าวถึงเที่ยวบินภายในประเทศที่จะทำการบินเป็นประจำในแต่ละวันตามแต่ที่ได้ทำการวางแผนไว้ ไปยังสนามบินต่างๆ ซึ่งในแผนการบินจะทำการกำหนดจุดหมายปลายทาง, เวลาออก และเวลาถึงของแต่ละเที่ยวบินไว้แล้ว แต่โดยส่วนมากของสายการบินภายในประเทศมีแบบเครื่องบินมากกว่าหนึ่งแบบจึงต้องมีการกำหนดแบบเครื่องที่ใช้ในแต่ละเที่ยวบินลงไปด้วย ตัวอย่างของตารางเวลาเที่ยวบินภายในประเทศสามารถแสดงให้เห็นได้ตามตารางที่ 1

ตารางที่ 1 ตัวอย่างตารางเวลาเที่ยวบินภายในประเทศ

หมายเลขเที่ยวบิน	ต้นทาง	ปลายทาง	เวลาออก	เวลาถึง	หมายเลขเครื่องบิน	แบบเครื่อง
0001	C01	C02	06:00	07:05	AC1	734
0002	C02	C01	07:45	08:50	AC1	734
0003	C01	C03	06:05	07:00	AC2	734
0004	C03	C01	07:40	08:25	AC2	734
0005	C01	C04	06:05	07:35	AC3	AB7
0006	C04	C01	08:20	09:50	AC3	AB7
0007	C01	C05	06:15	07:20	AC4	734
0008	C05	C01	08:00	09:00	AC4	734
0009	C01	C06	08:00	09:10	AC5	AB7
0010	C06	C01	09:55	11:05	AC5	AB7

0011	C01	C07	09:10	10:05	AC6	734
0012	C07	C01	10:45	11:40	AC6	734
…	…	…	…	…	…	…

จากตารางที่ 1 เป็นตัวอย่างของตารางเวลาเที่ยวบินภายในประเทศเพื่อแสดงหมายเลขเที่ยวบิน, จุดหมายปลายทางของเที่ยวบินนั้น, เวลาที่เที่ยวบินออกจากสนามบิน และเวลาที่เที่ยวบินไปถึงยังจุดหมายปลายทาง โดยมีการกำหนดหมายเลขเครื่องคือ ชื่อของเครื่องบินที่ใช้ในเที่ยวบินแต่ละเที่ยวบิน และแสดงแบบเครื่องของเครื่องบินที่ใช้ในเที่ยวบินเที่ยวนั้นๆ ด้วย เช่น 0001, 0002, และ 0003 เป็นหมายเลขของเที่ยวบิน C01, C02, และ C03 เป็นจุดหมายปลายทางของเที่ยวบิน 6:00, 7:45, 6:05 เป็นเวลาที่เที่ยวบินออกจากสนามบินต้นทาง 7:05, 8:50, และ 7:00 เป็นเวลาที่เที่ยวบินถึงสนามบินปลายทาง AC1, AC2, และ AC3 เป็นหมายเลขเครื่องบินที่ใช้ในเที่ยวบิน 734 และ AB7 เป็นแบบเครื่องของเครื่องบินที่ใช้เป็นต้น ดังนั้นจากตารางสามารถอ่านได้ว่า เที่ยวบินหมายเลข 0001 เดินทางจาก C01 เวลา 6:00 น. เพื่อไปยังจุดหมาย C02 ในเวลา 7:05 น. โดยสารเครื่องบินหมายเลข AC1 ซึ่งอยู่ในแบบเครื่อง Boeing 737-400 และเที่ยวบินหมายเลข 0002 เดินทางจาก C02 เวลา 7:45 น. เพื่อไปยังจุดหมาย C01 ในเวลา 8:50 น. โดยสารเครื่องบินหมายเลข AC1 ซึ่งอยู่ในแบบเครื่อง Boeing 737-400 เป็นต้น



รูปที่ 1 ตัวอย่างการแสดงเที่ยวบินในรูปของกำหนดการการหมุนเวียนเครื่องบิน

ทั้งนี้ตารางเวลาเที่ยวบินสามารถนำมาแสดงในรูปของกำหนดการการหมุนเวียนเครื่องบิน (Aircraft rotation schedule) ที่ใช้กันมากในสายการบินได้ดังรูปที่ 1 เพื่อแสดงเวลาของเที่ยวบินที่ทำการบินในแต่ละวัน โดยหัวข้อทางด้านซ้ายของกำหนดการจะแสดงหมายเลขเครื่องบิน (AC1 และ AC2) ด้านบนของกำหนดการแสดงเวลา 6, 7, 8… เป็นเวลา 6:00 น., 7:00 น., และ 8:00 น. ตามลำดับ ข้อมูลด้านในของกำหนดการแสดงเที่ยวบินที่ทำการบินโดยใช้หมายเลขเครื่องบินและเวลาตามที่แสดง เช่น จากข้อมูลเที่ยวบินอันดับแรกในกำหนดการ 0001 C01 C02 สามารถอ่านได้ว่า เที่ยวบินหมายเลข 0001 เดินทางจาก C01 เวลา 6:00 น. เพื่อไปยังจุดหมาย C02 ในเวลา 7:05 น. โดยสารเครื่องบินหมายเลข AC1 ซึ่งอยู่ในแบบเครื่อง Boeing 737-400

2.1.2 ปัญหาและการแก้ไขความผิดปกติในตารางเวลาเที่ยวบิน

ตารางเวลาเที่ยวบินจะถูกสร้างขึ้น และแสดงให้ผู้โดยสารรับทราบก่อนที่จะทำการบินจริง ซึ่งถ้าหากไม่เหตุผิดปกติใดๆ ผู้โดยสาร, สำภาระ และเครื่องบิน ก็จะออกเดินทางได้ตามกำหนด แต่เนื่องจากในความเป็นจริงแล้วไม่มีใครสามารถทราบได้ว่า ในวันที่จะทำการบินจริงนั้นจะเกิดเหตุผิดปกติหรือไม่ ความผิดปกติที่พบเช่น เครื่องบินขัดข้อง, สภาพอากาศเลวร้าย, เจ้าหน้าที่ป่วยกะทันหัน เป็นต้น ซึ่งเหตุการณ์เหล่านี้ไม่สามารถที่จะคาดเดาไว้ก่อนได้ จึงเป็นผลให้ต้องมีการปรับเปลี่ยนตารางเวลาเที่ยวบิน เพื่อให้ผู้โดยสารและสำภาระสามารถเดินทางไปยังจุดหมายให้ได้เร็วที่สุด ในส่วนของสายการบินเองก็ต้องการที่จะลดค่าใช้จ่ายให้เหลือน้อยที่สุดในการปรับตารางเวลาเที่ยวบิน โดยมีจุดประสงค์หลักคือเพื่อกลับไปยังแผนการบินที่ได้ทำการวางแผนไว้ก่อน

ในการแก้ปัญหาการปรับตารางเวลาเที่ยวบินเมื่อเกิดเหตุผิดปกตินั้นโดยทั่วไปจะอาศัยผู้เชี่ยวชาญในศูนย์ควบคุมการปฏิบัติการ (Operation Control Centre) [7] ที่จะเข้ามาเป็นหลักในแก้ไขปัญหา หลักการในการปรับตารางเวลาเที่ยวบินของสายการบินมีหลายแบบด้วยกันแต่ในที่นี้จะขอกล่าวเพียง 5 แบบคือ

1. การดีเลย์เที่ยวบิน (Delay flights)
2. การเปลี่ยนเครื่องบินในแบบเครื่องเดียวกัน
3. การเปลี่ยนเครื่องบินในแบบเครื่องที่ต่างกัน
4. การรวมและการยกเลิกเที่ยวบิน
5. การรวมและการเปลี่ยนเส้นทางเที่ยวบิน

โดยในการปรับตารางเวลาเที่ยวบินในแต่ละขั้นตอนนั้น จะต้องคำนึงถึงเงื่อนไขบังคับ (constraint) หลายอย่างด้วยกันไม่ว่าจะเป็น เงื่อนไขบังคับของตารางเวลาเที่ยวบิน, การบำรุงรักษาและกราวนด์ไทม์ (Ground time), ความเชื่อมต่อของผู้โดยสาร, ความจุผู้โดยสาร, ข้อจำกัดของเครื่องบิน และ/หรือสนามบิน, สภาพอากาศ และอื่นๆ [3]

จากยุทธศาสตร์และเงื่อนไขบังคับดังที่กล่าว การหาคำตอบที่เหมาะสมที่สุดให้ได้ในแบบทันกาลนั้นแทบจะเป็นไปไม่ได้ อันเนื่องมาจากขนาดของปัญหาเองที่เป็นอุปสรรคในการที่จะหาคำตอบที่เหมาะสมที่สุด เพราะฉะนั้นถ้าต้องการคำตอบแบบทันกาลสำหรับปัญหานี้จึงจำเป็นต้องพึ่งการแก้ปัญหาแบบศึกษาสำนึกเข้าช่วย [4] ซึ่งในขณะนี้สายการบินส่วนมากได้อาศัยผู้เชี่ยวชาญเพื่อที่จะแก้ปัญหาปัญหานี้อยู่

จุดมุ่งหมายของงานวิจัยนี้คือ เพื่อนำเสนอการประยุกต์ใช้การเรียนรู้แบบการเสริมกำลังในการแก้ปัญหาการปรับตารางเวลาเที่ยวบินเมื่อเกิดความผิดปกติขึ้น โดยตั้งอยู่บนยุทธศาสตร์หลักสองอย่างด้วยกันคือ การเปลี่ยนเครื่องบินในแบบเครื่องเดียวกัน และการเปลี่ยนเครื่องบินในแบบเครื่องที่ต่างกัน เงื่อนไขบังคับที่พิจารณาในงานนี้จะสนใจในสามส่วนหลักๆ คือ เงื่อนไขบังคับของตารางเวลาเที่ยวบิน, กราวนด์ไทม์ และความจุผู้โดยสาร ในส่วนของความผิดปกติที่เกิดขึ้นนั้นในที่นี้จะพิจารณาว่าเป็นความผิดปกติใดๆ ที่ทำให้เครื่องบินไม่สามารถใช้งานได้ในช่วงเวลาที่เกิดความผิดปกติ

2.1.3 การเรียนรู้แบบการเสริมกำลังและการเรียนรู้คิว

การเรียนรู้แบบการเสริมกำลังเรียนรู้ที่จะเลือกการกระทำที่เหมาะสมที่สุดเพื่อไปสู่จุดหมายในแต่ละครั้งที่ระบบกระทำการกระทำ a ในบางสถานะ s มันจะรับค่ารางวัล r กลับคืนมา การเรียนรู้เป็นการเรียนรู้นโยบายควบคุม $\pi : S \to A$ ที่จะให้ค่าผลรวมของค่ารางวัลมากที่สุด ค่ารางวัลสะสมคือ การรวมของค่ารางวัลทั้งหมดของนโยบาย π ที่เริ่มจากสถานะ s นโยบายเหมาะที่สุด (optimal policy) คือ นโยบายที่ทำให้ค่ารางวัลสะสมมากที่สุด [10]

ในงานการเรียนรู้การเรียนรู้แบบเสริมกำลังไม่สามารถที่จะเรียนรู้นโยบายเหมาะที่สุดได้โดยตรง ในทางตรงกันข้ามมันจะเรียนรู้นโยบายเหมาะที่สุดจากค่ารางวัลสะสมของนโยบายเหมาะที่สุดแทน ดังนั้นในการเรียนรู้นี้ มันเรียนรู้ที่จะเลือกการกระทำที่ดีที่สุด a ในแต่ละสถานะ s ที่ทำให้ค่ารางวัลสะสมมากที่สุด จากที่ได้กล่าวมานี้การเรียนรู้แบบเสริมกำลังจำเป็นต้องมีความรู้ที่สมบูรณ์ของการเปลี่ยนสถานะ เพื่อใช้ในการประยุกต์การกระทำa ในสถานะ s

การเรียนรู้คิวเป็นอีกรูปแบบหนึ่งของการเรียนรู้แบบการเสริมกำลังที่ทำการเรียนรู้ฟังก์ชันค่าการกระทำที่ได้จากผลของการกระทำในสถานะที่ให้ และตามด้วยค่านโยบายที่กำหนดหลังจากนั้น ประโยชน์ของการเรียนรู้คิวคือ มันสามารถเปรียบเทียบผลของการกระทำที่มีอยู่ โดยไม่จำเป็นต้องอาศัยแบบจำลองของสิ่งแวดล้อมที่มันเรียนรู้อยู่

ฟังก์ชันการประเมินค่าคิว (evaluation function Q) ถูกนิยามด้วยค่ารางวัลโดยตรง (immediate reward) ที่ได้รับหลังจากการกระทำ a กระทำการสำเร็จจากสถานะ s รวมกับค่านโยบายที่เหมาะสมที่สุดต่อจากนั้น หรือจะพูดได้ว่า ค่า Q เป็นค่ารางวัลโดยตรงรวมกับค่าสูงสุด

ของ Q ของสถานะถัดไป จากฟังก์ชันการประเมินค่าคิว มันจะสามารถคัดเลือกการกระทำที่เหมาะสมที่สุดได้แม้ว่าจะมีความรู้ไม่สมบูรณ์ก็ตาม

ขั้นตอนวิธีการเรียนรู้คิวเป็นการทำซ้ำปรับปรุง โดยใช้ $\hat{Q}$ เป็นการประมาณค่าหรือสมมติฐานของค่า Q จริง การฝึกสอนคือ การสอนการประมาณค่า $\hat{Q}$ เพื่อคาดเดาค่า Q จริงสำหรับในแต่ละสถานะ s และแต่ละการกระทำ a ทำการคำนวณการปรับปรุงค่ารางวัลได้ตามที่แสดง [11]

$$\hat{Q}(s,a) \leftarrow \hat{Q}(s,a) + \alpha\left[r + \gamma \max_{a'} \hat{Q}(s',a') - \hat{Q}(s,a)\right] \quad \text{------ (1)}$$

โดยค่า α จะเป็นค่าอัตราการเรียนรู้ (learning rate) มีค่า $0 \le \alpha \le 1$ และ γ เป็นค่าอัตราส่วนลด (discount rate) มีค่า $0 \le \gamma \le 1$ และขั้นตอนวิธีการเรียนรู้คิว [10] สามารถแสดงได้ดังรูปที่ 2

ขั้นตอนวิธีการเรียนรู้คิว

1. เลือกการกระทำในสถานะนั้น

2. รับรางวัลจากการกระทำที่ได้เลือก

3. สำรวจสถานะใหม่

4. ปรับปรุงค่า $\hat{Q}$ ตามที่แสดง

$$\hat{Q}(s,a) \leftarrow \hat{Q}(s,a) + \alpha\left[r + \gamma \max_{a'} \hat{Q}(s',a') - \hat{Q}(s,a)\right]$$

5. เปลี่ยนสถานะจากการกระทำที่ได้เลือกมา

6. กลับไป 1.

รูปที่ 2 ขั้นตอนวิธีการเรียนรู้คิว

2.2 การทบทวนวรรณกรรม และงานวิจัยที่เกี่ยวข้อง

ในปี ค.ศ.1995 Clarke, Michael [3] ได้ศึกษาปัญหา และผลกระทบเมื่อเกิดเหตุการณ์ไม่ปกติในการปฏิบัติการ (irregular operation) ในสายการบินที่เป็นผลให้ต้องมีการปรับตารางเวลาเที่ยวบิน ในงานวิจัยชิ้นนี้ได้บอกถึงเหตุการณ์ไม่ปกติที่อาจจะเกิดขึ้น เช่น ปัญหาทางด้านสภาพอากาศ, ความล่าช้าจากระบบควบคุมการจราจรทางอากาศ, เครื่องบินขัดข้อง และปัญหาทางภาคพื้นดินเป็นต้น และแนวคิดในการแก้ปัญหาการปรับตารางเวลาเมื่อเกิดเหตุการณ์ไม่ปกติรวมถึงสิ่งที่ต้องคำนึงถึงเมื่อเกิดเหตุผิดปกติขึ้น เช่น การปรับตารางเครื่องบินและการเปลี่ยนเส้นทาง (aircraft rescheduling and rerouting), การจัดตารางเวลาเจ้าหน้าที่ (crew scheduling) และการ

บำรุงรักษา (maintenance) ต่อมาในปี ค.ศ.1996 Clarke, Michael [4] นำเสนองานวิจัยที่ต่อเนื่องจากงานเดิมของตนเอง โดยให้ความสนใจไปในเรื่องของการปรับตารางเวลาเที่ยวบินเมื่อเกิดเหตุการณ์ไม่ปกติในการปฏิบัติการ และได้พัฒนาวิธีการแก้ปัญหา โดยการนำเสนอระบบสนับสนุนการตัดสินใจที่มีการนำกระบวนการและขั้นตอนวิธีศึกษาสำนึกเข้ามาช่วยในการจัดการปัญหา ซึ่งในงานนี้ครอบคลุมการแก้ปัญหาการปรับเปลี่ยนเครื่องบินในแบบเครื่องที่ต่างกัน, การทำให้ดีเลย์เครื่องบินและการยกเลิกเที่ยวบิน ในงานที่เกี่ยวกับระบบสนับสนุนการตัดสินใจยังมีงานของ Dennis F. X. Mathaisel [6] ในปี ค.ศ.1996 แต่ได้เน้นไปที่การนำเสนอรูปแบบของระบบสนับสนุนการตัดสินใจ และสิ่งที่ควรมีในการใช้งานระบบสนับสนุนการตัดสินใจสำหรับการปฏิบัติการทางการบินรวมถึงในส่วนของความไม่ปกติในการปฏิบัติการด้วย ในงานนี้จะกล่าวถึงลักษณะโครงสร้างของระบบสนับสนุนการตัดสินใจ, โครงสร้างของการเรียกข้อมูล และวิธีการแก้ปัญหาการปรับตารางการบินเมื่อเกิดเหตุการณ์ไม่ปกติขึ้น เพื่อให้เป็นมาตรฐานเดียวกันในการแก้ปัญหา งานวิจัยที่ได้กล่าวเกี่ยวกับปัญหาและผลกระทบที่อาจจะเกิดขึ้นรวมถึงแนวทางการแก้ไขปัญหาต่างๆ เมื่อเกิดเหตุการณ์ไม่ปกติหรือความผิดปกติเกิดขึ้นในสายการบินที่กล่าวถึงคืองานของ N. Kohl และคณะ [7] และงานของ J. Clausen และคณะ [8] โดยทั้งสองงานวิจัยนี้ได้รวบรวมปัญหา, ผลกระทบ และแนวทางการแก้ไขปัญหาต่างๆ ที่อาจจะเกิดขึ้นไว้ โดยเริ่มกล่าวตั้งแต่ที่มาของตารางเวลาเที่ยวบิน, ปัญหาที่อาจจะเกิดขึ้นได้เมื่อปฏิบัติการทางการบินจริง และวิธีการรับมือกับปัญหาที่เกิดขึ้น รวมถึงได้มีการทบทวนงานวิจัยที่เกี่ยวกับการแก้ปัญหาเมื่อเกิดความผิดปกติในตารางเวลาเที่ยวบินที่ผ่านมาอีกด้วย

ในส่วนงานวิจัยที่เกี่ยวกับการเรียนรู้แบบเสริมกำลังที่น่าสนใจเริ่มจากในปี ค.ศ. 1988 จากงานของ R. Sutton [12] ที่ได้กล่าวถึงการเรียนรู้แบบ $TD(\lambda)$ ซึ่งถือได้ว่าเป็นงานที่ได้กล่าวถึงมากงานหนึ่งในงานวิจัยที่เกี่ยวกับการเรียนรู้แบบเสริมกำลัง เนื่องจากเป็นแนวทางให้กับงานวิจัยที่สำคัญหลากหลายงานในปัจจุบัน ตัวอย่างเช่นงานวิจัยของ Tesauro [1] ในปี ค.ศ.1995 ที่ได้นำเสนอการใช้การเรียนรู้แบบเสริมกำลังในเกมแบคแกมมอน โดยการนำหลักการของ R. Sutton มาประยุกต์ใช้โดยให้ตัวเรียนรู้ทำการการเล่นเกมแบคแกมมอนเองแล้วทำการให้รางวัลกับตัวเรียนรู้เพื่อให้ได้ความรู้ในการเล่นเกม และใช้ข่ายงานประสาทแบบแพร่กระจายย้อนกลับมาช่วยในการทำความรู้ให้อยู่ในรูปทั่วไป สุดท้ายนำมาทำการทดสอบกับผู้เล่นที่เป็นมนุษย์ งานที่ใช้ $TD(\lambda)$ ร่วมกับข่ายงานประสาทแบบแพร่กระจายย้อนกลับยังมีงานของ Zhang และ T. G. Dietterich [2] ในปี ค.ศ.1995 ใช้การเรียนรู้แบบเสริมกำลังในการจัดกำหนดการงาน (job-shop scheduling) และนำไปในงานในองค์การอวกาศแห่งชาติของประเทศอเมริกา (NASA) ต่อมาได้เริ่มมีการใช้การเรียนรู้คิวในการแก้ปัญหาซึ่งในปี ค.ศ.1996 R. H. Crites and A. G. Barto [13] ประยุกต์การใช้งาน

การเรียนรู้คิวซึ่งเป็นรูปแบบหนึ่งของการเรียนรู้แบบเสริมกำลังเข้ากับการเพิ่มประสิทธิภาพการทำงานของลิฟต์ และในขั้นตอนการนำความรู้มาใช้ก็ได้มีการทำความรู้ที่เรียนได้มาทำให้อยู่ในรูปทั่วไปโดยอาศัยข่ายงานประสาทแบบแพร่กระจายย้อนกลับมาช่วย ผลจากงานวิจัยนี้แสดงให้เห็นถึงความสามารถในการใช้งานลิฟต์ได้อย่างมีประสิทธิภาพเพิ่มมากขึ้นเมื่อผ่านการเรียนรู้แล้ว นอกจากงานวิจัยที่ใช้ข่ายงานประสาทแบบแพร่กระจายย้อนกลับมาช่วยในการทำความรู้ให้อยู่ในรูปทั่วไปแล้วยังมีการใช้งานฟังก์ชันการประมาณอื่นอีกเช่น ในงานของ Gersmann และ Hammer [14] ในปี ค.ศ.2004 ได้มีการพัฒนาการเรียนรู้แบบเสริมกำลังโดยการนำ SVM เข้ามาทำหน้าที่เป็นฟังก์ชันการประมาณแทนส่วนของข่ายงานประสาท สุดท้ายในงานของ H. Cuayahuitl [15] ในปี ค.ศ.2010 ได้นำการเรียนรู้คิวมาประยุกต์ใช้งานกับระบบพูดสนทนาตอบโต้โดยทำการสอนบทพูดกับระบบ และทดสอบผลด้วยการพูดตอบโต้กับมนุษย์ จากผลการทดลองแสดงให้เห็นว่ามีค่าความผิดพลาดน้อยมากและยังเหมาะสมกับการนำไปใช้งานจริง

จากงานวิจัยที่ผ่านมา [1, 2, 13, 15] ทำให้เห็นได้ว่าการใช้งานการเรียนรู้แบบเสริมกำลังนั้นมีประสิทธิภาพ และเป็นที่น่าสนใจในการศึกษาว่าการเรียนรู้แบบเสริมกำลังนั้นสามารถเข้ามาช่วยแก้ปัญหาการปรับตารางเวลาการบินเมื่อเกิดเหตุการณ์ไม่ปกติในการปฏิบัติการได้หรือไม่ แล้วประสิทธิภาพของการเรียนรู้แบบเสริมกำลังนั้นจะให้ผลเป็นอย่างไร ดังนั้นในงานวิจัยนี้จะขอนำเสนอการเรียนรู้แบบเสริมกำลังในการแก้ปัญหาการปรับตารางเวลาเที่ยวบินเมื่อเกิดเหตุการณ์ไม่ปกติในการปฏิบัติการ และอาศัยข่ายงานประสาทแบบแพร่กระจายย้อนกลับมาช่วยในส่วนของฟังก์ชันการประมาณ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3
วิธีดำเนินงานวิจัย

ในบทนี้จะกล่าวถึงการประยุกต์การใช้งานการเรียนรู้คิวซึ่งเป็นรูปแบบหนึ่งของการเรียนรู้แบบเสริมกำลัง ในงานการแก้ปัญหาการปรับตารางเวลาเที่ยวบินเมื่อเกิดเหตุผิดปกติขึ้น ซึ่งจะกล่าวถึงขั้นตอนในการประยุกต์การเรียนรู้คิวเข้าไปในการแก้ปัญหาการปรับตารางเวลาเที่ยวบินได้อย่างไร การนิยามสถานะที่ใช้ในการเรียนรู้ การกำหนดการกระทำที่สามารถทำได้ในแต่ละสถานะของการปรับตารางเวลาเที่ยวบิน รวมถึงค่ารางวัลที่สมควรจะได้ในแต่ละการกระทำ การลู่เข้าของการเรียนรู้คิว และฟังก์ชันการประมาณก็จะถูกกล่าวถึงในบทนี้

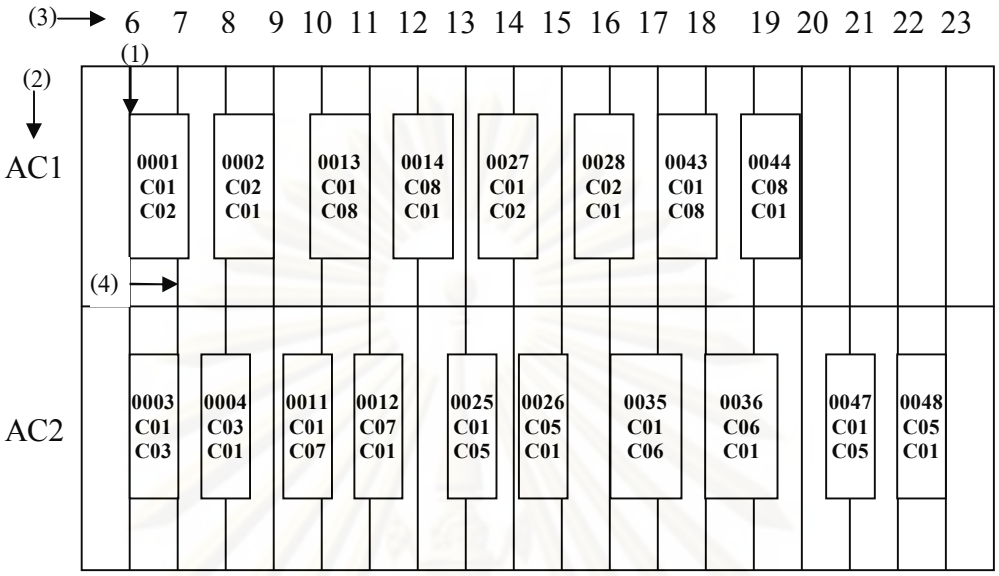
3.1 การประยุกต์การเรียนรู้คิว

ในส่วนแรกนี้จะทำการอธิบายเกี่ยวกับนิยามสถานะของงานวิจัยนี้ ซึ่งเกี่ยวข้องกับการปรับตารางเวลาเที่ยวบิน ซึ่งจะมีลายละเอียดเกี่ยวกับพารามิเตอร์ที่ใช้ในการนิยามสถานะทั้งในส่วนของตารางเวลาเที่ยวบิน และในส่วนของความผิดปกติที่เกิดขึ้นในตารางเวลาเที่ยวบินที่จะนำมาใช้ในการปรับตารางเวลาเที่ยวบินเมื่อเกิดเหตุผิดปกติต่อไป ในส่วนของสถานะจะแยกเป็นการนิยามสถานะในส่วนของตารางเวลาเที่ยวบิน และการนิยามสถานะในส่วนของความผิดปกติ

3.1.1 สถานะ

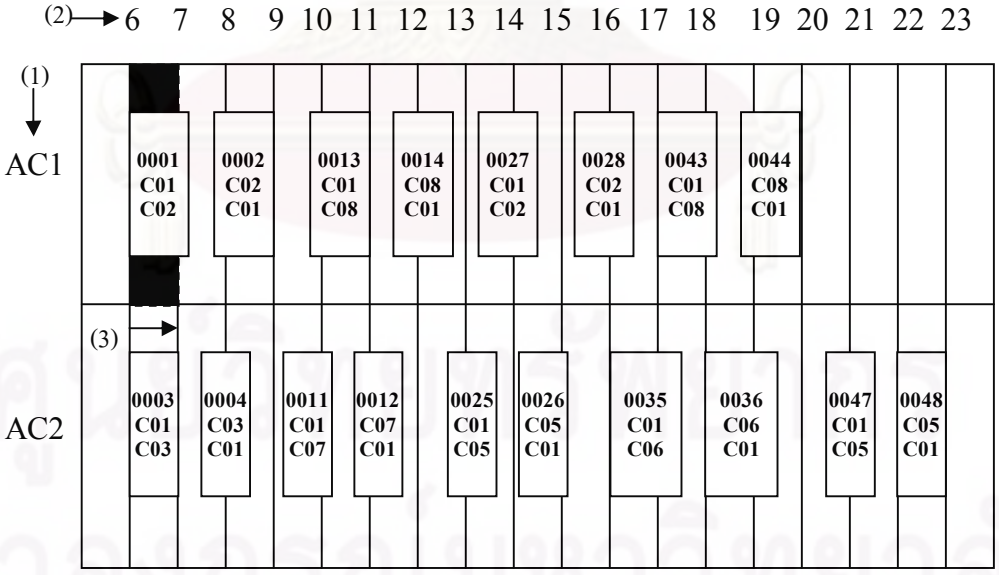
ในการนิยามสถานะนั้น สถานะจะต้องทำการอธิบายการปรับตารางเวลาในแต่ละขั้นได้ จากรูปที่ 1 ที่แสดงเที่ยวบินในรูปของกำหนดการ ถ้าต้องการบอกตำแหน่งของเที่ยวบินในตารางนั้นจำเป็นที่จะต้องทราบว่าเที่ยวบินเที่ยวนั้นเป็นเที่ยวบินที่เท่าใด, หมายเลขเครื่องบินที่ใช้ในการบินเที่ยวบินเที่ยวนั้น, เวลาออกจากสนามบิน และเวลาที่ใช้ในการบิน (รูปที่ 3) จากการบอกตำแหน่งที่กล่าวมานี้สามารถนำไปใช้บอกตำแหน่งของเที่ยวบินทุกๆ เที่ยวในตารางได้ และสามารถนำไปใช้ในการปรับตารางเวลาเที่ยวบินใหม่โดยไม่ต้องทำการเพิ่มเติมหรือแก้ไขแต่อย่างใด

ส่วนสำคัญอีกส่วนหนึ่งที่ต้องทำการบอกตำแหน่งในตารางเวลาให้ได้คือ ความผิดปกติที่เกิดขึ้นกับเครื่องบินว่ามีผลกระทบกับเครื่องบินหมายเลขเครื่องใด, เกิดขึ้นเมื่อใด และระยะเวลาที่มีความผิดปกติเกิดขึ้น (รูปที่ 4) เนื่องจากความผิดปกติที่เกิดขึ้นนั้นเป็นตัวอย่างที่นำมาใช้ในการฝึกสอนและทดสอบการเรียนรู้ และจะมีเพียงหนึ่งความผิดปกติต่อหนึ่งตารางเท่านั้น การบอกตำแหน่งจึงไม่มีความจำเป็นที่จะต้องใส่หมายเลขของความผิดปกติเข้าไป



รูปที่ 3 การกำหนดสถานะให้แต่ละเที่ยวบินในตารางเวลา

(1) หมายเลขเที่ยวบิน, (2) หมายเลขเครื่องที่ใช้, (3) เวลาที่เที่ยวบินออกจากสนามบิน และ (4) ระยะเวลาที่ใช้ในการบิน



รูปที่ 4 แถบสีดำแสดงความผิดปกติที่เกิดขึ้นกับเครื่องบินมีผลทำให้เครื่องบินหมายเลขเครื่องดังกล่าวไม่สามารถใช้งานได้

(1) หมายเลขเครื่องที่เกิดความผิดปกติ, (2) เวลาที่เกิดความผิดปกติ, (3) ระยะเวลาที่มีความผิดปกติเกิดขึ้น

ในที่นี้นิยามสถานะขึ้นมาจากลักษณะของตารางเวลาเที่ยวบินดังที่ได้กล่าว โดยจัดให้เป็นกลุ่มของเครื่องบินที่ถูกใส่ให้แต่ละเที่ยวบินในตารางเวลาเที่ยวบิน และความผิดปกติที่เกิดขึ้นในตารางเวลาเที่ยวบิน

[*กลุ่มของเครื่องบินที่ถูกใส่ให้แต่ละเที่ยวบิน, ความผิดปกติ*] ------ (2)

3.1.1.1 การนิยามกลุ่มของเครื่องบินที่ถูกใส่ให้แต่ละเที่ยวบิน

กลุ่มของเครื่องบินที่ถูกใส่ให้แต่ละเที่ยวบิน ในตารางเวลาเที่ยวบินที่ทำการพิจารณาในงานวิจัยนี้มีเที่ยวบินภายในประเทศทั้งหมดจำนวน 50 เที่ยวบินต่อการปฏิบัติงานทางการบินภายในหนึ่งวันทำการ โดยใช้เครื่องบินจำนวน 7 ลำในแบบเครื่องสองแบบเครื่องคือ Boeing 737-400 และ Airbus A300-600 ที่ได้รับมอบหมายงานการบินให้ในแต่ละวัน แต่เนื่องจากเหตุผลที่ได้กล่าวไว้ก่อนหน้านี้ที่ว่า มีการทราบการวางแผนตารางเวลาเที่ยวบินก่อนทำการปฏิบัติการทางการบินจริงเป็นเวลานาน ทำให้สามารถศึกษาตารางเวลาเที่ยวบินได้ก่อน จึงเป็นผลให้ทราบว่า

1. เที่ยวบินทุกเที่ยวบินจะทำการบินเป็นแบบประจำในแต่ละวันคือ ขึ้นบินในเวลาเดิมเพื่อไปยังจุดหมายเดิมเป็นประจำทุกวัน เพียงแต่การบินในแต่ละวันนั้นอาจจะไม่ได้ใช้เครื่องบินลำเดิมเท่านั้น เช่นวันแรกใช้เครื่องหมายเลข AC1 เพื่อเดินทางไปยังจุดหมาย C02 แต่วันที่สองอาจใช้เครื่องหมายเลข AC2 เพื่อเดินทางไปยังจุดหมาย C02 แทน ดังนั้นเพื่อให้เข้าใจไปในแนวทางเดียวกัน และสะดวกในการเรียนรู้ในงานวิจัยนี้จะนิยามสัญลักษณ์ aircraft_1, aircraft_2, aircraft_3, aircraft_4, aircraft_5, aircraft_6 และ aircraft_7 เพื่อใช้แทนหมายเลขเครื่องบิน AC1, AC2, AC3, AC4, AC5, AC6 และ AC7 ในแต่ละเที่ยวบิน และกำหนดให้สัญลักษณ์ aircraft_1, aircraft_2 และ aircraft_3 เป็นสัญลักษณ์แทนเครื่องบินในแบบเครื่อง Boeing 737-400 และให้สัญลักษณ์ aircraft_4, aircraft_5, aircraft_6 และ aircraft_7 เป็นสัญลักษณ์แทนเครื่องบินในแบบเครื่อง Airbus A300-600

2. เที่ยวบินภายในประเทศในไทยนั้นมีฐานการบินอยู่เพียงที่เดียว และจะทำการบินออกจากฐานการบินของตัวเองไปยังจุดหมายปลายทางต่างๆ และบินกลับมายังฐานการบินของตัวเองเสมอ ดังนั้นทำให้สามารถที่จะนำเที่ยวบินที่เป็นคู่ ไป-กลับ มาทำการปรับเปลี่ยนตารางเวลาใหม่ได้ ซึ่งทางสายการบินจะเรียกวิธีนี้ว่าการจับคู่ (pairing) เที่ยวบิน จากการจับคู่ที่ว่านี้จะเป็นการลดจำนวนเที่ยวบินที่ต้องนำไปเรียนรู้ และยังสามารถช่วยป้องกันการวนซ้ำในขั้นตอนการเรียนรู้ได้อีกด้วย ตัวอย่างการจับคู่เที่ยวบินในตารางเวลาเที่ยวบินจากรูปที่ 1 เช่น เที่ยวบิน

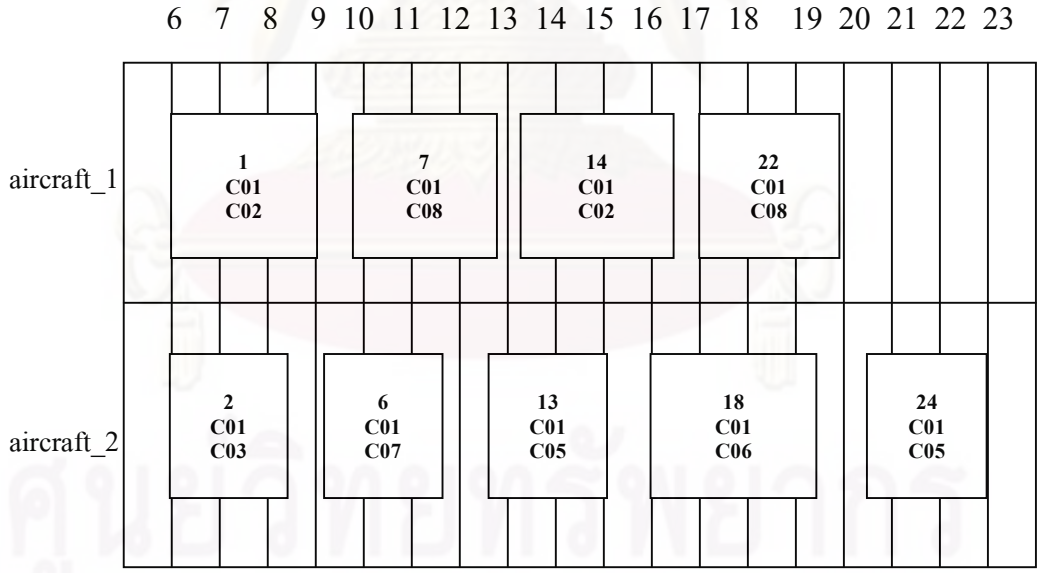
หมายเลข 0001 เดินทางจาก C01 เพื่อไปยังจุดหมาย C02 โดยสารเครื่องบินหมายเลข AC1 และเที่ยวบินหมายเลข 0002 เดินทางจาก C02 เพื่อไปยังจุดหมาย C01 โดยสารเครื่องบินหมายเลข AC1 เที่ยวบินทั้งคู่เดินทางไป-กลับอย่างต่อเนื่องกัน และใช้เครื่องบินหมายเลขเครื่องเดียวกันจึงสามารถจับคู่กันได้ ทั้งนี้หลังจากการจับคู่เที่ยวบินแล้ว ได้มีการนิยามหมายเลขคู่ของเที่ยวบินเพื่อใช้แทนหมายเลขเที่ยวบินของทั้งคู่ แต่เที่ยวบินหมายเลข 0001 เดินทางจาก C01 เพื่อไปยังจุดหมาย C02 โดยสารเครื่องบินหมายเลข AC1 กับเที่ยวบินหมายเลข 0028 ที่เดินทางจาก C02 เพื่อไปยังจุดหมาย C01 โดยสารเครื่องบินหมายเลข AC1 ถึงแม้จะไป-กลับจุดหมายปลายทางเดียวกัน แต่เที่ยวบินทั้งสองเที่ยวบินไม่ต่อเนื่องกันจึงไม่สามารถที่จะทำการจับคู่ได้เป็นต้น ดังนั้นจากตารางที่ 1 ทำให้สามารถนำมาเขียนตารางเวลาเที่ยวบินใหม่ โดยใช้หลักการที่ได้กล่าวไว้เป็นผลให้ได้ตารางเวลาเที่ยวบินใหม่ดังแสดงในตารางที่ 2 และรูปที่ 5

จากตารางที่ 2 หลังจากที่ได้ทำการจับคู่เที่ยวบิน และใช้สัญลักษณ์แทนเครื่องบินแล้วทำให้สามารถลดขนาดของตารางเที่ยวบินลงได้ครึ่งหนึ่ง ซึ่งวิธีการอ่านข้อมูลจากตารางที่ 2 ก็สามารถทำได้ในแบบเดียวกันกับตารางที่ 1 เพียงแต่ตารางที่ 2 นี้จะเป็นขนาดย่อและมีประโยชน์ต่อการเรียนรู้มากกว่าตารางที่ 1 จากเดิมที่เที่ยวบินในตารางเวลาจะแสดงเป็นเที่ยวบินหมายเลข 0001 เดินทางจาก C01 เวลา 6:00 น. เพื่อไปยังจุดหมาย C02 ในเวลา 7:05 น. โดยสารเครื่องบินหมายเลข AC1 ซึ่งอยู่ในแบบเครื่อง Boeing 737-400 และเที่ยวบินหมายเลข 0002 เดินทางจาก C02 เวลา 7:45 น. เพื่อไปยังจุดหมาย C01 ในเวลา 8:50 น. โดยสารเครื่องบินหมายเลข AC1 ซึ่งอยู่ในแบบเครื่อง Boeing 737-400 (ตารางที่ 1) หลังจากเปลี่ยนมาใช้วิธีทำการจับคู่เที่ยวบิน และใช้สัญลักษณ์แทนเครื่องบินแล้ว สามารถแสดงข้อมูลของเที่ยวบินได้สั้นกว่าคือ เที่ยวบินคู่ที่ 1 เดินทางจาก C01 เวลา 6:00 น. เพื่อไปยังจุดหมาย C02 และกลับมายัง C01 ในเวลา 8:50 น. โดยสารเครื่องบินหมายเลข AC1 ซึ่งอยู่ในแบบเครื่อง Boeing 737-400 การแสดงข้อมูลแบบเป็นคู่เที่ยวบินถือเป็นส่วนสำคัญในการสร้างสถานะเพื่อนำไปใช้ในการเรียนรู้ และยังช่วยลดเวลาที่จะต้องใช้ในการเรียนรู้ลงได้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 2 ตัวอย่างตารางเวลาเที่ยวบินภายหลังจากทำการจับคู่และใช้สัญลักษณ์แทนเครื่องบิน

หมายเลขคู่ของเที่ยวบิน	ต้นทาง	ปลายทาง	เวลาออก	เวลาถึง	สัญลักษณ์แทนเครื่องบิน	แบบเครื่อง
1	C01	C02	06:00	08:50	aircraft_1	734
2	C01	C03	06:05	08:25	aircraft_2	734
3	C01	C04	06:05	09:50	aircraft_4	AB7
4	C01	C05	06:15	09:00	aircraft_3	734
5	C01	C06	08:00	11:05	aircraft_5	AB7
6	C01	C07	09:10	11:40	aircraft_2	734
7	C01	C08	09:35	12:40	aircraft_1	734
…	…	…	…	…	…	…



รูปที่ 5 ตัวอย่างการแสดงตารางเวลาเที่ยวบินในรูปของกำหนดการภายหลังจากการใช้วิธีการจับคู่เที่ยวบินและใช้สัญลักษณ์แทนเครื่องบิน

กลุ่มของเครื่องบินที่ถูกใส่ให้แต่ละเที่ยวบิน ในตารางเวลาเที่ยวบินจะถูกนิยามโดยใช้พารามิเตอร์ 4 ตัวสำหรับเที่ยวบินแต่ละคู่คือ

1. *หมายเลขของเที่ยวบินแต่ละคู่* เป็นตัวบอกว่าเป็นเที่ยวบินคู่ที่เท่าใดในตาราง

2. *สัญลักษณ์แทนเครื่องบิน* เป็นตัวบอกว่าได้ทำการใช้เครื่องบินลำไหนในเที่ยวบินแต่ละคู่

3. *เวลาออกของเที่ยวบินแต่ละคู่* เป็นเวลาที่เที่ยวบินออกจากต้นทาง

4. *เวลาที่ใช้ในเที่ยวบิน* เป็นเวลาที่แต่ละคู่ใช้ในการบินไปกลับ (นาที)

[*หมายเลขของเที่ยวบินแต่ละคู่, สัญลักษณ์แทนเครื่องบิน, เวลาออกของเที่ยวบิน, เวลาที่ใช้ในเที่ยวบิน*] ------ (3)

ดังนั้นจากตารางที่ 2 เมื่อนำเที่ยวบินแต่ละคู่มาเขียนใหม่ตามพารามิเตอร์ที่ได้กำหนดขึ้นจะได้ว่า

[*1, aircraft_1, 6.00, 170*]

[*2, aircraft_2, 6.05, 140*]

และเมื่อมาเรียงต่อกันเพื่อบอกตำแหน่งของเที่ยวบินทั้งหมดในตารางเวลาเที่ยวบิน ในรูปแบบของ *กลุ่มของเครื่องบินที่ถูกใส่ให้แต่ละเที่ยวบิน* แล้วจะได้ดังนี้

[*1, aircraft_1, 6.00, 170, 2, aircraft_2, 6.05, 140, 3, aircraft_3, 6.05, 225…*]

โดย *1, 2, 3...* แสดงหมายเลขของเที่ยวบินแต่ละคู่

aircraft_1, aircraft_2, aircraft_3… แสดงสัญลักษณ์แทนเครื่องบิน

6.00, 6.05, 6.05… แสดงเวลาออกของเที่ยวบินแต่ละคู่

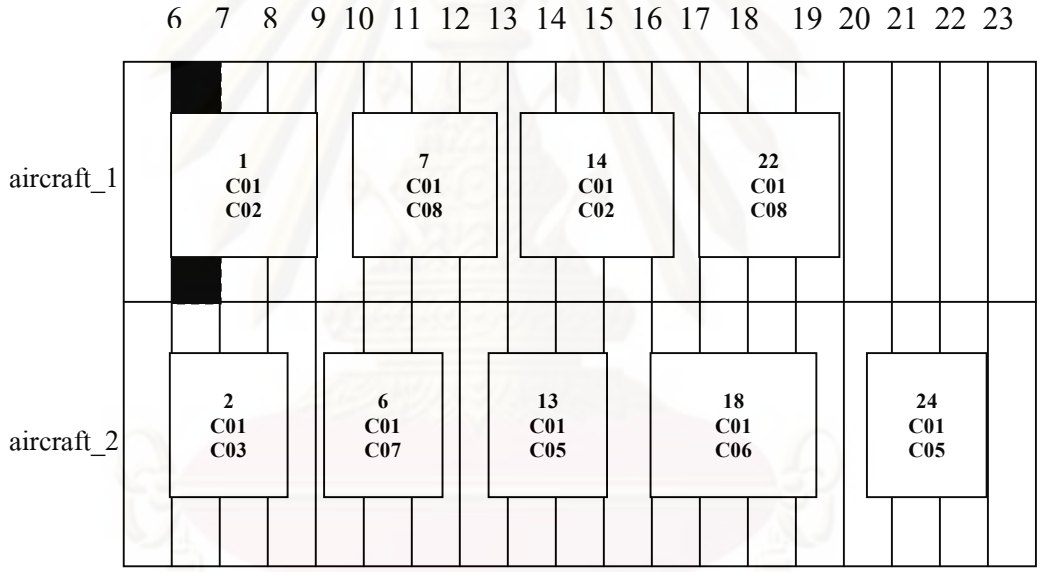
170,140,225… แสดงเวลาที่ใช้ในเที่ยวบินแต่ละคู่ (นาที)

3.1.1.2 การนิยามความผิดปกติ

ในส่วนของ *ความผิดปกติ* ที่เกิดขึ้นในตารางเวลาเที่ยวบินจะเป็นความผิดปกติที่ได้ทำการสร้างขึ้นมาในตารางเวลาเที่ยวบิน ซึ่งความผิดปกติที่ได้ทำการสร้างขึ้นนั้นจะพิจารณาว่าเป็นปัญหาใดๆ ที่ทำให้เครื่องบินไม่สามารถปฏิบัติการได้ในช่วงเวลาดังกล่าว วิธีที่ใช้ในการฝึกสอนในงานวิจัยนี้คือ จะทำการแทรกความผิดปกติลงไปในตารางเวลาเที่ยวบิน โดยจะสร้าง

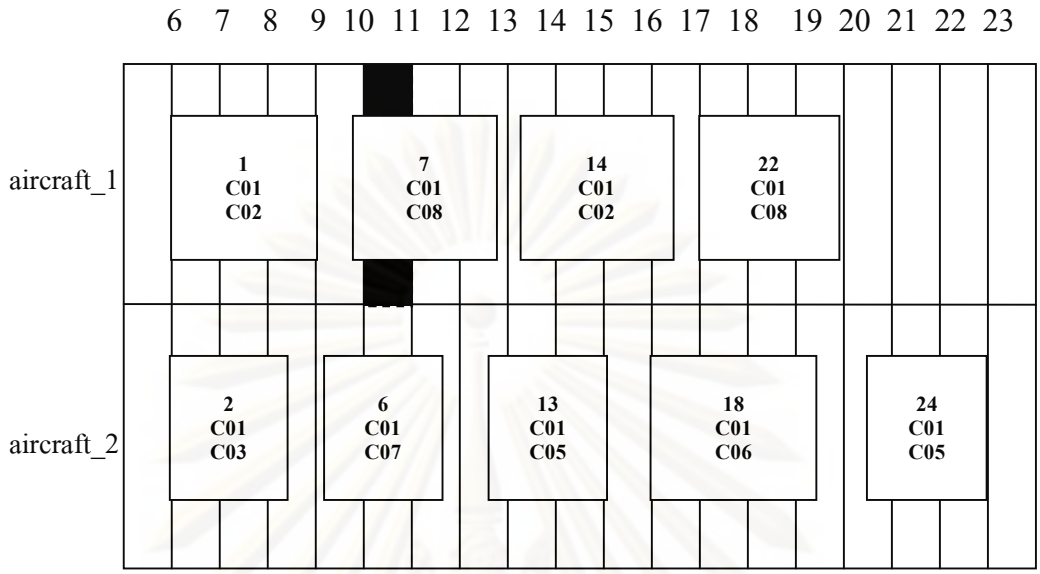
ความผิดปกติขึ้นมาทุกๆ ต้นของสองชั่วโมงตั้งแต่เวลา 6:00 น. จนถึงเวลา 20:00 น. โดยมีระยะเวลาของความผิดปกติหนึ่งๆนั้นยาวนานเป็นเวลาหนึ่งชั่วโมง และทำการแทรกลงในตารางเวลาเที่ยวบินตารางละหนึ่งความผิดปกติ (รูปที่ 6) ดังนั้นจำนวนความผิดปกติที่เกิดขึ้นจะมีจำนวน 56 ตัวอย่าง เพื่อนำมาเป็นข้อมูลการฝึก (training data) และข้อมูลทดสอบ (testing data)

จากรูปที่ 6.1 ความผิดปกติที่เกิดขึ้นทำให้เครื่องบินที่ใช้สัญลักษณ์ aircraft_1 ไม่สามารถใช้ปฏิบัติงานทางการบินได้ตั้งแต่เวลา 6:00 น. จนถึงเวลา 7:00 น. ในทำนองเดียวกันกับรูปที่ 6.2 ความผิดปกติที่เกิดขึ้นทำให้เครื่องบินที่ใช้สัญลักษณ์ aircraft_1 ไม่สามารถใช้ปฏิบัติงานทางการบินได้ตั้งแต่เวลา 10:00 น. จนถึงเวลา 11:00 น. ความผิดปกติที่เกิดขึ้นดังกล่าวเป็นผลทำให้ต้องมีการปรับตารางเวลาใหม่



6.1

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



6.2

รูปที่ 6 การแทรกความผิดปกติเข้าไปในตารางเวลาเที่ยวบินเพื่อใช้เป็นข้อมูลฝึกและข้อมูลทดสอบ

ความผิดปกติ ที่เกิดขึ้นในตารางเวลาเที่ยวบินนิยามโดยใช้พารามิเตอร์ 3 ตัว เพื่อแสดงความผิดปกติคือ

1. *เครื่องที่เกิดความผิดปกติ* จะแสดงว่าเครื่องบินลำไหนที่เกิดความผิดปกติในตารางเวลาเที่ยวบิน

2. *เวลาที่เกิดความผิดปกติ* เป็นเวลาที่เริ่มมีความผิดปกติเกิดขึ้นกับเครื่องบินในตารางเวลาเที่ยวบิน

3. *ระยะเวลาของความผิดปกติ* เป็นระยะเวลาที่เกิดความผิดปกติขึ้นกับเครื่องบิน (นาที)

[*เครื่องที่เกิดความผิดปกติ, เวลาที่เกิดความผิดปกติ, ระยะเวลาของความผิดปกติ*] ------ (4)

ดังนั้นจากรูปที่ 6.1 และ 6.2 เมื่อนำความผิดปกติที่เกิดขึ้นในตารางเวลาเที่ยวบินมาเขียนใหม่ตามพารามิเตอร์ที่ได้กำหนดขึ้นจะได้ว่า

[*aircraft_1, 6.00, 60*]

[*aircraft_1, 10.00, 60*]

และเมื่อนำความผิดปกติดังกล่าวมาทำการเขียนให้อยู่ในรูปของสถานะดังที่ได้นิยามไว้ข้างต้นเพื่อนำไปใช้ในการเรียนรู้จากรูปที่ 6.1 จะได้ตามที่แสดง

[*1, aircraft_1, 6.00, 170, 2, aircraft_2, 6.05, 140,*
6, aircraft_2, 9.10, 140 ... aircraft_1, 6.00, 60]

ในรูปที่ 6.2 ก็สามารถแสดงรูปแบบของสถานะได้ในทำนองเดียวกัน

[*1, aircraft_1, 6.00, 170, 2, aircraft_2, 6.05, 140,*
6, aircraft_2, 9.10, 140 ... aircraft_1, 10.00, 60]

โดยในส่วนแรกของสถานะแสดง *กลุ่มของเครื่องบินที่ถูกใส่ให้แต่ละเที่ยวบิน*

1, 2, 6... แสดงหมายเลขของเที่ยวบินแต่ละคู่
aircraft_1, aircraft_2… แสดงสัญลักษณ์แทนเครื่องบิน
6.00, 6.05, 9.10... แสดงเวลาออกของเที่ยวบินแต่ละคู่
170,140... แสดงเวลาที่ใช้ในเที่ยวบินแต่ละคู่ (นาที)

และใน *ความผิดปกติ* ซึ่งแสดงอยู่ในส่วนของสามตัวสุดท้ายของการแสดงสถานะแสดงถึงปัญหาที่เกิดขึ้นโดย

aircraft_1 แสดงเครื่องที่เกิดความผิดปกติ
6.00, 10.00 แสดงเวลาที่เกิดความผิดปกติ
60 แสดงระยะเวลาของความผิดปกติ (นาที)

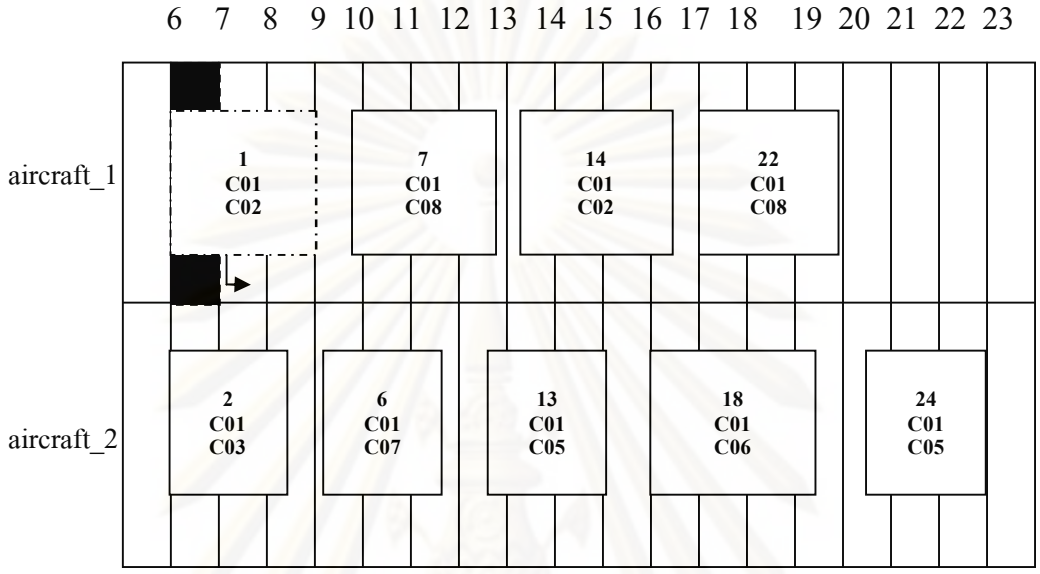
เป็นต้น

ในส่วนที่ถัดมานี้จะกล่าวถึง การกระทำที่สามารถทำได้ในแต่ละสถานะที่ตัวเรียนรู้สามารถกระทำได้ ซึ่งในส่วนนี้จะอธิบายถึงการกระทำที่ทำได้ ซึ่งในแต่ละสถานะจะถูกจำกัดด้วยข้อจำกัดต่างๆ ทำให้ในแต่ละสถานะมีการกระทำที่สามารถที่จะกระทำได้แตกต่างกันออกไป

3.1.2 การกระทำ

ในงานวิจัยนี้ได้ทำการนิยามการกระทำให้เป็นสัญลักษณ์แทนเครื่องบินที่สามารถกำหนดลงไปในตารางเวลาเที่ยวบินได้ในสถานะนั้น โดยในการพิจารณาว่าสัญลักษณ์แทนเครื่องบินลำไหนสามารถกำหนดลงไปได้นั้นจะพิจารณาจากเงื่อนไขบังคับที่ได้กล่าวไว้ข้างต้น 3 ข้อด้วยกันคือ

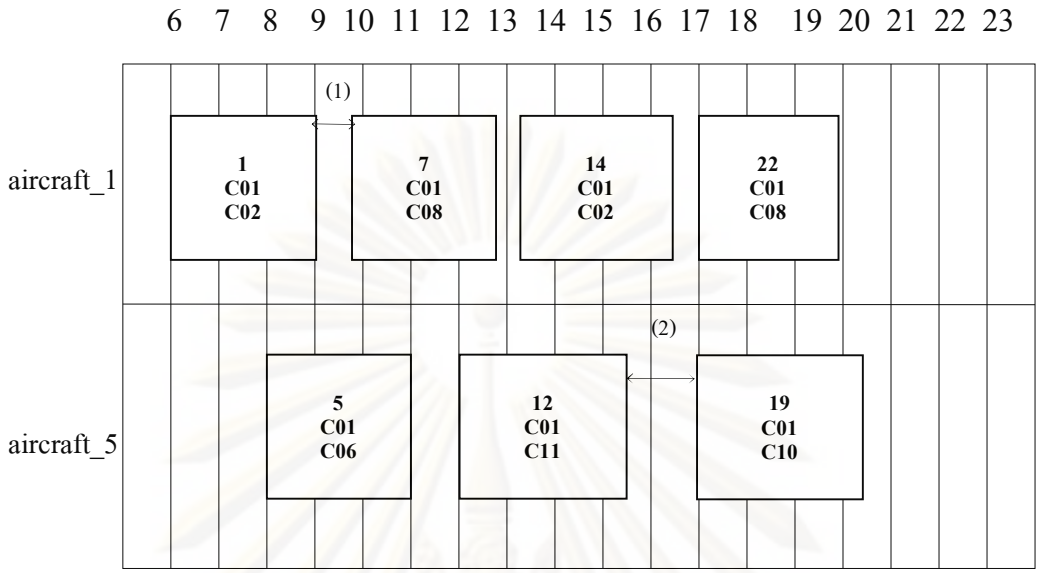
1. เงื่อนไขบังคับของตารางเวลาเที่ยวบินคือ ถ้าจะทำการปรับตารางเวลาเที่ยวบินใหม่ เที่ยวบินที่ทำการปรับนั้นจะต้องออกหลัง หรือออกเท่ากับเวลาที่ได้วางแผนไว้แล้วเท่านั้น



รูปที่ 7 ความผิดปกติที่เกิดขึ้นเป็นผลให้ต้องมีการปรับตารางเวลาใหม่ โดยพิจารณาจากเงื่อนไขบังคับของตารางเวลาเที่ยวบิน

จากรูปที่ 7 การปรับตารางเวลาใหม่นั้นจำต้องเลื่อนเที่ยวบินคู่ที่ 1 ให้ล่าช้าออกไป เพราะเกิดความผิดปกติเป็นผลให้เครื่องบินไม่สามารถใช้งานได้ในช่วงเวลาดังกล่าว แต่ไม่สามารถที่จะเลื่อนเที่ยวบินคู่ที่ 1 มาให้ออกก่อนเวลาที่ได้วางแผนไว้แล้วได้ เนื่องจากผู้โดยสารที่จะต้องโดยสารไปกับเที่ยวบินเที่ยวนั้นๆ ไม่อาจทราบได้ก่อนว่าเที่ยวบินที่ตนเองจะโดยสารจะทำการออกก่อนกำหนด ทำให้ในการปรับตารางเวลาใหม่จำเป็นที่จะต้องคำนึงถึงผู้โดยสารที่จะไปกับเที่ยวบินด้วย

2. กราวนด์ไทม์คือ เวลาขั้นต่ำที่เครื่องบินในแต่ละแบบเครื่องจะต้องจอดบนพื้นดินเมื่อมาถึงสนามบินเช่น เครื่องแบบ Boeing 737-400 จะต้องจอดบนพื้นดินเป็นเวลาขั้นต่ำ 40 นาที ส่วนเครื่องแบบ Airbus A300-600 จะต้องจอดบนพื้นดินเป็นเวลาขั้นต่ำ 50 นาที เป็นต้น (รูปที่ 8)

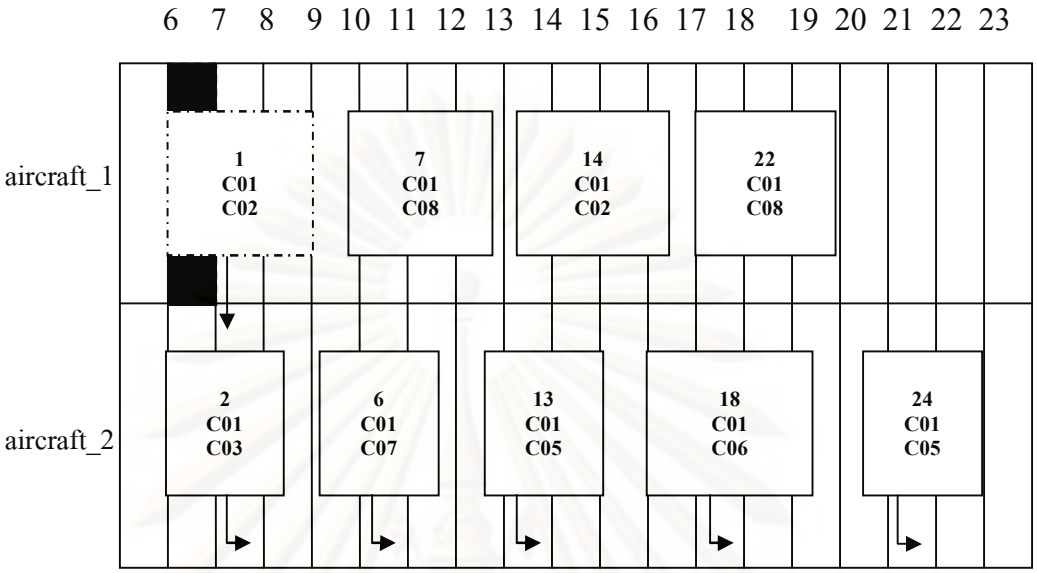


รูปที่ 8 กราวนด์ไทม์ของเครื่องบินในแต่ละแบบเครื่อง

(1) เป็นกราวนด์ไทม์ของเครื่องแบบ Boeing 737-400 และ (2) เป็นกราวนด์ไทม์ของเครื่องแบบ Airbus A300-600

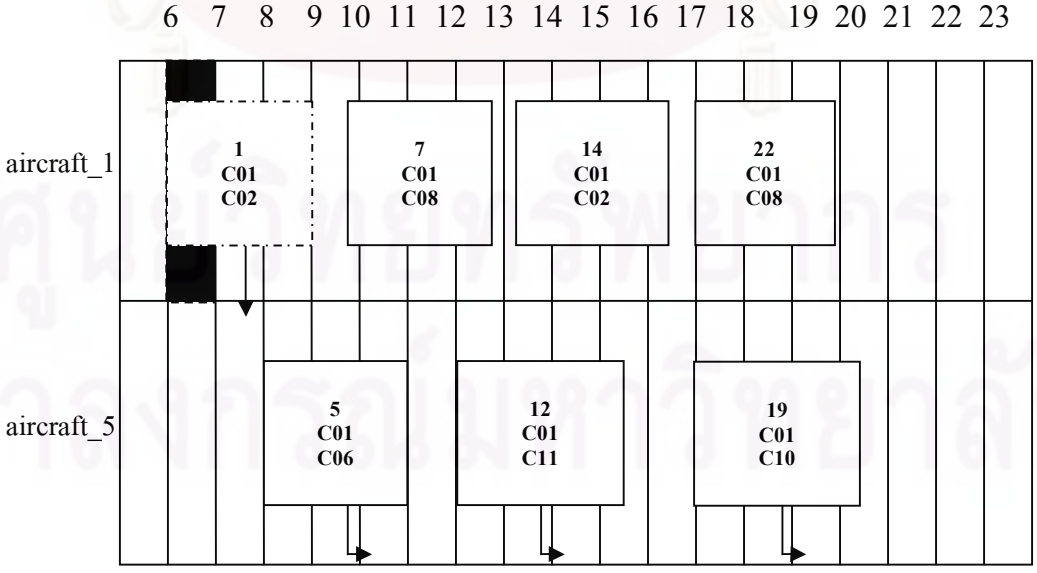
3. ความจุผู้โดยสารคือ ความจุของผู้โดยสารในแต่ละเครื่องบิน โดยความจุของผู้โดยสารในแต่ละเครื่องบินและแบบเครื่องนั้นไม่เท่ากัน เนื่องจากในงานวิจัยนี้มีความต้องการให้การเรียนสามารถรู้นำไปใช้ได้ในทุกกรณี ดังนั้นจึงมีการสมมติให้เครื่องบินในแบบเครื่องเดียวกันมีความจุผู้โดยสารเท่ากัน และการเปลี่ยนเครื่องในแบบเครื่องที่ต่างกันจะทำการเปลี่ยนขึ้นไปเป็นแบบเครื่องที่มีความจุของผู้โดยสารมากกว่าเสมอ ไม่สามารถทำการเปลี่ยนแบบเครื่องจากแบบเครื่องที่มีความจุของผู้โดยสารมากกว่ามาเป็นแบบเครื่องที่มีความจุของผู้โดยสารน้อยกว่าได้ ความจุผู้โดยสารของเครื่องแบบ Boeing 737-400 เท่ากับ 149 ที่นั่ง และความจุผู้โดยสารของเครื่องแบบ Airbus A300-600 เท่ากับ 247 ที่นั่ง

การกระทำที่สนใจในงานวิจัยนี้คือ การปรับตารางเวลาเที่ยวบินที่เครื่องบินจะทำการเปลี่ยนเครื่องบินภายในแบบเครื่องเดียวกัน และการเปลี่ยนเครื่องบินในแบบเครื่องที่ต่างกัน โดยพิจารณาถึงข้อจำกัดดังที่ได้กล่าวมาข้างต้น รูปที่ 9 เป็นตัวอย่างในการปรับตารางเวลาใหม่โดยพิจารณาการเปลี่ยนเครื่องบินภายในแบบเครื่องเดียวกัน เนื่องจากมีความผิดปกติเกิดขึ้นในตารางเวลาเที่ยวบินจึงเป็นผลทำให้ต้องมีการปรับตารางเวลาใหม่ เที่ยวบินคู่ที่ 1 ไม่สามารถใช้เครื่องบินสัญลักษณ์ aircraft_1 ปฏิบัติการทางการบินตามแผนที่วางไว้ได้ จึงทำการเปลี่ยนมาใช้เครื่องบินสัญลักษณ์ aircraft_2 ปฏิบัติการทางการบินแทน ภายหลังจากการเปลี่ยนเครื่องบินทำให้เที่ยวบินคู่ที่ 2 ไม่สามารถออกได้ตามเวลาที่ได้วางแผนจึงจำต้องล่าช้าเที่ยวบินคู่ที่ 2, 6, 13, 18, และ 24 เป็นต้น



รูปที่ 9 ตัวอย่างการปรับตารางเวลาใหม่โดยพิจารณาการเปลี่ยนเครื่องบินในแบบเครื่องเดียวกัน

ส่วนในรูปที่ 10 เป็นตัวอย่างในการปรับตารางเวลาใหม่โดยพิจารณาการเปลี่ยนเครื่องบินในแบบเครื่องที่ต่างกัน เนื่องจากมีความผิดปกติเกิดขึ้นในตารางเวลาเที่ยวบินจึงเป็นผลทำให้ต้องมีการปรับตารางเวลาใหม่ เที่ยวบินคู่ที่ 1 ไม่สามารถใช้เครื่องบินสัญลักษณ์ aircraft_1 ปฏิบัติการทางการบินตามแผนที่วางไว้ได้ จึงทำการเปลี่ยนมาใช้เครื่องบินสัญลักษณ์ aircraft_5 ซึ่งเป็นเครื่องในแบบที่ใหญ่กว่าปฏิบัติการทางการบินแทน ภายหลังจากการเปลี่ยนเครื่องบินทำให้เที่ยวบินคู่ที่ 5 ไม่สามารถออกได้ตามเวลาที่ได้วางแผนจึงจำต้องล่าช้าเที่ยวบินคู่ที่ 5, 12 และ 19 เป็นต้น



รูปที่ 10 ตัวอย่างการปรับตารางเวลาใหม่โดยพิจารณาการเปลี่ยนเครื่องบินในแบบเครื่องที่ต่างกัน

ส่วนต่อไปที่จะกล่าวถึงเป็นส่วนของค่ารางวัลที่จะทำการให้ในแต่ละสถานะที่ตัวเรียนรู้ได้เรียนรู้ ซึ่งถือว่าเป็นส่วนสำคัญอีกส่วนหนึ่งของการประยุกต์การเรียนรู้คิวเพื่อแก้ปัญหาการปรับตารางเวลาเที่ยวบินเมื่อเกิดเหตุผิดปกติเกิดขึ้น

3.1.3 รางวัล

รางวัลที่มีการให้ย่อมแตกต่างกันไปในแต่ละการประยุกต์ใช้งานแต่ว่ามีหลักการเดียวกัน คือ ให้รางวัลเมื่อสามารถทำผลงานได้ดี ดังเช่นในงานของ Tesauro [1] ให้รางวัลจากผลของเกมเมื่อตัวเรียนรู้เล่นเกมจบกระดาน หรือในงานของ Crites และ Barto [13] ที่พิจารณาค่ารางวัลจากเวลาในการรอใช้ลิฟต์ของผู้สาร

เนื่องจากในงานวิจัยนี้ต้องการหาตารางเวลาเที่ยวบินใหม่ที่มีเวลาล่าช้าน้อยๆ ด้วยเหตุนี้ ฟังก์ชันค่ารางวัล (reward function) จะถูกนิยามให้อยู่ในรูปของเวลาล่าช้า เวลาล่าช้าสามารถคำนวณได้จากของผลต่างของเวลาจากตารางเวลาที่ถูกปรับ t_r และเวลาจากตารางเวลาที่ได้วางแผนไว้ t_s ได้มีการนิยาม *DELAY_TIME* ให้เป็นเวลาล่าช้าที่เกิดขึ้นในเที่ยวบินแต่ละคู่ และนิยาม *TOTAL_DELAY_TIME* ให้เป็นเวลาล่าช้ารวมจากเที่ยวบินทุกคู่ *TOTAL_DELAY_TIME* จะแสดงจำนวนเวลาล่าช้ารวมหลังจากจบการปรับตารางเวลาเที่ยวบิน

$$DELAY_TIME(i) = t_r(i) - t_s(i) \quad \text{------ (5)}$$

$$TOTAL_DELAY_TIME = \sum_{i=1}^{n} DELAY_TIME(i) \quad \text{------ (6)}$$

เป้าหมายของการเรียนรู้คือ ต้องการค่ารางวัลมากในขณะที่ *TOTAL_DELAY_TIME* มีค่าน้อย และค่ารางวัลน้อยในขณะที่ *TOTAL_DELAY_TIME* มีค่ามาก ด้วยเหตุนี้จึงจำต้องเขียนฟังก์ชันค่ารางวัลตามนี้

$$r = \begin{cases} -TOTAL_DELAY_TIME & ;if \quad complete \quad schedule \\ Penalty & ;for \quad each \quad reschedule \quad step \end{cases} \quad \text{------ (7)}$$

จากฟังก์ชันค่ารางวัลนี้เห็นได้ว่าค่ารางวัล r คือ ค่าลบของ *TOTAL_DELAY_TIME* เมื่อกระบวนการปรับตารางเวลาจบลง และในแต่ละขั้นตอนของการปรับตารางเวลา จะมีการให้ค่ารางวัลที่เป็นค่าการลงโทษซึ่งจะเป็นค่าลบเล็กๆ เพื่อการเรียนรู้ไม่ติดอยู่กับเส้นทางหนึ่งใดมาก

จนเกินไปหรืออีกนัยหนึ่งคือ เพื่อให้ตัวเรียนรู้ได้เลือกเส้นทางอื่นที่ยังไม่เคยผ่านมาพิจารณาในการเรียนรู้ด้วย

ในส่วนต่อไปนี้จะกล่าวถึงการลู่เข้า (convergence) ที่ใช้งานวิจัยนี้ ซึ่งการลู่เข้าที่พบเห็นในงานวิจัยที่ผ่านมามีอยู่หลายวิธีซึ่งมีความสามารถที่แตกต่างกัน ในส่วนนี้จะอธิบายเกี่ยวกับการลู่เข้าที่เลือกใช้ในงานวิจัยนี้

3.1.4 การลู่เข้า

การลู่เข้าที่ใช้นั้นมีอยู่หลายวิธีเช่น การสุ่มค่าการกระทำไปเรื่อยๆ เพื่อให้สามารถเรียนรู้สถานะต่างๆ ได้อย่างอิสระ [1] หรือการลู่เข้าที่มีการใช้ความน่าจะเป็นเข้ามาช่วย [2, 13] ซึ่งในแต่ละวิธีที่กล่าวมาต่างมีข้อดีในตัวเอง แต่ในงานวิจัยนี้จะขอเลือกการลู่เข้าที่ใช้ความน่าจะเป็นเนื่องจากการลู่เข้าด้วยวิธีนี้สามารถควบคุมการลู่เข้าได้เป็นอย่างดี และยังสามารถกำหนดการลู่เข้าได้จากค่าของความน่าจะเป็น ดังนั้นจึงเป็นทางเลือกที่ดีที่น่าจะนำมาใช้ในงานวิจัยนี้

การลู่เข้าในงานวิจัยนี้ใช้วิธีการลู่เข้าที่ปรากฏใน Zhang และ T. G. Dietterich โดยใช้ยุทธศาสตร์การสำรวจ (exploration strategy) ด้วยความน่าจะเป็น β [2] ขั้นตอนวิธีจะทำการเลือกสัญลักษณ์แทนเครื่องบินแบบสุ่ม เพื่อมาใส่ในตารางเวลา นอกจากนั้นจะเลือกสัญลักษณ์แทนเครื่องบินที่มีค่า $\hat{Q}$ มากที่สุด และค่า β จะถูกตั้งไว้ที่ 1 และลดลงอย่างช้าๆในแต่ละรอบของการเรียนรู้

$$\pi(s) = \begin{cases} random \quad a \quad ; with \quad probability \quad \beta \\ \arg\max_a \hat{Q}(s,a) \quad ; with \quad probability \quad 1-\beta \end{cases} \qquad \text{------ (8)}$$

และในช่วงต้นของการฝึกสอนได้ทำการใช้วิธีเชิงละโมภ (greedy method) [2, 11] เพื่อช่วยให้การลู่เข้าเป็นไปได้อย่างรวดเร็วยิ่งขึ้น

จากที่ได้กล่าวมาข้างต้นนั้นสามารถนำมาเขียนเป็น แผนภูมิสายงานที่แสดงขั้นตอนการทำงานของการประยุกต์การเรียนรู้คิวในการแก้ปัญหาการปรับตารางเวลาเที่ยวบินใหม่ได้ดังรูปที่ 11

ส่วนสุดท้ายที่จะกล่าวถึงในการประยุกต์ใช้การเรียนรู้แบบเสริมกำลังใน การแก้ปัญหาการปรับตารางเวลาเที่ยวบินเมื่อเกิดความผิดปกติขึ้นคือ ฟังก์ชันการประมาณ (approximation function) ในส่วนนี้จะอธิบายถึงที่มาของการใช้ฟังก์ชันการประมาณ และฟังก์ชันการประมาณที่เลือกใช้ในงานวิจัยนี้

3.1.5 ฟังก์ชันการประมาณ

จากที่ได้กล่าวมาในบทก่อนหน้านี้ว่า ฟังก์ชันการประมาณเป็นส่วนที่สำคัญในการนำความรู้ไปใช้ ดังจากที่ได้เห็นในงานวิจัยที่ผ่านมาได้มีการนำฟังก์ชันการประมาณมาใช้เสมอ [1, 2, 13, 14] เพราะในความเป็นจริงแล้วการที่จะให้การเรียนรู้แบบเสริมกำลังทำการเรียนรู้สถานะทั้งหมดเพื่อที่จะนำความรู้ที่เรียนรู้ได้นั้นไปใช้นั้นแทบจะเป็นไปไม่ได้เลย เนื่องจากปัญหาที่ได้กล่าวถึงในงานวิจัยที่ผ่านๆ มานั้นมีขนาดใหญ่มากจนการเรียนรู้สถานะทั้งหมดนั้นยากมากหรือจนแทบจะเป็นไปไม่ได้เลย ดังนั้นจึงได้มีการนำฟังก์ชันการประมาณมาใช้ ฟังก์ชันการประมาณที่ถูกนำมาใช้และมีความสามารถในการประมาณค่าได้อย่างดีฟังก์ชันหนึ่งคือ ข่ายงานประสาทแบบแพร่กระจายย้อนกลับที่ได้รับการยอมรับในหลายๆ งานวิจัย [1, 2, 13] ในงานวิจัยนี้จึงขอยกข่ายงานประสาทแบบแพร่กระจายย้อนกลับ (backpropagation neural network) มาเป็นฟังก์ชันการประมาณในงานวิจัยนี้

ภายหลังจากการเรียนรู้เสร็จสิ้นสามารถนำนโยบายเหมาะสุด π^* ที่ได้จากการเรียนรู้มาใช้หาคำตอบในปัญหาที่ต้องการได้ในที่นี้คือการปรับตารางเวลาเที่ยวบินเมื่อเกิดเหตุไม่ปกติขึ้น โดยการพิจารณาแต่ละการกระทำที่สามารถทำได้ในสถานะนั้นและเลือกการกระทำที่ให้ค่า Q มากที่สุดในสถานะนั้น [10, 11] ตามที่แสดง

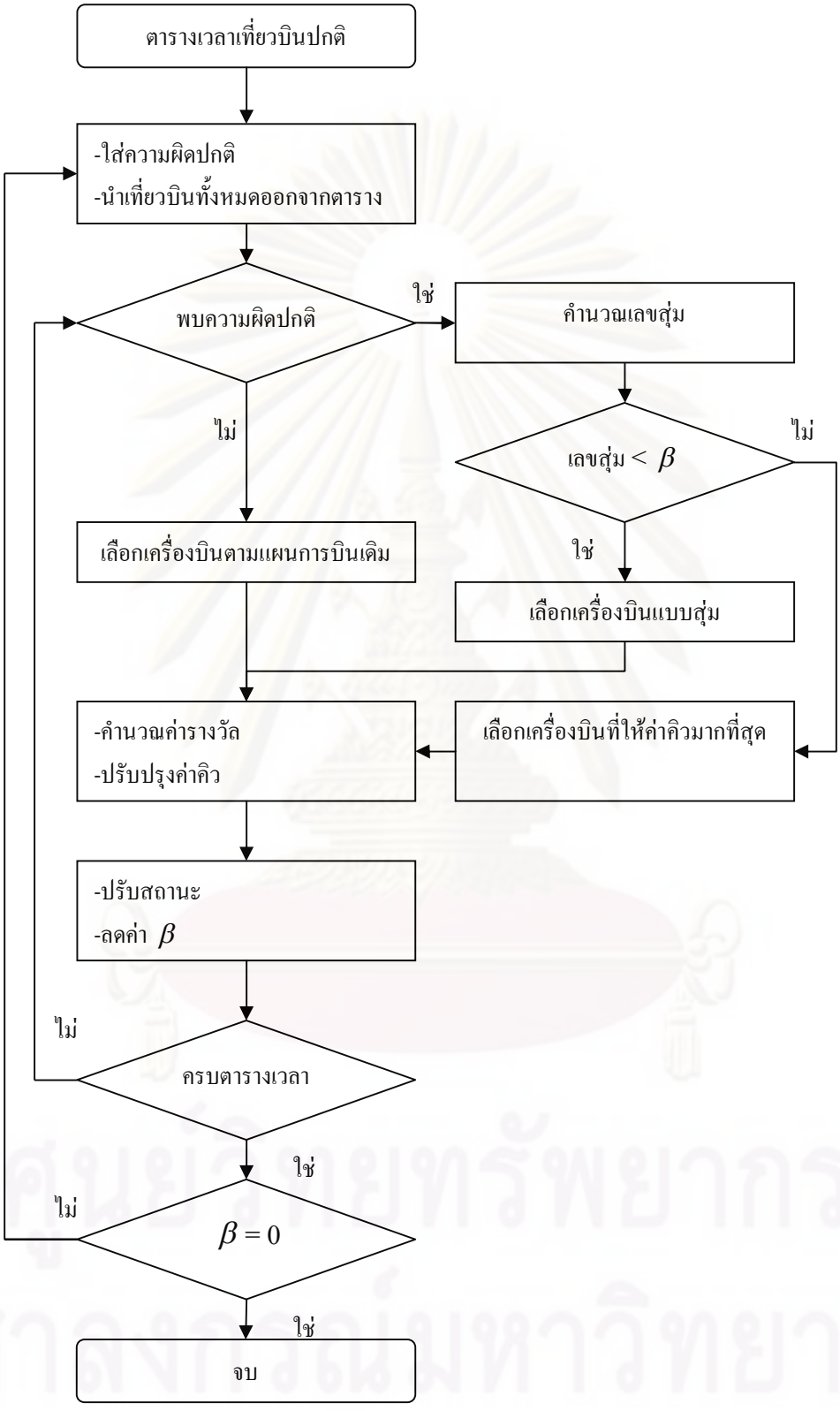
$$\pi^*(s) = \arg\max_a \hat{Q}(s,a) \qquad \text{------ (9)}$$

เพียงแต่ในความเป็นจริงแล้ว ปัญหาการปรับตารางเวลาเที่ยวบินใหม่เมื่อเกิดความผิดปกติขึ้นเป็นปัญหาที่มีสถานะที่เป็นไปได้เยอะมากจนถึงไม่สิ้นสุด ในการที่จะทำการเรียนรู้สถานะทั้งหมดนั้นแทบจะเป็นไปไม่ได้เลย ดังนั้นจึงจำเป็นที่จะต้องมีการปรับความรู้ที่ได้จากการเรียนรู้ให้อยู่ในลักษณะทั่วไปเสียก่อน ซึ่งการปรับความรู้ให้อยู่ในลักษณะทั่วไปนี้จะเห็นได้จากงานวิจัยที่ผ่านๆ มาที่ได้มีการนำฟังก์ชันการประมาณเข้ามาช่วย เพื่อที่จะประมาณค่าในสถานะที่ไม่เคยผ่านมาก่อน

ซึ่งในงานชิ้นนี้ได้นำข่ายงานประสาทแบบการแพร่กระจายย้อนกลับมาใช้เป็นฟังก์ชันการประมาณ ในขั้นตอนการปรับความรู้ให้อยู่ในลักษณะทั่วไปนั้นจะทำการสอนข่ายงานประสาทนี้จากผลของการปรับปรุงค่า $\hat{Q}(s,a)$ ที่ได้ภายหลังจากการเรียนรู้คิวเสร็จสิ้นมาเป็นตัวอย่างสอน โดยมีการนำค่าสถานะและการกระทำมาเป็นชั้นอินพุต ซึ่งจะเป็นชั้นที่รับข้อมูลเกี่ยวกับสถานะและการกระทำที่ได้เรียนรู้มาก่อนแล้วจากการเรียนรู้คิว จะทำการกำหนดตำแหน่งของบัพแต่ละบัพให้เป็นของแต่ละคู่บิน เพื่อความไม่สับสนของข่ายงานประสาท ส่วนถัดมาชั้นฮิดเดนหรือชั้นซ่อนเนื่องจากการใช้ข่ายงานประสาทแบบแพร่กระจายย้อนกลับที่จะมีชั้นซ่อนอยู่ด้วย แลสุดท้ายชั้น

เอาต์พุต ในชั้นนี้จะเป็นชั้นที่แสดงผลของค่าคิวที่เมื่อใส่สถานะและการกระทำเข้าไป จะทำการแสดงค่าคิวของสถานะและการกระทำที่เลือกและกำหนดค่าคิวให้บัพแต่ละบัพของเอาต์พุต ด้วยการประมาณจากข่ายงานประสาทนี้จะสามารถวางหลักให้ปัญหาปัญหานี้จากสถานะที่เคยได้เรียนรู้มาก่อน และคะเนค่าคิว สำหรับสถานะที่ยังไม่เคยพบมาก่อน ในส่วนของการตั้งค่าและการเข้ารหัสนั้นจะขอกล่าวต่อในการทดลอง

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 11 แผนภูมิสายงานแสดงการทำงานของการประยุกต์การเรียนรู้คิวเพื่อแก้ปัญหาการปรับตารางเวลาเที่ยวบิน

บทที่ 4
การทดลองและการวิเคราะห์ผล

บทนี้จะอธิบายถึงวิธีการทดลองการประยุกต์การเรียนรู้คิวในการแก้ปัญหาการปรับตารางเวลาเที่ยวบินใหม่, การทดลองปรับตารางเวลาใหม่เมื่อเกิดเหตุผิดปกติ, การทดลองปรับค่าพารามิเตอร์ที่ใช้ในการเรียนรู้ และส่วนสุดท้ายจะกล่าวถึงผลจากการวิจัยและการวิเคราะห์ผล

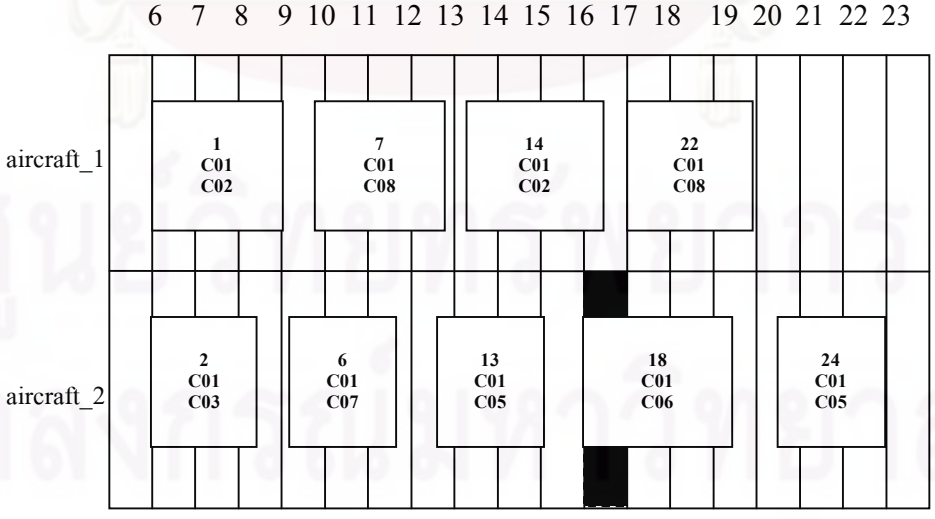
4.1 การทดลองการแสดงสถานะในตารางเวลาเที่ยวบิน

การทดลองนี้ทำเพื่อทำความเข้าใจกับสถานะในการเรียนรู้คิว เพื่อให้สามารถนำสถานะไปใช้ในการประยุกต์การเรียนรู้คิวเพื่อแก้ปัญหาการปรับตารางเวลาได้อย่างถูกต้อง รวมถึงความเข้าใจในพารามิเตอร์ต่างๆ ที่ได้ถูกนิยามไว้เพื่อใช้ในการแสดงสถานะ

สถานะถูกนิยามไว้โดยจัดให้เป็นกลุ่มของเครื่องบินที่ถูกใส่ให้แต่ละเที่ยวบินในตารางเวลาเที่ยวบิน และความผิดปกติที่เกิดขึ้นในตารางเวลาเที่ยวบิน จากนิยามที่ 2 ดังที่แสดง

[*กลุ่มของเครื่องบินที่ถูกใส่ให้แต่ละเที่ยวบิน, ความผิดปกติ*] ------ (2)

วิธีที่ใช้ในการทดลองนี้คือ ทำการทดลองสร้างสถานะจากตารางเวลาเที่ยวบินที่กำหนดไว้ให้ และทำการทดลองปรับตารางเวลาเที่ยวบินโดยใช้การบอกสถานะตามที่ได้นิยามไว้มาแสดงผลให้เห็นในแต่ละสถานะของการปรับตารางเวลาเที่ยวบิน



รูปที่ 12 ตารางเวลาเที่ยวบินของเครื่องบินสัญลักษณ์ aircraft_1 และ aircraft_2 พร้อมความผิดปกติในเครื่องบินสัญลักษณ์ aircraft_2 ที่เวลา 16:00 น.

การทดลองสร้างสถานะจากตารางเวลาเที่ยวบินที่ได้แสดงไว้ในรูปที่ 12

[*1, aircraft_1, 6.00, 170, 2, aircraft_2, 6.05, 140,*
6, aircraft_2, 9.10, 150, 7, aircraft_1, 9.35, 185,
13, aircraft_2, 12.25, 165, 14, aircraft_1, 13.25, 170,
18, aircraft_2, 16.10, 185, 22, aircraft_1, 17.05, 185,
24, aircraft_2, 20.20, 165, aircraft_2, 16.00, 60]

จากสถานะที่แสดงพบว่าเกิดความผิดปกติขึ้นกับเครื่องบิน aircraft_2 ที่เวลา 16.00 น. เป็นเวลาหนึ่งชั่วโมง ซึ่งเป็นผลให้เครื่องบิน aircraft_2 ใช้งานไม่ได้ในเวลาดังกล่าว ทำให้ต้องมีการปรับตารางเวลาเที่ยวบินเพื่อให้สามารถปฏิบัติการทางการบินต่อไปได้ ความผิดปกติที่เกิดขึ้นทำให้เที่ยวบินคู่ที่ 18 ไม่สามารถปฏิบัติการทางการบินได้ตามเวลาที่กำหนดจึงพิจารณาเปรียบเทียบระหว่าง aircraft_1 และ aircraft_2 ที่อยู่ในแบบเครื่องเดียวกันว่าเครื่องบินลำไหนสามารถปฏิบัติการทางการบินได้เร็วสุด เมื่อพิจารณาข้อจำกัดแล้วพบว่าเครื่องบิน aircraft_1 สามารถปฏิบัติการได้เร็วกว่า aircraft_2 จึงเลือกปรับตารางเวลาเที่ยวบินโดยใช้เครื่องบิน aircraft_1 ปฏิบัติการ และทำให้สถานะใหม่ที่เกิดขึ้นคือ

[*1, aircraft_1, 6.00, 170, 2, aircraft_2, 6.05, 140,*
6, aircraft_2, 9.10, 150, 7, aircraft_1, 9.35, 185,
13, aircraft_2, 12.25, 165, 14, aircraft_1, 13.25, 170,
18, aircraft_1, 16.55, 185, 0, 0, 0, 0,
0, 0, 0, 0, aircraft_2, 16.00, 60]

จากนั้นพิจารณาเที่ยวบินคู่ต่อไปคือ เที่ยวบินคู่ที่ 22 พบว่าในช่วงเวลาดังกล่าวเครื่องบินได้ถูกใช้งานโดยเที่ยวบินคู่ที่ 18 ไปแล้วแต่เมื่อพิจารณาความผิดปกติที่เกิดขึ้นพบว่าความผิดปกติได้สิ้นสุดลงในเครื่องบิน aircraft_2 ทำให้เที่ยวบินคู่ที่ 22 สามารถปฏิบัติการได้ตามเวลาที่ได้วางแผนไว้โดยใช้เครื่องบิน aircraft_2 แทนในการปฏิบัติการและสถานะใหม่ที่ได้คือ

[*1, aircraft_1, 6.00, 170, 2, aircraft_2, 6.05, 140,*
6, aircraft_2, 9.10, 150, 7, aircraft_1, 9.35, 185,

13, aircraft_2, 12.25, 165, 14, aircraft_1, 13.25, 170,
18, aircraft_1, 16.55, 185, 22, aircraft_2, 17.05, 185,
0, 0, 0, 0, aircraft_2, 16.00, 60]

สุดท้ายพิจารณาเที่ยวบินคู่ที่ 24 โดยการพิจารณาเปรียบเทียบระหว่าง aircraft_1 และ aircraft_2 ที่อยู่ในแบบเครื่องเดียวกันว่า เครื่องบินลำไหนสามารถปฏิบัติการทางการบินได้เร็วสุดเมื่อพิจารณาข้อจำกัดแล้ว พบว่าเครื่องบิน aircraft_1 สามารถปฏิบัติการได้เร็วกว่า aircraft_2 จึงเลือกปรับตารางเวลาเที่ยวบินโดยใช้เครื่องบิน aircraft_1 ปฏิบัติการ และทำให้สถานะใหม่ที่เกิดขึ้นคือ

[*1, aircraft_1, 6.00, 170, 2, aircraft_2, 6.05, 140,*
6, aircraft_2, 9.10, 150, 7, aircraft_1, 9.35, 185,
13, aircraft_2, 12.25, 165, 14, aircraft_1, 13.25, 170,
18, aircraft_1, 16.55, 185, 22, aircraft_2, 17.05, 185,
24, aircraft_1, 20.35, 165, aircraft_2, 16.00, 60]

จากการทดลองนี้ทำให้เข้าใจและสามารถอธิบายตารางเวลาเที่ยวบินจากรูปแบบของตารางเวลาเที่ยวบินให้มาอยู่ในรูปแบบของสถานะได้ การระบุสถานะนั้นถือเป็นส่วนสำคัญส่วนแรกในการประยุกต์การเรียนรู้คิวเพื่อให้ได้ผลการทดลองที่ถูกต้อง

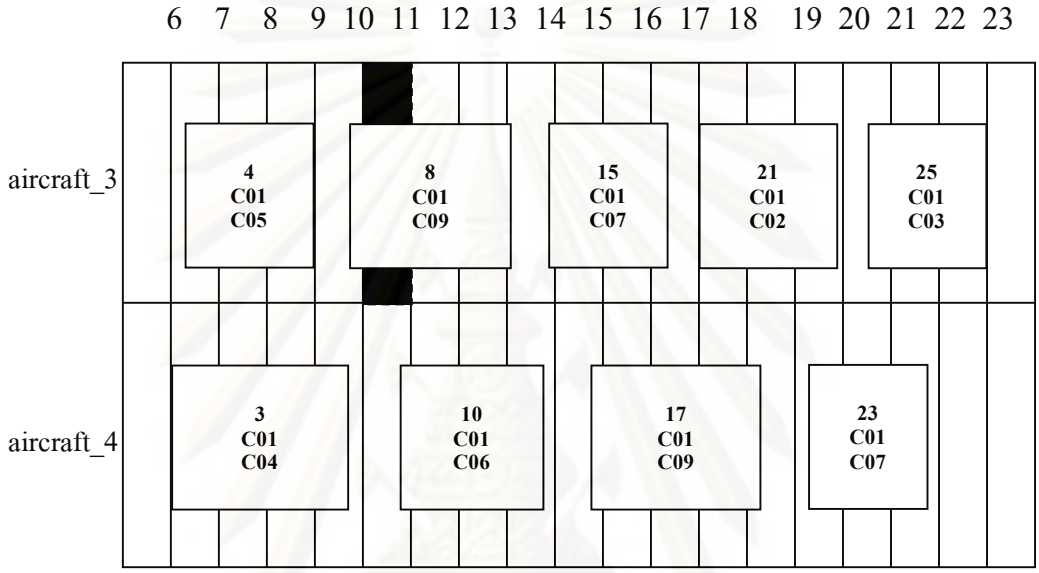
4.2 การทดลองการปรับตารางเวลาเที่ยวบินใหม่และคำนวณค่ารางวัล

การทดลองนี้ทำเพื่อทำความเข้าใจการกระทำและค่ารางวัลในการเรียนรู้คิว เพื่อให้สามารถนำการกระทำและค่ารางวัลไปใช้ในการประยุกต์การเรียนรู้คิวเพื่อแก้ปัญหาการปรับตารางเวลาได้อย่างถูกต้อง และเข้าใจถึงความสามารถในการกระทำที่สามารถทำได้ในแต่ละสถานะ รวมถึงเข้าใจความสัมพันธ์ระหว่างค่ารางวัลกับการปรับปรุงค่าคิว

การกระทำที่สนใจในงานวิจัยนี้คือ การปรับตารางเวลาเที่ยวบินที่เครื่องบินจะทำการเปลี่ยนเครื่องบินภายในแบบเครื่องเดียวกัน และการเปลี่ยนเครื่องบินในแบบเครื่องที่ต่างกัน โดยพิจารณาถึงข้อจำกัดสามส่วนหลักคือ เงื่อนไขบังคับของตารางเวลาเที่ยวบิน, กราวนด์ไทม์ และความจุผู้โดยสาร

ในการทดลองนี้จะทำการปรับตารางเวลาเที่ยวบินใหม่ตามตารางเวลาเที่ยวบินจากรูปที่ 13 ซึ่งในการปรับตารางเวลาเที่ยวบินนี้จะทำการกำหนดการกระทำในแต่ละสถานะไว้ แล้วให้ทำการคำนวณค่ารางวัลที่ได้ในแต่ละสถานะและนำมาปรับปรุงค่าคิวตามที่ได้กล่าวไว้ก่อนหน้าในนิยามที่

(1) ความผิดปกติที่เกิดขึ้นในตารางเวลาเที่ยวบินเป็นผลให้ต้องมีการปรับตารางเวลาเที่ยวบินใหม่ตามตารางที่ 3 และเนื่องจาก aircraft_3 มีความจุผู้โดยสารน้อยกว่า aircraft_4 จึงไม่สามารถที่จะทำการเปลี่ยนเครื่องจาก aircraft_4 มาเป็น aircraft_3 ได้แต่สามารถเปลี่ยนเครื่องจาก aircraft_3 ไปเป็น aircraft_4 ได้ ค่า *DELAY_TIME* ที่ได้ในตารางที่ 3 นั้นคำนวณจากเวลาที่คู่บินสามารถทำการบินได้โดยไม่ขัดกับเงื่อนไขบังคับเทียบกับเวลาที่ได้เคยวางแผนไว้



รูปที่ 13 ตารางเวลาเที่ยวบินของเครื่องบินสัญลักษณ์ aircraft_3 และ aircraft_4 พร้อมความผิดปกติในเครื่องบินสัญลักษณ์ aircraft_3 ที่เวลา 10:00 น.

ตารางที่ 3 การปรับตารางเวลาตามการกระทำที่กำหนดเพื่อคำนวณหาค่า *TOTAL_DELAY_TIME*

หมายเลขคู่บิน	การกระทำ	*DELAY_TIME* (นาที)	*TOTAL_DELAY_TIME* (นาที)
3	aircraft_4	0	0
4	aircraft_3	0	0
8	aircraft_3	75	75
10	aircraft_4	0	75
15	aircraft_4	45	120
17	aircraft_4	140	260
21	aircraft_3	0	0
23	aircraft_4	75	335
25	aircraft_3	0	335

ในส่วนของการคำนวณค่ารางวัลที่จะให้ในแต่ละรอบของการปรับปรุงค่า $\hat{Q}(s,a)$ จะเป็นไปตามฟังก์ชัน *r* ที่ได้นิยามไว้ในนิยามที่ (7) ก่อนหน้านี้ คือ

$$r = \begin{cases} -TOTAL_DELAY_TIME & ;if \quad complete \quad schedule \\ Penalty & ;for \quad each \quad reschedule \quad step \end{cases} \qquad \text{------ (7)}$$

ในที่นี้ค่า *Penalty* ถูกกำหนดไว้ที่ -10 สำหรับในแต่ละรอบการปรับปรุงค่า $\hat{Q}(s,a)$ ในแต่ละสถานะที่ยังปรับตารางเวลาเที่ยวบินไม่ครบทั้งตารางเวลา และรางวัลที่จะได้รับเมื่อปรับตารางเวลาเที่ยวบินครบทั้งตารางจะเท่ากับ *-TOTAL_DELAY_TIME* และในส่วนของการเรียนรู้คิวนั้นจะใช้การปรับปรุงค่า $\hat{Q}(s,a)$ ผ่านรูปแบบที่ได้นิยามไว้ในนิยามที่ (1) คือ

$$\hat{Q}(s,a) \leftarrow \hat{Q}(s,a) + \alpha\left[r + \gamma \max_{a'} \hat{Q}(s',a') - \hat{Q}(s,a)\right] \qquad \text{------ (1)}$$

และได้ทำการกำหนดค่าα ไว้ที่ 0.7 และค่า γ ไว้ที่ 0.3 ส่วนวิธีการคำนวณการปรับปรุงค่า $\hat{Q}(s,a)$ จากรอบที่ 10ในตารางที่ 4 สามารถแสดงได้ดังนี้

$\hat{Q}(s_1, aircraft_4)$ = -14.261+0.7(-10+0.3(-14.275)-(-14.261)) = -14.276
$\hat{Q}(s_2, aircraft_3)$ = -14.275+0.7(-10+0.3(-14.358)-(-14.275)) = -14.298
$\hat{Q}(s_3, aircraft_3)$ = -14.358+0.7(-10+0.3(-14.701)-(-14.358)) = -14.395
$\hat{Q}(s_4, aircraft_4)$ = -14.701+0.7(-10+0.3(-15.95)-(-14.701)) = -14.76
$\hat{Q}(s_5, aircraft_4)$ = -15.95+0.7(-10+0.3(-20.156)-(-15.95)) = -16.018
$\hat{Q}(s_6, aircraft_4)$ = -20.156+0.7(-10+0.3(-34.076)-(-20.156)) = -20.203
$\hat{Q}(s_7, aircraft_3)$ = -34.076+0.7(-10+0.3(-80.345)-(-34.076)) = -34.095
$\hat{Q}(s_8, aircraft_4)$ = -80.345+0.7(-10+0.3(-335)-(-80.345)) = -80.349
$\hat{Q}(s_9, aircraft_3)$ = -335+0.7(-335+0.3(0)-(-335)) = -335

โดย s_1 คือ *[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, aircraft_3, 10.00, 60]

s_2 คือ *[3, aircraft_4, 6.05, 225, 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, aircraft_3, 10.00, 60]

s_3 คือ *[3, aircraft_4, 6.05, 225, 4, aircraft_3, 6.15, 165,*
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, aircraft_3, 10.00, 60]

s_4 คือ *[3, aircraft_4, 6.05, 225, 4, aircraft_3, 6.15, 165,*
8, aircraft_3, 11.00, 200, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, aircraft_3, 10.00, 60]

s_5 คือ *[3, aircraft_4, 6.05, 225, 4, aircraft_3, 6.15, 165,*
8, aircraft_3, 11.00, 200, 10, aircraft_4, 10.40, 185,
0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, aircraft_3, 10.00, 60]

s_6 คือ *[3, aircraft_4, 6.05, 225, 4, aircraft_3, 6.15, 165,*
8, aircraft_3, 11.00, 200, 10, aircraft_4, 10.40, 185,
15, aircraft_4, 14.35, 150, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, aircraft_3, 10.00, 60]

s_7 คือ *[3, aircraft_4, 6.05, 225, 4, aircraft_3, 6.15, 165,*
8, aircraft_3, 11.00, 200, 10, aircraft_4, 10.40, 185,
15, aircraft_4, 14.35, 150, 17, aircraft_4, 17.05, 205,
0, 0, 0, 0, 0, 0, 0, 0,

0, 0, 0, 0, aircraft_3, 10.00, 60]

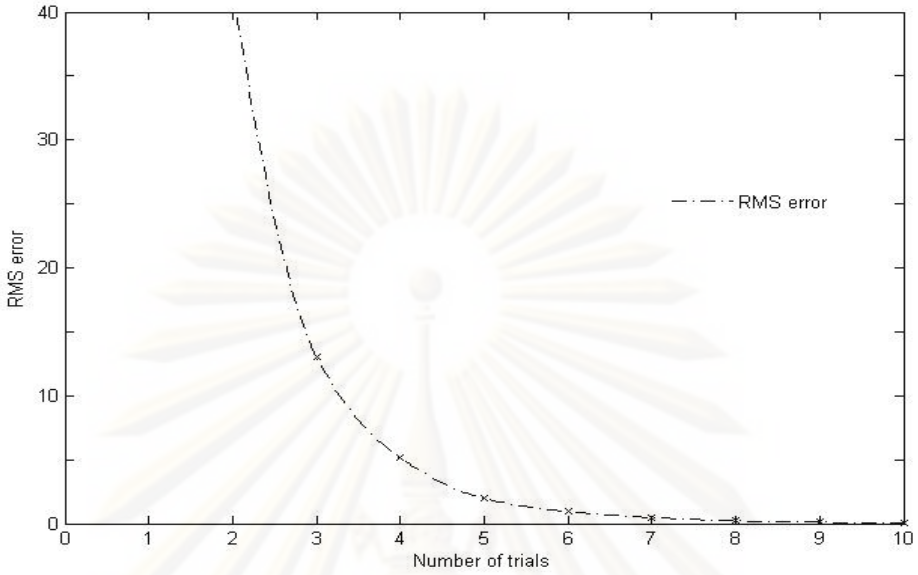
s_8 คือ *[3, aircraft_4, 6.05, 225, 4, aircraft_3, 6.15, 165,*
8, aircraft_3, 11.00, 200, 10, aircraft_4, 10.40, 185,
15, aircraft_4, 14.35, 150, 17, aircraft_4, 17.05, 205,
21, aircraft_3, 17.05, 170, 0, 0, 0, 0,
0, 0, 0, 0, aircraft_3, 10.00, 60]

s_9 คือ *[3, aircraft_4, 6.05, 225, 4, aircraft_3, 6.15, 165,*
8, aircraft_3, 11.00, 200, 10, aircraft_4, 10.40, 185,
15, aircraft_4, 14.35, 150, 17, aircraft_4, 17.05, 205,
21, aircraft_3, 17.05, 170, 23, aircraft_4, 20.30, 155,
0, 0, 0, 0, aircraft_3, 10.00, 60]

ในการทดลองนี้ที่ทำการแสดงการปรับปรุงค่า $\hat{Q}(s,a)$ ในสถานะต่างๆ นั้นก็เพื่อพิสูจน์ว่าการปรับปรุงค่า $\hat{Q}(s,a)$ จากฟังก์ชัน r ที่ได้นิยามขึ้นนั้นสามารถใช้งานได้จริงและสามารถคำนวณค่าได้ในแต่ละสถานะ และได้ทำการแสดงค่ารากกำลังสองเฉลี่ยไว้ในรูปที่ 14 ในการทดลองนี้เป็นเพียงส่วนย่อยของสถานะที่เป็นไปได้ทั้งหมดของปัญหาจากการทดลองนี้ ซึ่งถ้าทำการเรียนรู้สถานะทั้งหมดที่เป็นไปได้จากปัญหาที่เกิดขึ้นในการทดลองนี้จะพบว่ามีจำนวนสถานะที่เป็นไปได้มากมากและขนาดใหญ่มาก ดังนั้นการที่จะทำให้การประยุกต์การเรียนรู้คิวสำเร็จผลนั้นจำต้องมีส่วนเพิ่มเติมที่เข้ามาช่วยในการค้นหาสถานะซึ่งจะขอกล่าวในการทดลองถัดไป

ตารางที่ 4 $\hat{Q}(s,a)$ ที่ได้จากการปรับปรุงค่าในแต่ละรอบ

สถานะ	$\hat{Q}(s,a)$									
	รอบที่ 1	รอบที่ 2	รอบที่ 3	รอบที่ 4	รอบที่ 5	รอบที่ 6	รอบที่ 7	รอบที่ 8	รอบที่ 9	รอบที่10
s_1	-7	-11.4	-12.814	-13.535	-13.902	-14.09	-14.186	-14.235	-14.261	-14.276
s_2	-7	-11.4	-12.814	-13.535	-13.902	-14.09	-14.186	-14.239	-14.275	-14.298
s_3	-7	-11.4	-12.814	-13.535	-13.902	-14.09	-14.205	-14.303	-14.358	-14.395
s_4	-7	-11.4	-12.814	-13.535	-13.902	-14.18	-14.483	-14.606	-14.701	-14.76
s_5	-7	-11.4	-12.814	-13.535	-14.337	-15.045	-15.529	-15.807	-15.95	-16.018
s_6	-7	-11.4	-12.814	-15.605	-17.832	-19.12	-19.754	-20.037	-20.156	-20.203
s_7	-7	-11.4	-22.673	-29.29	-32.24	-33.42	-33.862	-34.021	-34.076	-34.095
s_8	-7	-58.345	-73.749	-78.369	-79.756	-80.17	-80.296	-80.334	-80.345	-80.349
s_9	-234.5	-304.85	-325.95	-334.99	-335	-335	-335	-335	-335	-335



รูปที่ 14 รากกำลังสองเฉลี่ยจากการทดลองที่ 4.2

4.3 การทดลองการลู่เข้า

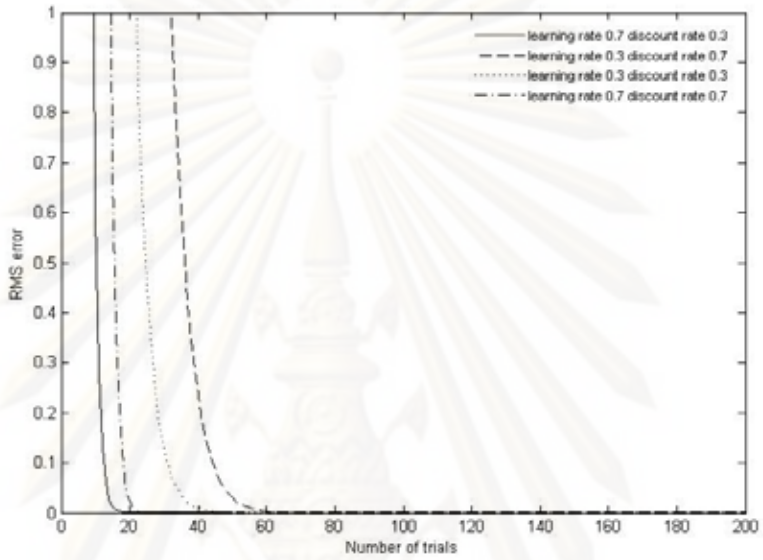
การทดลองเพื่อทำความเข้าใจการลู่เข้าโดยใช้ยุทธศาสตร์การสำรวจในการเรียนรู้คิว และสามารถนำการลู่เข้าไปใช้ในการประยุกต์การเรียนรู้คิวเพื่อแก้ปัญหาการปรับตารางเวลาได้อย่างถูกต้อง

จากที่ได้กล่าวมาในการทดลองก่อนหน้านี้ว่าการที่เราจะทำการเรียนรู้สถานะทั้งหมดนั้นยากมาก หรือแทบจะเป็นไปไม่ได้เลย ดังนั้นเราจำต้องหาวิธีการเรียนรู้ที่จะได้มาทั้งการเรียนรู้สถานะที่มากที่สุดโดยไม่ทิ้งนโยบายควบคุมของความรู้ และสามารถควบคุมการเรียนรู้ได้ วิธีที่เหมาะสุดสำหรับเหตุผลที่ได้กล่าวมาคือ การใช้การลู่โดยใช้ยุทธศาสตร์การสำรวจด้วยความน่าจะเป็น β [2] และใช้นโยบายในการเลือกการกระทำในแต่ละสถานะตามนิยามที่ (8) คือ

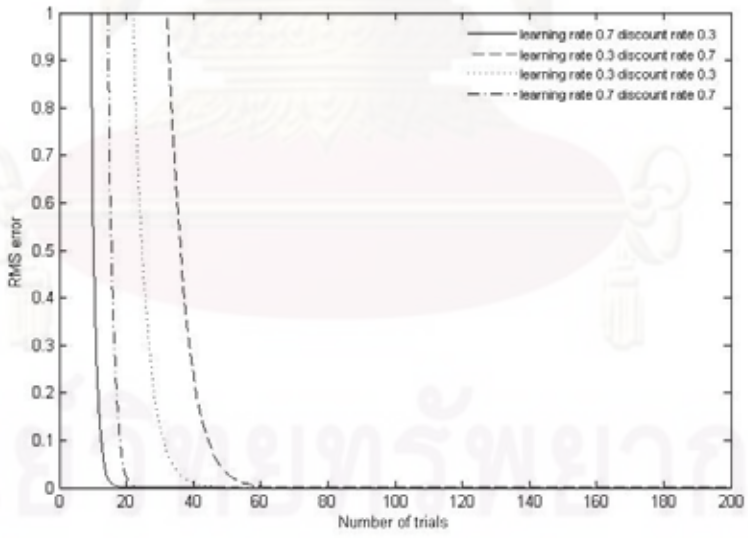
$$\pi(s) = \begin{cases} random \quad a & ;with \quad probability \quad \beta \\ \arg\max_a \hat{Q}(s,a) & ;with \quad probability \quad 1-\beta \end{cases} \qquad \text{------ (8)}$$

ในการทดลองนี้จะทำการทดลองเกี่ยวกับการลู่เข้า โดยจะทำการตั้งค่า β เริ่มต้นค่าไว้ที่ 1 และทำการลดลงเรื่อยๆครั้งละ 0.005 และทำการพิจารณาค่า α และ γ โดยจะทำการตั้งค่า α ไว้ที่ 0.7 และค่า γ ไว้ที่ 0.3, ค่า α ไว้ที่ 0.3 และค่า γ ไว้ที่ 0.7, ค่าα ไว้ที่ 0.3 และค่า γ ไว้ที่ 0.3, ค่า α ไว้ที่ 0.7 และค่า γ ไว้ที่ 0.7 เพื่อเปรียบเทียบความสามารถในการลู่เข้าของแต่ละกรณีเพื่อที่จะนำไปใช้ในการทดลองถัดไป โดยตัวแปรต่างๆ จะใช้ในงานขั้นตอนวิธีการเรียนรู้คิว โดยในการทดลองจะ

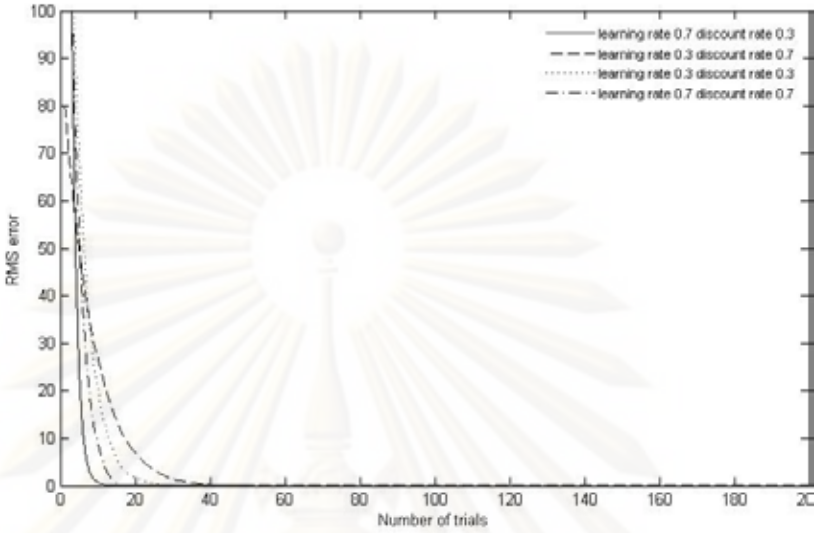
ทำการฝึกสอนทั้งหมดสามครั้งแล้วเลือกนำผลที่ให้ค่ารากกำลังสองเฉลี่ยที่ดีที่สุดมาใช้เพื่อทำการนำค่าพารามิเตอร์ที่ให้ผลดีที่สุดไปใช้ในการทดลองถัดไป



15.1



15.2



15.3

รูปที่ 15 รากกำลังสองเฉลี่ยจากการทดลองที่ 4.3

4.4 การทดลองประยุกต์การเรียนรู้คิวในตารางเวลาเที่ยวบิน

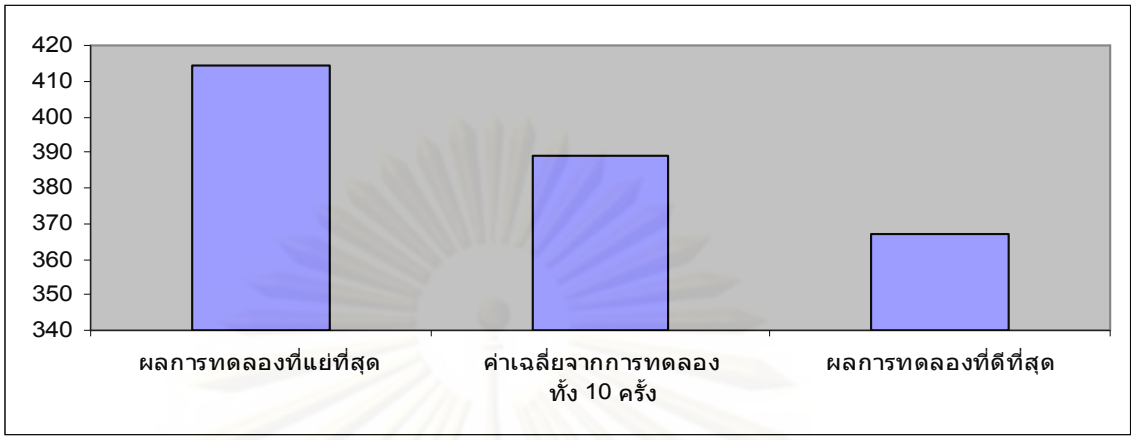
การทดลองเพื่อทำความเข้าใจการแก้ปัญหาความผิดปกติในตารางเวลาเที่ยวบินด้วยการเรียนรู้คิว และสามารถนำความรู้ที่ได้จากการเรียนรู้ไปใช้กับการแก้ปัญหาความผิดปกติอื่นๆ ที่ยังไม่ได้เรียนรู้ได้

ในการทดลองนี้จะทำการทดลองแก้ปัญหาการปรับตารางเวลาเที่ยวบินใหม่ โดยความผิดปกติที่ใช้ในการทดลองจะทำการสร้างขึ้นมาในตารางเวลาเที่ยวบินปกติที่ได้ทำการวางแผนไว้และสร้างความผิดปกติเข้าไปเริ่มตั้งแต่เวลา 6.00 น. จนถึงเวลา 20.00 น. ในทุกๆ สองชั่วโมง ซึ่งจะได้ความผิดปกติทั้งหมดจำนวน 56 ตัวอย่าง ความผิดปกติที่สร้างขึ้นนี้เป็นผลทำให้เครื่องบินเครื่องที่เกิดความผิดปกติไม่สามารถใช้งานได้จนกระทั่งความผิดปกติสิ้นสุดลง ในการทดลองจะทำการแบ่งตัวอย่างออกมาจำนวน 50 ตัวอย่างเพื่อใช้ทำการสอนการเรียนรู้คิว และที่เหลืออีก 6 ตัวอย่างใช้เพื่อทำการทดสอบประสิทธิภาพของการเรียนรู้คิวในการทดลองถัดไป โดยจะทำการตั้งค่า β เริ่มต้นค่าไว้ที่ 1 และทำการลดลงเรื่อยๆครั้งละ 0.00001 โดยจะทำการตั้งค่า α ไว้ที่ 0.7 และค่า γ ไว้ที่ 0.3

ตารางที่ 5 เปรียบเทียบค่า *TOTAL_DEAY_TIME* จากครั้งที่ให้ผลการทดลองดีที่สุดเทียบกับค่าเฉลี่ยของการทดลองทั้ง 10 ครั้ง

ความผิดปกติ	*TOTAL_DELAY_TIME* (นาที)	
	ผลการทดลองที่ดีที่สุด	ค่าเฉลี่ยจากการทดลอง
[*aircraft_1, 6.00, 60*]	695	733.5
[*aircraft_1, 8.00, 60*]	695	729
[*aircraft_1, 10.00, 60*]	385	387.5
[*aircraft_1, 12.00, 60*]	320	352
[*aircraft_1, 14.00, 60*]	460	516
[*aircraft_1, 18.00, 60*]	85	99.5
[*aircraft_1, 20.00, 60*]	85	88
[*aircraft_2, 6.00, 60*]	1575	1596.5
[*aircraft_2, 8.00, 60*]	1550	1596.5
[*aircraft_2, 10.00, 60*]	695	700
[*aircraft_2, 12.00, 60*]	185	201.5
[*aircraft_2, 16.00, 60*]	0	36.5
[*aircraft_2, 18.00, 60*]	85	100.5
[*aircraft_2, 20.00, 60*]	0	5
[*aircraft_3, 6.00, 60*]	1475	1506.5
[*aircraft_3, 8.00, 60*]	1430	1482.5
[*aircraft_3, 10.00, 60*]	430	434
[*aircraft_3, 12.00, 60*]	355	359
[*aircraft_3, 14.00, 60*]	235	290
[*aircraft_3, 16.00, 60*]	360	381.5
[*aircraft_3, 18.00, 60*]	85	102.5
[*aircraft_3, 20.00, 60*]	0	30
[*aircraft_4, 6.00, 60*]	795	802
[*aircraft_4, 8.00, 60*]	795	821

[*aircraft_4, 10.00, 60*]	605	638
[*aircraft_4, 12.00, 60*]	605	648
[*aircraft_4, 14.00, 60*]	190	214.5
[*aircraft_4, 16.00, 60*]	65	98.5
[*aircraft_4, 20.00, 60*]	0	8
[*aircraft_5, 6.00, 60*]	0	32
[*aircraft_5, 8.00, 60*]	695	730.5
[*aircraft_5, 10.00, 60*]	695	734
[*aircraft_5, 12.00, 60*]	430	464
[*aircraft_5, 14.00, 60*]	535	543
[*aircraft_5, 16.00, 60*]	0	35
[*aircraft_5, 20.00, 60*]	85	101.5
[*aircraft_6, 6.00, 60*]	0	0
[*aircraft_6, 8.00, 60*]	0	0
[*aircraft_6, 14.00, 60*]	390	434
[*aircraft_6, 16.00, 60*]	75	107.5
[*aircraft_6, 18.00, 60*]	0	0
[*aircraft_6, 20.00, 60*]	0	0
[*aircraft_7, 6.00, 60*]	0	0
[*aircraft_7, 8.00, 60*]	0	0
[*aircraft_7, 10.00, 60*]	0	0
[*aircraft_7, 12.00, 60*]	225	266
[*aircraft_7, 14.00, 60*]	640	655
[*aircraft_7, 16.00, 60*]	175	218
[*aircraft_7, 18.00, 60*]	85	95.5
[*aircraft_7, 20.00, 60*]	85	95



รูปที่ 16 เปรียบเทียบค่าเฉลี่ยของ *TOTAL_DELAY_TIME* จากการทดลองทั้ง 10 ครั้ง

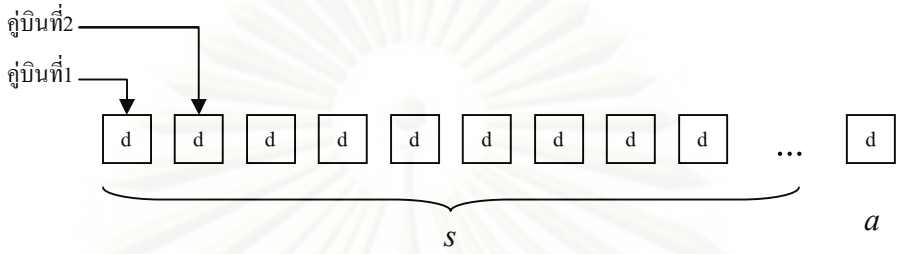
4.5 การทดลองเปรียบเทียบผลของการเรียนรู้คิว

ในการทดลองนี้จะนำผลการเรียนรู้ที่ดีที่สุดจากการทดลองที่ 4.4 นำมาทำความรู้ที่เรียนได้ให้อยู่ในรูปทั่วไปโดยการใช้ฟังก์ชันการประมาณ ซึ่งในการทดลองนี้จะใช้งานฟังก์ชันการประมาณด้วยข่ายงานประสาทแบบแพร่กระจายย้อนกลับ ในการใช้ประมาณด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับทำการฝึกสอน โดยข้อมูลจากการทดลองที่ผ่านมานั้นจะอยู่ในรูปของ $\hat{Q}(s,a)$ ในส่วนของรูปแบบของฟังก์ชันการประมาณที่ต้องการคือ เมื่อทราบสถานะและการกระทำแล้วจะต้องทราบค่าของคิวที่สอดคล้องกับสถานะนั้นๆ ดังนั้นจากข้อมูลที่ทราบจากการทดลองคือ $\hat{Q}(s,a)$ ที่เป็นค่าคิวที่จะได้ในแต่ละสถานะที่กระทำการกระทำลงไป จากที่กล่าวมานี้จึงทำให้ทราบได้ว่ามีข้อมูลในส่วนของสถานะที่ได้จาก s ใน $\hat{Q}(s,a)$, ข้อมูลในส่วนของการกระทำที่ได้จาก a ใน $\hat{Q}(s,a)$ และข้อมูลในส่วนของค่าคิวที่ได้จากตัว $\hat{Q}(s,a)$ นั่นเอง

ดังนั้นขั้นตอนในการที่จะสร้างข่ายงานประสาทแบบแพร่กระจายย้อนกลับให้ได้ตามที่ต้องการที่ว่า เมื่อใส่สถานะและการกระทำเข้าไปแล้วจะได้ค่าคิวออกมาจึงสามารถทำได้โดยมีขึ้นตอนการเข้ารหัส เพื่อแปลงข้อมูลที่ได้มาจากการทดลองก่อนหน้าให้ใช้ได้กับการสอนข่ายงานประสาท โดยเริ่มจาก

1. ชั้นอินพุต ซึ่งจะเป็นชั้นที่รับข้อมูลเกี่ยวกับสถานะและการกระทำที่ได้เรียนรู้มาก่อนแล้วจากการเรียนรู้คิว วิธีเข้ารหัสนั้นจะทำการกำหนดตำแหน่งของบัพแต่ละบัพให้เป็นของแต่ละคู่บินเพื่อความไม่สับสนของข่ายงานประสาท โดยจะทำหน้าที่รับข้อมูลของสถานะที่เรียงตามลำดับตามหมายเลขคู่บินข้อมูลที่ใส่คือเลข 1 ถึง 7 ที่เข้ารหัสแทนสัญลักษณ์แทนเครื่องบิน aircraft_1 ถึง aircraft_7 ที่ได้จากการเรียนรู้คิวแล้วละส่วนที่บอกเวลาออกของคู่บิน และระยะเวลาที่ใช้ในแต่ละคู่บินออก ในส่วนบัพสุดท้ายจะถูกเข้ารหัสให้เป็นบัพที่รับข้อมูลในส่วนของการกระทำในสถานะ

ดังกล่าว โดยเข้ารหัสเป็นเลข 1 ถึง 7 ใช้แทนการกระทำสำหรับสัญลักษณ์แทนเครื่องบิน aircraft_1 ถึง aircraft_7 ดังรูปที่ 17 จะเป็นการเรียงลำดับตามคู่บินส่วนบัพสุดท้ายใช้แทนการกระทำ ในส่วน d คือข้อมูลที่ใส่มีค่าระหว่าง 1 ถึง 7 ตามที่ได้กล่าว



รูปที่ 17 การเข้ารหัสชั้นอินพุต

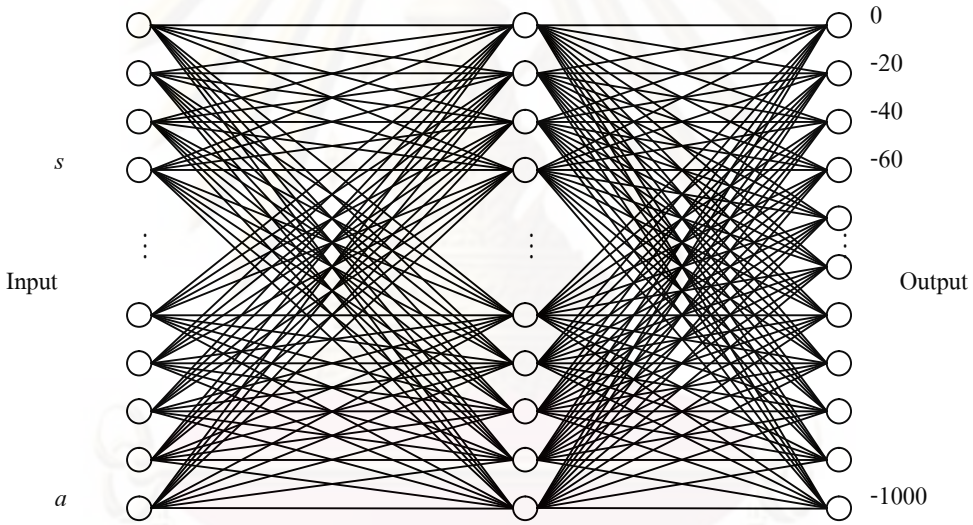
2. ชั้นฮิดเดนหรือชั้นซ่อน เนื่องจากการทดลองนี้ใช้ข่ายงานประสาทแบบแพร่กระจายย้อนกลับที่จะมีชั้นซ่อนอยู่ด้วย ดังนั้นในการทดลองนี้จะทำการกำหนดชั้นซ่อนนี้ให้มีจำนวน 20 บัพ ซึ่งน่าจะเพียงพอต่อการใช้งานแต่ว่าชั้นซ่อนที่ได้กำหนดไว้ที่ 20 บัพไม่ได้หมายถึงชั้นซ่อนที่เหมาะสุดสำหรับการใช้งาน

3. ชั้นเอาต์พุต ในชั้นนี้จะเป็นชั้นที่แสดงผลของค่าคิวที่เมื่อใส่สถานะและการกระทำเข้าไปจะทำการแสดงค่าคิวของสถานะและการกระทำที่เลือก จากการทดลองที่ผ่านมาค่าค่าคิวมีค่าระหว่าง 0 จนถึง -1000 ดังนั้นในการทดลองนี้จะทำการเข้ารหัสและกำหนดค่าคิวให้แต่ละบัพ เพื่อใช้แทนค่าคิวในค่าที่ต่างๆ กัน วิธีการเข้ารหัสคือ จะกำหนดให้เอาต์พุต์มีทั้งหมด 50 บัพแต่ละบัพใช้แทนค่าคิวที่ 0, -20, -40,...,-1000 และในส่วนของข้อมูลที่นำมาสอนก็จะทำการเข้ารหัสในทำนองเดียวกันคือ ถ้าค่าคิวอยู่ในช่วง 0 ถึงน้อยกว่า -20 ให้บัพใช้แทนค่าคิวที่ 0 มีค่าเป็น 0.9 ส่วนบัพที่เหลือกำหนดให้เป็น 0.1 ทั้งหมด เหตุที่ให้ค่าบัพในชั้นเอาต์พุตเป็น 0.1 และ 0.9 เนื่องจากข่ายงานประสาทในการทดลองได้ใช้ฟังก์ชันซิกมอยด์ (sigmoid function) ซึ่งยากที่จะเรียนรู้ค่าเอาต์พุตที่เป็น 0 กับ 1

รูปที่ 18 การเข้ารหัสชั้นเอาต์พุต

การสอนข่ายงานประสาทแบบแพร่กระจายย้อนกลับในการทดลองนี้ใช้ การเลื่อนลงตามความชัน (gradient descent) ในการปรับปรุงค่าและอัตราการเรียนรู้กำหนดไว้ที่ 0.01 จำนวนรอบการเรียนรู้ทั้งหมด 10,000 รอบ ในขั้นตอนของการทดลองจะทำการแบ่งข้อมูลออกเป็น 8 ส่วนเท่าๆ กันและนำมาทำการสอนจำนวน 7 ส่วน เหลืออีกส่วนไว้ทำการทดสอบกับวิธีการขยายและจำกัดเขต ทำซ้ำจนครบทั้ง 8 ส่วนของข้อมูล

ในส่วนของการสอนข่ายงานประสาทเนื่องจากเกิดข้อจำกัดทางด้านหน่วยความจำจึงไม่สามารถที่จะทำการทีเดียวทั้งหมดของชุดข้อมูลได้ จึงจำต้องแก้ปัญหาโดยการใส่ข้อมูลเข้าไปทีละส่วนและทำการสอนจำนวนหนึ่งรอบการสอนแล้วจึงใส่ข้อมูลในส่วนถัดๆ ไปวนซ้ำจนครบ 10,000 รอบแทน



รูปที่ 19 การใช้ข่ายงานประสาทแบบแพร่กระจายย้อนกลับในการประมาณค่าคิว

หลังจากเสร็จสิ้นการฝึกสอนฟังก์ชันการประมาณดังกล่าวแล้ว จะนำมาใช้ในการประมาณค่าคิวเมื่อใส่ความผิดปกติที่ไม่เคยฝึกสอนเข้าไปเพื่อทดสอบประสิทธิภาพการเรียนรู้ ในส่วนของฟังก์ชันในการจำกัดเขตของวิธีขยายและจำกัดเขตนั้น ในการทดลองใช้จำนวนเที่ยวบินที่ได้ทำการปรับปรุงคูณกับระยะเวลาที่ใส่ความผิดปกติเข้าไปในที่นี้คือ 60 นาที

ตารางที่ 6 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับเมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 1

ความผิดปกติ	*TOTAL_DELAY_TIME* (นาที)	
	วิธีการเรียนรู้คิว	วิธีขยายและจำกัดเขต
[aircraft_1 8.00 60]	715	695
[aircraft_2 10.00 60]	695	695
[aircraft_3 6.00 60]	1515	1515
[aircraft_3 10.00 60]	440	430
[aircraft_5 6.00 60]	50	0
[aircraft_6 8.00 60]	0	0
[aircraft_7 12.00 60]	280	225

ตารางที่ 7 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับเมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 2

ความผิดปกติ	*TOTAL_DELAY_TIME* (นาที)	
	วิธีการเรียนรู้คิว	วิธีขยายและจำกัดเขต
[aircraft_1 10.00 60]	390	385
[aircraft_2 8.00 60]	1585	1585
[aircraft_3 8.00 60]	1515	1455
[aircraft_4 10.00 60]	715	605
[aircraft_5 8.00 60]	740	730
[aircraft_6 14.00 60]	500	415
[aircraft_7 8.00 60]	0	0

ตารางที่ 8 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับเมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 3

ความผิดปกติ	*TOTAL_DELAY_TIME* (นาที)	
	วิธีการเรียนรู้คิว	วิธีขยายและจำกัดเขต
[aircraft_1 12.00 60]	365	320
[aircraft_2 20.00 60]	10	0
[aircraft_3 14.00 60]	260	235
[aircraft_4 8.00 60]	815	795
[aircraft_5 10.00 60]	740	705
[aircraft_6 8.00 60]	0	0
[aircraft_7 10.00 60]	0	0

ตารางที่ 9 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับเมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 4

ความผิดปกติ	*TOTAL_DELAY_TIME* (นาที)	
	วิธีการเรียนรู้คิว	วิธีขยายและจำกัดเขต
[aircraft_1 18.00 60]	90	85
[aircraft_2 12.00 60]	220	220
[aircraft_3 16.00 60]	395	405
[aircraft_4 12.00 60]	670	615
[aircraft_5 12.00 60]	500	430
[aircraft_6 6.00 60]	0	0
[aircraft_7 6.00 60]	0	0

ตารางที่ 10 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับเมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 5

ความผิดปกติ	*TOTAL_DELAY_TIME* (นาที)	
	วิธีการเรียนรู้คิว	วิธีขยายและจำกัดเขต
[aircraft_1 6.00 60]	695	695
[aircraft_2 6.00 60]	1620	1575
[aircraft_3 12.00 60]	355	355
[aircraft_4 14.00 60]	225	225
[aircraft_5 20.00 60]	110	85
[aircraft_6 16.00 60]	110	75
[aircraft_7 14.00 60]	730	650

ตารางที่ 11 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับเมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 6

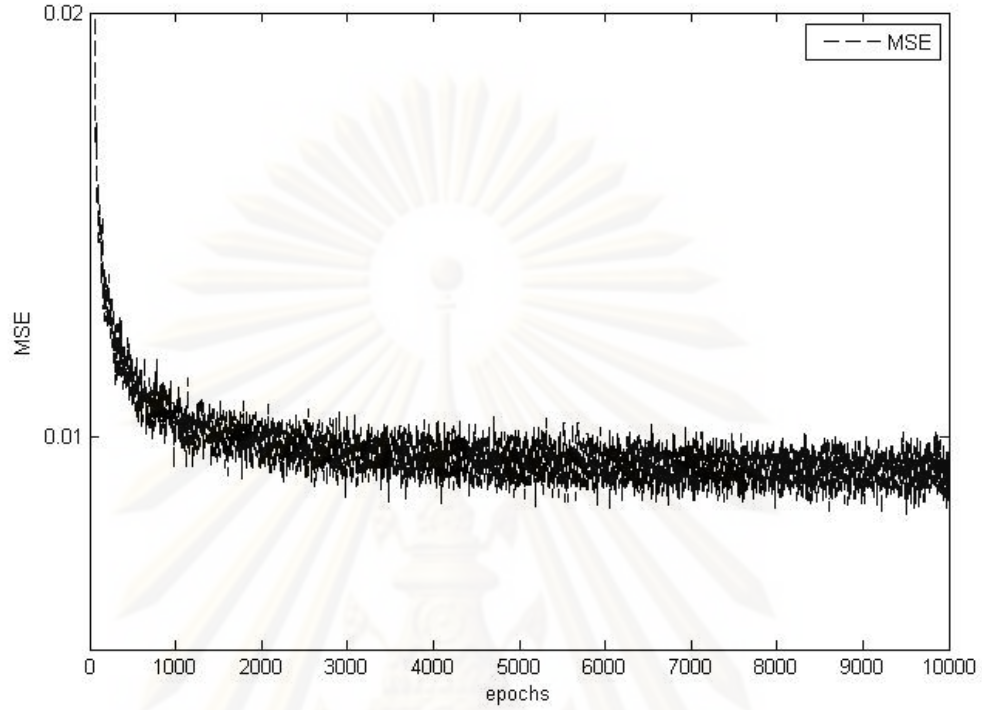
ความผิดปกติ	*TOTAL_DELAY_TIME* (นาที)	
	วิธีการเรียนรู้คิว	วิธีขยายและจำกัดเขต
[aircraft_1 20.00 60]	95	95
[aircraft_2 16.00 60]	35	90
[aircraft_3 18.00 60]	90	85
[aircraft_4 20.00 60]	5	0
[aircraft_5 14.00 60]	560	545
[aircraft_6 18.00 60]	0	0
[aircraft_7 16.00 60]	200	175

ตารางที่ 12 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับเมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 7

ความผิดปกติ	*TOTAL_DELAY_TIME* (นาที)	
	วิธีการเรียนรู้คิว	วิธีขยายและจำกัดเขต
[aircraft_1 14.00 60]	500	495
[aircraft_2 18.00 60]	100	185
[aircraft_3 20.00 60]	45	45
[aircraft_4 16.00 60]	120	65
[aircraft_5 16.00 60]	110	0
[aircraft_6 20.00 60]	0	0
[aircraft_7 18.00 60]	85	85

ตารางที่ 13 ผลการทดลองการประมาณค่าคิวด้วยข่ายงานประสาทแบบการแพร่กระจายย้อนกลับเมื่อพบความผิดปกติที่ไม่เคยพบมาก่อนเปรียบเทียบกับวิธีขยายและจำกัดเขต ด้วยชุดข้อมูลส่วนที่ 8

ความผิดปกติ	*TOTAL_DELAY_TIME* (นาที)	
	วิธีการเรียนรู้คิว	วิธีขยายและจำกัดเขต
[aircraft_1 16.00 60]	400	350
[aircraft_2 14.00 60]	500	500
[aircraft_4 18.00 60]	180	180
[aircraft_5 18.00 60]	125	125
[aircraft_6 10.00 60]	600	580
[aircraft_6 12.00 60]	460	460
[aircraft_7 20.00 60]	110	85



รูปที่ 20 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่ได้จากการทดลอง

ตารางที่ 14 เปรียบเทียบผลที่ได้จากวิธีการเรียนรู้คิวและวิธีขยายและจำกัดเขต

ผลการเปรียบเทียบการทดลอง	จำนวน (ตัวอย่าง)	คิดเป็นร้อยละ
1. การทดลองที่การเรียนรู้คิวให้ผลดีกว่า	3	5.35
2. การทดลองที่ให้ผลเท่ากัน	22	39.29
3. การทดลองที่การเรียนรู้คิวให้ผลแย่กว่าแบ่งเป็น		
-การทดลองที่การเรียนรู้คิวให้ผลแย่กว่าไม่เกิน 1 ส่วนเบี่ยงเบนมาตรฐาน	22	39.29
-การทดลองที่การเรียนรู้คิวให้ผลแย่กว่าเกิน 1 ส่วนเบี่ยงเบนมาตรฐาน	9	16.07
รวมการเปรียบเทียบการทดลองทั้งหมด	56	100

4.6 การทดลองหาเวลาที่ใช้ในการปรับตารางเวลาเที่ยวบิน

ในการทดลองนี้ทำการทดลองเกี่ยวกับการหาเวลาที่ใช้ไปในการปรับตารางเวลาเที่ยวบิน เมื่อได้นำความรู้มาผ่านฟังก์ชันการประมาณ ข่ายงานประสาทแบบแพร่กระจายย้อนกลับ เพื่อหาเวลาที่ใช้ในแต่ละครั้งที่ทำการปรับตารางเวลาเที่ยวบินใหม่

จากที่ได้กล่าวว่าเมื่อตัวเรียนรู้คิวได้เรียนรู้การปรับตารางเวลาเที่ยวบินใหม่แล้ว ก็จะนำความรู้ที่ได้มาใส่ในฟังก์ชันการประมาณ โดยในงานวิจัยนี้จะใช้ฟังก์ชันการประมาณที่เรียกว่าข่ายงานประสาทแบบแพร่กระจายย้อนกลับ ซึ่งเป็นฟังก์ชันการประมาณที่ให้ผลเป็นที่น่าพอใจและถูกกล่าวถึงในหลายงานวิจัยที่ผ่านมา จากการทดลองที่ 4.5 ทำให้ทราบว่าผลของการประมาณจากข่ายงานประสาทที่ได้เรียนรู้นั้นให้ผลที่ใกล้เคียงกับวิธีขยายและจำกัดเขต แต่ว่าข้อดีของการเรียนรู้คิวเมื่อทำการเรียนรู้แล้วนำมาประมาณด้วยฟังก์ชันการประมาณแล้วดำเนินการค้นหาตามนิยามที่ (9) คือ ความเร็วที่ได้จากการประมวลผล ซึ่งในการทดลองนี้จะทำการวัดเวลาในการค้นหาคำตอบในตัวอย่างที่ได้ทำการเลือกมา

$$\pi^*(s) = \arg\max_a \hat{Q}(s,a) \qquad \text{------ (9)}$$

ตารางที่ 15 เวลาที่ใช้ในการปรับตารางเวลาเที่ยวบินใหม่ด้วยการใช้ข่ายงานประสาทเป็นฟังก์ชันการประมาณ

ความผิดปกติ	เวลา (วินาที)
[aircraft_1 16.00 60]	10.350529
[aircraft_2 14.00 60]	10.217925
[aircraft_4 18.00 60]	10.211321
[aircraft_5 18.00 60]	10.249977
[aircraft_6 10.00 60]	10.212999
[aircraft_6 12.00 60]	10.260989
[aircraft_7 20.00 60]	10.230130
[aircraft_3 20.00 60]	10.215834
[aircraft_4 16.00 60]	10.209810
[aircraft_5 16.00 60]	10.205713

4.7 การวิเคราะห์ผล

จากการทดลองที่ 4.1 ในงานวิจัยนี้ที่ได้นำเสนอการแสดงสถานะจากตารางเวลาเที่ยวบินที่กำหนดให้ ซึ่งผลจากการทดลองก็ตรงตามที่ได้คาดหวังไว้คือ สถานะที่ได้นิยามไว้นั้นสามารถที่จะแสดงแทนสิ่งที่เกิดขึ้นในตารางเวลาเที่ยวบินได้ และยังมีความยืดหยุ่นในการปรับตารางเวลาเที่ยวบินที่จะสามารถแสดงสถานะของแต่ละช่วงในการปรับปรุงตารางเวลาเที่ยวบินได้อย่างอิสระ แต่ข้อดีที่ได้มาจากการที่ทำให้สถานะนั้นมีความอิสระสูงนั่นคือ ภาระที่จำต้องให้ตัวเรียนรู้คิวต้องเรียนรู้สถานะต่างๆ ดังที่ได้กล่าวซึ่งก็จะเห็นได้จากการทดลองที่ 4.2 ที่ได้ทำการทดลองการปรับตารางเวลาเที่ยวบิน แต่ได้มีข้อจำกัดไว้แล้วในการทดลอง เพื่อไม่ให้สร้างสถานะมามากจนเกินไป และผลที่ได้จากการทดลองที่สองนี้ก็คือ การกำหนดและความหมายของการเรียนรู้คิวที่ได้มุ่งเน้นในส่วนของการทำซ้ำปรับปรุงค่า ซึ่งก็เพื่อให้มั่นใจได้ว่าค่ารางวัลที่ได้นิยามและได้ให้นั้นมีความสามารถเพียงพอที่จะใช้ในการให้รางวัลกับตัวเรียนรู้ดังเห็นได้จากการทดลองที่รางวัลที่ได้ให้ไปนั้นสามารถตอบสนองกับตัวเรียนรู้ได้อย่างดี แต่ทว่าการตั้งค่าต่างๆ ให้สอดคล้องและให้ผลดีกับตัวเรียนรู้ที่สุดนั้นสมควรที่จะตั้งค่าไว้ที่เท่าใด จึงเป็นที่มาของการทดลองถัดไป

การทดลองที่ 4.3 ที่เกี่ยวกับการลู่เข้าของการเรียนรู้ โดยผลจากการเรียนรู้ทั้งสองครั้งของการทดลองที่สามนี้ได้ชี้อย่างชัดเจนว่าภายใต้เงื่อนไขและข้อจำกัดต่างๆ ในงานวิจัยนี้ค่า *learning rate* ที่ 0.7 และ *discount rate* ที่ 0.3 สามารถให้การลู่เข้าได้ดีที่สุดต่อการประยุกต์ใช้การเรียนรู้คิวในการแก้ปัญหาการปรับตารางเวลาเที่ยวบินเมื่อเกิดเหตุผิดปกติ เพราะจากการทดลองในสองครั้งแรก (รูปที่ 15.1 และ 15.2) ค่า *learning rate* ที่ 0.7 และ *discount rate* ที่ 0.3 สามารถให้ผลดีที่สุดและให้ผลที่แทบจะไม่แตกต่างกันในการทดลองทั้งสองครั้ง ดังนั้นในการทดลองครั้งที่สามที่ให้ผลแบบเดียวกันกับสองครั้งแรกจึงได้ขยายเขตของกราฟเพื่อให้ทราบว่าเหตุใดจึงเป็นเช่นนั้น จากการทดลองครั้งที่สามในรูปที่ 15.3 ทำให้ทราบว่าความจริงแล้วการตั้งค่า *learning rate* ที่ 0.3 และ *discount rate* ที่ 0.7 น่าจะสามารถให้ผลลัพธ์ของการลู่เข้าที่ดีกว่าการตั้งค่า *learning rate* ที่ 0.7 และ *discount rate* ที่ 0.3 เพราะเริ่มต้นให้ค่ารากกำลังสองเฉลี่ยที่ต่ำกว่าทุกๆ การตั้งค่า แต่เมื่อเทียบกับจำนวนรอบที่ต้องใช้ในการเรียนรู้เพื่อให้ค่ารากกำลังสองเฉลี่ยลดลงในการเรียนรู้แล้วกลับเป็นการตั้งค่าที่ให้ผลที่แย่ที่สุด เมื่อเปรียบเทียบกับการตั้งค่า *learning rate* ที่ 0.7 และ *discount rate* ที่ 0.3 ที่เริ่มต้นให้ค่ารากกำลังสองเฉลี่ยสูงที่สุดแต่กลับให้ค่าการลดลงของรากกำลังสองเฉลี่ยดีที่สุด

การทดลองที่ 4.4 ได้นำการตั้งค่า และข้อดีต่างๆ ที่ได้มาจากการทดลองก่อนหน้ามารวมกันเป็น การประยุกต์ในงานการเรียนรู้คิวในการแก้ปัญหาการปรับตารางเวลาเที่ยวบินเมื่อเกิดความผิดปกติขึ้น ผลที่ได้เมื่อทำการเปรียบเทียบการทดลองจากทั้ง 10 ครั้งแล้วนำผลการทดลองครั้งที่ดีที่สุด, ค่าเฉลี่ยจากการทดลองทั้ง 10 ครั้ง, และผลการทดลองครั้งที่แย่ที่สุดนำมาเปรียบเทียบกัน

พบว่า ผลจากการทดลองครั้งที่ดีที่สุดเมื่อเทียบกับผลจากการทดลองครั้งที่แย่ที่สุดปรากฏว่า ผลจากการทดลองครั้งที่ดีที่สุดให้ผลที่ดีกว่าผลจากการทดลองครั้งที่แย่ที่สุด 11.4% และเมื่อนำผลจากการทดลองครั้งที่ดีที่สุดเมื่อเทียบกับผลจากค่าเฉลี่ยจากการทดลองทั้ง 10 ครั้ง พบว่าผลจากการทดลองครั้งที่ดีที่สุดให้ผลที่ดีกว่าผลจากค่าเฉลี่ยจากการทดลองทั้ง 10 ครั้ง 5.7% ซึ่งผลที่ได้จากการทดลองนี้ภายหลังจากการนำไปเปรียบเทียบโดยจะทำการเลือกนำผลครั้งที่ดีที่สุดไปใช้ในการทดลองต่อไป การที่ต้องทำการเลือกผลครั้งที่ดีที่สุดไปใช้ในการทดลองถัดไปนั้น ก็เพื่อหวังว่าจะสามารถเป็นตัวแทนของความรู้ที่ได้เรียนมาจากตัวอย่างสอน ซึ่งผลที่ออกมาก็มีหลากหลายแต่เมื่อเปรียบเทียบจากค่าเฉลี่ยของแต่ละการทดสอบที่เทียบจากความล่าช้ารวมแล้ว ทำให้สามารถแยกแยะการทดลองที่ให้ผลที่ดีที่สุดออกมาได้และนำไปเป็นข้อมูลในการสอนในการทดลองที่ 4.5 ต่อไป

การทดลองที่ 4.5 นี้เป็นการนำข้อมูลชุดที่ทำการสอนแล้วให้ผลดีที่สุดที่ได้จากการทดลองที่สี่ได้ทำการทดลองมา และใช้ฟังก์ชันการประมาณข่ายงานประสาทแบบแพร่กระจายย้อนกลับทำความรู้ให้เป็นรูปแบบทั่วไปและเปรียบเทียบผลกับการใช้วิธีการขยายและจำกัดเขต ซึ่งผลของคำตอบที่ได้จากการทดลองเมื่อนำมาวัดผล ปรากฏว่าวิธีการขยายและจำกัดเขตสามารถให้ผลของคำตอบได้ดีกว่าวิธีการเรียนรู้คิวคิดเป็น 55.36% โดยแบ่งเป็นผลของคำตอบที่มีค่าไม่เกินหนึ่งส่วนมาตรฐานเท่ากับ 39.29% และผลของคำตอบที่มีค่าเกินหนึ่งส่วนมาตรฐานเท่ากับ 16.07% (ส่วนเบี่ยงเบนมาตรฐานของการเปรียบเทียบการทดลองซึ่งเท่ากับ 33.85 และค่าเฉลี่ยของการเปรียบเทียบการทดลองเท่ากับ 18.21) ส่วนผลของคำตอบจากวิธีการเรียนรู้คิวที่ให้ค่าเท่ากับหรือดีกว่าวิธีการขยายและจำกัดเขตคิดเป็น 44.64% เมื่อพิจารณาจากผลของคำตอบที่มีค่าไม่เกินหนึ่งส่วนมาตรฐานรวมกับผลของคำตอบจากการเรียนรู้คิวที่เท่ากับหรือดีกว่าแล้ว อาจจะพูดได้ว่าผลของคำตอบที่ได้จากวิธีการเรียนคิวนั้นใกล้เคียงกับผลของคำตอบจากวิธีการขยายและจำกัดเขตโดยคิดเป็น 83.93% ในการทดลองที่ 4.6 นั้นได้แสดงให้เห็นแล้วว่าวิธีการเรียนรู้คิวเมื่อใช้ฟังก์ชันการประมาณเข้ามาช่วยทำให้สามารถเพิ่มความเร็วในการปรับตารางเวลาเที่ยวบินได้ ซึ่งการค้นหาจะทำผ่านนโยบายตามนิยามที่ (9) ซึ่งจะทำการเลือกการกระทำที่ให้ค่าคิวมากที่สุดเพียงอย่างเดียว

ในการทดลองที่ได้กล่าวมานี้ได้แสดงให้เห็นถึง การประยุกต์ใช้งานการเรียนรู้คิวเพื่อแก้ปัญหาการปรับตารางเวลาเที่ยวบินเมื่อเกิดความผิดปกติขึ้น ซึ่งผลจากการทดลองต่างๆ ที่ได้ทำมานั้นก็ได้แสดงให้เห็นแล้วว่า การที่จะทำการประยุกต์ใช้งานการเรียนรู้คิวในปัญหาดังกล่าวนั้นสามารถที่จะกระทำได้และสามารถให้ผลลัพธ์ที่ใกล้เคียงกับวิธีที่ใช้อยู่ ถึงแม้ว่าอาจจะไม่ดีกว่าแต่ก็ถือเป็นทางเลือกที่ดีทางหนึ่งที่เหมาะสมกับปัญหาๆ นี้ เพราะด้วยเหตุผลทางด้านของตัวปัญหาเองที่ต้องมีการแก้ไขปัญหาที่เกิดขึ้นในเวลาอันสั้น และรูปแบบของโครงสร้างของปัญหาที่มักจะใช้ข้อมูลเดิมมาสร้างเป็นตารางเวลาเที่ยวบินใหม่ ทำให้การแก้ปัญหาด้วยการเรียนรู้คิวที่ได้กล่าวใน

งานวิจัยนี้นั้นสามารถหรือมีความเป็นไปได้ที่จะแก้ปัญหานี้ได้อย่างมีประสิทธิภาพในอนาคต ซึ่งเหตุผลและแนวทางในการแก้ไขปรับปรุงจะขอกล่าวในบทถัดไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

เป็นความจริงที่ว่าการปรับตารางเวลาเที่ยวบินเมื่อเกิดความผิดปกติขึ้นด้วยวิธีการเรียนรู้คิวสามารถให้ผลของคำตอบในเวลาที่เร็วกว่าวิธีการขยายและจำกัดเขต เนื่องจากการเรียนรู้คิวนั้นเมื่อได้ผ่านการสอนและทำการเรียนรู้แล้วจะนำความรู้ที่เรียนรู้ได้มาเก็บไว้ในรูปแบบของฟังก์ชันการประมาณ และการค้นหาคำตอบก็ทำการหาผ่านทางนโยบายที่ได้เรียนรู้มา ทำให้การหาคำตอบสามารถทำได้เร็วกว่า แต่วิธีการขยายและจำกัดเขตนั้นแม้จะให้ผลที่ดีกว่าแต่ก็ต้องใช้เวลาในการค้นหาคำตอบเป็นเวลาพอสมควร เนื่องจากในการแก้ปัญหาการปรับตารางเวลาเที่ยวบินนั้น ทางสายการบินไม่สามารถที่จะรอการประมวลผลที่ยาวนานเพื่อให้ได้ผลเฉลยที่ดีที่สุดได้ แต่จำต้องแก้ไขปัญหาเฉพาะหน้าให้ได้ทันท่วงที ที่ในงานวิจัยนี้ได้นำเสนอแนวทางการแก้ปัญหาการปรับตารางเวลาเที่ยวบินเมื่อเกิดความผิดปกติด้วยวิธีการเรียนรู้คิวนั้น ก็เพื่อให้เป็นทางเลือกในการแก้ปัญหาดังกล่าวถึงแม้ผลของการแก้ปัญหายังอาจจะสู้วิธีการขยายและจำกัดเขตไม่ได้ก็ตาม แต่ก็ได้แลกมาด้วยความเร็วในการแก้ปัญหาที่น่าจะมีความสำคัญมากกว่าในการแก้ปัญหานี้ ซึ่งข้อควรปรับปรุงหรือแนวทางที่จะทำให้ผลของการเรียนรู้มีประสิทธิภาพดีขึ้นนั้นขอขอกล่าวในส่วนของข้อเสนอแนะในส่วนถัดไป

ในงานวิจัยนี้ได้นำเสนอการประยุกต์ใช้งานการเรียนรู้คิวในงานปฏิบัติการทางการบินเมื่อมีเหตุการณ์ไม่ปกติเกิดขึ้น ได้ทำการปรับปรุงตารางเวลาเที่ยวบินเพื่อให้ง่ายและเหมาะกับการเรียนรู้พร้อมทั้งยังแสดงการสร้างสถานะ, การนิยามการกระทำที่ตัวเรียนรู้สามารถกระทำได้ในแต่ละสถานะที่ตัวเรียนรู้สำรวจ และประยุกต์ค่ารางวัลเพื่อให้เกิดความเหมาะสมกับงานปฏิบัติการทางการบิน ในงานวิจัยนี้ได้นำตารางเวลาเที่ยวบินภายในประเทศมาใช้เพื่อเป็นตัวอย่างในการสอนและทดสอบการเรียนรู้ โดยมีการสร้างความผิดปกติเข้าไปในแต่ละตัวอย่างสอนและนำไปสอนตัวเรียนรู้ ซึ่งผลที่ได้ถือว่ามีประสิทธิภาพอย่างมีนัยยะ และสามารถให้ผลลัพธ์ได้ในแบบทันกาล ซึ่งเป็นผลสืบเนื่องมาจากการนำฟังก์ชันการประมาณมาใช้ร่วมกับการเรียนรู้คิว ฟังก์ชันการประมาณที่นำมาใช้ในที่นี้คือ ข่ายงานประสาทแบบการแพร่กระจายย้อนกลับ การใช้งานฟังก์ชันการประมาณทำได้โดยการนำผลที่ได้จากการเรียนรู้คิวมาทำการสอนฟังก์ชันการประมาณอีกรอบ

แต่งานวิจัยนี้ยังมีข้อจำกัดในหลายเรื่องสืบเนื่องมาจากความสามารถของการเรียนรู้คิวนั่นเองคือ การที่จะให้ผลลัพธ์เป็นที่น่าพอใจนั้นจำเป็นที่จะต้องทำการเรียนรู้คิวเสียก่อนถึงจะสามารถนำความรู้ที่ได้จากการเรียนรู้ไปใช้งานในขั้นตอนต่อไปได้ และในขั้นตอนของการเรียนรู้

นั้นก็จำเป็นที่จะต้องใช้เวลานานพอสมควรถึงจะได้ความรู้ที่สามารถนำไปใช้ได้ ดังนั้นการนำการเรียนรู้คิวไปใช้นั้นจึงจำเป็นที่จะต้องเข้าใจข้อจำกัดนี้ด้วย ดังเช่นในงานวิจัยนี้สามารถนำการเรียนรู้คิวมาประยุกต์ใช้งานได้ เพราะว่าทราบตารางเวลาเที่ยวบินล่วงหน้าก่อนที่จะทำการปฏิบัติการทางการบินจริงเป็นเวลานานทำให้มีเวลาในการเรียนรู้ แต่ถ้าตารางเวลาเที่ยวบินไม่เหมือนเดิมหรือถูกปรับเปลี่ยนอย่างปัจจุบันทันด่วน การเรียนรู้คิวก็จะไม่สามารถให้ผลลัพธ์ได้ หรือถ้ามีเหตุผิดปกติที่เกินขึ้นภายหลังจากที่ได้ทำการปฏิบัติการทางการบินไปแล้ว ก็ไม่สามารถที่จะรองรับได้ เพราะในงานนี้สนใจเฉพาะความผิดปกติที่เกิดขึ้น และยังไม่ได้ทำการปฏิบัติการทางการบินจริง แต่ข้อจำกัดต่างๆ ที่กล่าวมาในส่วนของข้อจำกัดนี้อาจสามารถที่จะแก้ไขหรือพัฒนาต่อได้ซึ่งจะขอกล่าวในข้อเสนอแนะต่อไป

5.2 ข้อเสนอแนะ

สิ่งที่ควรคำนึงถึงในการแก้ปัญหาการประยุกต์ใช้การเรียนรู้คิวในการแก้ปัญหาการปรับตารางเวลาเที่ยวบินเมื่อเกิดความผิดปกติขึ้นนั้น มีสิ่งที่ควรพิจารณาอยู่ 3 ส่วนหลักคือ

1. การจำกัดหรือการนิยามปัญหา เพื่อที่จะให้การเรียนรู้คิวมีความสามารถในการบอกและเรียนรู้สถานะแต่ละสถานะได้ เพราะถ้าไม่มีการนิยามปัญหาให้ดี สถานะก็ไม่อาจจะนิยามได้และจะส่งผลต่อการเรียนรู้ทำให้ไม่สามารถที่จะเรียนรู้สถานะทั้งหมดได้ แต่ถ้าสามารถที่จะนิยามสถานะได้แล้วปัญหาต่อไปที่จะตามมาก็คือ ขนาดของปัญหา แต่เนื่องจากมีหลายงานวิจัย ดังเช่น งานวิจัยของ W. Zhang และ T. G. Dietterich [2] หรืองานวิจัยของ R. H. Crites และ A. G. Barto [13] ที่ได้มีการนำฟังก์ชันการประมาณมาใช้และก็ถือเป็นคำตอบที่ดีในเรื่องขนาดของปัญหาในการเรียนรู้แบบเสริมกำลัง

2. การกำหนดค่ารางวัล ในส่วนนี้ไม่มีฟังก์ชันการให้รางวัยที่แน่นอนสำหรับปัญหาแต่ละปัญหาที่เกิดขึ้น ขึ้นกับว่าในงานการเรียนรู้หรือการประยุกต์ใช้งานในปัญหานั้นๆ ให้ความสนใจที่จะพิจารณาในส่วนใดเป็นพิเศษหรือให้ความสนใจในส่วนใดมากกว่ากัน แต่ประเด็นสำคัญก็คือ ค่ารางวัลที่ให้จะต้องสะท้อนปัญหาในส่วนที่ให้ความสนใจในงานนั้นๆ ดังเช่นในงานวิจัยนี้ที่ต้องการค่ารางวัลมากเมื่อการเรียนรู้ปรับตารางเวลาเที่ยวบินมีความล่าช้าน้อย จึงได้ให้ค่ารางวัลเป็นฟังก์ชันของเวลาล่าช้า หรืองานของ R. H. Crites และ A. G. Barto [13] ที่ให้ความสนใจในเรื่องของเวลาในการรอใช้ลิฟต์จึงให้รางวัลตามฟังก์ชันของเวลาในการรอลิฟต์เป็นต้น

3. การลู่เข้า การลู่เข้าเป็นส่วนที่เกี่ยวข้องกับการประยุกต์การเรียนรู้คิวที่จะขอกล่าวถึงในส่วนนี้ เพราะเนื่องจากในงานการเรียนรู้ไม่ว่าจะเป็นการเรียนรู้แบบใดก็ตามจำเป็นที่จะต้องทราบว่าความรู้เหล่านั้นที่ได้เรียนรู้ไปได้ลู่เข้าไปสู่จุดใดจุดหนึ่ง ทั้งนี้เพื่อเวลาที่จะนำความรู้ที่ได้เรียนรู้มาใช้จะได้เข้าใจและทราบถึงที่มาของคำตอบ และทำให้คำตอบเหล่านั้นดูเป็นเหตุเป็น

ผลมากขึ้น การลู่เข้านั้นมีอยู่หลายวิธีด้วยกันแต่ที่นำมาใช้ในการประยุกต์การเรียนรู้คิวในการแก้ปัญหาการปรับตารางเวลาเที่ยวบินเมื่อเกิดความผิดปกติ เป็นการลู่เข้าโดยอาศัยหลังของความน่าจะเป็น β ซึ่งข้อดีของวิธีนี้คือ มีความสามารถในการควบคุมจำนวนรอบของการเรียนรู้ได้ ทั้งนี้เนื่องมาจากการกำหนดค่า β และอัตราในการลดทอนของค่า β ในแต่ละขั้นของการเรียนรู้ ข้อดีอีกประการหนึ่งคือ มีความสามารถในการค้นหาสถานะที่ยังไม่เคยได้สำรวจมาก่อน เพราะในการลู่เข้าด้วยวิธีนี้มักจะมีการกำหนดเงื่อนไขไว้ให้ด้วย ดังนั้นอาจจะสามารถทำการเพิ่มเงื่อนไขให้ละเอียดเพิ่มขึ้นหรือแบ่งช่วงของการสุ่มค่า β ให้เพิ่มขึ้น

แนวทางการทำวิจัยต่อในเรื่องนี้สามารถแบ่งออกได้เป็น 3 ทางด้วยกันคือ

1. ในแนวทางแรกเป็นการทำการวิจัยในส่วนของการปรับตารางเวลาเที่ยวบินโดยคำนึงถึงเงื่อนไขบังคับอื่นๆ หรือขอบเขตต่างๆ ที่ยังไม่ได้พิจารณาในงานวิจัยนี้ เช่นการเพิ่มความสามารถในการยกเลิกเที่ยวบินในบางเที่ยวเพื่อให้ได้เวลาล่าช้าน้อยที่สุด การเพิ่มข้อจำกัดของสนามบินและ/หรือนักบินเป็นต้น

2. ในแนวทางที่สองควรพิจารณาในส่วนของฟังก์ชันการประมาณ ดังเช่นในงานวิจัยของ Gersmann และ Hammer ที่ได้มีการนำ SVM เข้ามาทำหน้าที่เป็นฟังก์ชันการประมาณแทนส่วนของข่ายงานประสาท [14]

3. แนวทางที่สามคือในส่วนของตัวเรียนรู้เอง เช่นงานวิจัยของ C. Claus และ C. Boutilier ที่มีการนำตัวเรียนรู้มากกว่าหนึ่งตัวเพื่อเข้ามาช่วยในการหาคำตอบ [16] ซึ่งในปัจจุบันมีหลากหลายงานวิจัยที่ได้ทำการค้นคว้าและพัฒนาการเรียนรู้แบบนี้ หรือจะเป็นงานวิจัยของ J. A. Boyan และ A. W. Moore ที่มุ่งพัฒนาตัวเรียนรู้ให้มีประสิทธิภาพเพิ่มขึ้น [17, 18]

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รายการอ้างอิง

[1] Gerald Tesauro. Temporal difference learning and TD-Gammon. Communications of the ACM 58 (1995): 58-68.

[2] Wei Zhang and Thomas G. Dietterich. A reinforcement learning approach to job-shop scheduling. In Proceeding of the 14th International Joint Conference on Artificial Intelligence: 1995.

[3] Michael Clarke. Solving the Problem of Irregular Airline Operations. INFORMS New Orleans: 1995.

[4] Michael Clarke. Development of Heuristic Procedures for Flight Rescheduling in the Aftermath of Irregular Airline Operations. INFORMS Atlanta: 1996.

[5] Ahmad Jarrah, Yu, G., Krishnamurthy, N., and Rakshit, A. Decision support framework for airline flight cancellations and delays. Transportation Science 27 (1993).

[6] Dennis F. X. Mathaisel. Decision support for airline system operations control and irregular operations. Computers and Operations Research 23(1996):1083-1096.

[7] Niklas Kohl, Allan Larsen, Jesper Larsen, Alex Ross, and Sergey Tiourine. Airline disruption management—Perspectives, experiences and outlook. Journal of Air Transport Management 13 (2007): 149-162.

[8] Jens Clausen, Allan Larsen, Jesper Larsen, and Natalia J. Rezanova. Disruption Management in the airline industry—Concepts,models and methods. Computer & Operations Research 37 (2010): 809-821.

[9] Jay M. Rosenberger, Andrew J. Schaefer, David Goldsman, Ellis L. Johnson, Anton J. Kleywegt, and George L. Nemhauser. A stochastic model of airline operations. Transportation Science 36 (2002): 357-377.

[10] Tom M. Mitchell. Reinforcement learning. In Machine learning, pp. 367-387. McGraw-Hill: 1997.

[11] Stuart Russell and Peter Norvig. Reinforcement learning. In Artificial Intelligence a Modern Approach second edition, pp. 763-788. Prentice Hall: 2003.

[12] Sutton, R. Learning to predict by the methods of temporal differences. Machine Learning 3 (1988): 9-44.

[13] Robert H. Crites and Andrew G. Barto. Improving Elevator Performance Using Reinforcement Learning. In D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo (eds.), Advances in Neural Information Processing Systems, pp. 1017-1023. The MIT Press: 1996.

[14] Gersmann, K., and Hammer B. A reinforcement learning algorithm to improve scheduling search heuristics with the SVM. In Proceedings 2004 IEEE International Joint Conference on Neural Networks 2004, pp. 1811-1816. : 2004.

[15] Heriberto Cuayahuitl, Steve Renals, Oliver Lemon, and Hiroshi Shimodaira. Evaluation of a hierarchical reinforcement learning spoken dialogue system. Computer Speech & Language 24 (2010): 395-429.

[16] Caroline Claus and Craig Boutilier. The dynamics of reinforcement learning in cooperative multiagent systems. In Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence, pp. 746 - 752. : 1998.

[17] Justin A. Boyan and Andrew W. Moore. Learning Evaluation Functions for Global Optimization and Boolean Satisfiability. In AAAI/IAAI, pp. 3-10. : 1998.

[18] Justin A. Boyan and Andrew W. Moore. Learning evaluation functions to improve optimization by local search. Journal of Machine Learning Research 1 (2000): 77-112.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก ผลงานการตีพิมพ์

1. Opat Plernprapaporn and Proadpran Punyabukkana, Q-Learning Application for Flight Rescheduling. In Proceedings of First AUN/SEED-Net Regional Conference in Information and Communication Technology. Bangkok, Thailand, March 3-4 2009.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Q-LEARNING APPLICATION FOR FLIGHT RESCHEDULING

Opat Plernprapaporn, Proadpran Punyabukkana
Faculty of Engineering, Chulalongkorn University,
Bangkok THAILAND
opat.p@student.chula.ac.th, proadpran.p@chula.ac.th

ABSTRACT

While most researchers attempt to solve flight operation irregularity using mathematical models or network flow theory, only few researchers have tried applying Q-learning technique, a reinforcement learning method. Q-learning can learn without model of the surroundings, therefore it can give results in real time. This research presents Q-learning application for irregular flight re-scheduling.

Key Words: flight rescheduling, Q-learning, action, reward

1. INTRODUCTION

Reinforcement learning is a machine learning algorithm found to work effectively in many applications. Examples of its application include Tesauro's who introduced the use of reinforcement learning in backgammon game [1], and Zhang et al. who implemented it in job-shop scheduling task [2].

In general, reinforcement learning learns to choose optimal actions to achieve its objective. It formulates problems into states and actions. Reward is received when an action is applied to each state. The goal of this learning system is to find a control policy that maximizes the reward. Therefore, we describe our concept that considers using Q-learning, a more recent form of reinforcement learning that can be used online, to solve flight rescheduling problem when irregularity occurs.

In daily operation, airlines frequently face with various situations which have an effect on its operating plan. These situations are called irregularity that includes several unexpected situations such as terrible weather, aircraft failure, or personnel failure that demands a change in the planned schedule. When flight schedule has an irregularity, airlines needs to adjust its schedule by changing by canceling, reschedule, combine, re-route, or combine and re-route a flight or several flights. In this research, we are interested in flight rescheduling with the decision to change aircraft in a same fleet type or change aircraft in different fleet types. Delay time usually occurs with any adjustment. The objective of this work is to adjust flight schedule with minimum delay time for real time operation.

Most of the works in this area utilized mathematical model or network flow theory to solve a problem. Much less is known as to use reinforcement learning approach to this problem. For instance, Clarke has introduced this problem and affects of irregularity to the operation [3], and later, he showed a mathematic model that encompasses heuristic procedures to solve the problem [4]. Moreover, there are many researches related to this problem such as Jarrah et al. [5] and Mathaisel [6].

One of the reasons may be attributed to the fact that most machine learning techniques are not feasible when real-time solution is needed. However, with Q-Learning, it does not need a model of its environment, hence allows it to be used on-line. We found that we can apply Q-learning algorithm by formulating flight reschedules into state transitions which have feasible action in each state. Furthermore, we can apply Q-learning by mapping flight rescheduling problem into state transition which has an appropriate action in each state.

2. FIGHT SCHEDULE MANAGEMENT

Airlines normally plan flight schedule in advance, planning can be split into two seasons: summer and winter. For experimental purposes, we simulate a schedule with 50 daily flights, assuming that flights depart from its base airport, and 2 aircraft types.

Often, airline policy on rescheduling is to minimize passenger delay time in order to gain maximum reliability from their passengers. Flight rescheduling strategy can be divided into five issues.

1. Delay flights
2. Change aircraft in the same fleet type

Table 1 Example of domestic flight schedule

Flight Number	Departure	Arrival	Departure Time	Arrival Time	Aircraft tail no.	Type
0001	C01	C02	06:00	07:05	AC1	734
0002	C02	C01	07:45	08:50	AC1	734
0003	C01	C03	06:05	07:00	AC2	734
0004	C03	C01	07:40	08:25	AC2	734
0005	C01	C04	06:05	07:35	AC3	AB7
0006	C04	C01	08:20	09:50	AC3	AB7
0007	C01	C05	06:15	07:20	AC4	734
0008	C05	C01	08:00	09:00	AC4	734
0009	C01	C06	08:00	09:10	AC5	AB7
0010	C06	C01	09:55	11:05	AC5	AB7
0011	C01	C07	09:10	10:05	AC6	734
0012	C07	C01	10:45	11:40	AC6	734
…	…	…	…	…	…	…

3. Change aircraft in different fleet types

4. Combine and flight cancellation

5. Combine and reroute flights

Each rescheduling step will consider several constraints for example flight schedule constraint, maintenance and ground time, flight crew available, passenger connection, passenger capacity, aircraft or airport restriction, weather and so on.

From these strategies and constraints, optimization to this problem is hardly to be solved in real time because of the size of problem. Therefore in real-time performance, we have to find a heuristic approach to this problem.

The objective of this research is to introduce reinforcement learning application to solve flight rescheduling problem when irregularity occurs. We cover two strategies, changing aircraft in the same fleet type and changing aircraft in different fleet types, and three constraints, flight schedule constraint, ground time and passenger capacity.

3. REINFORCEMENT LEARNING AND Q-LEARNING

Reinforcement learning learns how to choose an optimal action to archive its goals. Each time the system performs an action, *a,* in some state, *s*, it receives a reward, *r*. Learning task is to learn a control policy $\pi : S \to A$, that maximizes the sum of the rewards. Cumulative reward is the sum of all rewards given policy π, starting from state *s*. An optimal policy is a policy that maximizes this cumulative reward [10].

Reinforcement learning reinforcement learning does not learn optimal policy directly. Rather, it learns an optimal policy from the cumulative reward of the optimal policy. Therefore in the learning task, we learn to choose best action *a* in each state *s* that maximizes the cumulative reward. As above-mentioned reinforcement learning must have a perfect knowledge of state transition by applying action *a* to state *s*.

Q-learning is a more recent form of reinforcement learning technique that learns an action-value function that gives the expected utility of taking a given action in a given state and following a fixed policy thereafter. An advantage of Q-learning is that it is able to compare the expected utility of the available actions without requiring a model of the environment.

The evaluation function Q is defined by the immediate reward received after executing action *a* from state *s* , plus the value of the optimal policy thereafter. In the other word, value of Q is the immediate reward plus the maximum value of Q of next state transition. From this evaluation function Q, It will be able to select optimal action even when there is no perfect knowledge.

Q-learning algorithm is an iteration update. $\hat{Q}$ use as an estimated value or hypothesis of the actual Q function. Training train this estimated $\hat{Q}$ to approximate the actual Q function. For each state *s* and each action *a*, we calculate an update to its expected discounted reward with the following expression [7]:

$$\hat{Q}(s,a) \leftarrow \hat{Q}(s,a) + \alpha\left[r + \gamma \max_{a'} \hat{Q}(s',a') - \hat{Q}(s,a)\right]$$

where α is learning rate $0 \le \alpha \le 1$

γ is discount rate $0 \le \gamma \le 1$

Table 2 pairing flight schedule and symbols

NO. pair	Departure	Arrival	Departure time	Arrival time	symbol	type
1	**C01**	**C02**	0600	0850	aircraft_1	734
2	**C01**	**C03**	0605	0825	aircraft_2	734
3	**C01**	**C04**	0605	0950	aircraft_3	AB7
4	**C01**	**C05**	0615	0900	aircraft_4	734
5	**C01**	**C06**	0800	1105	aircraft_5	AB7
6	**C01**	**C07**	0910	1140	aircraft_2	734
7	**C01**	**C08**	0935	1240	aircraft_1	734
…	…	…	…	…	…	…

4. THE APPLICATION

In this section, we describe variables used and the application of Q-learning to the problem of flight rescheduling when irregularity occurs. We must first understand the definition of state, action and reward.

4.1. State

For state definition, state must describe rescheduling in each step. State is defined from type of flight schedule which can be divided into group of aircraft in each flight.

[*Group of aircraft, irregularity*]

Group of aircraft takes into account 50 daily flights. Because we know flight schedule in advance, we know flight departure and arrival time which is the same for everyday. However, airlines may not use the same aircraft in each day they operate. In the learning process, we generalize aircraft tail number by using symbols to represent aircraft tail number in each flight: aircraft_1, aircraft_2, aircraft_3… aircraft_7 (Table 2).

Because flights always depart from the same base airport, we can pair the departure and return flights and call it pairing flight. These pairing flights will reduce the amount of flights Q-Learning has to learn and will avoid the repetition in the learning step.

Thereby, aircraft group in the table is defined with four variables for each pairing flight:

- Pairing flights number indicates that what pair is it in the table.
- Symbol represent aircraft, it shows what aircraft using in each pairing flight.
- Departure time of each pairing flight.
- Pairing flight length (min).

[*pairing flights number, symbol, departure time, pairing flight length*]

We then generate irregularities into normal flight schedule by inserting irregularities every two hours lasting one hour each. This is for the purpose of generating training and testing data set.

The irregularity occurred in flight schedule can be divided into three problem areas:

- Irregular aircraft: represents the aircraft that has irregular schedule.
- Irregular time: represents time the irregularity starts.
- Irregularity length: represents duration in which the irregularity lasts in minute.

[*irregular aircraft, irregular time, irregularity length*]

Thus, form table 2 we can write this into the following state.

[*1 aircraft_1 6.00 170, 2 aircraft_2 6.05 140, 3 aircraft_3 6.05 225… aircraft_1 10.00 60*]

where *1, 2, 3*... represent numbers of each pairing flights
aircraft_1, aircraft_2, aircraft_3… represent symbol for each aircrafts
6.00, 6.05, 6.05… represent a departure time of each pairing flights
170,140,225… represent time in which each pairing flight consume (min)

And the last three factors represent the irregularity status.
aircraft_1 represent the irregular aircraft
10.00 represent time that the irregularity occurred
60 represents time period that the irregularity occurred (min)

4.2. Action

We define an action as a symbol representing aircraft assigned to that particular flight. To determine the symbol, we consider three constraints.

1. The flight schedule constraint: is when we need to adjust a new flight schedule we have to delay or departure at the time we have already planned.

2. Ground time: is a minimum time that every aircraft must be on ground after the landing for each fleet type. For example, Boeing 737-400 ground time is 40 min whereas Airbus A300-600 ground time is 50 min.

3. Passenger seat capacity: each fleet type or even different aircraft offers different capacity.

4.3. Reward

With the objective to minimize delay, reward function will be defined in terms of delay time.

The function of delay can be computed from the difference of a reschedule flight time t_r and a scheduled flight time t_s. We define DELAY_TIME to be a delay time in each pairing flights and TOTAL_DELAY_TIME to be a total delay time in every pairing flight. The TOTAL_DELAY_TIME represents the total delay time when we complete flight reschedule process.

$$DELAY_TIME(i) = t_r(i) - t_s(i) \quad (1)$$

$$TOTAL_DELAY_TIME = \sum_{i=1}^{n} DELAY_TIME(i) \quad (2)$$

where i represents pairing flight number

n represents total number of pairing flights

To maximize rewards by minimizing TOTAL_DELAY_TIME, we can write the reward function as follow:

$$r = \begin{cases} -TOTAL_DELAY_TIME & ;if\ \ complete\ schedule \\ Penalty & ;for\ \ each\ \ rescheduling\ \ step \end{cases} \quad (3)$$

From equation (3) we can see that reward r is a negative of TOTAL_DELAY_TIME.

4.4. Convergence

In this work, we use an approach appeared in Zhang et al. by using exploration strategy with probability β [2]. The algorithm choose a randomly assigned symbol representative aircrafts, otherwise the algorithm choose assign the symbolic representative aircrafts which maximize Q value with probability β. We assigned β value to 1 and slowly decreases its value.

$$\pi^*(s) = \begin{cases} randomly\ \ assign\ \ aircraft & ;with\ \ probability\ \ \beta \\ \arg\max_a Q(s,a) & ;with\ \ probability\ \ 1-\beta \end{cases} \quad (4)$$

4.5. Approximation function

Practically, flight rescheduling problems are an infinite state but reinforcement learning, especially Q- learning is only convergent in a finite state. Therefore, in previous research, the approximation function was brought to generalize this problem from state that has been seen and estimate Q value for unseen states.

In this work, we employ back propagation Neural Network as an approximation function. First, we train Neural Network by using the state learning from Q to become an input and Q value in each state to become an output. Then can calculate from the Neural Network and estimate Q value for the status that we have never seen before.

5. RESULTS

We assigned α value to 0.7 and γ value to 0.3, while, β value starts from 1 and decreases gradually by 0.00001. In the early of the training iteration, we use greedy method. In this experiment, we train and estimate $\hat{Q}$ three times and select the best result from TOTAL_DELAY_TIME with the purpose to estimate with the back propagation Neural Network.

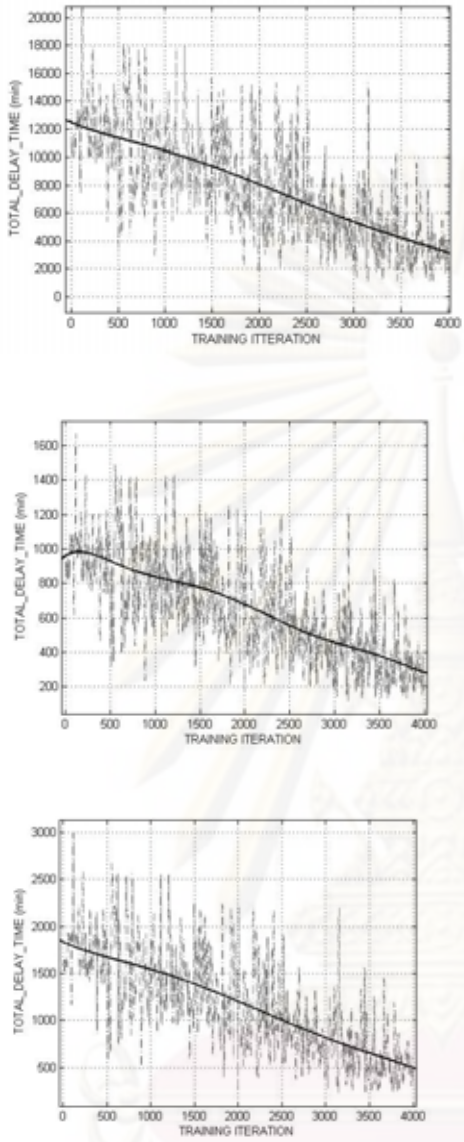


Figure 1 TOTAL_DELAY_TIME and convergence of Q-learning method in various training data set

From Figure 1, the y-axis represents TOTAL_DELAY_TIME from the training process of Q-learning. We found that in the early period of learning, TOTAL_DELAY_TIME fluctuated in wide range but later, TOTAL_DELAY_TIME decreased considerably because β value decreased. Thus, a result of this has the effect on the selection of symbol representing aircraft that have the highest Q value so TOTAL_DELAY_TIME falls gradually.

To estimate with back propagation Neural Network approach, we train an input by using the state from Q-learning. We fix the output at 50 nodes ranging from 0 to –1000 which was the result from Q value from the training step. Hidden layer has 20 nodes.

After completing training the approximation function, we can estimate Q value when we put the irregularity that we have never trained before into schedules. We do this because we want to test an effectiveness of our learning.

From table 3 we can see the TOTAL_DELAY_TIME from the irregularity that we have never trained before. The result represents the ability of this method in term of apply Q-learning. By using back propagation Neural Network to be the approximation function when we met the irregularity that we have never seen before. Results show a good performance compared to the solutions from the branch and bound method.

Table 3 Results of Q-learning with back propagation Neural Network approximation and Branch and Bound method

Irregularity	TOTAL_DELAY_TIME	
	Q-Learning method	Branch and Bound method
[aircraft_1 16.00 60]	400	350
[aircraft_2 14.00 60]	500	500
[aircraft_4 18.00 60]	180	180
[aircraft_5 18.00 60]	125	125
[aircraft_6 10.00 60]	600	580
[aircraft_6 12.00 60]	460	460

6. CONCLUSION AND DEVELOPMENT

Our aim is to implement Q-learning algorithm to solve flight rescheduling problems in airlines. During training, Q-learning spends moderate time to converge but planning for flight rescheduling was done before flight operation. Consequently, we can say that Q-learning is an effective method to solve flight scheduling problem in real time.

For future research, more constraints can be added and other approximation functions like SVM should be considered.

7. REFERENCES

[1] G. Tesauro, "Temporal difference learning and TD-Gammon, "*Communications of the ACM,* vol. 58, pp. 58-68, 1995.

[2] W. Zhang and T. G. Dietterich, "A reinforcement learning approach to job-shop scheduling," in *Proceeding of the 14th International Joint Conference on Artificial Intelligence*, 1995.

[3] M. Clarke, "Solving the Problem of Irregular Airline Operations," in *INFORMS* New Orleans, 1995.

[4] M. Clarke, "Development of Heuristic Procedures for Flight Rescheduling in the Aftermath of Irregular Airline Operations," in *INFORMS* Atlanta, 1996.

[5] A. Jarrah, G. Yu, N. Krishnamurthy, and A. Rakshit, "Decision support framework for airline flight cancellations and delays," *Transportation Science,* vol. 27, 1993.

[6] D. F. X. Mathaisel, "Decision support for airline system operations control and irregular operations," *Computers and Operations Research* vol. 23, pp. 1083-1096, 1996.

[7] T. M. Mitchell, "Reinforcement learning," in *Machine learning*: McGraw-Hill, 1997, pp. 367-387.

[8] K. Gersmann and B. Hammer, "A reinforcement learning algorithm to improve scheduling search heuristics with the SVM," in *2004 IEEE International Joint Conference on Neural Networks* vol. 3, 2004, pp. 1811-1816.

[9] J. A. Boyan and A. W. Moore, "Learning Evaluation Functions for Global Optimization and Boolean Satisfiability," in *AAAI/IAAI*, 1998, pp. 3-10.

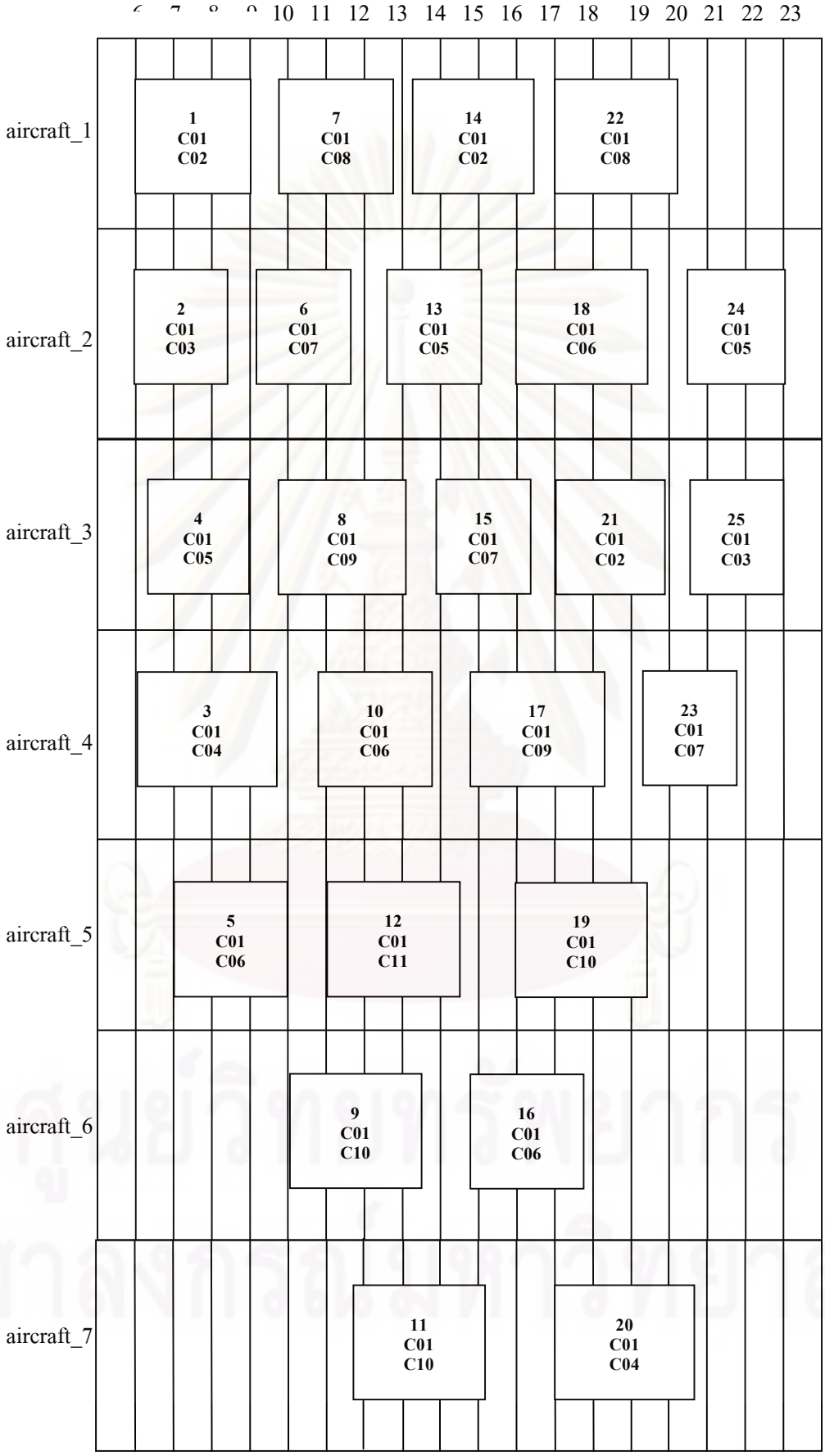
[10] R. H. Crites and A. G. Barto, "Improving Elevator Performance Using Reinforcement Learning," in *Advances in Neural Information Processing Systems*. vol. 8, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, Eds.: The MIT Press, 1996, pp. 1017-1023.

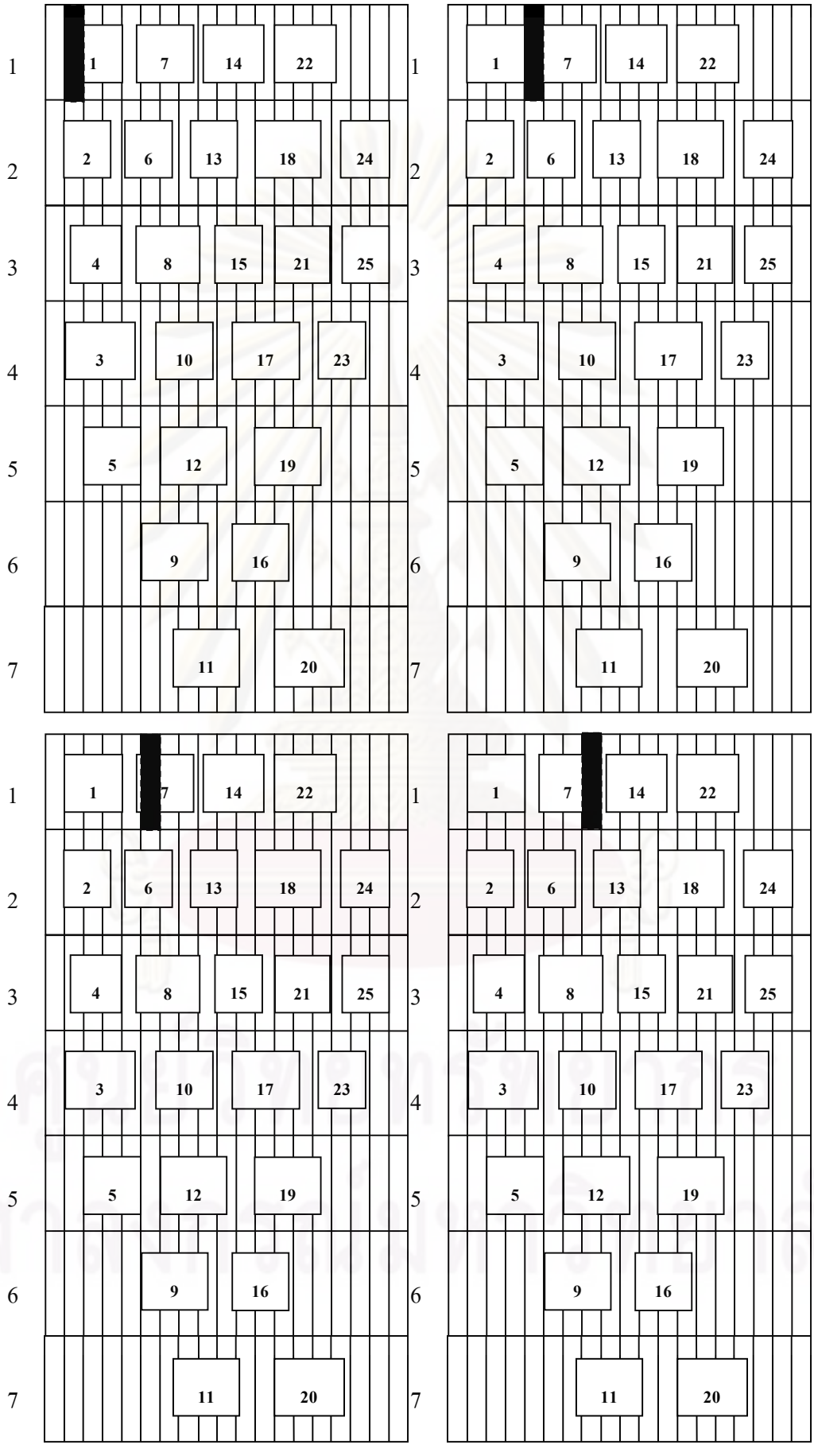
[11] R. Sutton, "Learning to predict by the methods of temporal differences," *Machine Learning,* vol. 3, pp. 9-44, 1988.

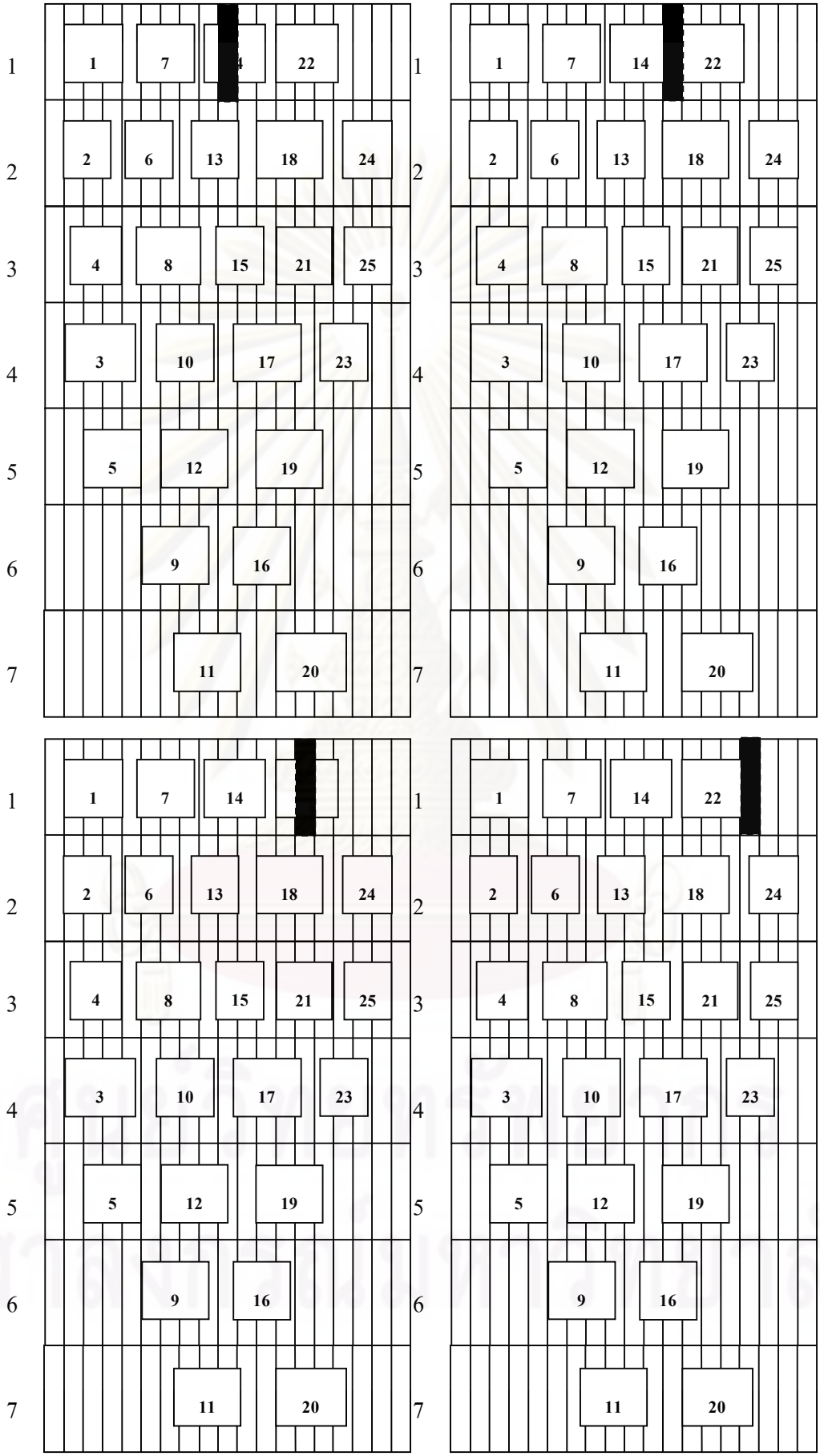
ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

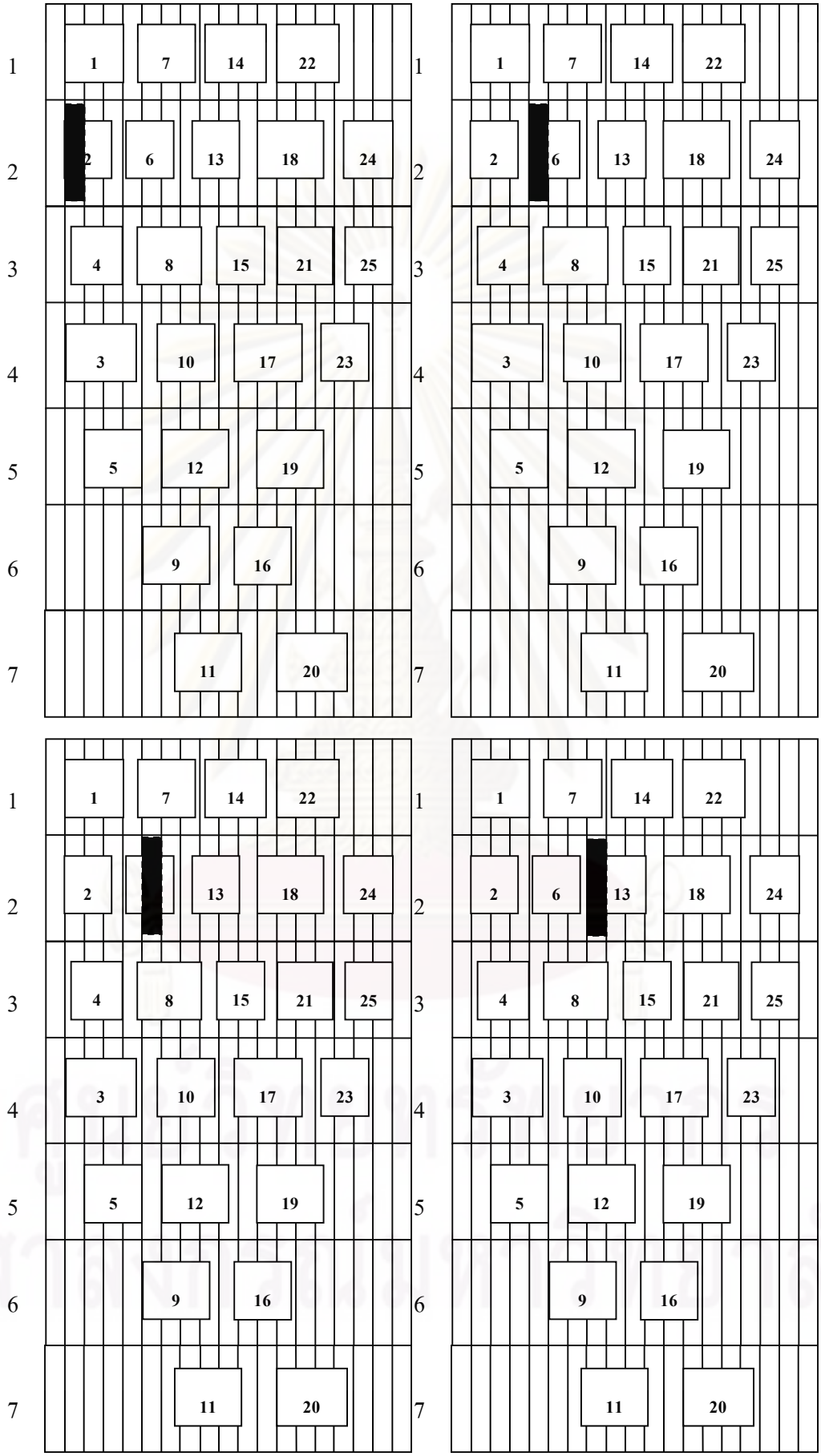
ภาคผนวก ข ข้อมูลตารางเวลาเที่ยวบินที่ใช้ในการเรียนรู้คิว

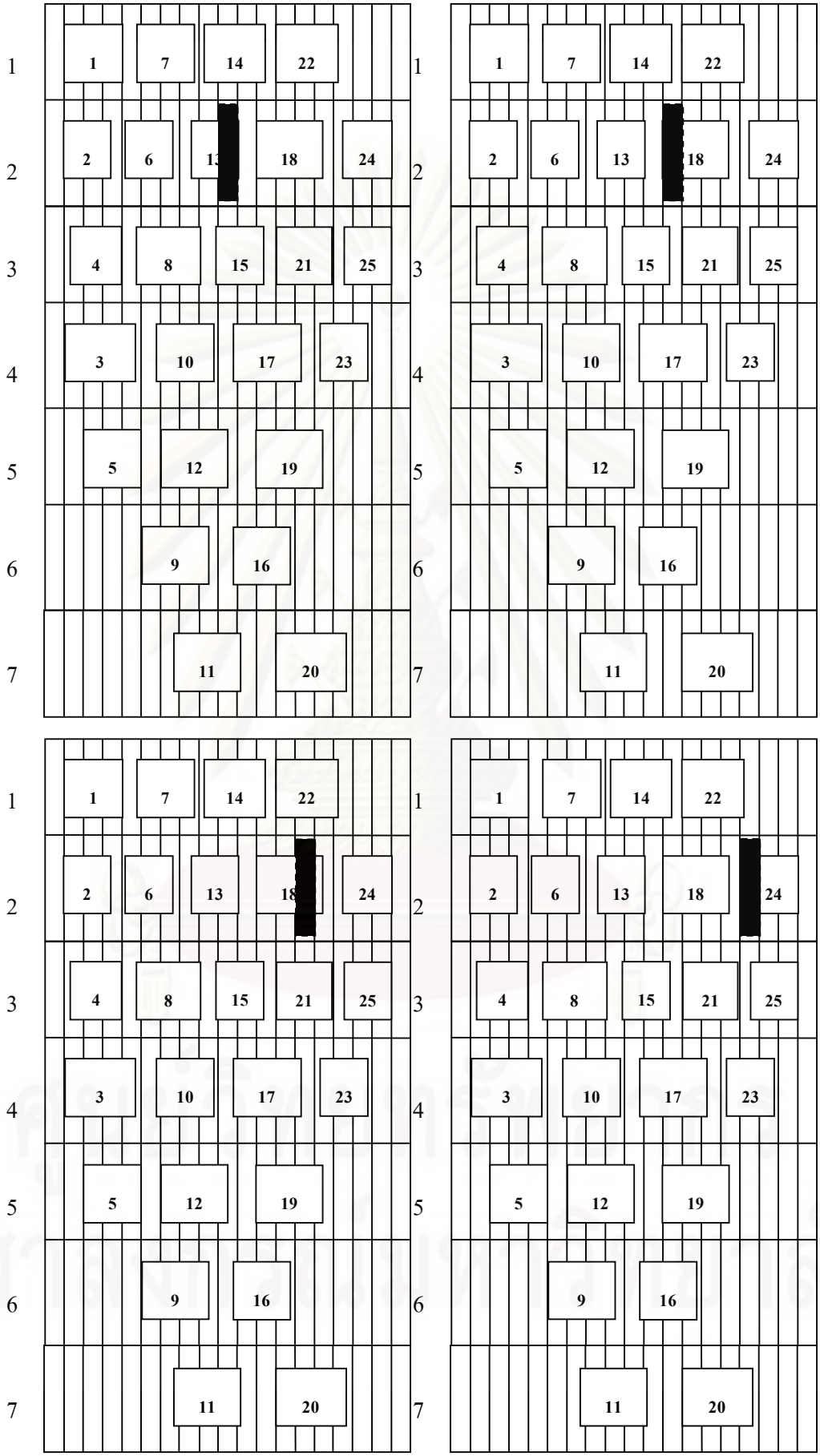
หมายเลขคู่ของเที่ยวบิน	ต้นทาง	ปลายทาง	เวลาออก	เวลาถึง	สัญลักษณ์แทนเครื่องบิน	แบบเครื่อง
1	C01	C02	06:00	08:50	aircraft_1	734
2	C01	C03	06:05	08:25	aircraft_2	734
3	C01	C04	06:05	09:50	aircraft_4	AB7
4	C01	C05	06:15	09:00	aircraft_3	734
5	C01	C06	08:00	11:05	aircraft_5	AB7
6	C01	C07	09:10	11:40	aircraft_2	734
7	C01	C08	09:35	12:40	aircraft_1	734
8	C01	C09	09:45	13:05	aircraft_3	734
9	C01	C10	10:05	13:35	aircraft_6	AB7
10	C01	C06	10:40	13:45	aircraft_4	AB7
11	C01	C10	11:40	15:10	aircraft_7	AB7
12	C01	C11	11:55	15:20	aircraft_5	AB7
13	C01	C05	12:25	15:10	aircraft_2	734
14	C01	C02	13:25	16:15	aircraft_1	734
15	C01	C07	13:50	16:20	aircraft_3	734
16	C01	C06	14:35	17:40	aircraft_6	AB7
17	C01	C09	14:45	18:10	aircraft_4	AB7
18	C01	C06	16:10	19:15	aircraft_2	734
19	C01	C10	16:50	20:25	aircraft_5	AB7
20	C01	C04	17:00	20:45	aircraft_7	AB7
21	C01	C02	17:05	19:55	aircraft_3	734
22	C01	C08	17:05	20:10	aircraft_1	734
23	C01	C07	19:15	21:50	aircraft_4	AB7
24	C01	C05	20:20	23:05	aircraft_2	734
25	C01	C03	20:40	23:00	aircraft_3	734

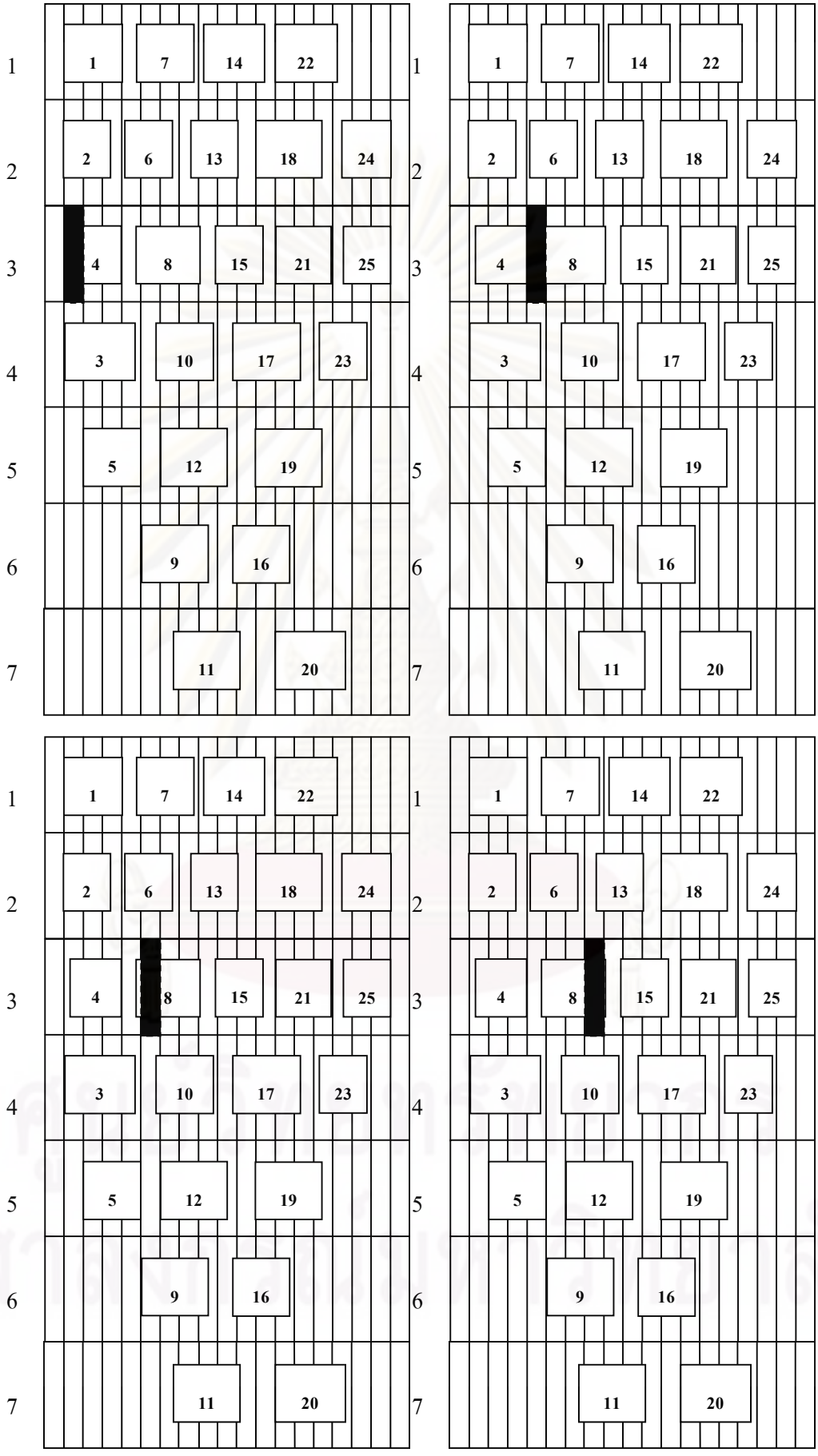

6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
aircraft_1
1 C01 C02
7 C01 C08
14 C01 C02
22 C01 C08
aircraft_2
2 C01 C03
6 C01 C07
13 C01 C05
18 C01 C06
24 C01 C05
aircraft_3
4 C01 C05
8 C01 C09
15 C01 C07
21 C01 C02
25 C01 C03
aircraft_4
3 C01 C04
10 C01 C06
17 C01 C09
23 C01 C07
aircraft_5
5 C01 C06
12 C01 C11
19 C01 C10
aircraft_6
9 C01 C10
16 C01 C06
aircraft_7
11 C01 C10
20 C01 C04

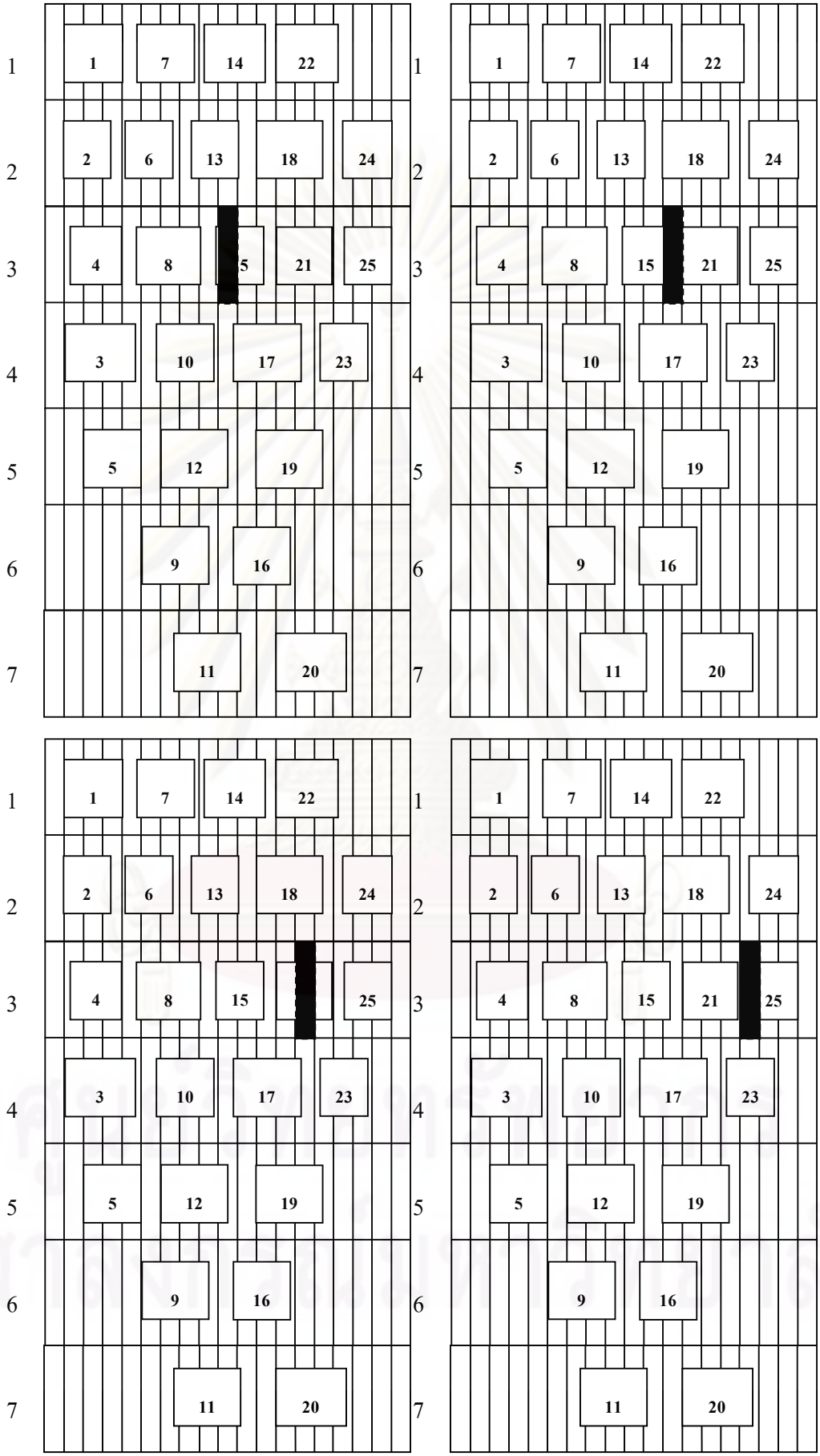

1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20

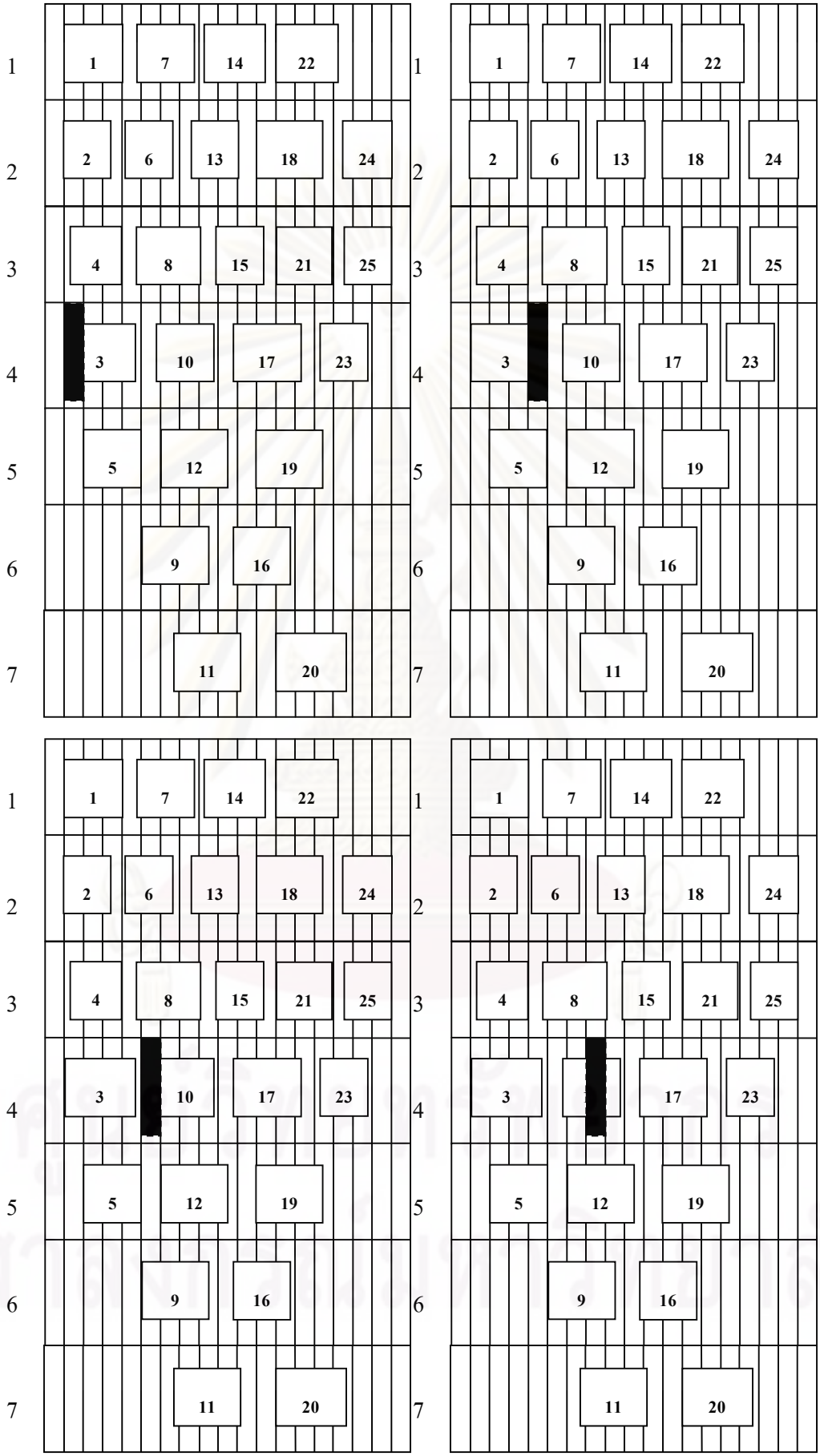

1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20

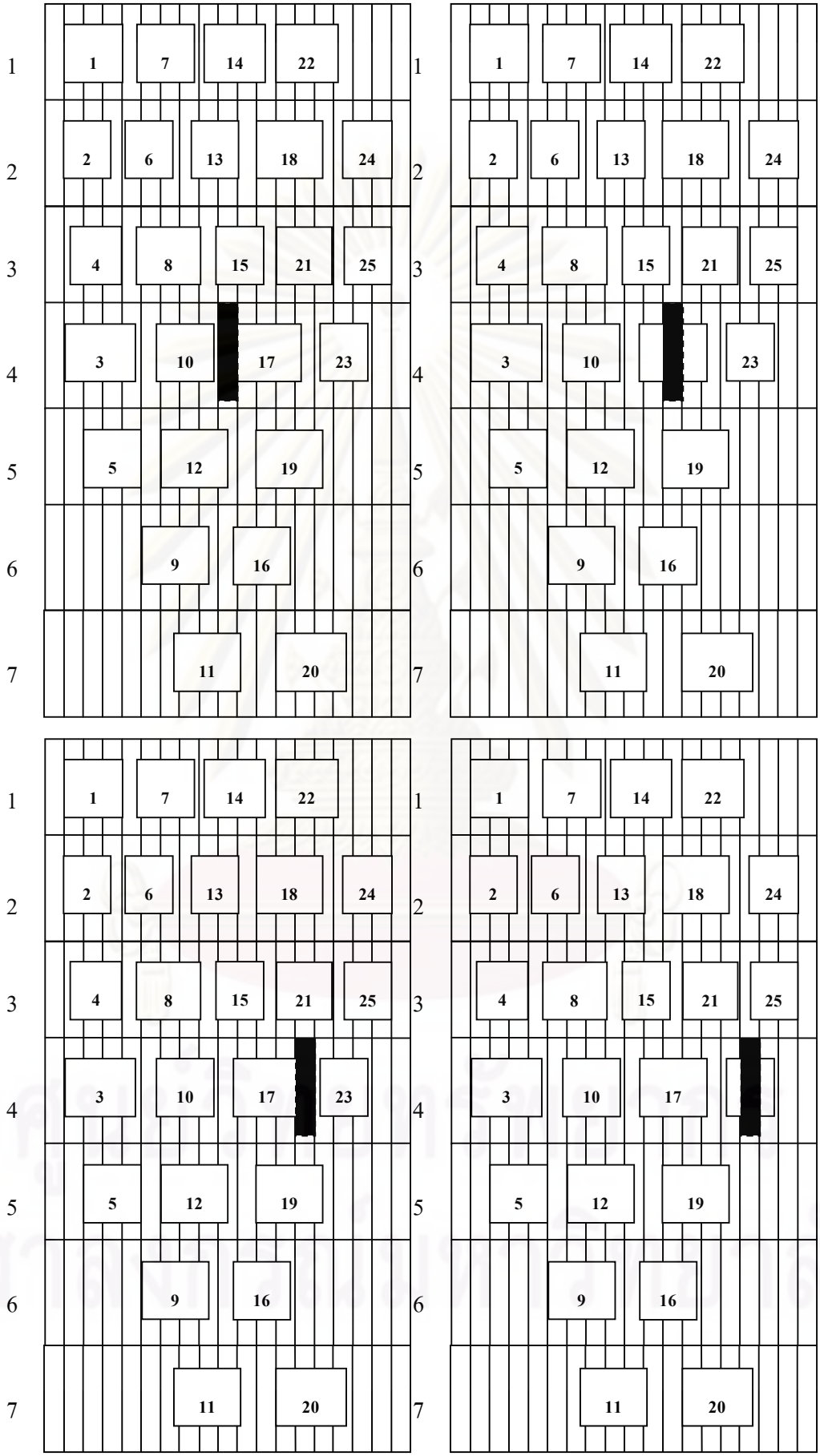

1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20

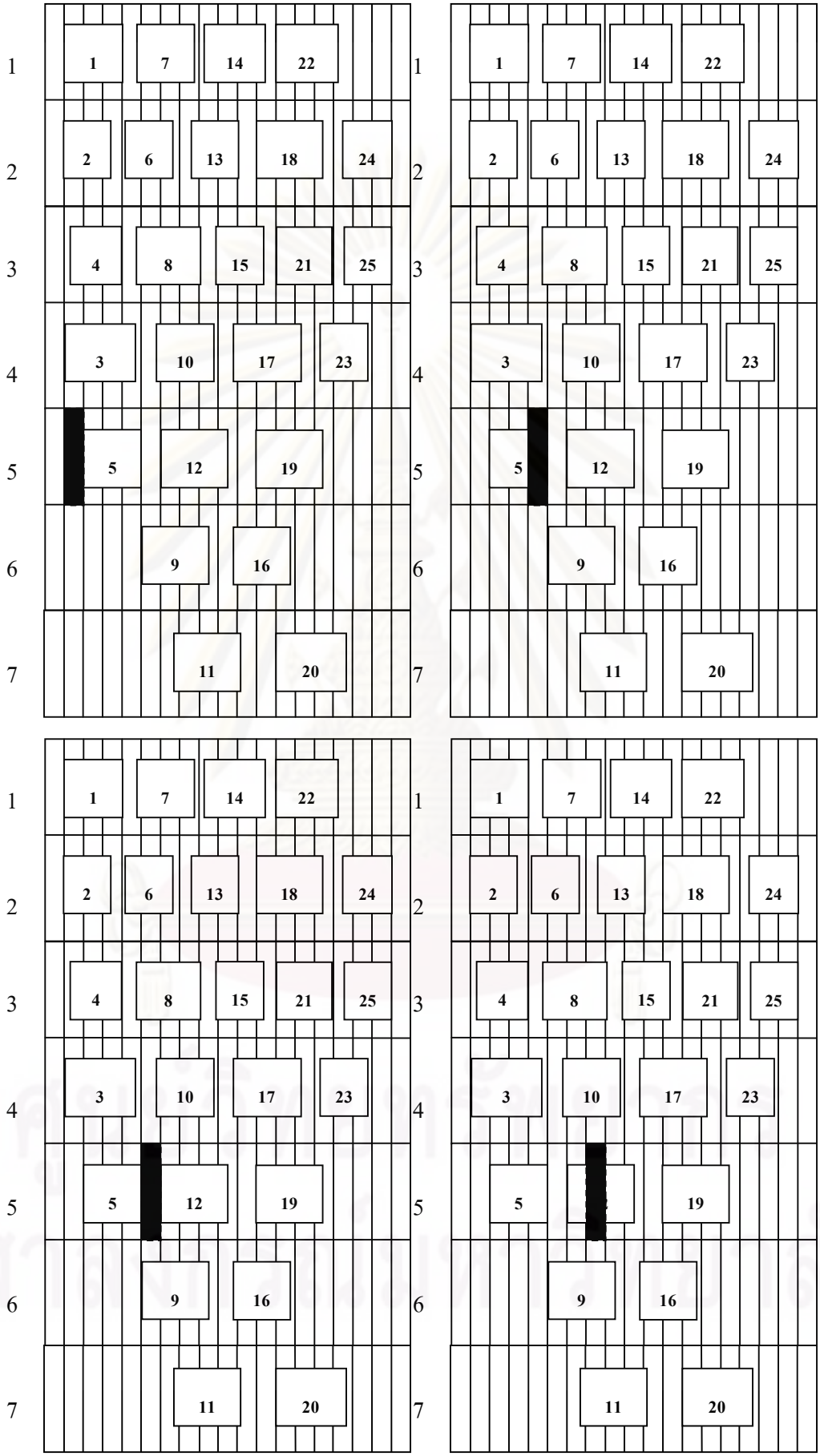

1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20

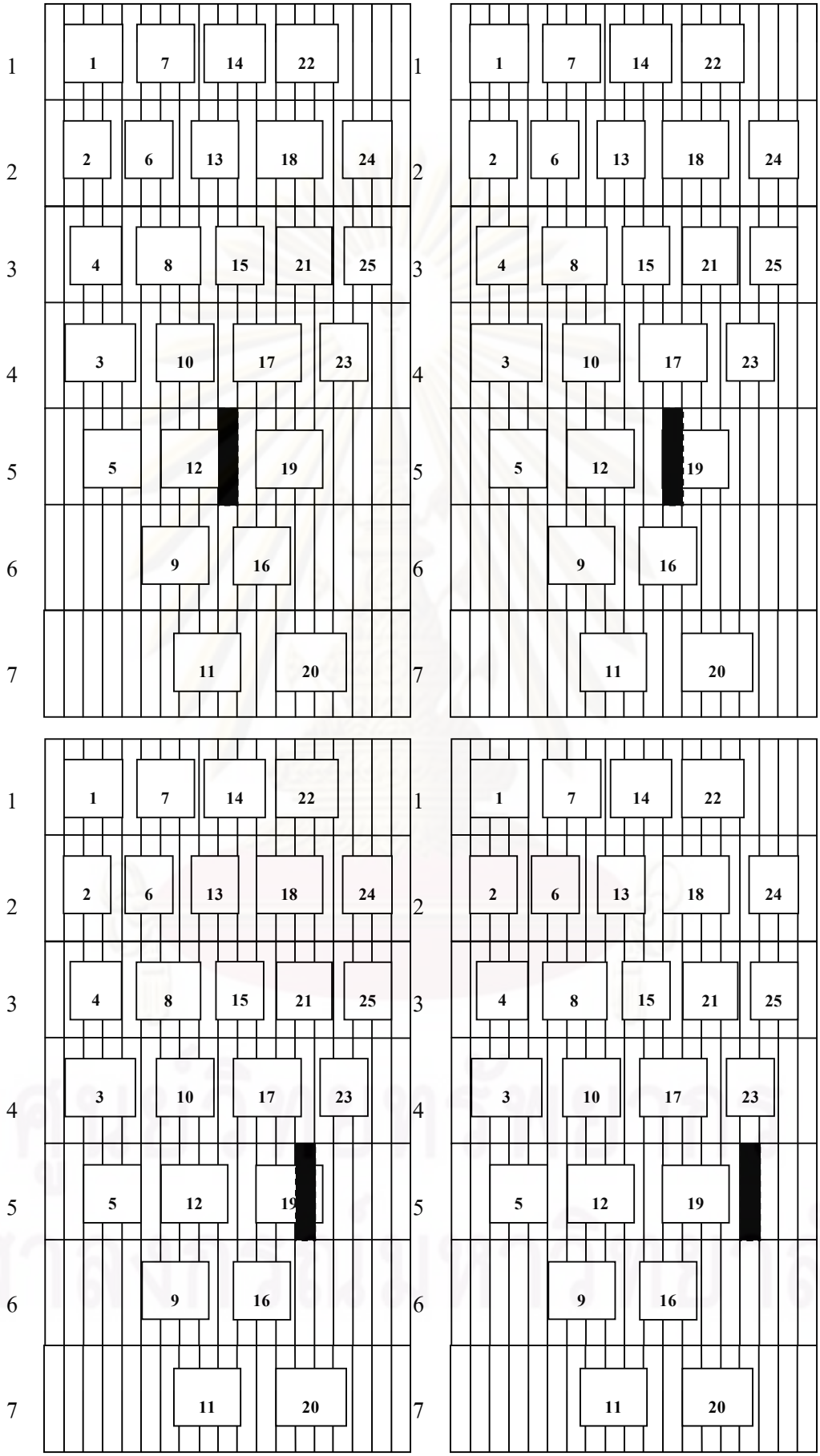

1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20

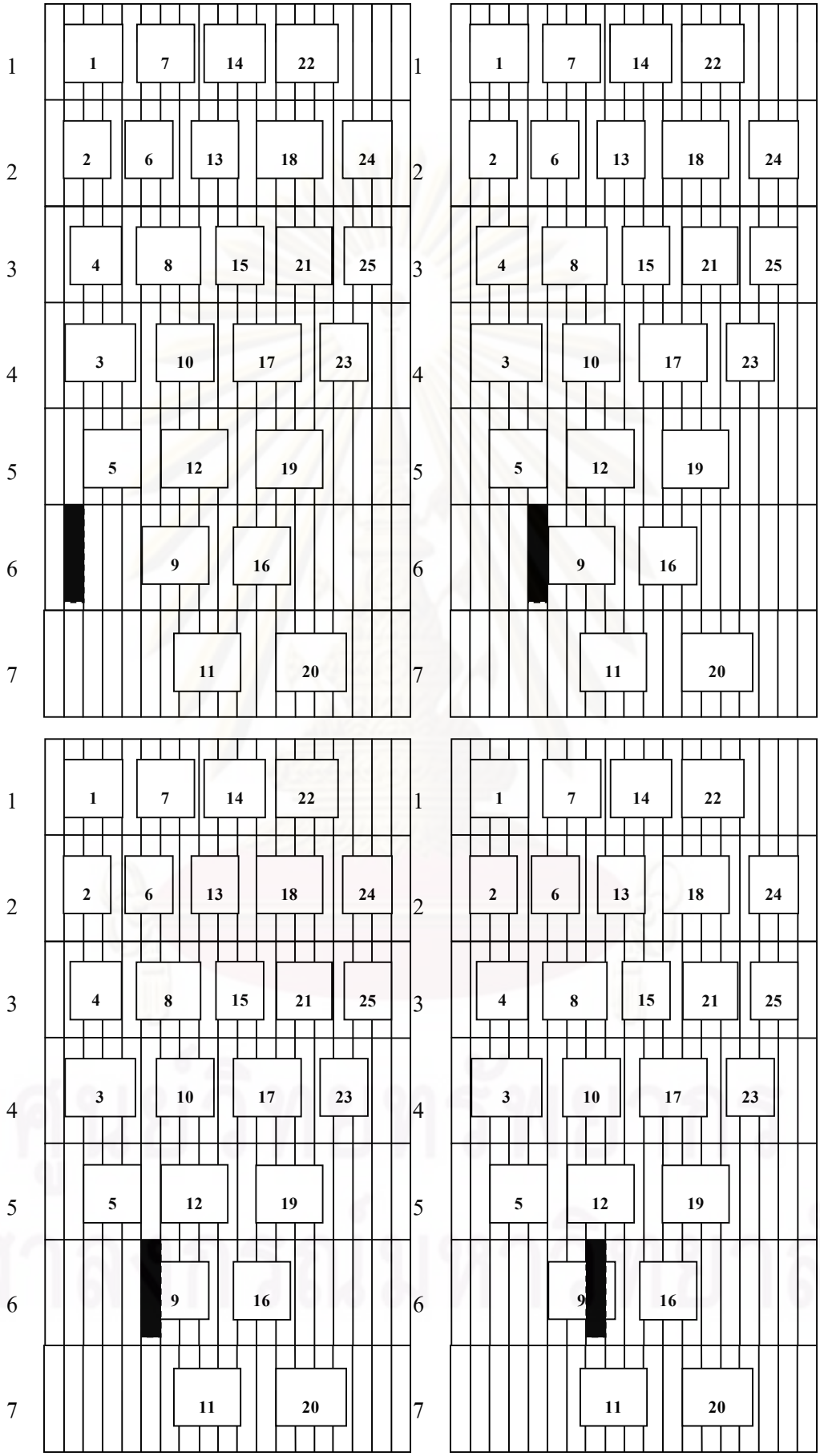

1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20

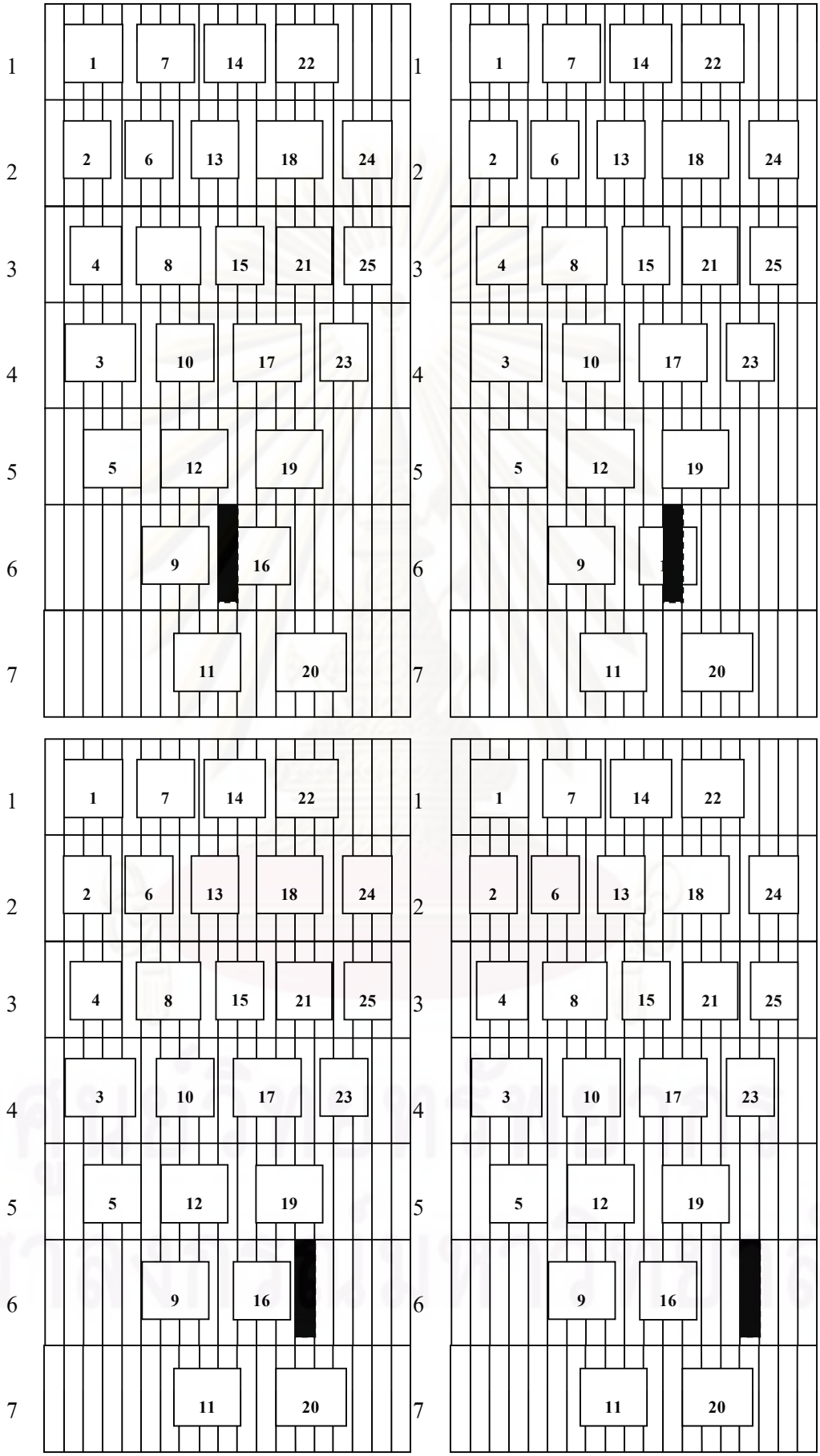

1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 17 23
5 12 19
9 16
11 20

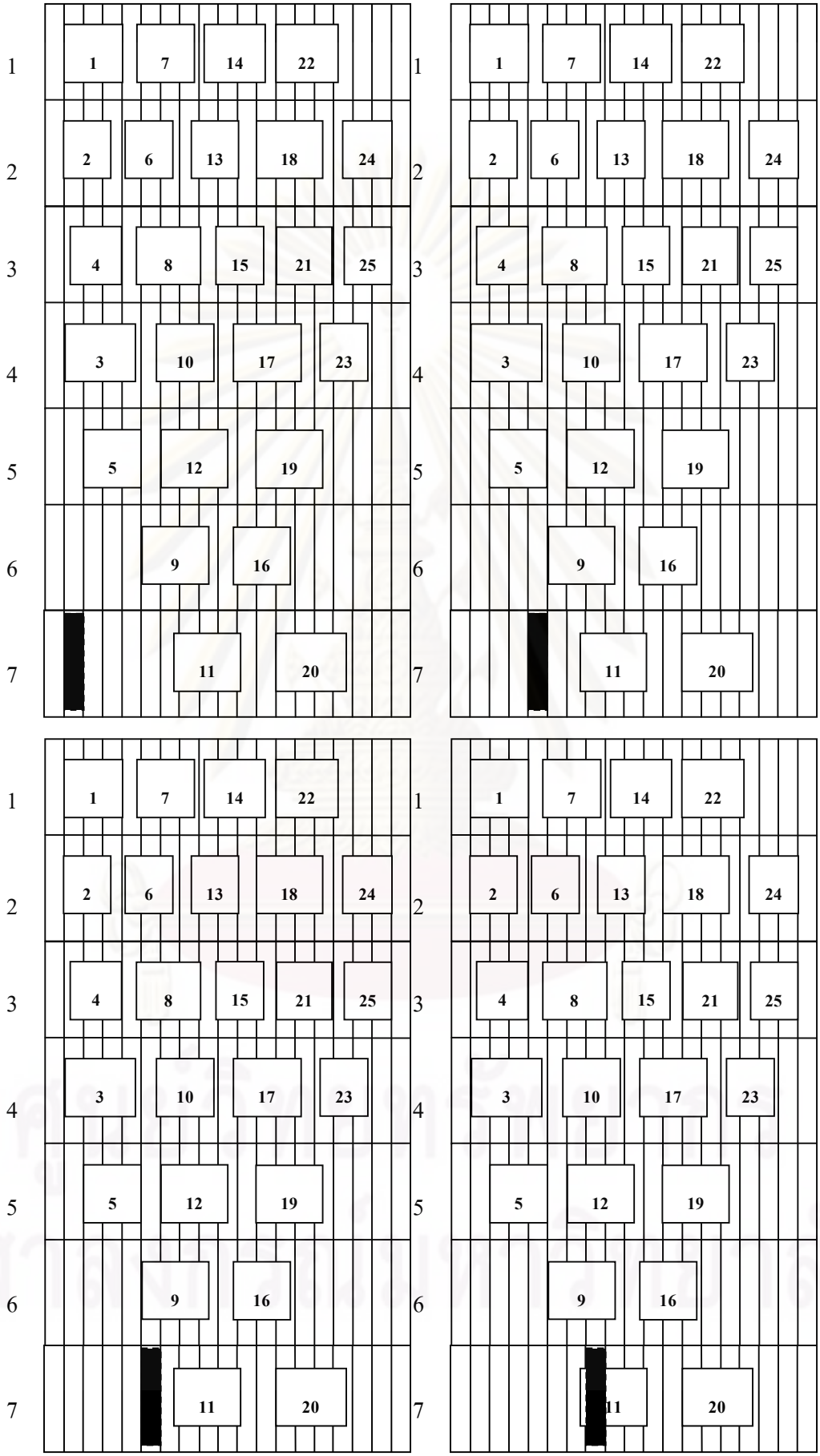

1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17
5 5 12 19
6 9 16
7 11 20

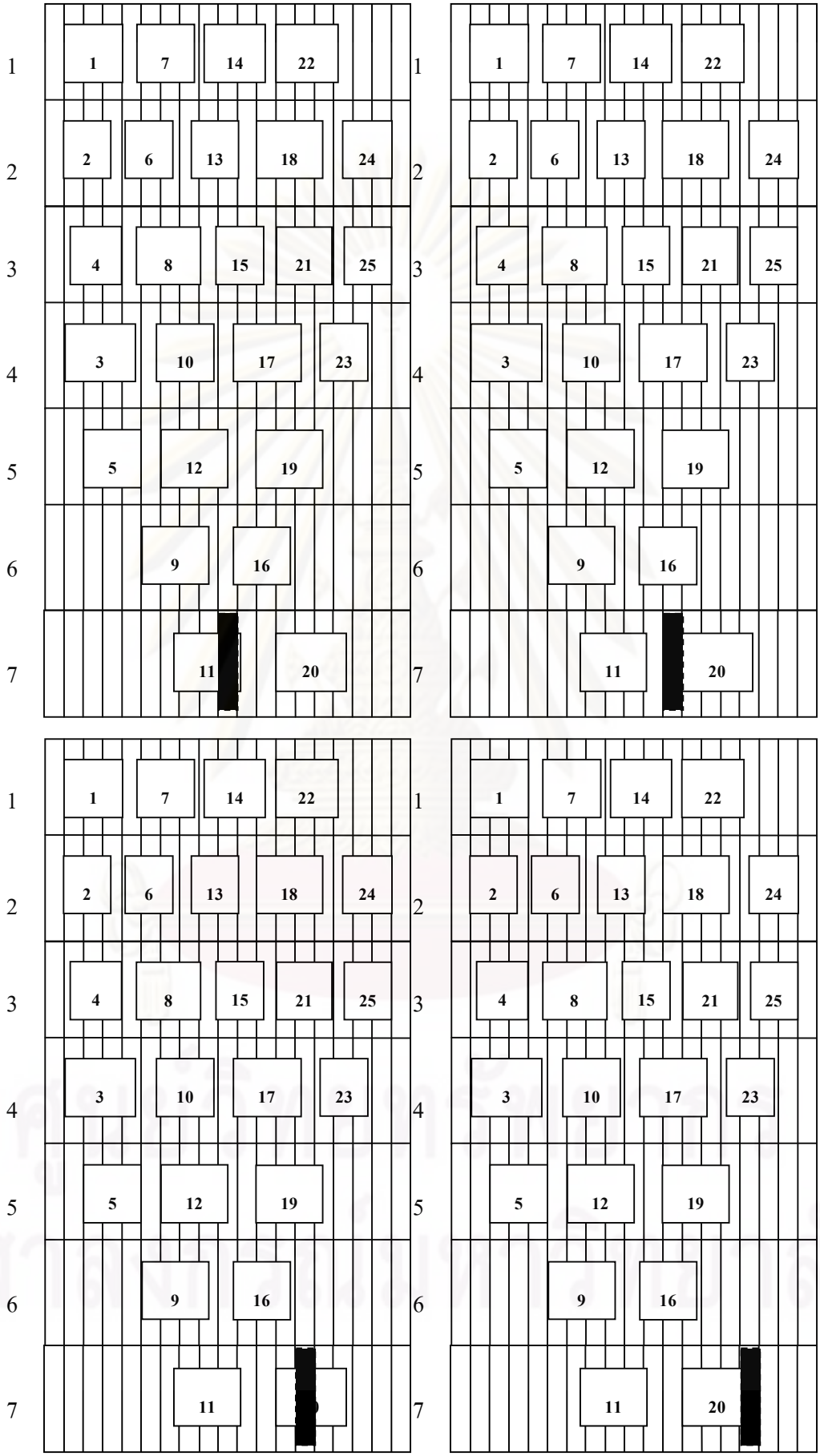

1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 19
6 9 16
7 11 20


1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20


1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20
1 1 7 14 22
2 2 6 13 18 24
3 4 8 15 21 25
4 3 10 17 23
5 5 12 19
6 9 16
7 11 20


1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9
11 20
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20


1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20

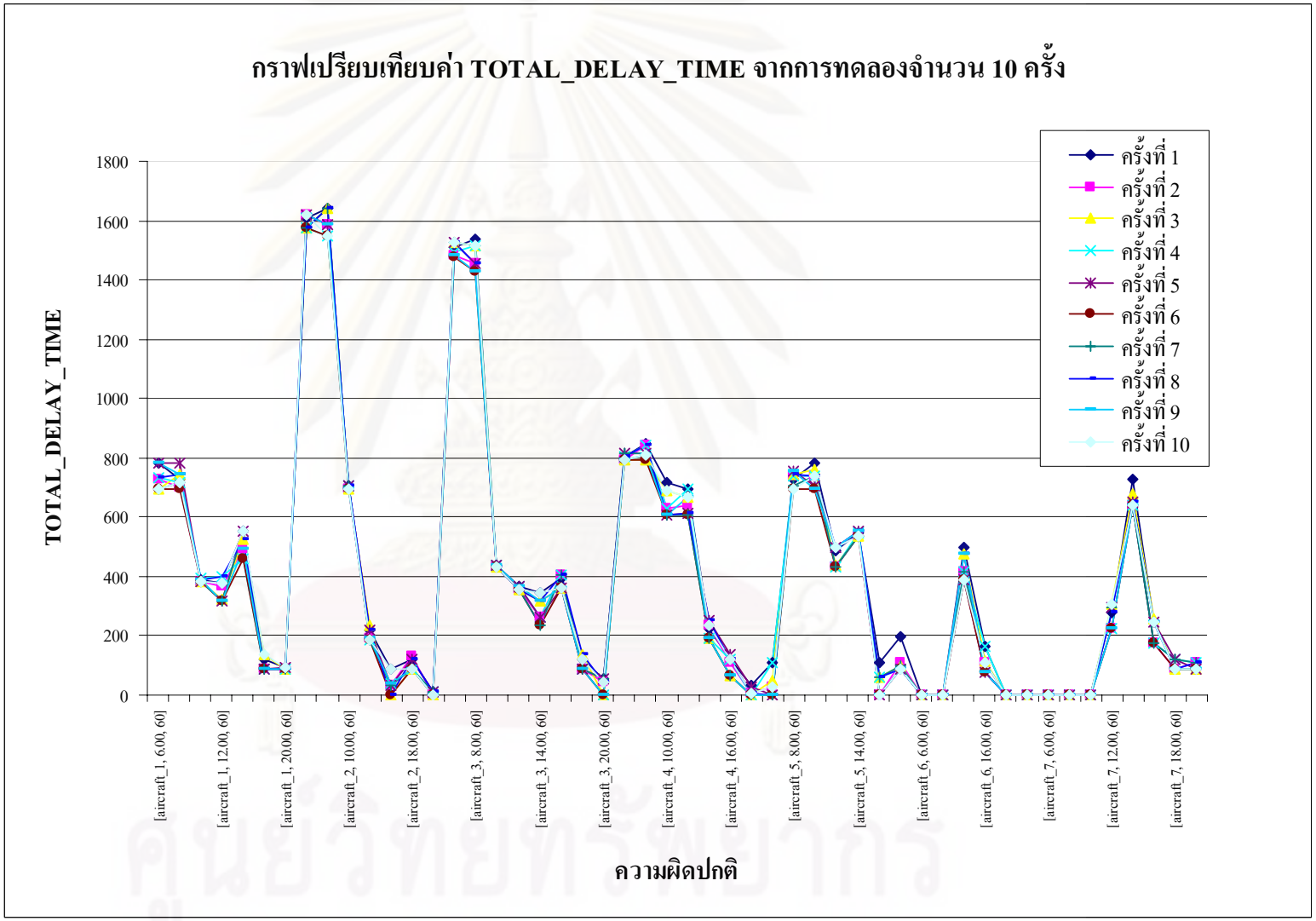

1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11
1 7 14 22
2 6 13 18 24
4 8 15 21 25
3 10 17 23
5 12 19
9 16
11 20

ภาคผนวก ค ตารางค่า TOTAL_DELAY_TIME จากการเรียนรู้คิว

ความผิดปกติ	TOTAL_DELAY_TIME									
	1	2	3	4	5	6	7	8	9	10
[*aircraft_1, 6.00, 60*]	785	730	695	735	785	695	695	735	785	695
[*aircraft_1, 8.00, 60*]	730	700	745	715	785	695	715	745	745	715
[*aircraft_1, 10.00, 60*]	390	385	385	395	385	385	385	390	390	385
[*aircraft_1, 12.00, 60*]	375	365	325	400	320	320	320	400	320	375
[*aircraft_1, 14.00, 60*]	550	495	525	460	550	460	550	525	495	550
[*aircraft_1, 18.00, 60*]	120	90	135	85	90	85	85	85	85	135
[*aircraft_1, 20.00, 60*]	95	85	85	85	95	85	85	85	85	95
[*aircraft_2, 6.00, 60*]	1610	1620	1575	1575	1620	1575	1575	1575	1620	1620
[*aircraft_2, 8.00, 60*]	1640	1585	1640	1550	1585	1550	1640	1640	1585	1550
[*aircraft_2, 10.00, 60*]	705	695	695	705	705	695	705	705	695	695
[*aircraft_2, 12.00, 60*]	220	190	235	185	220	185	185	220	190	185
[*aircraft_2, 16.00, 60*]	90	35	0	40	35	0	35	0	40	90
[*aircraft_2, 18.00, 60*]	120	130	90	85	120	85	85	120	85	85
[*aircraft_2, 20.00, 60*]	10	0	0	10	10	0	10	10	0	0
[*aircraft_3, 6.00, 60*]	1510	1480	1525	1495	1525	1475	1525	1525	1480	1525
[*aircraft_3, 8.00, 60*]	1540	1455	1515	1515	1455	1430	1515	1455	1430	1515
[*aircraft_3, 10.00, 60*]	440	430	430	440	440	430	430	430	440	430
[*aircraft_3, 12.00, 60*]	365	355	355	365	365	355	355	355	365	355
[*aircraft_3, 14.00, 60*]	345	260	320	260	260	235	235	320	320	345
[*aircraft_3, 16.00, 60*]	395	405	360	405	360	360	405	405	360	360
[*aircraft_3, 18.00, 60*]	120	90	135	85	85	85	85	135	85	120
[*aircraft_3, 20.00, 60*]	55	45	0	0	55	0	55	45	0	45
[*aircraft_4, 6.00, 60*]	805	795	795	805	815	795	815	805	795	795
[*aircraft_4, 8.00, 60*]	850	840	795	840	815	795	815	840	810	810
[*aircraft_4, 10.00, 60*]	715	630	690	630	605	605	605	605	605	690
[*aircraft_4, 12.00, 60*]	695	640	670	695	615	605	605	615	670	670
[*aircraft_4, 14.00, 60*]	225	235	190	190	250	190	190	250	190	235
[*aircraft_4, 16.00, 60*]	120	110	65	120	135	65	65	120	65	120

[*aircraft_4, 20.00, 60*]	35	5	0	0	30	0	0	5	0	5
[*aircraft_5, 6.00, 60*]	110	25	50	110	0	0	0	0	0	25
[*aircraft_5, 8.00, 60*]	730	740	745	740	755	695	705	745	755	695
[*aircraft_5, 10.00, 60*]	785	740	760	740	705	695	740	740	695	740
[*aircraft_5, 12.00, 60*]	485	430	435	430	500	430	430	500	500	500
[*aircraft_5, 14.00, 60*]	545	535	535	545	555	535	535	555	555	535
[*aircraft_5, 16.00, 60*]	110	0	60	60	0	0	60	60	0	0
[*aircraft_5, 20.00, 60*]	195	110	100	85	85	85	100	85	85	85
[*aircraft_6, 6.00, 60*]	0	0	0	0	0	0	0	0	0	0
[*aircraft_6, 8.00, 60*]	0	0	0	0	0	0	0	0	0	0
[*aircraft_6, 14.00, 60*]	500	415	475	415	390	390	415	475	475	390
[*aircraft_6, 16.00, 60*]	165	110	115	165	75	75	110	75	75	110
[*aircraft_6, 18.00, 60*]	0	0	0	0	0	0	0	0	0	0
[*aircraft_6, 20.00, 60*]	0	0	0	0	0	0	0	0	0	0
[*aircraft_7, 6.00, 60*]	0	0	0	0	0	0	0	0	0	0
[*aircraft_7, 8.00, 60*]	0	0	0	0	0	0	0	0	0	0
[*aircraft_7, 10.00, 60*]	0	0	0	0	0	0	0	0	0	0
[*aircraft_7, 12.00, 60*]	280	225	305	225	295	225	295	280	225	305
[*aircraft_7, 14.00, 60*]	730	640	680	640	650	640	640	650	640	640
[*aircraft_7, 16.00, 60*]	245	175	255	175	245	175	175	245	245	245
[*aircraft_7, 18.00, 60*]	85	85	85	120	120	85	120	85	85	85
[*aircraft_7, 20.00, 60*]	85	110	85	110	85	85	110	110	85	85
ค่าเฉลี่ย	414.1	384.4	393.2	388.6	391.5	367.1	384.1	395	383.1	392.7

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย


กราฟเปรียบเทียบค่า TOTAL_DELAY_TIME จากการทดลองจำนวน 10 ครั้ง
TOTAL_DELAY_TIME
1800
1600
1400
1200
1000
800
600
400
200
0
ครั้งที่ 1
ครั้งที่ 2
ครั้งที่ 3
ครั้งที่ 4
ครั้งที่ 5
ครั้งที่ 6
ครั้งที่ 7
ครั้งที่ 8
ครั้งที่ 9
ครั้งที่ 10
[aircraft_1, 6.00, 60]
[aircraft_1, 12.00, 60]
[aircraft_1, 20.00, 60]
[aircraft_2, 10.00, 60]
[aircraft_2, 18.00, 60]
[aircraft_3, 8.00, 60]
[aircraft_3, 14.00, 60]
[aircraft_3, 20.00, 60]
[aircraft_4, 10.00, 60]
[aircraft_4, 16.00, 60]
[aircraft_5, 8.00, 60]
[aircraft_5, 14.00, 60]
[aircraft_6, 6.00, 60]
[aircraft_6, 16.00, 60]
[aircraft_7, 6.00, 60]
[aircraft_7, 12.00, 60]
[aircraft_7, 18.00, 60]
ความผิดปกติ

84

ประวัติผู้เขียนวิทยานิพนธ์

นายโอภาส เพลินประภาพร เกิดวันที่ 28 ตุลาคม พ.ศ. 2525 ที่เขตบางแค จังหวัดกรุงเทพมหานคร สำเร็จการศึกษาระดับปริญญาวิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมคอมพิวเตอร์ จากภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยมหิดลในปีการศึกษา 2547 ภายหลังจบการศึกษาได้เข้าศึกษาต่อในหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2549

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย