การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง สำหรับทดสอบความรู้ทางไอซีที: การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์

ธีร์นวัช สุขวิลัยหิรัญ

ดุษฎีนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา
วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา
ธันวาคม 2561
ลิขสิทธิ์เป็นของมหาวิทยาลัยบูรพา

คณะกรรมการควบคุมดุษฎีนิพนธ์และคณะกรรมการสอบดุษฎีนิพนธ์ ได้พิจารณาดุษฎีนิพนธ์ของ ธีร์นวัช สุขวิลัยหิรัญ ฉบับนี้แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา ของมหาวิทยาลัยบูรพาได้

คณะกรรมการควบคุมดุษฎีนิพนธ์

.....อาจารย์ที่ปรึกษาหลัก
(รองศาสตราจารย์ ดร. เสรี ชัดแช้ม)

.....อาจารย์ที่ปรึกษาร่วม
(ดร. ปิยะทิพย์ ประดุจพรม)

คณะกรรมการสอบดุษฎีนิพนธ์

.....ประธาน
(รองศาสตราจารย์ ดร. กานดา พูนลาภทวี)

.....กรรมการ
(รองศาสตราจารย์ ดร. เสรี ชัดแช้ม)

.....กรรมการ
(ดร. ปิยะทิพย์ ประดุจพรม)

.....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. พูลพงศ์ สุขสว่าง)

.....กรรมการ
(ดร. กนก พานทอง)

วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญาอนุมัติให้รับดุษฎีนิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา ของมหาวิทยาลัยบูรพา

.....คณบดีวิทยาลัยวิทยาการวิจัย
(ผู้ช่วยศาสตราจารย์ ดร. ภัทราวดี มากมี) และวิทยาการปัญญา
วันที่ 16 เดือน ธันวาคม พ.ศ. 2561

ดุษฎีนิพนธ์นี้ “ได้รับการสนับสนุนการวิจัยแผนงานเสริมสร้างศักยภาพและพัฒนานักวิจัยรุ่นใหม่
ตามทิศทางยุทธศาสตร์การวิจัยและนวัตกรรม: ประเภทบัณฑิตศึกษา
จากสำนักงานคณะกรรมการวิจัยแห่งชาติ ประจำปี 2562”

ประกาศคุณูปการ

การทำดุษฎีนิพนธ์ฉบับนี้สำเร็จลุล่วงลงได้ด้วยความกรุณาจาก รองศาสตราจารย์ ดร.เสรี ชัดแช้ม อาจารย์ที่ปรึกษาหลัก และ ดร.ปิยะทิพย์ ประดุจพรม อาจารย์ที่ปรึกษาร่วม ที่กรุณาให้คำปรึกษาและแนะนำแนวทางที่ถูกต้อง ตลอดจนแก้ไขข้อบกพร่องต่าง ๆ ด้วยความละเอียดถี่ถ้วน และเอาใจใส่ด้วยดีเสมอมา ผู้วิจัยรู้สึกซาบซึ้งเป็นอย่างยิ่ง จึงกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณ ผู้อำนวยการและเจ้าหน้าที่สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ที่ได้อนุเคราะห์ผลการทดสอบ V-NET เพื่อจัดทำคลังข้อสอบที่มีคุณภาพ ส่วนด้านสถานที่ทดสอบโปรแกรมและจัดสรรนักศึกษาเข้าร่วมการทดสอบ ได้รับความอนุเคราะห์จาก ดร.ประเสริฐ กลิ่นชู ในด้านโปรแกรมที่ใช้ในการวิเคราะห์คุณภาพข้อสอบ ได้รับความอนุเคราะห์จาก ดร.มิ่ง เทพครเมือง ในด้านการจัดหาแหล่งวารสารวิจัยเพื่อตีพิมพ์บทความวิจัย ได้รับความอนุเคราะห์จาก รองศาสตราจารย์ ดร.ปณิตา วรรณพิรุณ ส่วนด้านวิธีการวิเคราะห์ข้อมูลทางสถิติและการอำนวยความสะดวกสถานที่ในการวิเคราะห์ข้อมูล ได้รับความอนุเคราะห์จาก พันเอก ดร.เอกพงศ์ แพ่งกุล และด้านวิธีการนำเสนอผลงาน ได้รับความอนุเคราะห์จาก ผู้ช่วยศาสตราจารย์ ดร.พูลพงศ์ สุขสว่าง ขอขอบคุณผู้เชี่ยวชาญทุกท่านที่ให้ความอนุเคราะห์ตรวจประสิทธิภาพโปรแกรมพร้อมให้คำแนะนำปรับปรุงให้โปรแกรมมีคุณภาพ และขอขอบคุณคณาจารย์ที่กรุณาประสาทวิชา มอบความรู้ กำลังใจ และเจ้าหน้าที่วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญาทุกท่านที่อำนวยความสะดวกในการทำวิจัย

ขอกราบขอบพระคุณ คุณพ่อประยงค์ คุณแม่วิลัย สุขวิลัย บุพการีผู้เป็นครูที่รักยิ่ง พี่สาว พี่ชาย ที่คอยให้ความช่วยเหลือ เป็นกำลังใจสำคัญในการต่อสู้กับปัญหาและอุปสรรคที่เกิดขึ้น

ขอกราบขอบพระคุณ สำนักงานคณะกรรมการวิจัยแห่งชาติ ที่สนับสนุนการวิจัยแผนงานเสริมสร้างศักยภาพและพัฒนานักวิจัยรุ่นใหม่ ตามทิศทางยุทธศาสตร์การวิจัยและนวัตกรรม: ประเภทบัณฑิตศึกษา ประจำปี 2562

คุณค่าและประโยชน์ของดุษฎีนิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นกตัญญูกตเวทิตาแด่บุพการี บูรพาจารย์ และผู้มีพระคุณทุกท่านทั้งในอดีตและปัจจุบัน ที่ส่งเสริมให้ข้าพเจ้าเป็นผู้มีการศึกษาและประสบความสำเร็จมาจนตราบเท่าทุกวันนี้

ธีร์นวัช สุขวิลัยหิรัญ

55810256: สาขาวิชา: การวัดและเทคโนโลยีทางวิทยาการปัญญา;
ปร.ด. (การวัดและเทคโนโลยีทางวิทยาการปัญญา)
คำสำคัญ: การควบคุมการใช้ข้อสอบ/ การควบคุมสมดุลเนื้อหา/ โหลดบาลานซ์ซิ่ง/
ความรู้ทางไอซีที/ การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์

ธีร์นวัช สุขวิลัยหิรัญ: การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง สำหรับทดสอบความรู้ทางไอซีที: การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (DEVELOPMENT OF THE ITEM EXPOSURE CONTROL WITH CONTENT BALANCE PROCEDURE USING LOAD BALANCING FOR ICT LITERACY: MULTIDIMENSIONAL COMPUTERIZED ADAPTIVE TESTING) คณะกรรมการควบคุมดุษฎีนิพนธ์: เสรี ชัดแช้ม, ค.ด., ปิยะทิพย์ ประดุจพรม, Ph.D., 210 หน้า. ปี พ.ศ. 2561.

การควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา เป็นวิธีการเพิ่มประสิทธิภาพในการคัดเลือกข้อสอบข้อถัดไปและเลือกใช้ข้อสอบครอบคลุมเนื้อหาทุกด้าน การวิจัยนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง (LB) 2) พัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา สำหรับทดสอบความรู้ทางไอซีที (ICT-MCAT) มี 2 วิธี คือ 2.1) วิธีโหลดบาลานซ์ซิ่ง (LB) และ 2.2) วิธี Item Eligibility (IE) ในรูปแบบเว็บแอพพลิเคชั่น 3) เปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา 2 วิธี ในด้าน 3.1) จำนวนข้อสอบที่ใช้ในการทดสอบ 3.2) เวลาที่ใช้ในการทดสอบ 3.3) อัตราการแสดงการใช้ข้อสอบสูงสุด และ 3.4) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ใช้กลุ่มตัวอย่างเป็นนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง ปีการศึกษา 2561 จำนวน 400 คน ใช้การสุ่มอย่างเป็นระบบ วิเคราะห์ข้อมูลด้วยค่าความคลาดเคลื่อนกำลังสองเฉลี่ยการทดสอบแมน-วิทนีย์ และการทดสอบค่าที และ 4) ศึกษาความสัมพันธ์ระหว่างค่าความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT) โดยใช้กลุ่มตัวอย่างจำนวน 40 คน ทดสอบสหสัมพันธ์แบบเพียร์สัน

ผลการวิจัยปรากฏว่า 1) วิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง (LB) มี 3 ขั้นตอน ได้แก่ การสำรวจและเคลื่อนย้ายข้อมูลการคำนวณภาระงานและการควบคุมการทำงาน 2) โปรแกรม ICT-MCAT มีประสิทธิภาพอยู่ในระดับมากที่สุด 3) วิธี LB มีประสิทธิภาพสูงกว่า วิธี IE ในด้านจำนวนข้อสอบที่ใช้ในการทดสอบ เวลาที่ใช้ในการทดสอบและอัตราการแสดงการใช้ข้อสอบสูงสุด และ 4) ค่าความสามารถของผู้ทดสอบจาก MCAT กับคะแนนรวมของผู้ทดสอบจาก CT มีความสัมพันธ์กันในทางบวกที่ระดับนัยสำคัญทางสถิติ .01

55810256: MAJOR: MEASUREMENT AND TECHNOLOGY IN COGNITIVE SCIENCE;
Ph.D. (MEASUREMENT AND TECHNOLOGY IN COGNITIVE SCIENCE)
KEYWORDS: ITEM EXPOSURE CONTROL/ CONTENT BALANCE/ LOAD BALANCING/
ICT LITERACY/ MULTIDIMENSIONAL COMPUTERIZED ADAPTIVE TESTING
TEENAWAT SUKVILAIHIRUN: DEVELOPMENT OF THE ITEM EXPOSURE CONTROL WITH CONTENT BALANCE PROCEDURE USING LOAD BALANCING FOR ICT LITERACY: MULTIDIMENSIONAL COMPUTERIZED ADAPTIVE TESTING. ADVISORY COMMITTEE: SEREE CHADCHAM, Ph.D., PIYATHIP PRADUJPROM, Ph.D. 210 P. 2018.

The item exposure control employing the content balancing procedure is a method for increasing the efficiency of test item selection, useful for covering all content areas. This research aimed 1) to develop an item exposure control using load balancing; 2) to develop a Multidimensional Computerized Adaptive Testing program (ICT-MACT) using item exposure controls, with the content balancing procedure for ICT literacy by applying 2.1) Load Balancing (LB) and Item Eligibility (IE) in the form of web applications; 3) to compare the efficiency between LB and IE methods in 3.1) test length, 3.2) test time, 3.3) Maximum Item Exposure Rate (MIER), and 3.4) Root Mean Square Error (RMSE). A sample of 400 vocational college students at the Diploma level in the academic year 2018 was systematically randomized. The Mann-Whitney test and *t*-test were used to analyze the data; and 4) to study the correlation of the ability estimation of the examinees via the Multidimensional Computerized Adaptive Testing (MCAT) and the observed score via the Computerized Testing (CT) with a sample of 40 students, with Pearson's product moment correlation used to analyze the data.

The results showed that: 1) the item exposure control with the content balancing procedure using load balancing had 3 steps in Migration Agent (MA), Load Agent (LA), and Channel Agent (CA); 2) the ICT-MCAT program had the highest efficiency level; 3) the LB method was more effective than IE as measured by test length, test time, and MIER; and 4) the ability estimation of the examinees via the MCAT and the observed score via CT were positively correlated ($p<.01$).

สารบัญ

หน้า

บทคัดย่อภาษาไทย..... จ

บทคัดย่อภาษาอังกฤษ..... ฉ

สารบัญ..... ช

สารบัญตาราง..... ฌ

สารบัญภาพ..... ฎ

บทที่

1 บทนำ..... 1

ความเป็นมาและความสำคัญของปัญหา..... 1

วัตถุประสงค์ของการวิจัย..... 6

กรอบแนวทางการวิจัย..... 7

สมมติฐานของการวิจัย..... 10

ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย..... 10

ขอบเขตของการวิจัย..... 10

นิยามศัพท์เฉพาะ..... 11

2 เอกสารและงานวิจัยที่เกี่ยวข้อง..... 15

ตอนที่ 1 โหลดบาลานซ์ซิ่ง และงานวิจัยที่เกี่ยวข้อง..... 15

ตอนที่ 2 วิธีการควบคุมการใช้ข้อสอบ วิธีการควบคุมสมดุลเนื้อหา และงานวิจัยที่เกี่ยวข้อง..... 23

ตอนที่ 3 ทฤษฎีตอบสนองข้อสอบแบบพหุมิติ..... 35

ตอนที่ 4 การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์..... 46

ตอนที่ 5 ทฤษฎีวงจรการพัฒนาโปรแกรม..... 54

3 วิธีดำเนินการวิจัย..... 65

ระยะที่ 1 การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง (LB)..... 66

ระยะที่ 2 การพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา..... 82

ระยะที่ 3 การเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา..... 104

สารบัญ (ต่อ)

บทที่ หน้า

ระยะที่ 4 การศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT)... 108

4 ผลการวิจัย..... 112

ตอนที่ 1 ผลการพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง (LB)..... 114

ตอนที่ 2 ผลการพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา..... 115

ตอนที่ 3 ผลการเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา..... 129

ตอนที่ 4 ผลการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT)..... 133

5 สรุปผลและอภิปราย..... 135

สรุปผลการวิจัย..... 135

อภิปรายผล..... 137

ข้อเสนอแนะ..... 141

บรรณานุกรม..... 143

ภาคผนวก..... 150

ภาคผนวก ก อัลกอริทึมโหลดบาลานซ์ซิ่ง..... 151

ภาคผนวก ข รายชื่อผู้เชี่ยวชาญ..... 157

ภาคผนวก ค แบบประเมินประสิทธิภาพโปรแกรม (Black box testing)..... 161

ภาคผนวก ง คู่มือการใช้โปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์..... 166

ภาคผนวก จ จริยธรรมการวิจัย..... 193

ภาคผนวก ฉ คลังข้อสอบความรู้ทางไอซีที..... 195

ภาคผนวก ช ผลการทดสอบวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา. 198

ภาคผนวก ซ ผลการทดสอบ MCAT กับ CT..... 207

ประวัติย่อของผู้วิจัย..... 209

สารบัญตาราง

ตารางที่		หน้า
2-1	ความสัมพันธ์ของขนาดเลขชี้กำลังกับจำนวนมิติจากโมเดลของ Mulaik (p=.5).....	40
2-2	ความสัมพันธ์ของขนาดเลขชี้กำลังกับจำนวนมิติจากโมเดลของ Sympson (p=.5)...	41
2-3	สัญลักษณ์ที่ใช้ในการสร้างแผนภาพกระแสข้อมูล.....	57
3-1	จำนวนข้อสอบจำแนกตามหมวดเนื้อหาและความรู้ทางไอซีที.....	69
3-2	จำนวนข้อสอบที่ผ่านการคัดเลือกจำแนกตามเนื้อหาและความรู้ทางไอซีที.....	69
3-3	การเปรียบเทียบวิธี LB กับ วิธี IE	75
3-4	ตัวอย่างการเก็บลักษณะข้อสอบ.....	77
3-5	ตัวอย่างการเก็บค่าเฉลี่ยถ่วงน้ำหนัก.....	79
3-6	ตัวอย่างการเรียงลำดับจากน้อยไปหามากเพื่อคัดเลือกข้อสอบ.....	81
3-7	รายละเอียดข้อมูลผู้ใช้โปรแกรม.....	89
3-8	รายละเอียดข้อมูลแบบทดสอบ.....	90
3-9	รายละเอียดข้อมูลความรู้ทางไอซีที.....	90
3-10	รายละเอียดข้อมูลหมวดไอซีที.....	90
3-11	รายละเอียดข้อมูลการทดสอบ.....	90
3-12	รายละเอียดข้อมูลข้อสอบ.....	91
3-13	รายละเอียดการทดสอบ.....	91
3-14	ลำดับการทดสอบวิธี LB กับวิธี IE.....	106
3-15	ลำดับการทดสอบโปรแกรม MCAT กับโปรแกรม CT.....	110
4-1	เกณฑ์การประเมินระดับค่าความสามารถของผู้ทดสอบ วิธี LB และวิธี IE.....	123
4-2	เกณฑ์การประเมินระดับคะแนนของผู้ทดสอบ วิธี CT.....	123
4-3	ผลการประเมินประสิทธิภาพโปรแกรมในภาพรวม.....	125
4-4	ผลการประเมินประสิทธิภาพด้านตรงตามความต้องการของโปรแกรม.....	126
4-5	ผลการประเมินประสิทธิภาพด้านการทำงานของโปรแกรม.....	126
4-6	ผลการประเมินประสิทธิภาพด้านการใช้งานโปรแกรม.....	127
4-7	ผลการประเมินประสิทธิภาพด้านความปลอดภัยของโปรแกรม.....	127
4-8	ผลการประเมินประสิทธิภาพด้านการใช้งานโปรแกรมของผู้ทดสอบ.....	128
4-9	ค่าสถิติพื้นฐานด้านจำนวนข้อสอบที่ใช้ในการทดสอบ.....	129

สารบัญตาราง (ต่อ)

ตารางที่		หน้า
4-10	ผลการเปรียบเทียบประสิทธิภาพด้านจำนวนข้อสอบที่ใช้ในการทดสอบ.....	129
4-11	ค่าสถิติพื้นฐานด้านเวลาที่ใช้ในการทดสอบ.....	130
4-12	ผลการเปรียบเทียบประสิทธิภาพด้านเวลาที่ใช้ในการทดสอบ.....	130
4-13	ผลอัตราการแสดงการใช้ข้อสอบในคลังข้อสอบความรู้ทางไอซีที.....	131
4-14	ผลการวิเคราะห์ความแตกต่างของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) จำแนกตามวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา.....	131
4-15	สรุปผลการเปรียบเทียบประสิทธิภาพระหว่างวิธี LB กับวิธี IE.....	132
4-16	ผลการทดสอบ MCAT กับ CT.....	133
4-17	ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันระหว่างค่าประมาณความสามารถของผู้ทดสอบกับคะแนนรวมของผู้ทดสอบ.....	134

สารบัญภาพ

ภาพที่		หน้า
1-1	กรอบแนวทางการวิจัยเรื่อง การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง สำหรับทดสอบความรู้ทางไอซีที	9
2-1	ความสามารถในการทำงานของโหลดบาลานซ์ซิ่ง	15
2-2	ความแตกต่างของการทดสอบโปรแกรม	62
3-1	ลำดับการดำเนินการวิจัยแบ่งตามระยะ	66
3-2	ขั้นตอนการพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา	67
3-3	ตัวอย่างการวิเคราะห์เนื้อหาข้อสอบความรู้ทางไอซีทีโดยผู้เชี่ยวชาญ	68
3-4	อัลกอริทึมการสำรวจและเคลื่อนย้ายข้อมูล	71
3-5	อัลกอริทึมการคำนวณภาระงาน	72
3-6	อัลกอริทึมการควบคุมการทำงาน	74
3-7	วิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง	76
3-8	วงจรการพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา	82
3-9	แผนภาพบริบทการพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา	84
3-10	แผนภาพกระแสข้อมูลการพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา	85
3-11	แผนภาพกระแสข้อมูลระดับ 1 การจัดการทดสอบ	87
3-12	แบบจำลองฐานข้อมูลโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา	88
3-13	การเข้าสู่โปรแกรม	92
3-14	หน้าแรกหลังจากเข้าสู่โปรแกรม	92
3-15	การจัดการข้อมูลผู้ใช้โปรแกรม	93
3-16	การจัดการข้อมูลแบบทดสอบ	94
3-17	การจัดการข้อมูลหมวดไอซีที	94
3-18	การจัดการข้อมูลความรู้ทางไอซีที	95
3-19	การจัดการข้อมูลข้อสอบ	95

สารบัญภาพ (ต่อ)

ภาพที่ หน้า

3-20 รายงานผลการทดสอบ..... 96
3-21 การจัดการข้อมูลลงทะเบียนผู้ทดสอบ..... 97
3-22 การเข้าสู่โปรแกรมผู้ทดสอบ..... 98
3-23 การจัดการข้อมูลการทดสอบ..... 98
3-24 รายงานผลการทดสอบเมื่อทดสอบเสร็จ..... 99
3-25 รายงานผลการทดสอบแต่ละแบบ..... 99
3-26 ขั้นตอนการพัฒนาโปรแกรมตามแนวคิดการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์..... 100
3-27 ขั้นตอนการเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับวิธีการควบคุมสมดุลเนื้อหา..... 104
3-28 บัตรทดสอบ ICT-MCAT แบบ LB กับ IE..... 106
3-29 ขั้นตอนการศึกษาความสัมพันธ์ที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับการทดสอบด้วยคอมพิวเตอร์ (CT)..... 108
3-30 บัตรทดสอบ ICT-MCAT แบบ LB กับ CT..... 110
4-1 การเข้าสู่โปรแกรม..... 115
4-2 หน้าแรกหลังจากเข้าสู่โปรแกรม..... 116
4-3 การจัดการข้อมูลผู้ใช้โปรแกรม..... 117
4-4 การจัดการข้อมูลแบบทดสอบ..... 118
4-5 การจัดการข้อมูลด้านไอซีที..... 118
4-6 การจัดการข้อมูลความรู้ทางไอซีที..... 119
4-7 การจัดการข้อมูลข้อสอบ..... 119
4-8 รายงานผลการทดสอบ..... 120
4-9 การจัดการข้อมูลลงทะเบียนผู้ทดสอบ..... 121
4-10 การเข้าสู่โปรแกรมผู้ทดสอบ..... 122
4-11 การจัดการข้อมูลการทดสอบ..... 122
4-12 รายงานผลการทดสอบเมื่อทดสอบเสร็จ..... 124
4-13 รายงานผลการทดสอบแต่ละแบบ..... 124

บทที่ 1
บทนำ

ความเป็นมาและความสำคัญของปัญหา

การพัฒนาความรู้ทางเทคโนโลยีสารสนเทศและการสื่อสาร หรือความรู้ทางไอซีที (ICT Literacy) เป็นจุดมุ่งหมายสำคัญของการจัดการศึกษาในปัจจุบัน เพื่อให้ผู้เรียนมีความรู้เท่าทันที่จำเป็นสำหรับการเรียนรู้ตลอดชีวิต ซึ่งสอดคล้องกับเครือข่ายองค์กรความร่วมมือเพื่อทักษะการเรียนรู้ในศตวรรษที่ 21 (Partnership for 21st Century Skills) ที่ได้กำหนดแนวคิดการพัฒนาทักษะแห่งศตวรรษที่ 21 โดยกำหนดให้นักศึกษาต้องมีความรู้ทางไอซีที เนื่องจากเป็นความรู้ที่ผู้เรียนควรได้รับการพัฒนาให้ประสบความสำเร็จในการใช้ชีวิต ความรู้นี้จัดอยู่ในกลุ่มของทักษะด้านสารสนเทศ สื่อ และเทคโนโลยี (Information, Media, and Technology Skill) เป็นทักษะที่จำเป็นต่อมนุษย์ในศตวรรษที่ 21 ที่ทุกคนจำเป็นต้องได้รับการพัฒนาทักษะเพราะโลกจะยิ่งเปลี่ยนแปลงเร็วขึ้นเรื่อย ๆ และมีความซับซ้อนมากขึ้น คนที่อ่อนแอในทักษะด้านการเรียนรู้และนวัตกรรมจะเป็นคนที่ตามโลกไม่ทัน เป็นคนอ่อนแอ ชีวิตก็จะยากลำบาก (วิจารณ์ พานิช, 2555, หน้า 16-21)

ความรู้ทางไอซีที เป็นความรู้และความสามารถของแต่ละบุคคลในการใช้เทคโนโลยีดิจิทัลและเครื่องมือการสื่อสาร ซึ่งมีสถาบันสร้างแบบทดสอบความรู้ทางไอซีทีนำเสนอกรอบความรู้ทางไอซีที เช่น สถาบัน Education Testing Service (ETS) (Educational Testing Service, 2007, p. 3) นำเสนอว่า บุคคลต้องมีความรู้ทางไอซีที ประกอบด้วย การเข้าถึง (Access) การจัดการ (Manage) การเชื่อมโยง (Integrate) การประเมิน (Evaluate) การจัดลำดับเนื้อหา (Construct) และการสื่อสาร (Communication) ต่อมามีการพัฒนากรอบความรู้ทางไอซีทีใหม่คือ NAP-ICT Literacy Assessment Framework Processes (NAP-ICT Literacy, 2014, p. 11) นำเสนอว่า บุคคลต้องมีความรู้ทางไอซีที 6 ด้าน ได้แก่ 1) การเข้าถึงสารสนเทศ (Accessing information) เป็นความสามารถระบุความต้องการที่จะค้นหาและสามารถเข้าถึงสารสนเทศและเรียกใช้สารสนเทศได้ 2) การจัดการสารสนเทศ (Managing information) เป็นความสามารถจัดระเบียนและจัดเก็บข้อมูลสารสนเทศในการเรียนและนำมาใช้ได้ 3) การประเมิน (Evaluating) เป็นการสะท้อนให้เห็นถึงกระบวนการที่ใช้ในการออกแบบและสร้างทางเลือกในการใช้ไอซีทีและการตัดสินเกี่ยวกับความซื่อสัตย์ ความเกี่ยวข้อง และประโยชน์ของสารสนเทศ 4) การสร้างสิ่งใหม่จากความเข้าใจ (Developing new understandings) เป็นการสร้างข้อมูลและความรู้โดยการสังเคราะห์ ปรับประยุกต์ใช้ การออกแบบ การประดิษฐ์ หรือการเขียน 5) การสื่อสาร (Communicating) เป็นการแลกเปลี่ยนข้อมูลด้วยการแบ่งปันความรู้และข้อมูลการสร้างผลิตภัณฑ์เพื่อให้เหมาะกับบริบทของผู้ใช้

และ 6) การใช้ไอซีทีอย่างเหมาะสม (Using ICT appropriately) เป็นการใช้ความคิดอย่างมีวิจารณญาณในการตัดสินใจใช้ไอซีที ที่มีความรับผิดชอบด้วยการพิจารณาข้อกฎหมายทางสังคมและปัญหาทางจริยธรรม ดังนั้นการมีความรู้ทางไอซีที จึงเป็นการผสมผสานทักษะการอ่าน การรู้สารสนเทศ การรู้เรื่องคอมพิวเตอร์เข้าด้วยกัน โดยใช้เทคโนโลยีสารสนเทศเป็นเครื่องมือ

สำหรับการทดสอบความรู้ทางด้านไอซีที ปัจจุบันมีนักวิชาการและหน่วยงานทางภาครัฐและเอกชนทั้งในประเทศและต่างประเทศให้ความสนใจเกี่ยวกับการสร้างเครื่องมือทดสอบความรู้ดังกล่าว สถาบันที่ทำวิจัยทางด้านความรู้เทคโนโลยีสารสนเทศและการสื่อการ Educational Testing Service (ETS) เป็นสถาบันที่พัฒนาแบบทดสอบทักษะด้านต่าง ๆ และหนึ่งในนั้นคือการพัฒนาการทดสอบความรู้ทางไอซีที ตั้งแต่ปี ค.ศ. 2001 และได้พัฒนาขึ้นเรื่อย ๆ มีลักษณะการทำข้อสอบเป็นแบบออนไลน์ (Online test) ชื่อแบบทดสอบ “iSkills Assessment” ส่วนประเทศไทยมีนักวิจัยและสถาบันที่สร้างแบบทดสอบเพื่อทดสอบความรู้ด้านเทคโนโลยีสารสนเทศ ตัวอย่างเช่น กระทรวงวิทยาศาสตร์และเทคโนโลยี ได้ทำแบบทดสอบความรู้ทางไอซีทีเพื่อใช้ทดสอบความรู้ของบุคลากรในหน่วยงาน สอดคล้องตามแผนแม่บทเทคโนโลยีสารสนเทศและการสื่อสาร พ.ศ. 2558 - 2562 ในปี พ.ศ. 2546 สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) หรือ สทศ. ได้จัดการทดสอบทางการศึกษาระดับชาติด้านอาชีวศึกษา (Vocational National Educational Test: V-NET) เพื่อวัดความรู้และทักษะของผู้เรียนระดับประกาศนียบัตรวิชาชีพชั้นสูง มีเนื้อหาหลักคือการประยุกต์ใช้คอมพิวเตอร์และอุปกรณ์โทรคมนาคม การสืบค้นและสื่อสารข้อมูลสารสนเทศเพื่อการจัดการ แต่ข้อสอบยังอยู่ในรูปแบบกระดาษ-ดินสอ (Paper and Pencil test: P&P) ดังนั้นในการเตรียมความพร้อมต่อการสอบวัดมาตรฐานวิชาชีพ ผู้วิจัยจึงนำเนื้อหาที่ใช้ในการออกข้อสอบมาวิเคราะห์พร้อมจัดกลุ่มความรู้ทางไอซีทีเพื่อนำไปสร้างคลังข้อสอบ

ปัจจุบัน มีการพัฒนารูปแบบการทดสอบโดยใช้เทคโนโลยีเข้ามาประยุกต์เพื่อสร้างแบบทดสอบรูปแบบต่าง ๆ ดังเช่น Thompson and Weiss (2011) ได้เสนอรูปแบบการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (Computerized Adaptive Testing: CAT) ซึ่งเป็นการทดสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบ วิธีนี้สามารถประมาณค่าคุณลักษณะแฝงหรือค่าความสามารถของผู้ทดสอบ (Proficiency) ด้วยข้อมูลที่ได้จากการตอบข้อสอบเป็นรายข้อ เช่น การวินิจฉัยทักษะการอ่านภาษาอังกฤษ ข้อมูลที่ได้รับจากการตอบข้อสอบสามารถนำไปวิเคราะห์เพื่อประมาณค่าความสามารถของผู้ทดสอบเพื่อหาระดับความสามารถในการอ่านภาษาอังกฤษ และเลือกข้อสอบข้อถัดไปที่มีความเหมาะสมกับความสามารถของผู้ทดสอบมากที่สุด ดังนั้น ในกระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ผู้ทดสอบแต่ละคนจึงได้รับข้อสอบที่แตกต่างกัน ซึ่งให้ข้อมูลสารสนเทศของข้อสอบสอดคล้องกับระดับความสามารถของผู้ทดสอบนั้นมากที่สุด ส่งผลให้จำนวนข้อสอบที่ใช้กับผู้ทดสอบแต่ละคนมีจำนวนลดลง ขณะที่ประสิทธิภาพในการทดสอบเพิ่มมากขึ้น เนื่องจากความ

แม่นยำในการประมาณค่าความสามารถ และสามารถยุติการทดสอบด้วยค่าความคลาดเคลื่อนมาตรฐาน (Standard Error of Estimation: SEE) ดังนั้นการทดสอบปรับเหมาะด้วยคอมพิวเตอร์สามารถช่วยลดความเหนื่อยล้าของผู้ทดสอบได้ (Gibbons et al., 2008)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ จึงเป็นทางเลือกหนึ่งที่ช่วยเพิ่มประสิทธิภาพในการทดสอบความรู้ทางไอซีที แต่ถ้านำการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT) มาใช้อาจจะเกิดข้อจำกัดของการสรุปความรู้ทางไอซีที เนื่องจากการทดสอบแบบปรับเหมาะส่วนใหญ่ ใช้คลังข้อสอบด้วยทฤษฎีโมเดลการตอบสนองข้อสอบแบบเอกมิติ (Unidimensional Item Response Theory Model: UIRT) ซึ่งมีข้อตกลงเบื้องต้นคือ คุณลักษณะแฝงที่จะทำการวัดไม่มีความสัมพันธ์กัน มี 2 ลักษณะคือ 1) ความเป็นเอกมิติรวม (Composite approach) ซึ่งมีลักษณะของข้อสอบวัดคุณลักษณะแฝงเดียว และ 2) ความเป็นเอกมิติแยกตามมิติ (Consecutive approach) มีลักษณะของข้อสอบวัดคุณลักษณะแฝงเดี่ยว แต่มีหลายคุณลักษณะแฝง ทั้งนี้แต่ละคุณลักษณะแฝงไม่มีความสัมพันธ์กัน (Frey & Seitz, 2009, pp. 89-94) แต่การทดสอบความรู้ทางไอซีทีมีคุณลักษณะแฝงแบบ 2 มิติ (ด้าน) ที่มีความสัมพันธ์กัน ได้แก่ มิติที่ 1 หมวดเนื้อหา และมิติที่ 2 ความรู้ทางไอซีทีในด้านมิติหมวดเนื้อหา เป็นรายวิชาเทคโนโลยีสารสนเทศและการสื่อสาร ระดับประกาศนียบัตรวิชาชีพชั้นสูง ประกอบด้วย 5 ด้าน ได้แก่ 1) เครือข่าย 2) ข้อมูล 3) ซอฟต์แวร์ 4) ฮาร์ดแวร์ และ 5) กระบวนการและผู้ใช้ ส่วนมิติความรู้ทางไอซีที ประกอบด้วย 6 ด้าน (NAP-ICT, 2014, pp. 7-15) ดังนั้นการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT) ด้วยโมเดลการตอบสนองข้อสอบแบบเอกมิติ (UIRT) ไม่เข้าเงื่อนไขข้อตกลงเบื้องต้น ด้วยปัญหาดังกล่าวจึงมีนักวัดผลได้ขยายแนวคิดของทฤษฎีการตอบสนองข้อสอบแบบเอกมิติ สู่โมเดลการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory Model: MIRT) (Reckase, 2009, p. 57) แล้วพัฒนาการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (Multidimensional Computerized Adaptive Testing: MCAT) (Yao, 2014)

การพัฒนา MCAT มีขั้นตอนในการดำเนินการแบบเดียวกับ CAT (Thompson & Sireci, 2011) ซึ่งมี 5 ขั้นตอนเพื่อเป็นกรอบในการพัฒนา ได้แก่ 1) การสร้างคลังข้อสอบ (Develop item bank) เป็นขั้นตอนในการสร้างข้อสอบความรู้ทางไอซีทีเพื่อใช้ในการทดสอบ ใช้เกณฑ์การตอบสนองข้อสอบแบบพหุมิติสำหรับวิเคราะห์คุณภาพข้อสอบ และเกณฑ์การคัดเลือกข้อสอบสำหรับเลือกข้อสอบเข้าคลังข้อสอบความรู้ทางไอซีที 2) การคัดเลือกข้อสอบข้อแรก (First item selection) เป็นขั้นตอนการเลือกข้อสอบความรู้ทางไอซีที่มีความยากระดับปานกลางให้ผู้ทดสอบ แล้วนำผลจากการทดสอบไปประมาณค่าความสามารถของผู้ทดสอบเพื่อใช้ในการเลือกข้อสอบข้อถัดไป 3) การคัดเลือกข้อสอบข้อถัดไป (Next item selection) เป็นขั้นตอนการคัดเลือกข้อสอบที่มีความสอดคล้องหรือเหมาะสมกับระดับความสามารถของผู้ทดสอบ 4) การประมาณค่าความสามารถของผู้ทดสอบ

(Calculate possible ability level) เป็นขั้นตอนการประมาณความสามารถของผู้ทดสอบ เพื่อนำค่าประมาณความสามารถไปเทียบกับการคัดเลือกข้อสอบข้อถัดไป และ 5) เกณฑ์ยุติการทดสอบ (Termination criterion) เป็นการกำหนดเกณฑ์สิ้นสุดของการทดสอบ เมื่อค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถน้อยกว่าหรือเท่ากับ 0.30 มีความตรงเชิงสภาพสูงสุด

การพัฒนา MCAT ขั้นตอนที่ 3 การคัดเลือกข้อสอบข้อถัดไปนับเป็นขั้นตอนที่สำคัญ เพราะการคัดเลือกข้อสอบข้อถัดไปให้มีความเหมาะสมกับความสามารถของผู้ทดสอบ จะช่วยลดจำนวนข้อสอบที่ใช้ในการทดสอบ พร้อมทั้งเพิ่มประสิทธิภาพการประมาณค่าความสามารถของผู้ทดสอบ (Luecht & Sireci, 2011) แต่จากการศึกษางานวิจัย ได้พบปัญหาการเลือกข้อสอบข้อถัดไปจะคัดเลือกข้อสอบที่มีค่าสารสนเทศของข้อสอบสูงสุด ณ ระดับความสามารถของผู้ทดสอบขณะนั้น ซึ่งข้อสอบที่มีค่าสารสนเทศของข้อสอบสูงคือข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง และมีค่าความยากของข้อสอบใกล้เคียงกับผู้ทดสอบจะถูกเลือกใช้ในการทดสอบบ่อยครั้งเกินไป (Overexposed item) ส่งผลให้ผู้ทดสอบรุ่นถัดมาสามารถคาดการณ์ข้อสอบล่วงหน้าได้ (Chang & van der Linden, 2003, pp. 107-120) จากปัญหาดังกล่าว จึงมีนักวิจัยได้เสนอวิธีการควบคุมการใช้ข้อสอบ (Item exposure control) โดย Georgiadou and Triantafillou (2007) ได้จำแนกวิธีการควบคุมการใช้ข้อสอบ แบ่งออกเป็น 5 กลุ่ม ได้แก่ 1) วิธีการคัดเลือกแบบสุ่ม (Randomization strategies) จะคัดเลือกข้อสอบที่เหมาะสม 1 กลุ่ม จากนั้นสุ่มเลือกข้อสอบข้อใดข้อหนึ่งจากกลุ่มข้อสอบนี้มาให้ผู้ทดสอบใช้ทดสอบ จึงเป็นวิธีการค่อนข้างง่าย ไม่ซับซ้อน แต่ไม่รับรองว่าจะสามารถควบคุมการใช้ข้อสอบได้ตามต้องการ 2) วิธีการคัดเลือกแบบมีเงื่อนไข (Conditional selection strategies) จะควบคุมความน่าจะเป็นของข้อสอบที่ได้รับเลือกแล้วให้มีการใช้ข้อสอบเป็นไปตามที่กำหนด วิธีนี้รับรองว่าอัตราการใช้ข้อสอบจะไม่สูงเกินกว่าค่าที่กำหนด โดยควบคุมอัตราการใช้ข้อสอบผ่านค่าพารามิเตอร์ควบคุมการใช้ข้อสอบ ข้อเสียคือมีความยุ่งยากซับซ้อนมาก 3) วิธีคัดเลือกแบบตามระดับชั้น (Stratified strategies) จะแบ่งข้อสอบในคลังข้อสอบเป็นชั้นตามค่าอำนาจจำแนกของข้อสอบ และควบคุมการใช้ข้อสอบโดยให้เลือกข้อสอบจากทุกชั้นมาใช้ในการทดสอบ วิธีนี้จะเพิ่มโอกาสให้ข้อสอบที่มีค่าสารสนเทศต่ำมีโอกาสถูกเลือกมาใช้สอบเพิ่มมากขึ้น 4) วิธีคัดเลือกแบบผสม (Combined strategies) เป็นการรวมวิธีการควบคุมการใช้ข้อสอบ 2 วิธี ที่นำจุดเด่นของแต่ละวิธีมาผสานการทำงานร่วมกัน และ 5) วิธีคัดเลือกแบบปรับเหมาะหลายขั้นตอน (Multiple stage adaptive test designs) จะดำเนินการร่วมกันระหว่างการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับการทดสอบแบบเขียนตอบบนกระดาษ โดยใช้ผลการทดสอบแบบเขียนตอบบนกระดาษมาเป็นข้อมูลในการคัดเลือกข้อสอบข้อถัดไป แล้วหาความสามารถของผู้ทดสอบด้วยการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ วิธีการประเภทนี้ เช่น Multiple forms structures (Armstrong, Jones, Koppel, & Pashley, 2004, pp. 147-164) ได้ควบคุมการใช้ข้อสอบใน CAT

วิธีการควบคุมการใช้ข้อสอบได้มีการขยายขีดความสามารถในการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ เช่น วิธี Item Eligibility (IE) (van der Linden & Veldkamp, 2004) เป็นวิธีการควบคุมการใช้ข้อสอบที่ขยายขีดความสามารถมาจาก Sympson - Hetter method (SH) (Hetter & Sympson, 1997) เพื่อแก้ไขปัญหาความล่าช้าของเวลาในการนำข้อสอบมาแสดง วิธีการของ IE จะคัดเลือกข้อสอบที่มีโอกาสได้รับเลือกก่อน แล้วค่อยส่งข้อสอบดังกล่าวไปยังวิธี SH ซึ่งจะลดความล่าช้าของเวลาได้ ดังนั้น วิธี IE มีขั้นตอนที่ยุ่งยาก มีการกำหนดอัตราการใช้ข้อสอบสูงสุดแบบคงที่ r^{max} และสัดส่วนของเนื้อหาอาจไม่สมดุลกันในแต่ละด้าน เมื่อทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ โปรแกรมอาจมีโอกาสที่จะดึงข้อสอบแต่ละด้านมาไม่ครบหรือเน้นหนักไปด้านใดด้านหนึ่งหรือข้อสอบซ้ำกัน จากการศึกษางานวิจัย ปรากฏว่า วิธีการควบคุมความสมดุลเนื้อหาที่ใช้สอบแต่ละองค์ประกอบให้เท่าเทียมกัน (Content balance) มีเทคนิคการจัดสมดุลเนื้อหาที่ได้รับความนิยม 3 วิธี ได้แก่ 1) วิธี Constrained CAT (CCAT) เป็นวิธีเลือกข้อสอบที่เหมาะสมที่สุดจากขอบเขตเนื้อหา โดยพิจารณาอัตราการแสดงข้อสอบครั้งหลังสุด และห่างจากเปอร์เซ็นต์ของการจัดข้อสอบตามกำหนดมากที่สุด (Kingsbury & Zara, 1989, pp. 359-375) 2) วิธี Modified Multinomial Model (MMM) เป็นวิธีที่ให้เนื้อหามีความสมดุลกัน เมื่อขอบเขตเนื้อหาไปถึงเปอร์เซ็นต์กำหนด จะมีการสร้างการแจกแจงมัลติโนเมียลเพื่อปรับเปอร์เซ็นต์ข้อสอบที่ยังไม่ถูกนำไปใช้ให้เต็มขอบเขตเนื้อหาที่เหลืออยู่ (Chen, Ankenmann, & Spray, 2003, pp. 129-145) และ 3) Modified Constrained CAT (MCCAT) เป็นวิธีประยุกต์มาจากการทดสอบแบบปรับเหมาะควบคุมสมดุลเนื้อหา เพื่อลดการคาดเดาการเรียงเนื้อหาและให้การควบคุมสมดุลเนื้อหาให้ดียิ่งขึ้น ข้อสอบเลือกมาจากขอบเขตเนื้อหาทั้งหมดที่ยังไม่ถูกใช้ เพื่อขจัดผลจากการเรียงลำดับ (Leung, Chang, & Hau, 2000) ซึ่งงานวิจัยส่วนใหญ่แยกระหว่างวิธีการควบคุมการใช้ข้อสอบกับวิธีการควบคุมสมดุลเนื้อหาออกจากกัน

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวกับวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาของการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ ปรากฏว่า วิธีดังกล่าวเป็นการแทรกการทำงานระหว่างการเลือกข้อสอบข้อถัดไปกับคลังข้อสอบ และยังไม่มีการนำวิธีโหลดบาลานซ์ซิ่ง (Load Balancing: LB) มาเป็นวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา การประยุกต์การทำงานของอัลกอริทึมสำหรับการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา 3 ขั้นตอน (Singh, Juneja, & Malhotra, 2015) ได้แก่ 1) การสำรวจและเคลื่อนย้ายข้อมูล เป็นการค้นหาพร้อมทั้งการจัดเก็บลักษณะของข้อสอบ 2) การคำนวณภาระงาน เป็นการคำนวณค่าเฉลี่ยถ่วงน้ำหนักของการใช้ข้อสอบในแต่ละด้านเพื่อควบคุมความสมดุลเนื้อหาแล้วนำค่าเฉลี่ยถ่วงน้ำหนักมากที่สุดในด้านนั้น มาเป็นตัวแทนของด้านเนื้อหา และ 3) การควบคุมการทำงานเป็นการควบคุมการใช้ข้อสอบโดยการเรียงลำดับจากน้อยไปหามากระหว่างผลต่างของค่าประมาณ

ความสามารถกับค่าสารสนเทศสูงสุดของข้อสอบและความถี่สะสมของการใช้ข้อสอบ อัลกอริทึมคัดเลือกข้อสอบข้อแรกจากการเรียงลำดับเพื่อนำมาเป็นข้อสอบข้อถัดไป

จากเหตุผลดังกล่าวข้างต้น ทำให้ผู้วิจัยสนใจในการพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่งซึ่งเป็นวิธีใหม่ แล้วนำวิธีที่พัฒนาขึ้นไปพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ เพื่อทดสอบความรู้ทางไอซีทีของนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง พร้อมทั้งเปรียบเทียบประสิทธิภาพวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาที่พัฒนาขึ้น ในด้านจำนวนข้อสอบที่ใช้ในการทดสอบ (Test length) เวลาที่ใช้ในการทดสอบ (Test time) (Chang & Ying, 1999, pp. 211-222) อัตราการแสดงการใช้ข้อสอบสูงสุด (Maximum Item Exposure Rate: MIER) (Ozturk & Dogan, 2015) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) (Zheng, Chang, & Chang, 2013, pp. 491-499, Ozturk & Dogan, 2015) ที่มีค่าความคลาดเคลื่อนต่ำสุดเข้าใกล้ 0 และศึกษาความสัมพันธ์ของคะแนนที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (Multidimensional Computerized Adaptive Testing: MCAT) กับการทดสอบด้วยคอมพิวเตอร์ (Computer Test: CT) คือ ค่าความสามารถของผู้ทดสอบกันค่าคะแนนรวม ดังนั้นวิธีโหลดบาลานซ์ซิ่งเป็นวิธีที่ช่วยควบคุมการใช้ข้อสอบเพื่อลดจำนวนข้อสอบ ลดเวลาที่ใช้ในการทดสอบ ข้อสอบที่ใช้แล้วไม่ถูกนำกลับมาใช้อีก และเหมาะกับผู้ที่สนใจนำไปพัฒนา พร้อมทั้งวิธีการนี้ยังสามารถควบคุมสมดุลเนื้อหาให้ผู้ทดสอบได้รับเนื้อหาของแบบทดสอบครบทุกด้านอย่างเท่าเทียม

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง (LB)

2. เพื่อพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา 2 วิธี คือ 1) วิธีโหลดบาลานซ์ซิ่ง (LB) และ 2) วิธี Item Eligibility (IE)

3. เพื่อเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา 2 วิธี คือ 1) วิธีโหลดบาลานซ์ซิ่ง (LB) และ 2) วิธี Item Eligibility (IE) ในด้าน 1) จำนวนข้อสอบที่ใช้ในการทดสอบ 2) เวลาที่ใช้ในการทดสอบ 3) อัตราการแสดงการใช้ข้อสอบสูงสุด และ 4) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย

4. เพื่อศึกษาความสัมพันธ์ระหว่างค่าความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT)

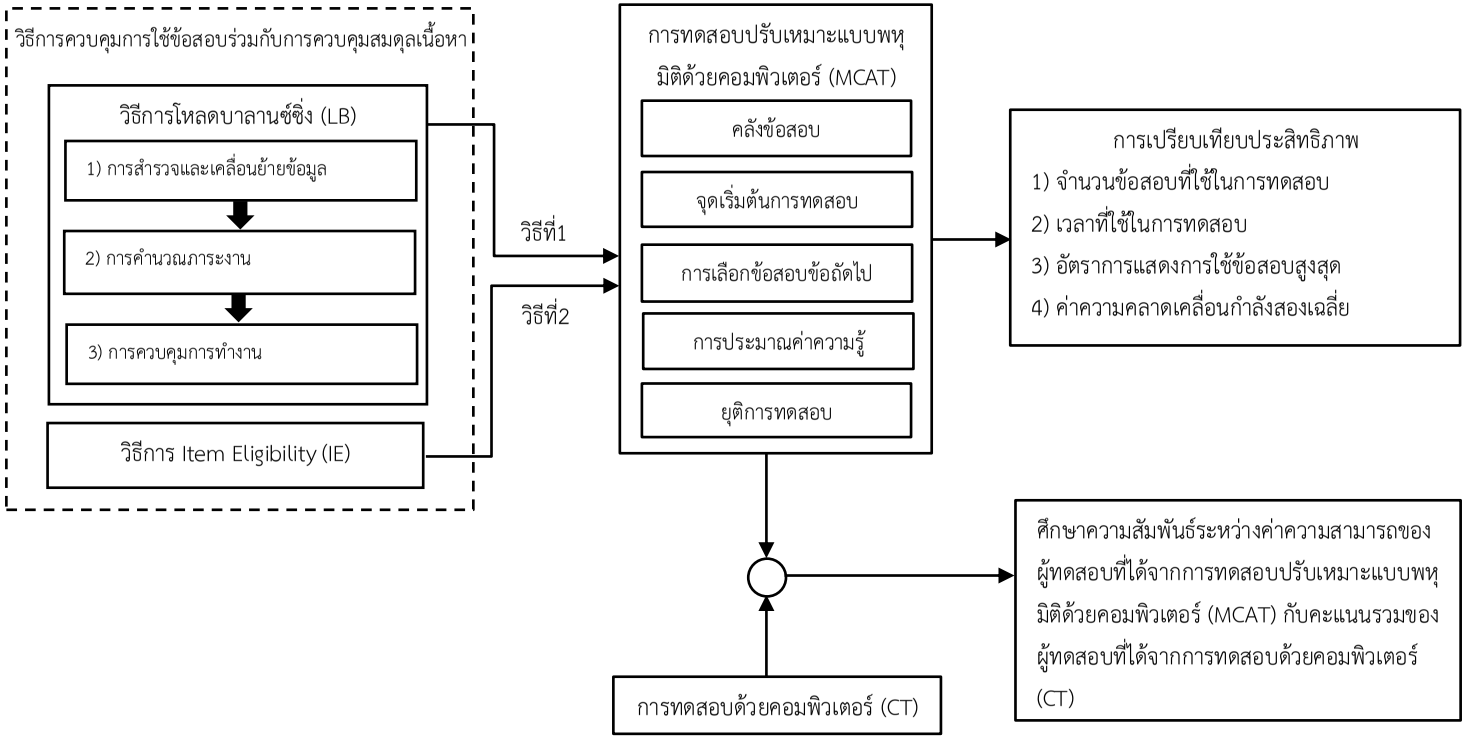
กรอบแนวทางการวิจัย

ในการพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง สำหรับทดสอบความรู้ทางไอซีทีของการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ มีแนวทางการวิจัยดังนี้

ผู้วิจัยได้ศึกษาวิธีโหลดบาลานซ์ซิ่ง (LB) มาประยุกต์ในวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาพร้อมกัน เป็นกระบวนการแทรกระหว่างวิธีการคัดเลือกข้อสอบข้อถัดไปกับคลังข้อสอบ ซึ่งวิธีโหลดบาลานซ์ซิ่ง มี 3 ขั้นตอน ได้แก่ 1) การสำรวจและเคลื่อนย้ายข้อมูล เป็นการค้นหาพร้อมทั้งการจัดเก็บคุณลักษณะของข้อสอบ 2) การคำนวณภาระงาน เป็นการคำนวณค่าเฉลี่ยถ่วงน้ำหนักของการใช้ข้อสอบในแต่ละด้านเพื่อควบคุมความสมดุลเนื้อหา และ 3) การควบคุมการทำงาน เป็นการควบคุมการใช้ข้อสอบโดยการเรียงลำดับความถี่ของการใช้ข้อสอบจากน้อยไปมาก แล้วใช้ฟังก์ชันค่าสารสนเทศสูงสุด (MIC) คำนวณข้อสอบทุกข้อ แล้วนำมาประมาณค่าใกล้เคียงกับความสามารถของผู้ทดสอบ (ML) เพื่อเป็นข้อสอบข้อถัดไป

วิธีการควบคุมข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่งที่พัฒนาขึ้น นำไปพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ ใช้กรอบการพัฒนาทฤษฎีวงจรการพัฒนาโปรแกรม (System Development Life Cycle: SDLC) (Singh et al., 2015) ซึ่งอาศัยหลักการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (Thompson & Weiss, 2011) เป็นขั้นตอนการสร้างกระบวนการทดสอบ ทางด้านการสร้างคลังข้อสอบ เป็นการนำข้อสอบวิชาเทคโนโลยีสารสนเทศและการสื่อสาร ระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) ที่ได้มาจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) มาวิเคราะห์แบ่งเป็น 2 มิติ ได้แก่ 1) ด้านหมวดเนื้อหา ประกอบไปด้วย 5 หมวด คือ 1.1) เครือข่าย 1.2) ข้อมูล 1.3) ซอฟต์แวร์ 1.4) ฮาร์ดแวร์ และ 1.5) กระบวนการและผู้ใช้ และ 2) ด้านความรู้ทางไอซีที ประกอบไปด้วย 6 ด้าน คือ 2.1) การเข้าถึงสารสนเทศ 2.2) การจัดการสารสนเทศ 2.3) การประเมิน 2.4) การสร้างสิ่งใหม่จากความเข้าใจ 2.5) การสื่อสาร และ 2.6) การใช้ไอซีที่อย่างเหมาะสม และใช้การวิเคราะห์ข้อสอบโมเดลการตอบสนองข้อสอบแบบพหุมิติ (MIRT) 3 พารามิเตอร์ ได้แก่ ค่าอำนาจจำแนกของข้อสอบ (a) ค่าความยากของข้อสอบ (b) และค่าโอกาสการเดาของข้อสอบ (c) (Reckase, 2009, p. 102) คัดเลือกข้อสอบที่มีคุณภาพ โดยมีค่าความยากของข้อสอบตั้งแต่ -2.50 ถึง 2.50 ค่าอำนาจจำแนกของข้อสอบตั้งแต่ 0.50 ถึง 2.50 และค่าโอกาสการเดาของข้อสอบไม่เกิน 0.30 เมื่อได้ข้อสอบที่มีคุณภาพจึงนำเข้าไปบันทึกไว้ในคลังข้อสอบ ส่วนโปรแกรมการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาจะพัฒนา 2 วิธี คือ 1) วิธีโหลดบาลานซ์ซิ่ง (LB) เป็นวิธีใหม่ และ 2) วิธี Item Eligibility (IE) เป็นวิธีเก่า ซึ่งทั้ง 2 วิธีนี้อยู่ภายใต้การทำงานของวิธีการเลือกข้อสอบ MIC

การตรวจสอบประสิทธิภาพวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาระหว่าง 2 วิธี ใน 5 ด้าน ได้แก่ 1) จำนวนข้อสอบที่ใช้ในการทดสอบ 2) เวลาที่ใช้ในการทดสอบ 3) อัตราการแสดงการใช้ข้อสอบสูงสุด และ 4) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย จากนั้น นำไปใช้ในการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับคะแนนรวมที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT) แสดงดังภาพที่ 1-1



ภาพที่ 1-1 กรอบแนวทางการวิจัยเรื่อง การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่งสำหรับทดสอบความรู้ทางไอซีที

สมมติฐานของการวิจัย

1. จำนวนข้อสอบที่ใช้ในการทดสอบของวิธี LB น้อยกว่า วิธี IE

2. เวลาที่ใช้ในการทดสอบของวิธี LB น้อยกว่า วิธี IE

3. อัตราการแสดงการใช้ข้อสอบสูงสุดของวิธี LB น้อยกว่า วิธี IE

4. ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธี LB น้อยกว่า วิธี IE

5. ค่าความสามารถของผู้ทดสอบที่ได้จาก MCAT กับค่าคะแนนรวมที่ได้จาก CT มีความสัมพันธ์ทางบวก

ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. ได้วิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง สำหรับนำไปใช้กับการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ เพื่อลดจำนวนข้อสอบที่ใช้ในการทดสอบ ลดเวลาที่ใช้ในการทดสอบ ข้อสอบที่ถูกนำไปใช้ทดสอบแล้วไม่นำกลับมาใช้ทดสอบอีก และลดซ้ำซ้อนในการนำไปพัฒนาโปรแกรม

2. ได้โปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยใช้วิธีโหลดบาลานซ์ซิ่งในการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา สำหรับทดสอบความรู้ทางไอซีที สามารถนำไปใช้ทดสอบนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง

3. ได้คลังข้อสอบสำหรับการทดสอบความรู้ทางไอซีที ซึ่งประกอบด้วยข้อสอบที่มีคุณภาพสำหรับนำไปใช้ทดสอบความรู้ทางไอซีทีนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง

ขอบเขตของการวิจัย

การวิจัยนี้มุ่งเน้นพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง แล้วนำมาพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ สำหรับทดสอบความรู้ทางไอซีที ในการพัฒนาโปรแกรมฯ แบ่งการพัฒนา 2 วิธี คือ 1) วิธี LB และ 2) วิธี IE แล้วนำ 2 วิธีมาเปรียบเทียบประสิทธิภาพ ขอบเขตในการวิจัยมีดังนี้

1. ด้านเนื้อหา เป็นวิชาเทคโนโลยีสารสนเทศและการสื่อสาร (ICT) ที่ใช้ในการเรียนการสอนระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) ซึ่งเป็นข้อมูลแบบทุติยภูมิ (Secondary data) ที่ได้มาจากการทดสอบมาตรฐานวิชาชีพ (V-NET) จากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ระหว่างปี พ.ศ.2556-2560 เนื้อหา 5 หมวด ได้แก่ 1) เครือข่าย 2) ข้อมูล 3) ซอฟต์แวร์ 4) ฮาร์ดแวร์ 5) กระบวนการและผู้ใช้ เป็นด้านที่ 1 หรือมิติที่ 1 แล้วกำหนดความรู้ทางไอซีที 6 ด้าน ได้แก่ 1) การเข้าถึงสารสนเทศ 2) การจัดการสารสนเทศ 3) การประเมิน 4) การสร้างสิ่งใหม่จากความเข้าใจ 5) การสื่อสาร และ 6) การใช้ไอซีที่อย่างเหมาะสม เป็นด้านที่ 2 หรือมิติที่ 2

2. ด้านประชากร คือนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) วิทยาลัยเทคโนโลยีภาคตะวันออก (อี.เทค) ที่ลงทะเบียนเรียนวิชาเทคโนโลยีสารสนเทศและการสื่อสาร ประจำปีการศึกษา 2561

3. ด้านตัวแปร แบ่งออกเป็น 2 ส่วน ดังนี้

ส่วนที่ 1 ตัวแปรของวัตถุประสงค์การวิจัยข้อที่ 3 มีดังนี้

ตัวแปรอิสระ เป็นวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา 2 วิธี ได้แก่ 1) วิธีโหลดบาลานซ์ซิ่ง (LB) และ 2) วิธี Item Eligibility (IE)

ตัวแปรตาม เป็นประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา มี 4 ด้าน ดังนี้

1) จำนวนข้อสอบที่ใช้ในการทดสอบ (หน่วยวัดเป็นจำนวนข้อ)

2) เวลาที่ใช้ในการทดสอบ (หน่วยวัดเป็นวินาที)

3) อัตราการแสดงการใช้ข้อสอบสูงสุด (หน่วยวัดเป็นจำนวนข้อ)

4) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (หน่วยวัดเป็นคะแนนมาตรฐาน)

ส่วนที่ 2 ตัวแปรของวัตถุประสงค์การวิจัยข้อที่ 4 มีดังนี้

1) ค่าความสามารถของผู้ทดสอบที่ได้จาก MCAT (หน่วยวัดเป็นคะแนนมาตรฐาน)

2) ค่าคะแนนรวมที่ได้จาก CT (หน่วยวัดเป็นคะแนนดิบ)

นิยามศัพท์เฉพาะ

การควบคุมการใช้ข้อสอบ (Item exposure control) หมายถึง การจัดสรรข้อสอบให้กับผู้ทดสอบ เพื่อไม่ให้ข้อสอบข้อหนึ่งใดถูกนำไปใช้ทดสอบมากเกินไป และนำกลับมาใช้ซ้ำของผู้ทดสอบคนเดียวกัน

การควบคุมสมดุลเนื้อหา (Content balance) หมายถึง การจัดสรรข้อสอบให้กับผู้ทดสอบให้มีเนื้อหาครอบคลุมทุกด้าน เพื่อไม่ให้ข้อสอบถูกนำไปใช้ทดสอบด้านหนึ่งด้านใดมากเกินไป หรือข้อสอบบางด้านไม่ถูกนำไปใช้ทดสอบ

วิธีโหลดบาลานซ์ซิ่ง (Load Balancing: LB) หมายถึง วิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา มีการทำงานหลัก 3 ขั้นตอน ได้แก่ 1) การสำรวจและเคลื่อนย้ายข้อมูล เป็นการค้นหาพร้อมทั้งการจัดเก็บคุณลักษณะของข้อสอบ 2) การคำนวณภาระงาน เป็นการควบคุมสมดุลเนื้อหาโดยการคำนวณค่าเฉลี่ยถ่วงน้ำหนักของการใช้ข้อสอบในแต่ละด้าน และ 3) การควบคุมการทำงาน เป็นการควบคุมการใช้ข้อสอบโดยการเรียงลำดับจากน้อยไปหามากแล้วเลือกข้อที่มีลำดับแรกเป็นข้อสอบถัดไป โดยเรียงลำดับระหว่างผลต่างของค่าประมาณความสามารถกับค่าสารสนเทศสูงสุดและค่าความถี่สะสม

วิธี Item Eligibility (IE) หมายถึง วิธีการควบคุมการใช้ข้อสอบ โดยใช้เงื่อนในการกำหนดอัตราการใช้ข้อสอบสูงสุดแบบคงที่ r^{max} แล้วนำข้อที่มีค่าอัตราการใช้ข้อสอบที่น้อยกว่า r^{max} มาเปรียบเทียบผลต่างน้อยที่สุดของค่าสารสนเทศสูงมาเป็นข้อสอบข้อถัดไป

ความรู้ทางไอซีที (ICT Literacy) หมายถึง ความรู้ที่สามารถเข้าถึงสารสนเทศ (Accessing information) การจัดการสารสนเทศ (Managing information) การประเมิน (Evaluating) การสร้างสิ่งใหม่จากความเข้าใจ (Developing new understandings) การสื่อสาร (Communicating) และการใช้ไอซีทีอย่างเหมาะสม (Using ICT appropriately)

การเข้าถึงสารสนเทศ (Accessing information) หมายถึง การระบุข้อมูลที่จำเป็นและรู้วิธีการค้นหาข้อมูลและนำข้อมูลกลับมาใช้

การจัดการสารสนเทศ (Managing information) หมายถึง การจัดระเบียบและการจัดเก็บข้อมูลสำหรับการเรียกใช้และนำมาใช้

การประเมิน (Evaluating) หมายถึง กระบวนการเก็บข้อมูล ตีความ บันทึก จากการสร้างโซลูชั่นไอซีที และประโยชน์ของข้อมูล

การสร้างสิ่งใหม่จากความเข้าใจ (Developing new understandings) หมายถึง การสร้างข้อมูลและความรู้โดยการสังเคราะห์ การปรับ การประยุกต์ใช้ การออกแบบ การประดิษฐ์ หรือการเขียน

การสื่อสาร (Communicating) หมายถึง การแลกเปลี่ยนข้อมูลด้วยการแบ่งปันความรู้และการสร้างผลิตภัณฑ์ข้อมูลเพื่อให้เหมาะกับผู้ใช้งาน

การใช้ไอซีทีอย่างเหมาะสม (Using ICT appropriately) หมายถึง การตัดสินใจต่อการใช้ไอซีที และมีความรับผิดชอบด้วยการพิจารณาทางสังคม กฎหมาย และจริยธรรม

คลังข้อสอบ (Item bank) หมายถึง แหล่งจัดเก็บข้อสอบความรู้ทางไอซีทีที่มีคุณภาพ ประกอบด้วย อำนาจจำแนกข้อสอบ (a) ค่าความยากของข้อสอบ (b) ค่าโอกาสการเดาของข้อสอบ (c) และค่าถี่สะสมของการใช้ข้อสอบ

ค่าอำนาจจำแนกของข้อสอบ (Discrimination power parameter: a) หมายถึง ค่าที่เป็นสัดส่วนโดยตรงกับความชันของโค้งคุณลักษณะของข้อสอบ ณ จุดเปลี่ยนโค้งมีค่าระหว่าง – ถึง + สำหรับค่าอำนาจจำแนกของข้อสอบที่เป็นลบ (-) แสดงว่า ข้อสอบไม่ดี ไม่สามารถจำแนกผู้ทดสอบได้ ต้องตัดข้อสอบข้อนั้นทิ้ง หรือปรับปรุงข้อสอบใหม่ ค่าอำนาจจำแนกของข้อสอบที่เป็น (+) แสดงว่าข้อสอบดี สามารถจำแนกผู้ทดสอบได้ ในคลังข้อสอบมีค่าอำนาจจำแนกตั้งแต่ 0.50 ถึง 2.50

ค่าความยากของข้อสอบ (Difficulty parameter: b) หมายถึง ค่าที่แสดงถึงความสามารถของผู้ทดสอบที่จุดเปลี่ยนโค้งคุณลักษณะของข้อสอบ ซึ่งมีค่าอยู่ระหว่าง – ถึง + สำหรับค่าความยาก

ของข้อสอบมีค่าเท่ากับ -2.50 แสดงว่า ข้อสอบง่ายมาก และค่าความยากของข้อสอบมีค่าเท่ากับ 2.50 แสดงว่า ข้อสอบยากมาก

ค่าโอกาสการเดาของข้อสอบ (Guess parameter: c) หมายถึง ค่าความน่าจะเป็นของผู้ทดสอบที่ไม่มีความสามารถในการตอบข้อสอบ แต่ตอบข้อสอบข้อนั้นได้ถูกต้อง หรือนั่นคือค่าที่แสดงโอกาสการตอบข้อสอบถูก โดยผู้ที่มีความรู้ในเรื่องนั้น ๆ สำหรับค่าการเดาของข้อสอบมีค่าตั้งแต่ 0 ถึง 1 ในงานวิจัยนี้จะมีค่าการเดาของข้อสอบไม่เกิน 0.30

การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (Multidimensional Computerized Adaptive Testing: MCAT) หมายถึง กระบวนการทดสอบด้วยคอมพิวเตอร์ ซึ่งประกอบด้วยข้อสอบจำนวนหนึ่งที่ผ่านกระบวนการสร้าง และคัดเลือกตามเกณฑ์คุณภาพที่กำหนดไว้ โดยผู้ทดสอบแต่ละคนจะได้รับข้อสอบต่างกัน เริ่มสอบจากข้อสอบที่มีค่าความยากปานกลางก่อน หากผู้ทดสอบตอบถูก ให้จัดข้อสอบข้อถัดไปที่มีค่าความยากสูงขึ้น แต่หากผู้ทดสอบตอบผิด ให้จัดข้อสอบข้อถัดไปที่มีค่าความยากน้อยลง ทำเช่นนี้จนกว่าจะสิ้นสุดกาทดสอบ

การทดสอบด้วยคอมพิวเตอร์ (Computer Test: CT) หมายถึง กระบวนการทดสอบที่นำคอมพิวเตอร์มาใช้ในการทดสอบแทนการทดสอบบนกระดาษดินสอ ผู้ทดสอบทุกคนจะได้รับแบบทดสอบชุดเดียวกัน ซึ่งประกอบด้วยข้อสอบจำนวน 40 ข้อ ที่ผ่านกระบวนการสร้างและคัดเลือกตามเกณฑ์คุณภาพที่กำหนดไว้ คะแนนที่ได้คือจำนวนข้อสอบที่ผู้ทดสอบตอบถูก

วงจรการพัฒนาโปรแกรม (Software Development Life Cycle: SDLC) หมายถึง โครงร่างหรือแนวทางวิธีการเพื่ออธิบายขั้นตอนการพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ สำหรับผู้พัฒนาใช้เป็นกรอบในการพัฒนาโปรแกรม มีจำนวน 6 ขั้นตอน

การเปรียบเทียบประสิทธิภาพ (Efficiency comparison) หมายถึง การเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา ประกอบด้วย 4 ด้าน ได้แก่ 1) จำนวนข้อสอบที่ใช้ในการทดสอบ (Test length) เป็นการเปรียบเทียบประสิทธิภาพของการนำข้อสอบไปใช้ทดสอบน้อยที่สุด 2) เวลาที่ใช้ในการทดสอบ (Test time) เป็นการเปรียบเทียบประสิทธิภาพการใช้เวลาในการทดสอบน้อยที่สุด 3) อัตราการแสดงการใช้ข้อสอบสูงสุด (Maximum Item Exposure Rate: MIER) เป็นอัตราการใช้ข้อสอบในคลังสูงสุด วิธีการใดมีอัตราการใช้สูงสุด แสดงว่า ประสิทธิภาพการควบคุมการใช้ข้อสอบวิธีนั้นต่ำ และ 4) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) เป็นค่าความถูกต้องระหว่างค่าประมาณความสามารถจริงที่ได้จากการจำลอง (Simulator) กับค่าประมาณความสามารถที่ได้จากผู้ทดสอบ ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าเข้าใกล้ 0 แสดงว่า ค่าประมาณความสามารถของผู้ทดสอบมีค่าใกล้เคียงกับค่าความสามารถที่แท้จริงของผู้ทดสอบ

ค่าประมาณความสามารถของผู้ทดสอบ (Ability estimate) หมายถึง ค่าที่ใช้สำหรับประมาณค่าความสามารถของผู้ทดสอบ (True ability) แสดงถึงระดับความสามารถของผู้ทดสอบ โดยคำนวณจากผลการตอบข้อสอบในทุกข้อที่ผ่านมา ตามทฤษฎีการตอบสนองข้อสอบ ค่าความสามารถที่แท้จริงของผู้ทดสอบมีค่าอยู่ระหว่าง -3.00 ถึง 3.00 ถ้าค่าประมาณความสามารถของผู้ทดสอบมีค่าเท่ากับ -3.00 แสดงว่า ผู้ทดสอบมีความสามารถในระดับต่ำ แต่ถ้าค่าประมาณความสามารถของผู้ทดสอบมีค่าเท่ากับ 3.00 แสดงว่า ผู้ทดสอบมีความสามารถสูงมาก โดยใช้วิธีการประมาณค่าแบบภาวะความน่าจะเป็นสูงสุด (Maximum Likelihood Estimation: ML)

ค่าคะแนนรวมของผู้ทดสอบ (Observed score) หมายถึง ค่าคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ หรือคะแนนที่สังเกตุได้ของผู้ทดสอบ คำนวณจากจำนวนข้อสอบทั้งหมดที่ผู้ทดสอบตอบถูกในการทดสอบด้วยคอมพิวเตอร์ ซึ่งมีหน่วยวัดเป็นคะแนน

ค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าความสามารถของผู้ทดสอบ (Maximum Information Criterion: MIC) หมายถึง วิธีการคัดเลือกข้อสอบข้อถัดไป โดยใช้พารามิเตอร์ค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าโอกาสการเดาของข้อสอบ มาร่วมในการคำนวณ ซึ่งการเลือกข้อสอบข้อถัดไปจะพิจารณาเลือกข้อสอบที่มีค่าสารสนเทศสูงสุดที่ใกล้เคียงกับค่าประมาณความสามารถของผู้ทดสอบ

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยเรื่องการพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง สำหรับทดสอบความรู้ทางไอซีที: การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ นำเสนอแนวคิดหรือทฤษฎีพื้นฐานเกี่ยวกับการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) (Reckase, 2009) ทฤษฎีวงจรการพัฒนาโปรแกรม (SDLC) (Singh et al., 2015) ทฤษฎีโหลดบาลานซ์ซิ่ง (Singh et al., 2015) วิธีการควบคุมการใช้ข้อสอบ วิธีการควบคุมสมดุลเนื้อหา และงานวิจัยที่เกี่ยวข้องซึ่งแบ่งออกเป็น 5 ตอน ดังนี้

ตอนที่ 1 โหลดบาลานซ์ซิ่ง และงานวิจัยที่เกี่ยวข้อง

ตอนที่ 2 วิธีการควบคุมการใช้ข้อสอบ วิธีการควบคุมสมดุลเนื้อหา และงานวิจัยที่เกี่ยวข้อง

ตอนที่ 3 ทฤษฎีตอบสนองข้อสอบแบบพหุมิติ

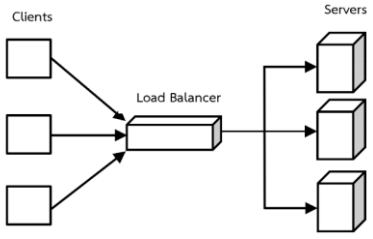
ตอนที่ 4 การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์

ตอนที่ 5 ทฤษฎีวงจรการพัฒนาโปรแกรม

ตอนที่ 1 โหลดบาลานซ์ซิ่ง และงานวิจัยที่เกี่ยวข้อง

1. โหลดบาลานซ์ซิ่ง (Load Balancing: LB)

โหลดบาลานซ์ซิ่งเป็นวิธีการที่มาช่วยสนับสนุนการทำงานระหว่างผู้ใช้งานกับเครื่องแม่ข่าย ถ้าผู้ใช้งานร้องขอข้อมูลจากเครื่องแม่ข่ายเครื่องที่ 1 มาเป็นจำนวนมาก ทำให้เครื่องแม่ข่ายเครื่องที่ 1 เริ่มมีภาระงานมากเกินไป โหลดบาลานซ์ซิ่งจะแชร์การใช้งานจากเครื่องแม่ข่ายเครื่องที่ 1 ไปยังเครื่องแม่ข่ายเครื่องที่ 2 ทันที โหลดบาลานซ์นี้จะติดต่อกับเครื่องแม่ข่ายหลายตัวพร้อมกันได้ เพื่อจัดสรรเครื่องแม่ข่ายให้กับผู้ใช้งานหลาย ๆ คนสามารถใช้งานได้พร้อมกัน ดังแสดงในภาพที่ 2-1 ซึ่งได้นำอัลกอริทึมของโหลดบาลานซ์ซิ่งมาควบคุมความสมดุลของเนื้อหาในการใช้สอบและการควบคุมการนำข้อสอบมาใช้ สำหรับเป็นการขยายขีดความสามารถในการเลือกข้อสอบข้อถัดไป แสดงดังภาพที่ 2-1



ภาพที่ 2-1 ความสามารถในการทำงานของโหลดบาลานซ์ซิ่ง

จากภาพที่ 2-1 เครื่องลูกข่าย (Clients) คือ ผู้ร้องขอบริการใช้งานเครื่องแม่ข่าย (Servers) โดยมีโหลดบาลานซ์เซอร์ (Load balancer) เป็นตัวกลางในการจัดสรรการทำงานของเครื่องแม่ข่าย ไม่ให้เครื่องใดเครื่องหนึ่งของเครื่องแม่ข่ายมีการทำงานมากเกินไป การใช้โปรแกรมโหลดบาลานซ์ซิ่ง จะช่วยให้การใช้งานที่ต้องเชื่อมต่อกับเครื่องแม่ข่ายมีความเร็วและเสถียรมากยิ่งขึ้น ลักษณะการทำงานของโหลดบาลานซ์ซิ่ง จะแบ่งวิธีการทำงานออกเป็น 3 แบบดังนี้

1) การทำงานแบบ Round robin เป็นการจัดแจงผู้ใช้งานให้ใช้งานกับเครื่องแม่ข่ายแต่ละเครื่องให้มีปริมาณเท่ากัน โหลดบาลานซ์ซิ่งจะจัดการผู้ใช้งานให้ไปใช้งานเครื่องแม่ข่ายที่เชื่อมต่อโหลดบาลานซ์ซิ่ง โดยคำนวณความสมดุลของจำนวนเครื่องแม่ข่ายไปเรื่อย ๆ ไม่ให้เครื่องแม่ข่ายใดเครื่องแม่ข่ายหนึ่งมีการใช้งานที่มากเกินไป

2) การทำงานแบบ Sticky เป็นการจดจำการใช้งานของผู้ใช้งานว่าเคยใช้เครื่องแม่ข่ายเครื่องใดมาก่อนเมื่อตรวจสอบ เช่น ผู้ใช้งานคนที่ 1 เคยใช้เครื่องแม่ข่าย1 มาก่อนโหลดบาลานซ์ซิ่งก็จะส่งผู้ใช้งานคนที่ 1 ไปยังเครื่องแม่ข่าย1

3) การทำงานแบบ Work load เป็นการแบ่งเบาภาระในการทำงานของเครื่องแม่ข่ายแต่ละเครื่อง โดยโปรแกรมโหลดบาลานซ์ซิ่งจะตรวจสอบเครื่องแม่ข่ายแต่ละเครื่อง ถ้าเครื่องใดมีการใช้งานหรือภาระงานมากและสุ่มเสี่ยงว่าจะแฮงค์ (เครื่องใช้งานผิดปกติ) โหลดบาลานซ์ซิ่งจะแบ่งผู้ใช้งานไปยังเครื่องแม่ข่ายที่ 2 และที่ 3 ต่อไปเรื่อย ๆ เพื่อให้ผู้ใช้งานสามารถใช้งานได้อย่างต่อเนื่อง ข้อดีของโปรแกรมโหลดบาลานซ์ซิ่งจะสามารถช่วยเพิ่มความเร็วในการใช้งานอินเทอร์เน็ตได้อย่างมาก เพราะการโหลดเว็บไซต์หรือข้อมูลในอินเทอร์เน็ตแต่ละครั้งจะแชร์การดาวน์โหลดจากทุกเครื่องแม่ข่ายที่มีการเชื่อมต่ออยู่ ซึ่งทำให้เครื่องแม่ข่ายแต่ละเครื่องทำงานไม่หนักเกินไป และที่สำคัญโปรแกรมโหลดบาลานซ์ซิ่งสามารถใช้งานกับผู้ใช้งานจำนวนมาก แต่มีข้อจำกัดคือ ถ้าโหลดบาลานซ์ซิ่งล่มขึ้นมาผู้ใช้งานทั้งโปรแกรมก็จะไม่สามารถใช้งานอินเทอร์เน็ตได้ (Faynberg et al., 2016)

1.1 คุณลักษณะของโหลดบาลานซ์ซิ่ง (Load balancing features) การพัฒนาโหลดบาลานซ์ซิ่ง ได้มีการประยุกต์กับงานหลายประเภท เพื่อลดภาระการทำงานของสิ่งที่ทำงานหนักเกินไป ซึ่งปัจจุบันได้นำมาประยุกต์กับการประมวลผลแบบแบ่งปันทรัพยากรผ่านเครือข่าย เป็นกระบวนทัศน์แบบกระจายบนอินเทอร์เน็ตที่ออกแบบมาสำหรับการแชร์การใช้ทรัพยากรและบริการต่าง ๆ เช่น การเก็บข้อมูลความสามารถในการคำนวณและแอพพลิเคชันอื่น ๆ ที่มีความน่าเชื่อถือสูงผ่านเครือข่ายขนาดใหญ่ อย่างไรก็ตามเนื่องจากข้อมูลคำขอขาเข้าเครื่องแม่ข่ายแบบไดนามิกต้องมีการจัดสรรทรัพยากรแบบไดนามิกในโปรแกรม โดยธรรมชาติการคำนวณแบบไดนามิกในโปรแกรมการประมวลผลแบบแบ่งปันทรัพยากรผ่านเครือข่ายจำเป็นต้องมีกลไกการปรับสมดุลภาระงานที่มีประสิทธิภาพ โหลดบาลานซ์ซิ่งช่วยลดความกังวลเกี่ยวกับการกระจายทรัพยากรระหว่างผู้ใช้หรือคำ

ขอในรูปแบบที่เหมือนกัน เพื่อให้โหนดไม่มีการโอเวอร์โหลด (Over Load) หรือโหนดไม่ได้ใช้งาน เช่นเดียวกับงานคอมพิวเตอร์แบบกระจายงานทางอินเทอร์เน็ตอื่น ๆ การปรับสมดุลภาระงานเป็นสิ่งสำคัญในโปรแกรมการประมวลผลแบบแบ่งปันทรัพยากรผ่านเครือข่าย ในกรณีที่ไม่มีข้อกำหนดสมดุลภาระงาน ประสิทธิภาพของโหนดที่โอเวอร์โหลดบางส่วนอาจลดลงอย่างมากในช่วงเวลาที่นำไปสู่การใช้งานในคอมพิวเตอร์รูปแบบกระจายการประมวลผลแบบขนานและสภาพแวดล้อม โหลดบาลานซ์ซิ่งอัลกอริทึมจะแบ่งเป็น 3 แบบ ได้แก่ 1) แบบคงที่ 2) แบบไดนามิก และ 3) แบบผสม ซึ่งมีลักษณะดังนี้

1.1.1 โหลดบาลานซ์ซิ่งอัลกอริทึมแบบคงที่ (Static) เหมาะสำหรับสภาพแวดล้อมแบบกระจายขนาดเล็กที่มีความเร็วสูงและการสื่อสารที่ล่าช้าในการสื่อสาร

1.1.2 โหลดบาลานซ์ซิ่งอัลกอริทึมแบบไดนามิก (Dynamic) มุ่งเน้นไปที่การลดความล่าช้าในการสื่อสารและเวลาในการดำเนินการ ดังนั้นจึงเหมาะสำหรับสภาพแวดล้อมแบบกระจายขนาดใหญ่

1.1.3 โหลดบาลานซ์ซิ่งอัลกอริทึมแบบผสม (Mixed) เน้นการแจกจ่ายแบบสมมาตรของงานคอมพิวเตอร์ที่ได้รับมอบหมายและลดภาระงานการสื่อสารของโหนดคอมพิวเตอร์แบบกระจาย

1.2 องค์ประกอบของโหลดบาลานซ์ซิ่ง (Elements of Load Balancing) เทคนิคโหลดบาลานซ์ซิ่ง สามารถนำมาประยุกต์กับวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาในการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ วิธีโหลดบาลานซ์ซิ่งมีองค์ประกอบ ดังนี้

1.2.1 การเคลื่อนย้าย (Transfer policy) เป็นการเคลื่อนย้ายการทำงานจากโหนดหนึ่งไปยังโหนดอื่น ๆ

1.2.2 การคัดเลือก (Selection policy) เป็นการคัดเลือกโหนดที่จะประมวลผลการทำงาน

1.2.3 การเตรียมความพร้อม (Location policy) เป็นการเตรียมความพร้อมพื้นที่การทำงานของโหนดที่จะประมวลผล

1.2.4 การแสดงข้อมูล (Information policy) เป็นการแสดงข้อมูลภาระงานที่โหนดนั้น ๆ กำลังดำเนินการอยู่

1.2.5 การประเมินภาระงาน (Load estimation policy) เป็นการประเมินว่าโหนดที่แสดงข้อมูลมานั้น ควรส่งภาระงานให้ทำอีกหรือไม่

1.3 ขั้นตอนการทำงานของโหลดบาลานซ์ซิ่ง (Process of load balancing) Singh et al. (2015) นำเสนอขั้นตอนการทำงานของโหลดบาลานซ์ซิ่งเพื่อลดภาระการทำงานของเครื่องแม่ข่ายเป็นการประมวลผลแบบแบ่งปันทรัพยากรผ่านเครือข่าย โดยมี 3 ขั้นตอนดังนี้

ขั้นตอนที่ 1 การสำรวจและเคลื่อนย้ายข้อมูล (Migration Agent: MA) เป็นการสำรวจคุณลักษณะ ของเครื่องแม่ข่ายทุกเครื่องที่อยู่ในวงเครือข่ายเดียวกัน ซึ่งลักษณะการทำงานมีดังนี้

1) การค้นหาเครื่องแม่ข่าย เป็นการค้นหาจำนวนเครื่องแม่ข่ายทั้งหมดในวงเครือข่าย เพื่อให้ได้มาซึ่งคุณลักษณะของเครื่องแม่ข่าย และในเครื่องแม่ข่ายจะมีภาระการทำงานของเครื่องแม่ข่ายเสมือน (Virtual Machine: VM) เพื่อจัดสรรการทำงาน

2) การเก็บคุณลักษณะเครื่องแม่ข่าย เป็นการเก็บข้อมูลคุณลักษณะของเครื่องแม่ข่าย ได้แก่ ภาระการประมวลผลของซีพียู ภาระการใช้หน่วยความจำสำรอง ภาระการใช้เส้นทางบนบัสของแมนบอร์ด เป็นต้น เพื่อนำไปคำนวณร้อยละทั้งหมดของเครื่องแม่ข่ายที่ยังไม่ใช้งาน

3) การเคลื่อนย้ายข้อมูล เป็นการนำข้อมูลที่ได้ในข้อ 2 ไปเก็บในขั้นตอนที่ 2 คือการคำนวณภาระงาน

ขั้นตอนที่ 2 การคำนวณภาระงาน (Load Agent: LA) เป็นวิธีการคำนวณจำนวนภาระงานของเครื่องแม่ข่ายเสมือน (VM) ที่อยู่ภายใต้การทำงานของเครื่องแม่ข่าย และสถานะนำไปใช้งาน ซึ่งเป็นข้อมูลคุณลักษณะของเครื่องแม่ข่าย ที่ได้มาจากขั้นตอนที่ 1 การสำรวจและเคลื่อนย้ายข้อมูล (Migration Agent: MA) ดังสมการ 1

$$\mu_{available} = \mu_{total} - \mu_{used} \tag{1}$$

$$v(\%) = \frac{\mu_{available}}{\mu_{total}} \times 100 \tag{2}$$

เมื่อ $\mu_{available}$ แทน จำนวนเครื่องแม่ข่ายที่ยังว่าง

μ_{total} แทน จำนวนเครื่องแม่ข่ายทั้งหมด

μ_{used} แทน จำนวนเครื่องแม่ข่ายที่ใช้งาน

$v(\%)$ แทน จำนวนเปอร์เซ็นต์เครื่องแม่ข่ายที่ว่าง

จากสมการ 2 คำนวณจำนวนเครื่องแม่ข่ายทั้งหมดที่ยังไม่นำไปใช้ เมื่อคำนวณเสร็จแล้ว สมการ 3 นำผลสถานะจากเครื่องแม่ข่ายที่ยังไม่ได้ใช้ในแต่ละเครื่องมาคำนวณร้อยละทั้งหมด เพื่อใช้ในการเปรียบเทียบตามเงื่อนไขที่กำหนด

$$v = \begin{cases} \leq 25\% \; Critical \; Allocation \\ > 25\% \; Normal \; Allocation \end{cases} \tag{3}$$

จากสมการ 3 เปรียบเทียบเงื่อนไขของเปอร์เซ็นต์ต่อการใช้เครื่องแม่ข่าย โดยมีเงื่อนไขดังนี้

1) ถ้าร้อยละของเครื่องแม่ข่ายน้อยกว่าหรือเท่ากับ 25% จะแสดงสถานะไม่ปกติ

2) ถ้าร้อยละของเครื่องแม่ข่ายมากกว่า 25% จะแสดงสถานะปกติ

ขั้นตอนที่ 3 การควบคุมการทำงาน (Channel Agent: CA) เป็นวิธีการประเมินภาระงาน (Load Estimation policy) เพื่อเคลื่อนย้ายภาระงาน จากเครื่องแม่ข่ายหนึ่งไปอีกเครื่องแม่ข่ายหนึ่ง (Transfer policy) เมื่อเข้าเงื่อนไขในสมการ 21 อัลกอริทึมจะเลือกเครื่องแม่ข่ายตามเงื่อนไข (Selection policy) เพื่อเตรียมความพร้อมเครื่องแม่ข่ายในการประมวลผล โดยใช้การสำรวจและการเคลื่อนย้ายข้อมูล เป็นช่องทางในการดึงข้อมูลจากการคำนวณภาระงาน

1.4 การวัดประสิทธิภาพของโหลดบาลานซ์ซิ่ง (Performance measurement of load balancing)

1.4.1 ความน่าเชื่อถือ (Reliability) อัลกอริทึมต้องมีความน่าเชื่อถือ เนื่องจากความล้มเหลวของกระบวนการในขณะที่การถ่ายโอนงานจากเครื่องแม่ข่ายที่หนึ่งไปยังที่เครื่องอื่นอาจทำให้เวลารอที่เพิ่มขึ้น

1.4.2 การปรับเหมาะ (Adaptability) อัลกอริทึมต้องมีความสามารถในการปรับเปลี่ยนคำขอผู้ใช้ที่มีการเปลี่ยนแปลงแบบไดนามิกและมอบการจัดสรรงานในช่วงเวลาที่น้อยที่สุด

1.4.3 ความผิดพลาด (Fault tolerance) อัลกอริทึมต้องให้ความทนทานต่อความผิดพลาดดังนั้นกรณีที่เกิดปัญหาในโปรแกรม กลไกการปรับสมดุลภาระสมบูรณ์จะไม่หยุดทำงาน

1.4.4 ปริมาณการใช้งาน (Throughput) อัลกอริทึมต้องทำให้มั่นใจได้ว่าจะมีปริมาณใช้งานเพิ่มขึ้นเมื่อมีทรัพยากรของเครื่องแม่ข่ายจำกัด ถ้าอัลกอริทึมไม่สามารถเพิ่มปริมาณงานของโปรแกรมได้จะไม่บรรลุจุดประสงค์

1.4.5 เวลาที่รอ (Waiting time) อัลกอริทึมควรลดระยะเวลารอของงานเพื่อจัดสรรทรัพยากรให้เครื่องแม่ข่าย

2. งานวิจัยที่เกี่ยวข้องกับวิธีโหลดบาลานซ์ซิ่ง

Hou, Thomas, and Varadharajan (2014) ได้เสนออัลกอริทึมการกระจายโหลดบาลานซ์แบบไดนามิกสำหรับ Hadoop MapReduce อัลกอริทึมได้สร้างความสมดุลของภาระงานในระหว่างชั้น ขณะที่งานก่อนหน้าจะพยายามโหลดความสมดุลระหว่างโหนดแต่ละโหนดใน MapReduce การเพิ่มประสิทธิภาพของระบบไม่มีทางใดที่จะทำให้ Hadoop สามารถรับประกันได้ว่าชั้นความสามารถที่สูงขึ้นจะมีปริมาณภาระงานมากขึ้นกว่าชั้นที่มีความสามารถที่ต่ำกว่า กล่าวอีกนัยหนึ่ง เมื่อกำหนดภาระงานให้โหนดแต่ละโหนด ความสามารถในการประมวลผลไม่เกี่ยวข้องกัน อัลกอริทึมจะทำงาน 2 รูปแบบ ได้แก่ 1) อัลกอริทึมให้ความสำคัญกับการปรับสมดุลภาระงานระหว่างชั้น และ 2) ใช้ Software Defined Network (SDN) เพื่อปรับปรุงการถ่ายโอนข้อมูล จาก

การทดลอง ปรากฏว่า การเคลื่อนย้ายภาระงานจากชั้นที่มีภาระงานมากไปเป็นภาระงานที่น้อยใช้ระยะเวลาลดลง

Singh, Thakur, and Chaudhary (2015) ได้เสนออัลกอริทึมการสร้างสมดุลภาระงานใหม่บนการประมวลผลแบบแบ่งปันทรัพยากรผ่านเครือข่าย ซึ่งเป็นวิธีการกระจายการทำงานของภาระงานแต่ละโหนดให้มีความสมดุลกัน โดยอัลกอริทึมมี 3 ขั้นตอน ได้แก่ 1) การสำรวจและเคลื่อนย้ายข้อมูล 2) การคำนวณภาระงาน และ 3) การควบคุมการทำงาน ในการทดลองเปรียบเทียบ 3 รูปแบบ ได้แก่ 1) รูปแบบการทำงานแบบปกติ (Under loaded VM) 2) รูปแบบการทำงานแบบเกินพิกัดและควบคุมด้วย A2LB (Overloaded VM) (with A2LB Algorithm) และ 3) รูปแบบการทำงานแบบเกินพิกัดและไม่ควบคุมด้วย A2LB (Overloaded VM) (without A2LB Algorithm) จากการทดลองปรากฏว่า เปรียบเทียบเรื่องเวลา รูปแบบที่ 2 ให้เวลาดีที่สุด รองลงมาคือรูปแบบที่ 1 และ รูปแบบที่ 3 ตามลำดับ

Chien, Son, and Loc (2016) ได้เสนออัลกอริทึมการสร้างสมดุลภาระงานโดยอาศัยวิธีการประมาณเวลาสิ้นสุดของเวลาให้บริการ ในอัลกอริทึมนี้พิจารณาถึงการประมวลผลทันทีของ VM และขนาดของงภาระงานที่ได้รับมอบหมาย อัลกอริทึมนี้รวมสองวิธีในการประเมินเวลาสิ้นสุดของการบริการใน VM ได้แก่ 1) VM ที่เลือกสามารถทำงานได้เสร็จสิ้นเร็วที่สุด และ 2) มีคำขอการจัดสรรครั้งต่อไป โหลดบาลานซ์อัลกอริทึมมีการประมาณการเวลาเข้าคิวของงานทั้งหมดและงานที่เข้ามาทำอย่างสมบูรณ์ในทุก VM และตรงกับเวลาที่เร็วที่สุดจะถูกเลือกเพื่อแจกจ่ายงาน ผลการทดลองปรากฏว่า อัลกอริทึมที่พัฒนาใช้เวลาประมวลผลอยู่ในระดับดี

Keshvadi and Faghih (2016) ได้เสนอระบบจัดสมดุลภาระงานแบบหลายระบบในสภาพแวดล้อมการประมวลผลแบบแบ่งปันทรัพยากรผ่านเครือข่าย IaaS กลไกของกระบวนการทำงาน ได้ดำเนินจากการเริ่มโดยผู้ที่รับและผู้ส่งให้มีความสมดุลของภาระงาน เพื่อลดเวลาในการรอคอยของงานและเป็นไปตามข้อตกลงเบื้องต้นในระดับการให้บริการ (SLA) กลไกที่นำเสนอในบทความนี้ ประกอบด้วยตัวแทน 3 ส่วน ได้แก่ 1) Virtual Machine Monitor (VMM) 2) Datacenter Monitor (DM) และ 3) Ant Negotiator (NA) ตัวแทน VMM รวบรวมข้อมูลซีพียู หน่วยความจำและแบนด์วิดท์ของแต่ละ VM ที่โฮสต์ของงานประเภทต่าง ๆ เพื่อตรวจสอบข้อมูลการโหลดในตารางสำหรับจัดเก็บสถานะของ VM ตัวแทน DM ดำเนินนโยบายข้อมูลในข้อมูลกลางโดยตรวจสอบข้อมูลจาก VMM ตัวแทนนี้ได้รับการสนับสนุนข้อมูลจากตารางที่เก็บรักษาข้อมูลเกี่ยวกับสถานะและลักษณะของ VM ทั้งหมด ผลการทดลอง ปรากฏว่า ระบบจัดสมดุลภาระงานแบบหลายระบบในสภาพแวดล้อมการประมวลผลแบบแบ่งปันทรัพยากรผ่านเครือข่าย IaaS มีการจัดสรรเวลาอยู่ในระดับดี

Ghoneem and Kulkarni (2017) ได้เสนอเทคนิคการตั้งเวลาแบบปรับตัวสำหรับกำหนดการ MapReduce เพื่อเพิ่มประสิทธิภาพเมื่อใช้ในสภาพแวดล้อมที่ต่างกัน ในรูปแบบนี้กำหนดการจัดเตรียมทรัพยากรและความต้องการของงานโดยการจัดตารางเวลาด้วยอัลกอริทึมการจัดหมวดหมู่ อัลกอริทึมนี้แบ่งประเภทงานออกเป็นสองประเภทของการปฏิบัติการ จากนั้นกำหนดงานที่ปฏิบัติการให้กับโหนดที่เหมาะสมได้ดำเนินการสำเร็จโดยไม่ล้มเหลว ซึ่งเป็นการเพิ่มเวลาในการดำเนินการของงาน ตัวกำหนดเวลานี้จะแก้ไขปัญหาของการกำหนดงานก่อนหน้า เช่น การประสานการทำงานเพื่อความสมดุลของโหนดแต่ละโหนดอย่างเป็นธรรม ตัวจัดตารางเวลาแบบปรับตัวช่วยเพิ่มประสิทธิภาพของโมเดล MapReduce ในสภาพแวดล้อมที่ต่างกัน ในขณะที่ลดค่าโหนดต้นแบบและการรับส่งข้อมูลเครือข่าย

Vanitha and Marikkannu (2017) ได้ศึกษาเรื่องการใช้ประโยชน์ทรัพยากรที่มีประสิทธิภาพในสภาพแวดล้อมแบบการประมวลผลแบบแบ่งปันทรัพยากรผ่านเครือข่าย โดยใช้โหลดบาลานซ์ซิ่งอัลกอริทึมแบบไดนามิกที่มีการจัดระเบียบแบบไดนามิกสำหรับการใช้เครื่องจำลอง การประมวลผลแบบแบ่งปันทรัพยากรผ่านเครือข่ายเป็นส่วนหนึ่งที่ใช้งานทุกด้านผ่านการใช้แอ็พพลิเคชันทำให้ผู้ใช้อินเทอร์เน็ตสามารถใช้บริการและแหล่งข้อมูลที่มีอยู่ได้ง่ายขึ้น แหล่งข้อมูลเหล่านี้ต้องมีการป้องกันอย่างดีเนื่องจากมีการใช้งานโดยสมาชิกจำนวนมาก ดังนั้นจึงจำเป็นต้องปรับโหลดในเครื่องแม่ข่ายต่าง ๆ เพื่อหลีกเลี่ยงความแออัดในเครือข่ายและเพื่อลดการเพิ่มค่าใช้จ่ายหรือทรัพยากร โดยการใช้โหลดบาลานซ์ซิ่งอัลกอริทึมซึ่งประยุกต์เจเนติกอัลกอริทึม (Genetic Algorithm: GA) ช่วยแก้ไขปัญหาในการกระจายโหลดการทำงานบนเครื่องจำลอง (VM) ที่มีอยู่ในเครือข่าย ซึ่งเป็นอัลกอริทึมที่พัฒนาขึ้นมาใหม่ เรียกได้ว่าเป็นอัลกอริทึมการโหลดบาลานซ์ที่มีการจัดระเบียบแบบไดนามิก ผลการทดลอง ปรากฏว่า อัลกอริทึมมีประสิทธิภาพสำหรับการลดภาระงานที่ใช้ในโปรแกรมการประมวลผลแบบแบ่งปันทรัพยากรผ่านเครือข่ายอยู่ในระดับดี

Kumar and Sharma (2017) ได้เสนออัลกอริทึมสร้างความสมดุลภาระงานคือวิธีการของอัลกอริทึมควบคุมความสมดุลภาระงานรับ-ส่งแบบไดนามิก โดยมีวัตถุประสงค์เพื่อลดเวลาในการประมวลผลและเพิ่มอัตราการใช้ทรัพยากรโดยเฉลี่ยในการประมวลผลแบบแบ่งปันทรัพยากรผ่านเครือข่าย สร้าง N จำนวนงานและ M จำนวนเครื่องเสมือนที่มีกำลังประมวลผลแตกต่างกัน และจัดเรียงงานรวมทั้งเครื่องเสมือนเพื่อลดลำดับความยาวงานและความเร็วในการประมวลผลต่อการเริ่มต้นจัดสรรงานไปยังเครื่องเสมือนในคำสั่ง (First Come First Served Scheduling: FCFS) สร้างหมายเลข ID งานที่จัดสรรไว้ของเครื่องเสมือนแต่ละเครื่อง เมื่องานได้รับมอบหมายให้เครื่องเสมือนแล้ว จะเริ่มต้นการโหลดสมดุลการทำงานที่เครื่องเสมือนและหาจำนวนความยาวของแต่ละงานที่กำหนดให้ คำนวณภาระงานในเครื่องเสมือนมาไว้ในข้อมูลส่วนกลาง คำนวณสมรรถนะของเครื่องเสมือนจากข้อมูลส่วนกลาง และตรวจสอบเงื่อนไขว่า เครื่องเสมือนที่กำลังทำงานอยู่ มีภาระงานที่

บรรจุอยู่ในข้อมูลส่วนกลางน้อยกว่างานที่มีอยู่ หากเงื่อนไขนี้เป็นเท็จ การดำเนินการโหลดบาลานซ์ไม่สามารถทำได้ในเครื่องเสมือนที่มีอยู่ เครื่องกำหนดเวลาในการประมวลผลแบบแบ่งปันทรัพยากรผ่านเครือข่ายจะบูตเครื่องเสมือนใหม่ โดยใช้แนวคิดการทำงานแบบไดนามิก หากเงื่อนไขเป็นจริงการปรับสมดุลภาระงานเป็นไปได้ตามพื้นฐานการประมวลผลแบบแบ่งปันทรัพยากรผ่านเครือข่ายที่มีอยู่ การทดลองจะเปรียบเทียบประสิทธิภาพ 4 อัลกอริทึม ได้แก่ 1) FCFS 2) SJF 3) Min-Max Priority Load Balancing Algorithm (MMLB) และ 4) Proposed LBA ซึ่งเป็นวิธีใหม่ ในการทดลองจะใช้เวลา 2 ชั่วโมง ในสถานการณ์ที่แตกต่างกัน แต่ละสถานการณ์มีความยาวในการประมวลผล 10 ถึง 50 สถานการณ์ โดยใช้วิธีการสุ่มสถานการณ์ ผลการทดสอบเวลารวม ปรากฏว่า อัลกอริทึม Proposed LBA ให้ผลรวมเวลาดีที่สุด

Demirci and Korçak (2018) ได้เสนอการนำ Cell breathing algorithms ของโหลดบาลานซ์มาประยุกต์ในการถ่ายโอนข้อมูลในเครือข่าย Wi-Fi ได้อย่างมีประสิทธิภาพในพื้นที่ที่มีกำลังแรงของคลื่น นอกจากนี้ยังมีกลยุทธ์การปรับตัวที่มีประสิทธิภาพสำหรับสภาพแวดล้อมแบบไดนามิกการกระจายของผู้ใช้ชั่วคราวที่เปลี่ยนแปลงไป ในการทดลองจะเปรียบเทียบประสิทธิภาพความเร็วในการโอนถ่ายข้อมูล 5 อัลกอริทึม ได้แก่ 1) Congestion Aware Load Balancing (CALB) 2) Congestion Aware Load Balancing-S (CALB-S) 3) Congestion Aware Load Balancing-A (CALB-A) 4) SSF-AC และ 5) O-CALB จากการทดลอง ปรากฏว่า อัลกอริทึมที่นำเสนอ CALB มีประสิทธิภาพในการถ่ายโอนข้อมูลเสถียรกว่าอัลกอริทึมอื่น สร้างความพึงพอใจสูงสุดให้กับผู้ใช้และประสบการณ์การใช้งานที่ดีขึ้นในเครือข่ายไร้สาย Wi-Fi หรือเซลลูลาร์

จากการศึกษางานวิจัยวิธีโหลดบาลานซ์ซิ่ง ซึ่งเป็นวิธีการของเครือข่าย ปรากฏว่า งานวิจัยแบ่งออกเป็น 3 แบบ คือ 1) การสร้างอัลกอริทึมขึ้นมาใหม่ เป็นการพัฒนาซอสโค้ดให้มีคุณสมบัติตามวัตถุประสงค์ที่ผู้วิจัยต้องการ เพื่อใช้แก้ปัญหาทางด้านเครือข่ายที่เกิดขึ้น 2) การเปรียบเทียบประสิทธิภาพแต่ละอัลกอริทึม เป็นการศึกษาเทคนิคหรือวิธีของโหลดบาลานซ์ซิ่งเพื่อใช้เปรียบเทียบประสิทธิภาพในสถานการณ์จำลอง แล้วสรุปประสิทธิภาพแต่ละด้าน และ 3) การนำอัลกอริทึมมาประยุกต์ เป็นการนำเทคนิคหรือวิธีของโหลดบาลานซ์ซิ่งมาประยุกต์ในการแก้ไขปัญหา เช่น การแก้ไขปัญหาการใช้เครือข่าย การแก้ไขปัญหาโอนถ่ายข้อมูล การแก้ไขปัญหาโอนถ่ายสื่อมีเดีย การแก้ไขปัญหาเครื่องแม่ข่ายที่ทำงานหนักเกินไป เป็นต้น จากการศึกษางานวิจัยดังกล่าว ผู้วิจัยจึงสนใจนำวิธีโหลดบาลานซ์ซิ่งมาประยุกต์ในการควบคุมการใช้ข้อสอบโดยการเรียงลำดับความถี่ของการใช้ข้อสอบจากน้อยไปมากร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้ค่าเฉลี่ยถ่วงน้ำหนัก ให้มีประสิทธิภาพมากขึ้น เนื่องจากข้อสอบความรู้ทางไอซีทีมีหมวดเนื้อหาหลายมิติที่ต้องควบคุมความสมดุลเนื้อหา ซึ่งอัลกอริทึมมีการพัฒนาจากเดิมให้สามารถประยุกต์กับวิธีการควบคุมการใช้ข้อสอบร่วมกับวิธีควบคุมสมดุลเนื้อหาได้

ตอนที่ 2 วิธีการควบคุมการใช้ข้อสอบ วิธีการควบคุมสมดุลเนื้อหา และงานวิจัยที่เกี่ยวข้อง

1. การควบคุมการใช้ข้อสอบ (Item exposure control)

การควบคุมการใช้ข้อสอบ การจัดข้อสอบที่ผ่านการคัดเลือกแล้วให้ผู้ทดสอบ โดยเป็นส่วนหนึ่งของขั้นตอนการคัดเลือกข้อสอบข้อถัดไป ในการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ ขั้นตอนการคัดเลือกข้อสอบข้อถัดไป จะคัดเลือกข้อสอบที่มีความเหมาะสมกับระดับความสามารถของผู้ทดสอบซึ่งเป็นข้อสอบที่ให้ค่าสารสนเทศสูงสุด ณ ระดับความสามารถของผู้ทดสอบขณะนั้น

ค่าสารสนเทศของข้อสอบมีค่าเพิ่มขึ้นใน 2 กรณี คือ 1) ค่าอำนาจจำแนกของข้อสอบมีค่าเพิ่มขึ้น หรือ 2) ค่าความยากของข้อสอบมีค่าใกล้เคียงกับค่าความสามารถของผู้ทดสอบ ด้วยเหตุนี้ในขั้นตอนการคัดเลือกข้อสอบข้อถัดไป ส่วนใหญ่จะคัดเลือกข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูงนำไปใช้ในการทดสอบ ทำให้ข้อสอบบางข้อถูกนำไปใช้ในการทดสอบบ่อยครั้งเกิดไป ขณะเดียวกันข้อสอบบางข้อถูกนำไปใช้ในการทดสอบน้อย หรือไม่เคยถูกนำไปใช้ในการทดสอบ การมีข้อสอบที่ถูกนำไปใช้ในการทดสอบบ่อยครั้งเกินไป จะส่งผลกระทบต่อความปลอดภัยของการทดสอบ กล่าวคือ เมื่อมีข้อสอบที่ถูกนำไปใช้ในการทดสอบบ่อยครั้งเกินไปจนรู้กันในกลุ่มผู้ทดสอบ เป็นผลทำให้ผู้ทดสอบในรุ่นถัดมาสามารถคาดการณ์ล่วงหน้าได้ว่าตนเองจะได้รับข้อสอบข้อใด ทำให้สามารถตอบข้อสอบข้อดังกล่าวได้โดยไม่ได้ใช้ความสามารถของตนเอง คะแนนสอบจึงขาดความถูกต้อง ผู้ทดสอบสามารถทำคะแนนได้มากเกินความสามารถของตนเอง ซึ่งส่งผลกระทบต่อความปลอดภัยของการทดสอบและประสิทธิภาพของการใช้ข้อสอบในคลังข้อสอบ (Item bank) กรณีที่ข้อสอบถูกจัดให้ผู้ทดสอบใช้น้อยหรือไม่ถูกจัดให้ผู้ทดสอบใช้เลย ทำให้เสียค่าใช้จ่ายในการจัดเตรียมคลังข้อสอบ

ดังนั้นจึงมีการศึกษาเกี่ยวกับวิธีการควบคุมการใช้ข้อสอบ โดย Georgiadou and Triantafillou (2007) ได้จำแนกวิธีการควบคุมการใช้ข้อสอบ 5 กลุ่ม ดังนี้

1. วิธีการคัดเลือกแบบสุ่ม (Randomization strategies) วิธีการควบคุมการใช้ข้อสอบในกลุ่มนี้ จะคัดเลือกข้อสอบที่เหมาะสมมา 1 กลุ่ม หลังจากนั้นจะสุ่มเลือกข้อสอบข้อใดข้อหนึ่งจากกลุ่มข้อสอบนี้มาให้ผู้ทดสอบใช้ จึงมีการดำเนินการค่อนข้างง่าย ไม่ซับซ้อน แต่ไม่รับรองว่าจะสามารถควบคุมการใช้ข้อสอบได้ตามต้องการ วิธีการในกลุ่มนี้เป็นวิธีการคัดเลือกข้อสอบข้อถัดไป พร้อมทั้งเป็นการควบคุมการใช้ข้อสอบอีกด้วย วิธีการในกลุ่มนี้ได้แก่

1.1 วิธี 5-4-1 เป็นวิธีที่เสนอโดย MCBride and Martin (1983) วิธีนี้ในขั้นแรกจะเลือกข้อสอบที่มีค่าสารสนเทศสูงสุดมา 5 ข้อ ข้อสอบข้อแรกที่จะจัดให้ผู้ทดสอบจะเป็นข้อสอบที่สุ่มจากข้อสอบที่มีค่าสารสนเทศสูงสุด 5 ข้อนี้ ถัดมาข้อสอบข้อที่ 2 จะสุ่มเลือกมาจากข้อสอบที่มีค่าสารสนเทศสูงสุดที่เหลืออีก 4 ข้อ จะดำเนินการลักษณะนี้จนถึงข้อที่ 5 หลังจากนั้นข้อที่ 5 ก็จะเลือกข้อสอบที่ให้ค่าสารสนเทศสูงสุด ณ ค่าประมาณความสามารถขณะนั้นมาใหม่อีก 5 ข้อ แล้วข้อสอบข้อ

ที่ 6 ก็จะดำเนินการตั้งแต่ขั้นตอนแรกอีกครั้ง กระบวนการนี้จะกระทำจนกระทั่งสิ้นสุดการทดสอบ ข้อดีของวิธีการนี้เป็นกระบวนการที่ไม่ยุ่งยาก แต่มีข้อจำกัดคือ อาจจะไม่สามารถควบคุมการใช้ข้อสอบได้ดีพอ

1.2 วิธี Randomesque strategy เสนอโดย Kingsbury and Zara (1989) เป็นวิธีการที่คล้ายกับวิธี 5-4-1 โดยเริ่มแรกจะเลือกข้อสอบที่มีค่าสารสนเทศสูงสุดมาเป็นจำนวนที่สุ่มจากเลขไม่เกิน 10 ข้อ หลังจากนั้นจะสุ่มเลือกข้อสอบเพียงข้อเดียวจากกลุ่มข้อสอบที่มีค่าสารสนเทศสูงสุดนี้มาเป็นข้อสอบข้อถัดไปให้กับผู้ทดสอบ โดยวิธีการนี้จะช่วยลดอัตราการใช้ข้อสอบของกลุ่มผู้ทดสอบที่มีความสามารถใกล้เคียงกัน

1.3 วิธี Within 0.1 Logits strategy เสนอโดย Lunz and Stahl (1998) เป็นวิธีควบคุมการใช้ข้อสอบ โดยสนใจกลุ่มข้อสอบที่มีความยากของข้อสอบใกล้เคียงกับค่าความสามารถของผู้ทดสอบ กล่าวคือ วิธีการนี้จะเลือกข้อสอบที่มีค่าความยากของข้อสอบเท่ากับค่าความสามารถของผู้ทดสอบขณะนั้น แต่ถ้าเลือกไม่ได้ จะสุ่มเลือกข้อสอบข้อใดข้อหนึ่งจากกลุ่มข้อสอบที่มีค่าความยากของข้อสอบแตกต่างจากค่าความสามารถของผู้ทดสอบขณะนั้น ไม่เกิน 0.1

2. วิธีการคัดเลือกแบบมีเงื่อนไข (Conditional selection strategies) วิธีการควบคุมการใช้ข้อสอบกลุ่มนี้ เป็นวิธีการควบคุมความน่าจะเป็นของข้อสอบที่ได้รับเลือก และให้มีการใช้ข้อสอบเป็นไปตามที่กำหนด วิธีนี้รับรองว่าอัตราการใช้ข้อสอบจะไม่สูงเกิดกว่าค่าที่กำหนด โดยควบคุมอัตราการใช้ข้อสอบผ่านค่าพารามิเตอร์ควบคุมการใช้ข้อสอบ แต่มีข้อจำกัดคือ วิธีการมีความยุ่งยากซับซ้อนมาก เนื่องจากการคำนวณค่าพารามิเตอร์ควบคุมการใช้ข้อสอบ ต้องทำให้สถานการณ์จำลองที่มีการทำซ้ำจำนวนมาก และเมื่อองค์ประกอบต่าง ๆ ในการทดสอบเปลี่ยนแปลงไป เช่น มีการเพิ่ม-ลดข้อสอบในคลังข้อสอบ จำเป็นต้องคำนวณค่าพารามิเตอร์ควบคุมการใช้ข้อสอบใหม่ทุกครั้ง วิธีการในกลุ่มนี้ได้แก่

2.1 วิธีซิมสัน-เฮตเทอร์ (Sympson-Hetter procedure: SH) เสนอโดย Sympson and Hetter (1985) โดยใช้ความถี่ของข้อสอบที่จัดให้กับกลุ่มผู้ทดสอบจำนวนมากในสถานการณ์จำลอง เปรียบเทียบกับอัตราการแสดงข้อสอบเป้าหมายเพื่อหาพารามิเตอร์การแสดงข้อสอบ กระบวนการนี้ทำซ้ำ ๆ กันจนพารามิเตอร์การแสดงข้อสอบแต่ละข้อมีค่าระหว่าง 0 ถึง 1 พารามิเตอร์นี้ใช้สำหรับการเลือกข้อสอบในการทดสอบสถานการณ์จริงโดยสร้างตัวเลขสุ่มจากการแจกแจงเหมือนกัน (0 ถึง 1) เปรียบเทียบกับพารามิเตอร์การแสดงข้อสอบ ถ้าพารามิเตอร์การแสดงข้อสอบมากกว่าตัวเลขสุ่มจะจัดข้อสอบให้ แต่ถ้าพารามิเตอร์การแสดงข้อสอบน้อยกว่าตัวเลขสุ่มจะเลือกข้อสอบข้อใหม่แทน ข้อดีของวิธีนี้คือ ยอมให้มีการกำหนดอัตราการแสดงข้อสอบล่วงหน้า เพื่อยืนยันว่าการแจกแจงอัตราการแสดงข้อสอบควรมีลักษณะเดียวกัน วิธีนี้ยังเป็นวิธีพื้นฐานประยุกต์ไปสู่

วิธีการเลือกอย่างมีเงื่อนไขอื่น ๆ อีกด้วย (Sympson & Hetter, 1985 cited in Stocking & Lewis, 2000)

2.2 วิธีซิมสัน-เฮตเทอร์แบบมีเงื่อนไข (Conditional Sympson-Hetter Procedure) พัฒนามาจากวิธีซิมสัน-เฮตเทอร์ เสนอโดย Stocking and Lewis (1995) โดยใช้พารามิเตอร์การแสดงข้อสอบได้จากการประมาณระดับค่าความสามารถแทนการแจกแจงค่าความสามารถของผู้ทดสอบ ระหว่างขั้นตอนสถานการณ์จำลอง จะมีการสร้าง n × m เมทริกซ์ เมื่อแถวนอน n คือ จำนวนข้อสอบในคลังข้อสอบ และแถวตั้ง m คือ จำนวนค่าความสามารถที่กระจายบนช่วงการแจกแจงค่าความสามารถ ดังนั้น ข้อสอบแต่ละข้อจะมีพารามิเตอร์การแสดงข้อสอบที่แต่ละค่าความสามารถของผู้ทดสอบ

2.3 วิธีดาเวย์-พาร์เชลล์ (Davey-Parshall procedure) เสนอโดย Davey and Parshall (1995) เน้นการลดการทับซ้อนข้อสอบระหว่างกลุ่มผู้ทดสอบ โดยสร้างพารามิเตอร์การแสดงข้อสอบจาก n × n เมทริกซ์ เมื่อ n คือ จำนวนข้อสอบในคลังข้อสอบ เงื่อนไขพารามิเตอร์ในไดอะโกนอลเมทริกซ์ (Diagonal matrix) ควบคุมความถี่ของข้อสอบจากการเปรียบเทียบรายคู่ (Pairwise comparisons) ในสถานการณ์จำลองเพื่อควบคุมการทับซ้อนข้อสอบ ไดอะโกนอลเมทริกซ์มีพารามิเตอร์การแสดงข้อสอบปฏิบัติคล้ายพารามิเตอร์การแสดงข้อสอบในวิธีซิมสัน-เฮตเทอร์

2.4 วิธีโดยรวมเงื่อนไข (Triconditional procedure) เสนอโดย Parshall, Hogerty and Kromrey (1999) วิธีนี้รวมวิธีซิมสัน- เฮตเทอร์ วิธีซิมสัน-เฮตเทอร์แบบมีเงื่อนไข และ วิธีดาเวย์-พาร์เชลล์ เข้าไว้ด้วยกัน

2.5 วิธีมัลติโนเมียลแบบไม่มีเงื่อนไข (Unconditional multinomial procedure) เสนอโดย Stocking and Lewis (1995) วิธีนี้พารามิเตอร์การแสดงข้อสอบได้จากวิธีซิมสัน-เฮตเทอร์ในช่วงระยะแรก ระยะที่ 2 ใช้รูปแบบมัลติโนเมียลในการเลือกข้อสอบ วิธีนี้ประยุกต์ไปสู่การควบคุมการแสดงข้อสอบมัลติโนเมียลแบบมีเงื่อนไข

2.6 วิธีจำกัดสารสนเทศสูงสุด (Restricted maximum information procedure) เสนอโดย Revuelta and Ponsoda (1998) วิธีนี้กำหนดอัตราการแสดงข้อสอบสูงสุดไว้ล่วงหน้าและจำกัดจำนวนครั้งการจัดข้อสอบให้ผู้ทดสอบในชุดแบบทดสอบทั้งหมด การเลือกข้อสอบพิจารณาจากสารสนเทศสูงสุด ข้อสอบที่มีอัตราการแสดงข้อสอบครั้งหลังสุดต่ำกว่าอัตราการแสดงข้อสอบสูงสุดจะนำไปรวมไว้ในคลังข้อสอบ ถ้าข้อสอบมีอัตราการแสดงข้อสอบมากกว่าอัตราการแสดงข้อสอบสูงสุด ข้อสอบจะถูกคัดออกจากคลังข้อสอบและไม่นำมาพิจารณาในการเลือกข้อสอบ

2.7 วิธี Item Eligibility (IE) เสนอโดย van der Linden and Veldkamp, (2004) เป็นวิธีการควบคุมการใช้ข้อสอบที่ขยายขีดความสามารถมาจาก (Hetter & Sympson, 1997) เพื่อ

แก้ไขปัญหาความล่าช้าของเวลาในการนำข้อสอบมาแสดง วิธีการของ IE จะคัดเลือกข้อสอบที่มีโอกาสได้รับเลือกก่อน แล้วค่อยส่งข้อสอบดังกล่าวไปยังวิธี SH ซึ่งจะลดความล่าช้าของเวลาได้ ดังนั้น วิธีการ IE มีขั้นตอนที่ยุ่งยาก มีการกำหนดอัตราการใช้ข้อสอบสูงสุดแบบคงที่ r^{max} แสดงดังสมการ 4

$$k_i^{(j+1)} = \begin{cases} 1 & if\ p^{(1\ldots j)}(A_i)/k_i^{(j)} \leq r^{max} \\ r^{max} k_i^{(j)} / p^{(i\ldots j)}(A_i) & if\ p^{(1\ldots j)}(A_i)/k_i^{(j)} > r^{max} \end{cases} \quad (4)$$

เมื่อ $k_i^{(j)}$ แทน พารามิเตอร์ความคุมการใช้ข้อสอบข้อที่ i

r^{max} แทน อัตราการใช้ข้อสอบสูงสุดที่กำหนดคือ 0.2

A_i แทน จำนวนผู้ทดสอบคนที่ i

$p^{(1\ldots j)}(A_i)$ แทน ความน่าจะเป็นของ A_i คำนวณจากผู้ทดสอบคนที่ 1...j

3. วิธีคัดเลือกแบบตามระดับชั้น (Stratified strategies) วิธีการควบคุมการใช้ข้อสอบในกลุ่มนี้ มีแนวคิดแบ่งข้อสอบในคลังข้อสอบเป็นชั้น ๆ ตามค่าอำนาจจำแนกของข้อสอบ และควบคุมการใช้ข้อสอบโดยให้เลือกข้อสอบจากทุกชั้นมาใช้ในการทดสอบ วิธีการนี้จะเพิ่มโอกาสให้ข้อสอบที่มีค่าสารสนเทศต่ำมีโอกาสถูกเลือกมาใช้สอบเพิ่มมากขึ้น วิธีการในกลุ่มนี้ได้แก่

3.1 วิธีเลือกตามระดับชั้นของค่าอำนาจจำแนก (a-Stratified Method: a-STR) วิธีนี้เสนอโดย Chang and Ying (1996) ประยุกต์มาจากรูปแบบปรับระดับชั้นของ Weiss (1974) เป็นวิธีการเลือกข้อสอบและให้ประสิทธิภาพในการควบคุมการแสดงข้อสอบโดยอัตโนมัติ เริ่มจากการแบ่งชั้นคลังข้อสอบตามค่าอำนาจจำแนกของข้อสอบ และแบ่งแบบทดสอบในแต่ละชั้นของข้อสอบ ชั้นแรกบรรจุข้อสอบอำนาจจำแนกต่ำสุด ชั้นต่อไปบรรจุข้อสอบอำนาจจำแนกสูงกว่าชั้นแรก ทำเช่นนี้เรื่อย ๆ จนกระทั่งชั้นสุดท้ายบรรจุข้อสอบค่าอำนาจจำแนกสูงสุด การทดสอบจะดำเนินจากชั้นแรกจนถึงชั้นสุดท้าย วิธีนี้กระจายการใช้ข้อสอบในคลังข้อสอบให้เท่าเทียมกัน

3.2 วิธีเลือกตามระดับชั้นของค่าอำนาจจำแนกและค่าความยาก (a-Stratified Method with b Blocking : ab-STR) วิธีนี้นำเสนอโดย Chang, Qian and Ying (2001) ประยุกต์มาจากวิธี a-SRT สำหรับวิธี ab-STR คลังข้อสอบจะถูกแบ่ง 4 ระดับเหมือนกัน แต่จัดบล๊อค 90 บล็อกด้วยกลุ่มค่าความยาก (b) เหมือนกัน แต่ละกลุ่มประกอบด้วยข้อสอบ 4 ข้อที่มีค่า b เหมือนกันมากที่สุด ดังนั้น กลุ่มที่ 1 ประกอบด้วยข้อสอบ 4 ข้อที่มีค่า b ต่ำที่สุดและกลุ่มที่ 90 บรรจุข้อสอบ 4 ข้อที่มีค่า b สูงสุด ข้อสอบค่า a ต่ำสุดมีอยู่ในทุก ๆ กลุ่มระดับแรก ข้อสอบค่า a ต่ำเป็นอันดับ 2 มีอยู่ในทุก ๆ กลุ่มระดับที่ 2 ค่า a อื่น ๆ ถูกสร้างในทำนองเดียวกัน

3.3 วิธีการออกแบบ 0-1-SRT เสนอโดย van der Linden and Chang (2003) ใช้วิธีการที่แตกต่างกันไปที่อยู่การแบ่งชั้นคลังข้อสอบ ขึ้นอยู่กับเทคนิคของการเขียนโปรแกรมเชิงเส้น

0-1 (LP) นี้ วิธี Stratifies item bank ได้อย่างดีที่สุดโดยร่วมกับการรวมกันของ ค่าอำนาจจำแนก (a) และ ค่าความยาก (b) โดยเฉพาะการกำหนดค่าเป้าหมายของแต่ละชั้นซึ่งมีการกำหนดเช่นเดียวกับค่าภายในแต่ละชั้น

4. วิธีคัดเลือกแบบผสม (Combined strategies) วิธีการควบคุมการใช้ข้อสอบในกลุ่มนี้เป็นการนำวิธีการควบคุมใช้ข้อสอบตั้งแต่ 2 วิธีการขึ้นไป มาใช้ร่วมกัน ซึ่งจะนำจุดเด่นของแต่ละวิธีมาใช้ วิธีการในกลุ่มนี้ได้แก่

4.1 วิธีจำกัดความก้าวหน้า (Progressive restricted procedure) เสนอโดย Revuelta and Ponsoda (1998) โดยรวมวิธีจำกัดสารสนเทศสูงสุด (Restricted maximum information procedure) และวิธีสุ่มเลือกแบบก้าวหน้า (Progressive procedure) ไว้ด้วยกัน วิธีจำกัดสารสนเทศสูงสุดนำมาใช้ เพื่อไม่ให้ข้อสอบมีอัตราการแสดงข้อสอบเกินอัตราการแสดงข้อสอบสูงสุดขณะที่วิธีสุ่มเลือกแบบก้าวหน้าใช้สำหรับเลือกข้อสอบให้ผู้ทดสอบ

4.2 วิธีเลือกตามระดับชั้นของค่าอำนาจจำแนกร่วมกับวิธีซิมสัน-เฮตเทอร์ (a-Stratified with Sympson-Hetter Method: a-STR-SH) เสนอโดย Leung, Chang, and Hau (2002) วิธีนี้จัดการกับปัญหาที่เกิดจากขั้นตอนการแบ่งชั้น โดยจำลองการศึกษาแล้ว ปรากฏว่า วิธี a-SRT-SH มีประสิทธิภาพมากกว่าการใช้วิธี SH หรือวิธี a-SRT อย่างเดียว โดยอัตราการทับซ้อนของข้อสอบและการควบคุมการแสดงข้อสอบที่มากเกินไปลดลง

4.3 วิธีซิมสันเฮตเทอร์ร่วมกับการหมุนคลังข้อสอบ (SH with rotating item banks) เสนอโดย Barrada, Olea, and Abda (2008) โดยรวมขั้นตอน SH และกลยุทธ์ของการหมุนคลังข้อสอบ โดยอัตราการเปิดรับรายการสูงสุดในการหมุนคลังข้อสอบ แสดงให้เห็นว่าวิธีการที่ร่วมกันดำเนินการในทำนองเดียวกันนี้อย่างใดอย่างหนึ่งกับขั้นตอน SH หรือกลยุทธ์การหมุนคลังข้อสอบแสดงผลดีในแง่ของอัตราอัตราการทับซ้อนกันลดลง (ขึ้นอยู่กับการกระจายของค่าพารามิเตอร์คือค่าอำนาจจำแนก (a) และค่าความยาก (b) ในคลังข้อสอบ)

4.4 วิธีขยายช่วงระดับชั้น (Enhanced stratified procedure) เสนอโดย Leung, Chang and Hua (2003) เป็นการประยุกต์รวมวิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ตามระดับชั้นของค่าอำนาจจำแนกและวิธีควบคุมการแสดงข้อสอบซิมสัน-เฮตเทอร์เข้าไว้ด้วยกัน

4.5 วิธีการแบ่งระดับชั้นข้อสอบร่วมกับวิธีข้อสอบที่มีสิทธิ์รับเลือก (MSTR-IE-MCAT) เสนอโดย Huebner et al. (2016) เป็นวิธีการเลือกข้อสอบและให้ประสิทธิ์ภาพในการควบคุมการแสดงข้อสอบโดยอัตโนมัติ เริ่มจากการแบ่งชั้นคลังข้อสอบตามค่าอำนาจจำแนกของข้อสอบ และแบ่งแบบทดสอบเข้าในแต่ละชั้นคลังข้อสอบ ชั้นแรกบรรจุข้อสอบอำนาจจำแนกต่ำสุด ชั้นต่อไปบรรจุข้อสอบอำนาจจำแนกสูงกว่าชั้นแรก ทำเช่นนี้เรื่อย ๆ จนกระทั่งชั้นสุดท้ายบรรจุข้อสอบค่าอำนาจจำแนกสูงสุด การทดสอบจะดำเนินจากชั้นแรกจนถึงชั้นสุดท้าย วิธีนี้กระจายการใช้ข้อสอบในคลัง

ข้อสอบให้เท่าเทียมกัน แต่โอกาสที่ข้อสอบที่มีค่าอำนาจจำแนกน้อยไม่ถูกนำไปใช้ทดสอบ จึงนำวิธีข้อสอบที่มีสิทธิ์รับเลือก (IE) มาร่วมในการควบคุม

5. วิธีคัดเลือกแบบปรับเหมาะหลายขั้นตอน (Multiple stage adaptive test designs) วิธีการในกลุ่มนี้ถูกพัฒนาขึ้น โดยรวมลักษณะของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์และการทดสอบแบบดั้งเดิม (ทดสอบบนกระดาษ) เข้าด้วยกัน คือใช้กระบวนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ในการหาความสามารถของผู้ทดสอบและใช้ผู้ที่ได้จากการทดสอบเบื้องต้นบนกระดาษหรือการทดสอบคู่ขนาน มาเป็นข้อมูลในการคัดเลือกข้อสอบข้อถัดไปที่เหมาะสมกับผู้ทดสอบ วิธีการนี้ได้แก่ วิธี Computerize Adaptive Sequential Testing (CAST) เสนอโดย Luecht et al. (1996)

2. การควบคุมสมดุลเนื้อหา (Content balancing)

การควบคุมสมดุลเนื้อหาเป็นสิ่งสำคัญมาก จะบอกถึงประสิทธิภาพการเลือกข้อสอบการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ ถึงแม้ผู้ทดสอบจะได้รับชุดแบบทดสอบแตกต่างกัน แต่จำเป็นต้องครอบคลุมขอบเขตเนื้อหาตามกำหนด เพื่อให้แน่ใจว่าผู้ทดสอบแต่ละคนได้รับการทดสอบด้วยแบบทดสอบคู่ขนานกัน และผลการทดสอบสามารถนำไปเปรียบเทียบกันได้ การเลือกข้อสอบมักเลือกข้อสอบสารสนเทศสูงสุดใกล้เคียงค่าประมาณความสามารถของผู้ทดสอบ การเลือกข้อสอบวิธีนี้จะไม่ได้รับสมดุลเนื้อหาตามกำหนด ส่งผลให้ผู้ทดสอบที่มีความสามารถน้อยหรือไม่มีความรู้ในการตอบข้อสอบตรงส่วนเนื้อหาที่ขาดไป อาจได้รับการประมาณค่าความสามารถสูงเกินความสามารถจริง ในขณะที่ผู้ทดสอบความสามารถสูงหรือมีความรู้ในการตอบข้อสอบตรงส่วนเนื้อหาที่ขาดไปจะได้รับการประมาณค่าความสามารถต่ำกว่าความสามารถจริง ดังนั้น การเลือกข้อสอบในการทดสอบแบบปรับเหมาะนอกจากจะต้องรักษาประสิทธิภาพการประมาณค่าความสามารถของผู้ทดสอบและควบคุมการใช้ข้อสอบแล้ว จำเป็นต้องควบคุมสมดุลเนื้อหาไปพร้อม ๆ กัน

การควบคุมสมดุลเนื้อหาไม่เป็นการควบคุมทางสถิติ (Non-Statistical constraint) หรือการควบคุมเนื้อหา (Content constraint) แนวคิดเกี่ยวกับการควบคุมสมดุลเนื้อหานี้ Stocking and Swanson (1993) อธิบายลักษณะของข้อสอบไม่ใช้สถิติสำหรับการจำแนกและเลือกข้อสอบ ได้แก่ 1) คุณสมบัติภายในข้อสอบ (Intrinsic item properties) กล่าวถึงลักษณะของข้อสอบด้านขอบเขตเนื้อหา 2) การควบคุมการทับซ้อน (Overlap constraints) กล่าวถึงข้อสอบมุ่งวัดสิ่งอื่นนอกเหนือจากลักษณะที่ต้องการวัด 3) การควบคุมชุดข้อสอบ (Item set constraints) กล่าวถึงการยอมให้กลุ่มข้อสอบแบ่งสิ่งเล้าหรือจัดสิ่งมุ่งหมายร่วมกัน ซึ่งจากการศึกษาวิธีการควบคุมสมดุลเนื้อหาในงานวิจัยต่าง ๆ พบได้ไม่มากนัก ได้แก่ 1) การทดสอบแบบปรับเหมาะแบบควบคุมสมดุลเนื้อหา (Constrained Computerized Adaptive Testing: CCAT) เสนอโดย Kingsbury and Zara, 1989 cited in Leung, Chang & Hau, 2003) เป็นวิธีเลือกข้อสอบที่เหมาะสมที่สุดจากขอบเขต

เนื้อหา โดยพิจารณาอัตราการแสดงข้อสอบครั้งหลังสุดและห่างจากเปอร์เซ็นต์ของการจัดข้อสอบตามกำหนดมากที่สุด 2) การทดสอบแบบปรับเหมาะแบบควบคุมสมดุลเนื้อหาประยุกต์ (Modified Constrained Computerized Adaptive Testing: MCCAT) เสนอโดย Leung, Chang and Hau (2003) เป็นวิธีประยุกต์มาจากการทดสอบแบบปรับเหมาะแบบควบคุมสมดุลเนื้อหา เพื่อลดการคาดเดาการเรียงเนื้อหาและให้การควบคุมสมดุลเนื้อหาให้ดียิ่งขึ้น ข้อสอบเลือกมาจากขอบเขตเนื้อหาทั้งหมดที่ยังไม่ถูกใช้เพื่อขจัดผลจากการเรียงลำดับ และ 3) รูปแบบมัลติโนเมียลประยุกต์ (Modified Multinomial Model : MMM) เสนอโดย Ankenmann and Spray (Cheng, Ankenmann & Spray, 1999 cited in Leung, Chang & Hau, 2003) เป็นวิธีการทดสอบแบบปรับเหมาะแบบควบคุมสมดุลเนื้อหาสามารถคาดเดาอันดับการเรียงเนื้อหา จึงได้พัฒนารูปแบบมัลติโนเมียลประยุกต์เพื่อต้องการให้เนื้อหามีความสมดุลกัน เมื่อขอบเขตเนื้อหามาถึงเปอร์เซ็นต์กำหนด จะมีการสร้างการแจกแจงมัลติโนเมียลเพื่อปรับเปอร์เซ็นต์ที่ยังไม่เต็มของขอบเขตเนื้อหาที่เหลืออยู่

3. การประเมินประสิทธิภาพการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

การหาประสิทธิภาพวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา จากการศึกษางานวิจัย มีการประเมินประสิทธิภาพ 5 ด้าน ดังนี้

3.1 จำนวนข้อสอบที่ใช้ในการทดสอบ (Test length) เป็นการหาผลรวมของจำนวนข้อสอบที่ผ่านทดสอบ (Chang & Ying, 1999) โดยคำนวนจากจำนวนข้อสอบของผู้ทดสอบแต่ละคนแล้วนำผลที่ได้มาค่า Medium ของแต่ละวิธี เนื่องจากเป็นข้อมูลแบบไม่ต่อเนื่อง จึงใช้สถิติ Mann-Whitney test ในการเปรียบเทียบเพื่อหาความต่างของจำนวนข้อสอบที่ใช้ในการทดสอบ

3.2 เวลาที่ใช้ในการทดสอบ (Test time) เป็นการหาค่าเฉลี่ยของเวลาที่ทดสอบแต่ละคน (Chang & Ying, 1999) โดยคำนวณจากการหาส่วนต่างระหว่างเวลาเริ่มต้นและเวลาสิ้นสุดของการทดสอบ เมื่อได้หน่วยเวลาจึงนำมาคูณกับ 1,440 เพื่อแปลงเป็นวินาที จึงใช้สถิติ t-test ในการเปรียบเทียบเพื่อหาความต่างของเวลาที่ใช้ในการทดสอบ

3.3 อัตราการแสดงการใช้ข้อสอบสูงสุด (Maximum item exposure rate: MIER) (Ozturk & Dogan, 2015) ถ้าข้อสอบข้อใดในคลังมีอัตราการใช้สูง แสดงถึงแนวโน้มข้อสอบนั้นมีการนำไปใช้ในการทดสอบมากเกินไป ส่งผลกระทบต่อความปลอดภัยของข้อสอบ ขาดสมดุลการใช้ข้อสอบในคลังข้อสอบ และสะท้อนถึงประสิทธิภาพวิธีเลือกข้อสอบไม่เหมาะสม วิธีการหาค่าอัตราการแสดงการใช้ข้อสอบสูงสุด คำนวณได้จาก สมการ (5) และ (6)

$$ko_j = \frac{number\ of\ times\ the\ j\ item\ id\ used}{m} \tag{5}$$

เมื่อ ko_j แทน อัตราการใช้ข้อสอบ
m แทน จำนวนผู้ทดสอบทั้งหมด

$$MIER = Max(ko_j) \tag{6}$$

เมื่อ $MIER$ แทน อัตราการแสดงการใช้ข้อสอบสูงสุด
$Max(ko_j)$ แทน ค่าอัตราการแสดงการใช้ข้อสอบสูงสุดข้อที่ ko_j

การแจกแจงอัตราการแสดงข้อสอบ (Item exposure rate distribution) เป็นลักษณะการกระจายอัตราการแสดงของข้อสอบในคลังข้อสอบ บ่งบอกถึงประสิทธิภาพการใช้ข้อสอบในการควบคุมการใช้ข้อสอบโดยรวม พิจารณาจากสถิติไคสแควร์ (Ozturk & Dogan, 2015) สามารถคำนวณได้จากสมการ 7

$$\chi^2 = \sum_{j=1}^{n}\left(\frac{ko_j - \overline{ko_j}}{\overline{ko}_j}\right)^2 \tag{7}$$

เมื่อ ko_j แทน อัตราการใช้ข้อสอบ
$\overline{ko_j}$ แทน ค่าเฉลี่ยของอัตราการใช้ข้อสอบ

การเปรียบเทียบการแจกแจงอัตราการแสดงข้อสอบ 2 วิธี ได้แก่ วิธีที่ 1 กับวิธีที่ 2 โดยใช้อัตราส่วน F (Ratio) จะพิจารณาจากอัตราส่วนของค่า χ^2 วิธีที่ 1 หารด้วย χ^2 วิธีที่ 2 ดังนั้น

$$F_{1.2} = \frac{\chi_1^2}{\chi_2^2} \tag{8}$$

เมื่อ $F_{1.2}$ แทน อัตราส่วนการแจกแจงการแสดงข้อสอบวิธีที่ 1 หารด้วยวิธีที่ 2
χ_1^2 แทน การแจกแจงอัตราการแสดงข้อสอบวิธีที่ 1
χ_2^2 แทน การแจกแจงอัตราการแสดงข้อสอบวิธีที่ 2
ถ้า $F < 1$ แสดงว่า วิธีที่ 1 มีการควบคุมการใช้ข้อสอบโดยรวมดีกว่าวิธีที่ 2

3.4 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) คือค่าความถูกต้องของค่าประมาณความสามารถของผู้ทดสอบจำนวนหนึ่ง ซึ่งมีค่าเท่ากับผลต่างกำลังสองเฉลี่ย ระหว่างค่าประมาณความสามารถมีค่าเข้าใกล้ 0 แสดงว่า ค่าประมาณความสามารถของผู้ทดสอบมีค่าใกล้เคียงกับค่าความสามารถที่แท้จริงของผู้ทดสอบ (Ozturk & Dogan, 2015) สามารถคำนวณได้จากสมการ 9

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{\theta}_i - \theta_i)^2}{n}} \quad (9)$$

เมื่อ $\hat{\theta}_i$ แทน ค่าประมาณความสามารถของผู้ทดสอบคนที่ i

θ_i แทน ค่าความสามารถที่แท้จริงของผู้ทดสอบคนที่ i

n แทน จำนวนผู้ทดสอบทั้งหมด

ค่าความลำเอียงเฉลี่ย (Average Bias) เป็นค่าขอกความเที่ยงตรงของค่าประมาณความสามารถที่แท้จริงของผู้ทดสอบ แสดงเป็นผลต่างเฉลี่ยระหว่างค่าประมาณความสามารถและค่าความสามารถที่แท้จริงของผู้ทดสอบ สามารถบอกทิศทางการประมาณค่า ให้ผลสูงหรือต่ำกว่าค่าความสามารถที่แท้จริงของผู้ทดสอบ ถ้าค่าความลำเอียงเฉลี่ยเข้าใกล้ศูนย์ จะสะท้อนให้เห็นถึงประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบ (Ozturk & Dogan, 2015) ดังสมการ 10

$$Bias = \frac{\sum_{i=1}^{n}(\hat{\theta}_i - \theta_i)}{n} \quad (10)$$

เมื่อ n แทน จำนวนผู้ทดสอบทั้งหมด

$\hat{\theta}_i$ แทน ค่าประมาณความสามารถของผู้ทดสอบคนที่ i

θ_i แทน ค่าความสามารถที่แท้จริงของผู้ทดสอบคนที่ i

4. งานวิจัยที่เกี่ยวข้องของวิธีการควบคุมการใช้ข้อสอบและวิธีความคุมสมดุลเนื้อหา

Lee, Ip and Fuh (2008) ได้ศึกษาวิธีในการควบคุมการใช้ข้อสอบซ้ำในการทดสอบปรับเหมาะแบบพหุมิติ โดยนำเสนอพื้นฐานของการแบ่งชั้นของข้อสอบด้วยฟังก์ชันของเวกเตอร์กับค่าอำนาจจำแนกของข้อสอบ การศึกษาครั้งนี้เปรียบเทียบอัตราการใช้ข้อสอบซ้ำ ซึ่งใช้การควบคุมการใช้ข้อสอบ 2 วิธี ได้แก่ 1) วิธี a-Stratified Adaptive Testing (ASTR) และ 2) วิธี D-optimality โดยใช้เกณฑ์การคัดเลือกข้อสอบด้วยวิธี Match-b ซึ่งเปรียบเทียบประสิทธิภาพ 2 ด้าน ได้แก่ 1) ความยาวของข้อสอบ 30 ข้อ และ 60 ข้อ และ 2) ความสัมพันธ์ระหว่างวิธีการควบคุมการใช้ข้อสอบ 2 วิธี

กับวิธีการคัดเลือกข้อสอบ ปรากฏว่า ผลการคัดเลือกข้อสอบวิธี Match-b ที่ใช้กับวิธี ASTR มีประสิทธิภาพมากกว่าที่ใช้กับวิธี D-optimality ในด้านความยาวของข้อสอบ โดยพิจารณาจากค่า Mean Square Error (MSE) ซึ่งมีค่าน้อยกว่า ทางด้านความสัมพันธ์ของทั้ง 2 วิธี มีความสัมพันธ์ทางบวก และความยาวของข้อสอบไม่มีผลต่อความสัมพันธ์ สรุปว่า วิธีการคัดเลือกข้อสอบ Match-b สามารถนำไปประยุกต์ใช้กับวิธีการควบคุมข้อสอบโดยความยาวของข้อสอบไม่มีผลต่อวิธีการนี้

Belov and Armstrong (2009) ได้เสนอวิธีการใหม่สำหรับการควบคุมเนื้อหา (Content Constraints) โดยเปรียบเทียบวิธี Shadow CAT กับวิธี Monte Carlo CAT ที่ใช้ควบคุมเนื้อหา การศึกษาในครั้งนี้ ใช้คลังข้อสอบจากการสอบเข้าศึกษาต่อในด้านกฎหมาย (Law School Admission Test: LSAT) ผลการศึกษาปรากฏว่า 1) Monte Carlo CAT มีอัตราการใช้ข้อสอบซ้ำสูงสุดต่ำกว่า และมีอัตราการใช้ข้อสอบสูงสุดของคลังข้อสอบดีกว่า โดยเฉพาะอย่างยิ่งอัตราการใช้ข้อสอบซ้ำสูงสุดสำหรับ Monte Carlo CAT มีค่าต่ำกว่าวิธี Shadow CAT 2) เมื่อค่าโอกาสการเดาทั้ง Monte Carlo CAT และ Shadow CAT มี Bias เพิ่มขึ้นทั้งคู่ และ ส่วน MSE ของ Monte Carlo CAT แทบจะไม่มีการเปลี่ยนแปลง ขณะที่ของ Shadow CAT เพิ่มขึ้น และ 3) Monte Carlo CAT จำนวนผู้ทดสอบทั้งหมด 2,723 คน ทำข้อสอบพร้อม ๆ กัน ไม่เกิน 5 นาทีในการนำข้อสอบออกมาใช้ ขณะที่ Shadow CAT ใช้เวลา มากกว่า 20 นาที กับผู้ทดสอบเพียง 134 คน โดยสรุปแล้ว Monte Carlo CAT ให้ผลดีกว่าอย่างเห็นได้ชัด ในเรื่องอัตราการใช้ข้อสอบซ้ำ ดังนั้น Monte Carlo CAT จึงเหมาะสมกับการนำไปใช้ในทางปฏิบัติเพื่อบริหารการทดสอบ

Cheng and Chang (2009) ได้เสนอวิธีดัชนีลำดับความสำคัญสูงสุด (Maximum Priority Index: MPI) สำหรับการคัดเลือกข้อสอบที่มีข้อบังคับอย่างเข้มงวด การวิจัยครั้งนี้ศึกษาผ่านการจำลองข้อมูลเพื่อเปรียบเทียบวิธีการจัดการข้อบังคับ 2 วิธี ได้แก่ 1) วิธีดัชนีลำดับความสำคัญสูงสุด (Maximum Priority Index: MPI) และ 2) วิธี Weighted Deviation Modeling (WMD) โดยใช้วิธีการคัดเลือกข้อสอบข้อถัดไปแบบ Maximum Information (MI) และวิธี Randomized เป็นฐานในการเปรียบเทียบ โดยใช้ข้อมูลจากคลังข้อสอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ในวิชาพื้นฐานพีชคณิต คลังข้อสอบประกอบด้วยข้อสอบจาก 3 กลุ่ม เนื้อหาวิชาแต่ละกลุ่มแบ่งออกเป็น 4 ถึง 10 ในขอบเขตเนื้อหา ผลการศึกษา ปรากฏว่า 1) วิธี MPI และวิธี WDM ลดการฝ่าฝืนข้อบังคับในการเปรียบเทียบกับวิธี MI และวิธี Randomized และผลการทดสอบแสดงให้ว่า MPI มีประสิทธิภาพดีกว่า WDM ในทุกด้าน โดยมีอัตราการใช้ข้อสอบซ้ำสูงสุด การใช้ข้อสอบซ้ำเกินที่กำหนด และอัตราการทับซ้อนของแบบทดสอบต่ำกว่าวิธี WDM นอกจากนี้ทุกวิธียกเว้นวิธี Randomized มีสัดส่วนของการข้อสอบที่ไม่ได้นำไปใช้สูงมากกว่า 50% แสดงให้เห็นว่าข้อสอบในคลังข้อสอบ ถูกใช้อย่างไม่มีประสิทธิภาพ ซึ่งปัญหานี้สามารถทำให้ลดลงได้โดยการนำไปรวมเข้ากับวิธี a-Stratified Design ของ Chang and Ying (1999) จากข้อค้นพบของการศึกษาครั้งนี้ วิธี MPI

สามารถปรับให้เข้ากับข้อบังคับที่ไม่ใช่ทางสถิติ แต่สามารถควบคุมได้หลายอย่างพร้อมกัน เช่น ความสมดุลของเนื้อหา การควบคุมการใช้ข้อสอบซ้ำ ความสมดุลของคำตอบ วิธีนี้สามารถนำไปปรับใช้กับโปรแกรมการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่มีอยู่ในปัจจุบันได้สะดวก โดยไม่จำเป็นต้องปรับค่าน้ำหนักความสัมพันธ์ระหว่างข้อบังคับและข้อสอบที่อยู่ในคลังข้อสอบ

Deng, Ansley, and Chang (2010) ได้ศึกษาเปรียบเทียบประสิทธิภาพการควบคุมการใช้ข้อสอบ 3 วิธี ได้แก่ 1) วิธี a-Stratified Adaptive Testing (ASTR) 2) วิธีการคัดเลือกอย่างสุ่ม (RAN) และ 3) วิธี The Low a-Stratified (USTR) ซึ่งพิจารณาค่าความแปรปรวนของความคลาดเคลื่อน (MSE) ในการประมาณค่าความสามารถของผู้ทดสอบ และค่าความน่าเชื่อถือของค่าประมาณความสามารถของผู้ทดสอบ ผลจากการศึกษา ปรากฏว่า วิธี USTR สามารถลดความแปรปรวนของค่าความคลาดเคลื่อนได้ เมื่อจำนวนข้อสอบในคลังมีขนาดเล็ก และประสิทธิภาพการประมาณค่า 3 วิธีนี้ไม่แตกต่างกัน

Murphy, Dodd, and Vaughn (2010) ได้ศึกษาเปรียบเทียบประสิทธิภาพของเทคนิคการคัดเลือกข้อสอบ 3 วิธี คือ 1) Maximum Fisher's Information (MFI) 2) Maximum Posterior Weighted Information (MPWI) และ 3) Minimum Expected Posterior Variance (MEPV) ใน การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ที่ข้อสอบถูกจัดกลุ่มเป็นแบบทดสอบย่อยภายใต้โมเดล IRT และ TRT ดำเนินการวิจัยผ่านการศึกษาในสถานการจำลองโดยใช้วิธีมอนติคาร์โลเพื่อสร้างชุดข้อมูลแบบแผนการตอบจำนวน 10 กลุ่ม แต่ละกลุ่มมีผู้ทดสอบจำนวน 1,000 คน ข้อมูลทั้ง 10 กลุ่ม สำหรับทดลองซ้ำ 10 ครั้ง ในแต่ละเงื่อนไขการศึกษา ผลการศึกษาปรากฏว่า เมื่อใช้โมเดล IRT กับข้อมูลที่ฝ่าฝืนข้อตกลงเบื้องต้นเรื่อง local item dependence (LID) ผลการประมาณค่าความสามารถมีแนวโน้มที่จะมีค่ามากกว่าค่าที่เป็นจริง (overestimate) เมื่อเปรียบเทียบประสิทธิภาพของการคัดเลือกข้อสอบทั้งสามวิธี ปรากฏว่า ไม่แตกต่างกัน

Leroux, Lopez, Hembry, and Dodd (2013) ได้เปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบระหว่างวิธีการ Progressive-Restricted Standard Error (PR-SE) เกับวิธีการควบคุมการใช้ข้อสอบทั่วไป คือ 1) วิธีการ Randomesque 2) วิธีการ Sympson-Hetter (SH) และ 3) วิธีการที่ไม่มีการควบคุมการใช้ข้อสอบภายใต้โมเดลโลจิสติก 3 พารามิเตอร์ และเปรียบเทียบในกรณีคลังข้อสอบขนาดเล็กเทียบกับขนาดใหญ่ ผลจากการศึกษา ปรากฏว่า วิธีการ PR-SE สามารถเข้าควบคุมการใช้ข้อสอบของคลังได้เกือบทุกข้อ แต่วิธีการ Randomesque กับวิธี SH สามารถเข้าควบคุมการใช้ข้อสอบของข้อสอบในคลังได้เพียง 52% และ 80% ในกรณีที่คลังข้อสอบมีขนาดเล็กและขนาดใหญ่

Ozturk and Dogan (2015) ได้เปรียบเทียบประสิทธิภาพการประมาณค่าความสามารถของผู้ทดสอบ และการควบคุมการใช้ข้อสอบระหว่างวิธีการควบคุมการใช้ข้อสอบ 3 วิธี ได้แก่ 1) วิธี

Randomesque 2) วิธี Sympson-Hetter และ 3) วิธี Fade-Away โดยใช้คลังข้อสอบจำลองด้วยการสุ่มค่าอำนาจจำแนกของข้อสอบ (a) จากการแจกแจงในช่วง 0.50-2.00 สุ่มค่าการเดาของข้อสอบ (c) จากการแจกแจงในช่วง 0.05-0.20 ในส่วนของการจำลองค่าความยากของข้อสอบ (b) แบ่งเป็น 2 ส่วน ได้แก่ 1) คลังข้อสอบที่มีค่าความยากปานกลางจากการสุ่มค่า (b) ในช่วง -3.00-3.00 และ 2) คลังข้อสอบที่มีค่าความยากสูงจากการสุ่มค่า (b) ผลการศึกษาพบว่า วิธีการ Fade-Away สามารถควบคุมการใช้ข้อสอบได้ดีกว่า 2 วิธี และประสิทธิภาพการประมาณค่าความสามารถของผู้ทดสอบ 3 วิธี ไม่มีความแตกต่างกัน

Huebner, Wang, Quinlan, and Seubert (2016) ได้เปรียบเทียบประสิทธิภาพการควบคุมการใช้ข้อสอบ 3 วิธี ได้แก่ 1) วิธี D-optimality 2) วิธี a-Stratification (MSTR) และ 3) วิธี a-Stratification (MSTR) ร่วมกับ Item Eligibility (IE) หรือ MSTR-IE โดยใช้ความยาวข้อสอบ 20, 30 และ 60 ค่าความสามารถ 0.30, 0.60 และ 0.80 ตามลำดับ ประมาณค่าความสามารถใช้วิธี MLE ปรากฏว่า ความยาวของข้อสอบที่ 20 และ 30 MSRT มีประสิทธิภาพด้าน Bias, Maximum item exposure rate (MIER), item with exposure rate over (OEXP) ส่วนความยาวข้อสอบที่ 30 ข้อ วิธี MSTR-IE จะมีประสิทธิภาพกว่า D-optimality และ MSRT

จากการศึกษาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาของการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ ปรากฏว่า ผู้วิจัยส่วนใหญ่ได้ศึกษาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา ใน 3 ลักษณะ ได้แก่ 1) การพัฒนาวิธีการใหม่ 2) การเปรียบเทียบแต่ละวิธีการเพื่อหาประสิทธิภาพ และ 3) การนำวิธีการที่มีอยู่ไปประยุกต์

วิธีการควบคุมการใช้ข้อสอบ (Item Exposure) จากงานวิจัยของ Georgiadou et al., (2007) ได้แบ่งวิธีการควบคุมการใช้ข้อสอบเป็น 5 กลุ่ม ได้แก่ 1) วิธีการคัดเลือกแบบสุ่ม (Randomization Strategies) 2) วิธีการคัดเลือกแบบมีเงื่อนไข (Conditional Selection Strategies) 3) วิธีคัดเลือกแบบตามระดับชั้น (Stratified Strategies) 4) วิธีคัดเลือกแบบผสม (Combined Strategies) และ 5) วิธีคัดเลือกแบบปรับเหมาะหลายขั้นตอน (Multiple Stage Adaptive Test Designs) วิธีโหลดบาลานซ์ซิ่งอยู่ในกลุ่มวิธีการคัดเลือกแบบมีเงื่อนไข

วิธีการควบคุมสมดุลเนื้อหา (Content balancing) จากการศึกษางานวิจัย ได้นำเสนอวิธีการควบคุมสมดุลเนื้อหา 3 วิธี ได้แก่ 1) วิธีการทดสอบแบบปรับเหมาะแบบควบคุมสมดุลเนื้อหา (Constrained Computerized Adaptive Testing: CCAT) 2) วิธีการทดสอบแบบปรับเหมาะแบบควบคุมสมดุลเนื้อหาประยุกต์ (Modified Constrained Computerized Adaptive Testing: MCCAT) และ 3) วิธีรูปแบบมัลติโนเมียลประยุกต์ (Modified Multinomial Model: MMM)

ตอนที่ 3 ทฤษฎีตอบสนองข้อสอบแบบพหุมิติ

1. โมเดลการตอบสนองข้อสอบแบบพหุมิติ

การเลือกทฤษฎีการตอบสนองข้อสอบมาวิเคราะห์ข้อสอบความรู้ทางไอซีที ซึ่งเป็นรายวิชาเทคโนโลยีสารสนเทศ ในระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) เป็นวิชาที่มีการเรียนการสอนในประเภทวิชาบริหารธุรกิจ ประเภทวิชาเทคโนโลยีสารสนเทศ และประเภทวิชาช่างอุตสาหกรรม จำเป็นต้องเข้าใจถึงแนวคิดพื้นฐานของลักษณะข้อสอบที่มีรูปแบบเป็นเอกมิติหรือพหุมิติ โดยมีการพิจารณา 2 ประเด็นคือ 1) แนวคิดพื้นฐานของลักษณะเอกมิติและลักษณะพหุมิติ และ 2) การแปลความหมายตามลักษณะเอกมิติและลักษณะพหุมิติ ดังนี้

1.1 แนวคิดพื้นฐานของลักษณะเอกมิติและลักษณะพหุมิติ

1.1.1 แนวคิดพื้นฐานของลักษณะเอกมิติ โดยลักษณะโมเดลพื้นฐานของความเป็นเอกมิติมีข้อตกลงเบื้องต้นคือ คุณลักษณะแฝงที่จะวัดไม่มีความสัมพันธ์กัน ซึ่งมีอยู่ 2 ลักษณะคือ 1) ความเป็นเอกมิติ (Composite approach) มีลักษณะของข้อสอบวัดคุณลักษณะแฝงเดียว 2) ความเป็นเอกมิติแบบต่อเนื่อง (Consecutive approach) มีลักษณะของข้อสอบวัดคุณลักษณะแฝงเดียว แต่มีหลายคุณลักษณะแฝง ทั้งนี้แต่ละคุณลักษณะแฝงไม่มีความสัมพันธ์กัน

1.1.2 แนวคิดพื้นฐานของลักษณะพหุมิติ โดยลักษณะโมเดลพื้นฐานของความเป็นพหุมิติ มีข้อตกลงเบื้องต้นคือ คุณลักษณะที่จะวัดมีความสัมพันธ์กัน ซึ่งมีอยู่ 2 ลักษณะคือ 1) ความเป็นพหุมิติแบบคุณลักษณะแฝง มีความสัมพันธ์กัน (Between-items multidimensionality) มีลักษณะของข้อสอบวัดคุณลักษณะแฝงเดียว ซึ่งคุณลักษณะแฝงเดียวมีมากกว่า 1 คุณลักษณะแฝง ทั้งนี้แต่ละคุณลักษณะแฝงมีความสัมพันธ์กัน 2) ความเป็นพหุมิติแบบข้อสอบวัดหลายคุณลักษณะแฝงและคุณลักษณะแฝงมีความสัมพันธ์กัน (Within-items multidimensionality) มีลักษณะของข้อสอบวัดหลายคุณลักษณะแฝง ซึ่งคุณลักษณะแฝงเดียวมีมากกว่า 1 คุณลักษณะแฝง ทั้งนี้แต่ละคุณลักษณะแฝงมีความสัมพันธ์กันแบบข้ามคุณลักษณะแฝงกันภายใน

1.2 การแปลความหมายตามลักษณะเอกมิติและลักษณะพหุมิติ

1.2.1 การแปลความหมายตามลักษณะเอกมิติ มีการแปลความหมาย 2 แบบ คือ 1) การแปลความหมายเอกมิติรวม และ 2) แปลความหมายเอกมิติแยกตามมิติ เช่น การวัดความรู้ทางด้านไอซีที ประกอบไปด้วย 6 ด้าน (มิติ) แต่ละด้านจะต้องไม่มีความสัมพันธ์กันตามข้อตกลง ถ้าแปลความหมายเอกมิติรวมจะอธิบายโดยรวม 6 ด้านเข้าด้วยกัน แต่ถ้าอธิบายตามความเหมาะสมแบบเอกมิติแยกตามมิติ ต้องแยกอธิบายแต่ละด้าน

1.2.2 การแปลความหมายตามลักษณะพหุมิติ มีการแปลความหมาย อธิบายแต่ละด้านโดยมีข้อตกลงพื้นฐานที่ว่า แต่ละด้านมีความสัมพันธ์กัน และจะแยกอธิบายความหมายแต่ละด้านหรือรวมอธิบายถึงความสัมพันธ์กันก็ได้ ตามลักษณะของเนื้อหาของข้อสอบที่จะแปลความ

แนวคิดของทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory Model: MIRT) เป็นแนวคิดที่ขยายขอบเขตมาจากโมเดลการตอบสนองข้อสอบแบบเอกมิติ (Unidimensional Item Response Theory Model: UIRT) เพื่อแก้ไขข้อจำกัดของข้อตกลงเบื้องต้นเกี่ยวกับความเป็นเอกมิติ ซึ่งทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ มาจากแนวคิดพื้นฐานที่สำคัญ 2 กลุ่ม ได้แก่ กลุ่มที่ 1 เป็นการศึกษาโมเดลการตอบสนองข้อสอบแบบพหุมิติ มีที่มาจากการวิเคราะห์องค์ประกอบ (Factor Analysis: FA) ซึ่งมีลักษณะการวิเคราะห์ความสัมพันธ์ขององค์ประกอบกับการวิเคราะห์โมเดลการตอบสนองข้อสอบแบบพหุมิติ (Spearman, 1928 cited in Reckase, 2009, p. 60) และกลุ่มที่ 2 เป็นการศึกษาโมเดลการตอบสนองข้อสอบแบบพหุมิติ ซึ่งได้รับอิทธิพลมาจากโมเดลการตอบสนองข้อสอบแบบเอกมิติ เช่นเดียวกับการวิเคราะห์องค์ประกอบแต่แนวคิดเกี่ยวกับโมเดลการตอบสนองข้อสอบแบบเอกมิติ ในช่วงแรกยังไม่เป็นที่นิยมเหมือนเช่นปัจจุบัน (Lazarsfeld, 1950 cited in Reckase, 2009, p. 60) จากสองแนวคิด โมเดลการตอบสนองข้อสอบแบบพหุมิติ สรุปได้ว่า เป็นการวิเคราะห์ข้อสอบเชิงสถิติหลายตัวแปร โดยเฉพาะการวิเคราะห์องค์ประกอบหรือโมเดลสมการเชิงโครงสร้าง (Structural Equation Modelling: SEM) หรือเป็นแนวคิดที่แผ่ขยายมาจากทฤษฎีตอบสนองข้อสอบแบบเอกมิติ (Reckase, 2009, p. 63)

1.3 การวิเคราะห์องค์ประกอบ (Factor analysis) จากการศึกษาของ Reckase (2009, p. 63) กล่าวว่า มีนักวิชาการหลายท่านได้ศึกษาและเห็นตรงกันว่าการวิเคราะห์องค์ประกอบ (Factor analysis) และทฤษฎีการตอบสนองข้อสอบ (MIRT) เป็นวิธีทางด้านสถิติที่พยายามจะระบุระดับสมมติฐาน เพื่อถอดแบบข้อมูลอันจะนำไปสู่การวิเคราะห์ข้อมูล โดยมาตรวัด (Scale) การวิเคราะห์องค์ประกอบจะต้องกำหนดจุดกำเนิด (Origin) และหน่วยการวัด (Unit of measurement) เพื่อพิจารณาถึงความแตกต่างของลักษณะตัวแปรตั้งต้น เช่น ความแตกต่างของค่าเฉลี่ย (Mean) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) เป็นต้น รวมถึงความเที่ยง (Reliability) ของการวิเคราะห์ข้อมูล เช่น การกำจัดตัวแปรแทรกซ้อนหรือตัวแปรที่ไม่จำเป็นออกจากโมเดลการวิเคราะห์โดยการปรับค่าสถิติให้เป็นมาตรฐาน โดยจะให้ความสำคัญกับเมทริกซ์ความสัมพันธ์ แต่ไม่ให้ความสำคัญกับความแตกต่างของค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของตัวแปรที่ศึกษา ต่อมา Herman (1976 Cited in Reckase, 2009, p. 664) ได้อธิบายลักษณะของการวิเคราะห์องค์ประกอบว่ามีหลักการเบื้องต้นของการวิเคราะห์องค์ประกอบ เป็นกระบวนการจัดกระทำกับข้อมูลหรือตัวแปรเพื่อจัดกลุ่มตัวแปรในการจำแนกประเภทหรือองค์ประกอบ ให้ได้จำนวนน้อยที่สุด

วิธีดังกล่าวสามารถทำได้โดยการศึกษาความสัมพันธ์ระหว่างตัวแปร โดยไม่ได้สนใจกับค่าพารามิเตอร์ข้อสอบที่มีลักษณะความสัมพันธ์กับข้อสอบโดยตรง เช่น ค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และค่าโอกาสการเดาของข้อสอบ เป็นต้น เมื่อพิจารณาตัวแปรแทรกซ้อนด้านอื่น ๆ ที่ตัดออกจากโมเดลการวิเคราะห์เพื่อแม่นยำจากการวิเคราะห์เพิ่มขึ้น (Carroll, 1945

cited in Reckase, 2009, p. 64) ต่อมาในปี ค.ศ. 1981 (Bock & Aitkin, 1981 Cited in Reckase, 2009, pp. 67-68) ได้นำเสนอแนวคิดที่คล้ายกันระหว่างการวิเคราะห์องค์ประกอบกับทฤษฎีการตอบสนองข้อสอบ และผลที่ได้จากการวิเคราะห์โมเดลการตอบสนองข้อสอบแบบพหุมิติ โดยระบุโมเดลปกติสะสม (Normal ogive model) สำหรับการวัดคุณลักษณะพหุมิติ รวมทั้งคุณลักษณะของข้อสอบ ในโมเดลของการวิเคราะห์องค์ประกอบและทฤษฎีการตอบสนองข้อสอบ ซึ่งส่วนใหญ่จะนำเสนอค่าอำนาจจำแนกของข้อสอบและค่าความยากของข้อสอบ เพื่อเป็นตัวแทนของค่าจุดแทนแกน (Intercept) และความชัน (Slopes) แต่ยังขาดในส่วนของโมเดลเต็มรูปแบบของโมเดลการตอบสนองข้อสอบแบบพหุมิติ แล้วใช้ความหมายของค่าพารามิเตอร์ข้อสอบว่าเป็นปฏิสัมพันธ์ระหว่างผู้ทดสอบกับข้อสอบ การตีความหมายยังเป็นลักษณะของการวิเคราะห์องค์ประกอบ นั่นคือ พิจารณาในลักษณะขององค์ประกอบ ซึ่ง Bock and Aitkin ได้ทำวิจัยขยายแนวคิดในลักษณะ 2 มิติ ของโมเดลปกติสะสมแบบ 2 พารามิเตอร์ ไปสู่โมเดลการตอบสนองข้อสอบแบบพหุมิติ โดย Bock and Aitkin ได้แสดงให้เห็นว่า การวิเคราะห์องค์ประกอบและทฤษฎีการตอบสนองข้อสอบต่างมุ่งศึกษาในสิ่งเดียวกันโดยผลลัพธ์สุดท้ายคือทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ อย่างไรก็ตาม แม้ว่าโมเดลที่นำเสนอโดย Bock and Aitkin จะมีประโยชน์สำหรับการขยายแนวคิดของโมเดลการตอบสนองข้อสอบแบบพหุมิติ แต่จุดมุ่งหมายหลักยังคงมุ่งไปที่การวิเคราะห์องค์ประกอบเป็นสำคัญ ซึ่งคล้ายคลึงกับการศึกษาของ Bock, Gibbons & Muraki (1988 cited in Reckase, 2009, p. 67) ที่เน้นไปที่การระบุจำนวนองค์ประกอบมากกว่าที่จะพิจารณาปฏิสัมพันธ์ของผู้ทดสอบและข้อสอบ

1.4 ทฤษฎีการตอบสนองข้อสอบ เมื่อพิจารณาแนวคิดของทฤษฎีการตอบสนองข้อสอบและการวิเคราะห์องค์ประกอบ จะเห็นได้ว่ามีจุดมุ่งหมายในการวิเคราะห์ที่แตกต่างกัน โดยการวิเคราะห์องค์ประกอบจะมุ่งเน้นที่การกำหนดองค์ประกอบให้น้อยที่สุด ส่วนโมเดลของทฤษฎีการตอบสนองข้อสอบ จะศึกษาความสัมพันธ์ระหว่างผู้ทดสอบและข้อสอบ ซึ่งการวัดแบบพหุมิติในมุมมองของทฤษฎีการตอบสนองข้อสอบ มีนักวิจัยได้นำเสนอแนวคิดดังนี้

Reckase (2009, pp. 68-69) ได้เสนอโมเดลในรูปแบบเวกเตอร์ (Vector) มากกว่าค่าสเกลาร์ (Scalar) ซึ่งเป็นที่ยอมรับโดยทั่วไป สามารถคำนวณได้ดังสมการ 11

$$P(\mu_{ij}/\theta_j,\eta_i)=\frac{1}{\gamma(\theta_j,\eta_i)}e^{f(u_{ij})'\theta_j+g(u_{ij})'\eta_i+\theta_j'h(u_{ij})\eta_i+\ell(u_{ij})} \qquad (11)$$

เมื่อ f, g, h และ ℓ แทน ฟังก์ชันของคะแนนในแต่ละข้อ

μ	แทน	ค่าสถิติที่ได้จากการสังเกตของค่าพารามิเตอร์ผู้ทดสอบ
η	แทน	เวกเตอร์ค่าพารามิเตอร์ของข้อสอบ
$\gamma(\theta_j,\eta_i)$	แทน	ฟังก์ชัน Normalizing ค่าที่อยู่ในฟังก์ชันระหว่าง 0 และ 1

เมื่อพิจารณาโมเดลดังกล่าว มีการระบุระดับความสามารถของผู้ทดสอบในรูปของเวกเตอร์ค่าพารามิเตอร์ โดยการประมาณค่าพารามิเตอร์ของผู้ทดสอบและข้อสอบมีความเป็นอิสระจากกัน เมื่อพิจารณาส่วนประกอบอื่น ๆ เช่น f,g,h ซึ่งเป็นฟังก์ชันของการให้คะแนน จะต้องทราบข้อมูลในส่วนนี้ก่อนวิเคราะห์ด้วยโมเดลราส์ช ซึ่งไม่สามารถประมาณค่าความสามารถได้จากข้อมูลในการตอบข้อสอบของผู้ทดสอบ โดยที่การให้คะแนนมี 2 ค่า คือ 0 กับ 1 สำหรับในกรณีของข้อมูลที่มีลักษณะเป็นเอกมิติ โดยไม่ได้คำนึงถึงลักษณะข้อสอบ ดังนั้น การพิจารณาเฉพาะจำนวนที่ตอบข้อสอบได้ถูกต้อง ถือเป็นค่าสถิติที่เพียงพอสำหรับการพิจารณา

ถ้าการตอบสนองข้อสอบในมิติที่ต่างกันสองด้าน นั่นคือค่าคะแนนการตอบข้อสอบในมิติที่ 1 และ มิติที่ 2 คือ μ_1 และ μ_2 ดังนั้นค่าสถิติสำหรับการประมาณค่า θ_1 และ θ_2 คือ $\eta\mu_1$ และ $\eta\mu_2$ เมื่อ n คือจำนวนข้อที่ตอบข้อสอบได้ถูกต้อง อย่างไรก็ตามเนื่องจาก μ_1 และ μ_2 เป็นค่าคงที่ของข้อสอบ ดังนั้น คะแนนที่ได้แต่ละมิติเป็นฟังก์ชันของ n เท่านั้น และค่า θ_1 และ θ_2 ที่ประมาณได้ จะมีความสัมพันธ์กันอย่างสมบูรณ์ นั่นคือโมเดลที่ได้ยังคงเป็นโมเดลแบบเอกมิติ จากปัญหาดังกล่าว มีนักวิจัยหลายท่านพยายามที่จะแบ่งคะแนนให้มากกว่า 2 ค่า เช่น จัดกระทำกับข้อสอบจากลักษณะ Dichotomous items แปลงให้เป็น Single polytomous item เพื่อที่จะนำไปสู่การวิเคราะห์แบบ Polytomous Item (Reckase, 2009, pp. 68-69)

Reckase (2009, pp. 69-70) เสนอโมเดลพื้นฐานทฤษฎีการตอบสนองข้อสอบแบบพหุมิติที่แสดงถึงนิยามของมิติคุณลักษณะแฝงของข้อสอบ และข้อตกลงเบื้องต้นของความเป็นอิสระของข้อสอบ หมายถึงคุณลักษณะของกลุ่มผู้ทดสอบแต่ละกลุ่ม มีค่าความสามารถเหมือนกันในแต่ละมิติ $\theta_1,\theta_2,...,\theta_k$ โดยค่าคะแนนในผู้ทดสอบแต่ละคนไม่มีความสัมพันธ์กัน เมื่อ k คือ จำนวนของมิติ ส่วน θ ซึ่งมีลักษณะเป็นเวกเตอร์แบบสมบูรณ์ และยังแสดงให้เห็นถึงความสัมพันธ์ระหว่างโมเดลปกติสะสมตามทฤษฎีตอบสนองข้อสอบแบบเอกมิติมากขึ้น ต่อมา Samejima (1974 Cited in Reckase, 2009, p. 70) ได้เสนอวิธีการตอบที่มีลักษณะการตรวจให้คะแนน 2 ค่า หรือมากกว่า 2 ค่าเป็นผลมาจากการแบ่งเป็นส่วนย่อย ๆ จากการตอบของตัวแปรต่อเนื่อง ซึ่ง Samejima ได้พัฒนาโมเดลการตอบสนองข้อสอบแบบพหุมิติ สำหรับข้อสอบซึ่งมีลักษณะเป็นการตอบแบบต่อเนื่อง ซึ่งคำนวณได้ดังสมการ 12 (Reckase, 2009, p. 70)

$$P_{z_i}(\theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a_i'(\theta_j - b_i)} e^{-\frac{t^2}{2}} dt \qquad (12)$$

เมื่อ z_i แทน ค่าการตอบสนองแบบต่อเนื่องต่อข้อสอบข้อที่ i

P แทน ค่าความน่าจะเป็นที่จะได้รับคะแนน z หรือดีกว่าจากข้อสอบข้อที่ i

b_i แทน เวกเตอร์ของพารามิเตอร์ความยาก พร้อมกับค่าพารามิเตอร์อื่น ๆ

จากโมเดลของ Samejima เป็นโมเดลต้นแบบที่ทำให้เกิดแนวคิดเกี่ยวกับโมเดลการตอบสนองข้อสอบแบบพหุมิติ เหมาะสำหรับการทดสอบเกี่ยวกับทักษะการปฏิบัติ แต่ไม่สามารถนำไปประยุกต์ได้ในทางปฏิบัติจริง เนื่องจากการตอบข้อสอบเป็นตัวแปรแบบต่อเนื่อง ซึ่งไม่สอดคล้องกับบริบทของแบบสอบทั้งในทางด้านการศึกษาและทางด้านจิตวิทยา (Reskase, 2009, p. 70)

2. พัฒนาการของโมเดลการตอบสนองข้อสอบแบบพหุมิติ

ในระหว่างปี ค.ศ.1970 ถึง ปี ค.ศ.1980 กลุ่มนักวิจัยได้เริ่มพัฒนาแนวคิดเกี่ยวกับโมเดลการตอบสนองข้อสอบแบบพหุมิติ ให้สามารถนำไปใช้ได้ในทางปฏิบัติ ดังจะเห็นจากการศึกษาเกี่ยวกับโมเดลราช์สแบบพหุมิติ โดย Reckase (1972) ต่อมาก็มีการนำเสนอแนวคิดเกี่ยวกับความสัมพันธ์ระหว่างผู้ทดสอบกับข้อสอบโดย Mulaik (1972), Sympsom (1978) และ Whitely (1980) ซึ่ง Mulaik (1972) ได้นำเสนอโมเดลการตอบสนองข้อสอบแบบพหุมิติ คำนวณได้ดังสมการ 13 (Reckase, 2009, p. 71)

$$P(u_{ij}|\theta_j, \eta_i) = \frac{\sum_{k=1}^{m} e^{(\theta_{jk} - \eta_{ik})u_{ij}}}{1 + \sum_{k=1}^{m} e^{(\theta_{jk} - \eta_{ik})u_{ij}}} \qquad (13)$$

จากสัญลักษณ์ เมื่อ u_{ij} = 0 หรือ 1 คุณสมบัติที่น่าสนใจของโมเดลคือ มีการกำหนดให้ค่าของเลขชี้กำลังและความน่าจะเป็นในการตอบข้อสอบถูกเพิ่มขึ้นเมื่อจำนวนของมิติเพิ่มขึ้น เช่น ถ้าผู้ทดสอบและพารามิเตอร์ของข้อสอบเท่ากันและกำหนดให้เลขชี้กำลังเท่ากับ 0 ความน่าจะเป็นในการตอบได้ถูกมีค่าเท่ากับ 0.5 เมื่อ m = 1 และจะเพิ่มขึ้นเป็น 0.67 เมื่อ m = 2 และจะเพิ่มขึ้นเป็น 0.75 เมื่อ m = 3 โดยทั่วไปถ้าทุกมิติมีเลขชี้กำลังเป็น 0 โอกาสของการตอบถูก คือ m/(m+1) ซึ่งถือได้ว่าเป็นคุณสมบัติของโมเดล ถ้าหากความน่าจะเป็นของการตอบถูกได้กำหนดไว้แล้ว ค่าของตัวเลขชี้กำลังจะมีการเปลี่ยนแปลงตามจำนวนของมิติที่เพิ่มขึ้น ดังตัวอย่างใน ตาราง 2-1 กรณีนี้ กำหนดให้ค่าของเลขชี้กำลังเท่ากับ x เมื่อ e^x มีค่าเท่ากันในแต่ละมิติของ m และความน่าจะเป็นของการตอบ

ข้อสอบถูกเท่ากับ 0.5 แสดงว่าเลขชี้กำลังมีการเปลี่ยนแปลงเล็กน้อยตามจำนวนที่เพิ่มขึ้นของมิติ (Reckase. 2009, pp. 71-72)

ตารางที่ 2-1 ความสัมพันธ์ของขนาดเลขชี้กำลังกับจำนวนมิติจากโมเดลของ Mulaik (*p*=.5)

จำนวนมิติ	เลขชี้กำลังของ e (Exponent of e)
1	0
2	-.69
3	-1.10
4	-1.39
5	-1.61
6	-1.79
7	-1.95
8	-2.08
9	-2.20
10	-2.30

ต่อมา Sympson (1978) และ Whitely (1980) ได้นำเสนอโมเดลความสัมพันธ์ระหว่างจำนวนมิติและขนาดเลขชี้กำลังที่มีทิศทางตรงกันข้ามกับโมเดลของ Mulaik (1972) ซึ่งเป็นโมเดลที่กำหนดค่าของเลขชี้กำลังและโอกาสการตอบถูกลดลงเมื่อจำนวนมิติเพิ่มขึ้น สามารถแสดงได้ดังสมการ 14 (Reckase, 2009, p. 72)

$$P\left(u_{ij}=1\middle|\theta_j,a_i,b_i,c_i\right)=c_i+\left(1-c_i\right)\prod_{k=1}^{m}\frac{e^{a_{ik}\left(\theta_{jk}-b_{ik}\right)}}{1+e^{a_{ik}\left(\theta_{jk}-b_{ik}\right)}} \qquad (14)$$

เมื่อ $\boldsymbol{a_i}$ แทน สเกลาร์ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ

$\boldsymbol{b_i}$ แทน สเกลาร์ค่าพารามิเตอร์ความยากของข้อสอบ

$\boldsymbol{c_i}$ แทน สเกลาร์พารามิเตอร์โอกาสการเดาของข้อสอบ

เมื่อเลขชี้กำลังทั้งหมดเท่ากับ 0 ความน่าจะเป็นในการตอบถูกเท่ากับ $c_i+\left(1-c_i\right)(.5)^m$ เช่น ในกรณีที่มิติ m เพิ่มขึ้น ค่าที่ได้จะมีค่าเข้าใกล้ $\boldsymbol{c_i}$ (Reckase, 2009, p. 72) โมเดลของ Mulaik (1972) จะประมาณค่าพารามิเตอร์ของข้อสอบเปลี่ยนแปลงไปตามจำนวนมิติ ดังตาราง 2-2 แสดงค่าเลขชี้กำลังของ e เมื่อจำนวนมิติเพิ่มขึ้น ในกรณีที่ $\boldsymbol{c_i}$ = 0 และเลขชี้กำลังทั้งหมดเท่ากัน

เนื่องจากความน่าจะเป็นในการตอบถูกขึ้นอยู่กับจำนวนของ m และจำนวนทั้งหมดมีค่าน้อยกว่า 1.0 ซึ่งจำนวนต่าง ๆ จะต้องเพิ่มค่าขึ้นด้วยถ้าเพิ่มจำนวนของ m ทำให้เกิดโอกาสในการตอบถูกเท่ากัน (Reckase, 2009, p. 773)

ตารางที่ 2-2 ความสัมพันธ์ของขนาดเลขชี้กำลังกับจำนวนมิติจากโมเดลของ Sympson (*p*=.5)

จำนวนมิติ	เลขชี้กำลังของ e (Exponent of e)
1	0.0
2	0.9
3	1.3
4	1.7
5	1.9
6	2.1
7	2.3
8	2.4
9	2.5
10	2.6

ในปี ค.ศ. 1982 ได้มีศึกษาความแปรเปลี่ยนของโมเดลทั่วไปของราสช์ (General Rasch Model) มาเป็นโมเดลโลจิสติกที่เป็นที่นิยมในปัจจุบันของ McKinley and Reckase (1982) ซึ่งสามารถแสดงได้ดังสมการ 15 (Reckase, 2009, p. 73)

$$P\left(u_{ij}=1\middle|\theta_j,a_i,d_i\right)=\frac{e^{\sum_{k=1}^{m}a_{ik}\theta_{jk}+d_i}}{1+e^{\sum_{k=1}^{m}a_{ik}\theta_{jk}+d_i}} \qquad (15)$$

เมื่อ $\boldsymbol{a}_i$ แทน เวกเตอร์ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบ

$\boldsymbol{d}_i$ แทน ค่าพารามิเตอร์ความยากของข้อสอบ

จากสมการเป็นโมเดลแบบพหุมิติ ที่ขยายแนวคิดมาจากโมเดลโลจิสติกแบบ 2 พารามิเตอร์ ซึ่งมีชื่อเรียกเฉพาะว่า Compensatory model เพราะการที่ค่า $\boldsymbol{\theta}$ มีค่าความสามารถลดลงในมิติใดมิติหนึ่ง สามารถไปทดแทนให้ค่า $\boldsymbol{\theta}$ ในมิติอื่น ๆ สูงขึ้นได้ ต่อมา Spry, Reckase, Ackerman, and Carlson (1990) มีการพัฒนาโมเดลที่ใช้ทั่วไป (Generalized model) โดยเฉพาะ Compensatory

model และ Partial compensatory model ให้มีประสิทธิภาพในการประมาณค่าความสามารถให้มีความแม่นยำเพิ่มขึ้น

3. ประเภทของโมเดลการตอบสนองข้อสอบแบบพหุมิติ

โมเดลการตอบสนองข้อสอบแบบพหุมิติ ประกอบด้วยค่าพารามิเตอร์ของผู้ทดสอบตั้งแต่ 2 ค่าขึ้นไป โดยการศึกษาโมเดลการตอบสนองข้อสอบแบบพหุมิติ ทำให้โมเดลมีความสอดคล้องกับข้อมูลมากขึ้น สำหรับผู้ทดสอบที่มีความแตกต่างกันของพื้นฐานความรู้หรือทักษะที่ผ่านมา ชุดทดสอบที่ได้รับมีความยากที่แตกต่างกันออกไป จึงอาจไม่สอดคล้องกับความสามารถของผู้ทดสอบ โมเดลหลายโมเดลที่เป็นแบบพหุมิติ จะพิจารณาค่าอำนาจจำแนกของข้อสอบ ส่งผลกระทบต่อมิติในแต่ละชุดข้อสอบ สามารถแบ่งโมเดลการตอบสนองข้อสอบแบบพหุมิติ ออกเป็น 2 ประเภท โดยขึ้นอยู่กับคะแนนที่ได้จากแบบวัดเป็นเกณฑ์ คือ คะแนนที่ได้จากแบบวัดที่มี 2 ค่า (Dichotomously scored) และคะแนนที่ได้จากแบบวัดที่มีหลายค่า (Polytomously scored) (Ackerman, 1994; Wilson, Adam & Wang, 1997 Cited in Briggs & Wilson, 2003) รายละเอียดมีดังนี้

3.1 ประเภทคะแนนที่ได้จากแบบวัดมี 2 ค่า (Dichotomously scored) โมเดลการวัดตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Items Response Theory Model: MIRT) ซึ่งเป็นการรวมกันของชุดตัวแปรแฝง (Combination of latent variables) ได้ถูกจำแนกออกเป็น 2 โมเดล เพื่อใช้อธิบายคะแนนจากการทดสอบที่ถูกแบ่งเป็น 2 ค่า คือ โมเดลการวัดมิติที่ไม่สามารถทดแทนกันได้ (Compensatory model) และ โมเดลการไม่ชดเชย (Non-compensatory model) (Ackerman, 1994; Yao & Boughton, 2007) ดังนี้

3.1.1 โมเดลการวัดมิติที่สามารถทดแทนกันได้ (Compensatory model) การวัดตามทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ ความน่าจะเป็นในการตอบข้อสอบข้อที่ i ถูกต้อง อธิบายโดยใช้ Multidimensional Compensatory Two-Parameter Logistic Model: MC2PL (Reckase, 1985 cited in Ackerman, 1994) โมเดลดังกล่าว มีพารามิเตอร์ค่าอำนาจจำแนกของข้อสอบในแต่ละมิติที่กำหนดขึ้น แต่มีพารามิเตอร์ค่าความยากของข้อสอบเพียงค่าเดียว และพารามิเตอร์ความสามารถ (θ) ที่ขาดหายไปในมิติใดมิติหนึ่ง สามารถเพิ่มเติมในมิติอื่น ๆ ให้สูงขึ้นได้ (Ackerman, 1994; Ackerman, Gierl & Walker, 2003) ซึ่งคำนวณได้ดังสมการ 16

$$P(X_i = 1) = \frac{e^{\sum_{k=1}^{m} a_{ik}\theta_k + d_i}}{1.0 + e^{\sum_{k=1}^{m} a_{ik}\theta_k + d_i}} \qquad (16)$$

เมื่อ X_i แทน คะแนน (0, 1) ของข้อสอบข้อที่ i

a_{ik} แทน เวกเตอร์ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i

d_i แทน สเกลาร์ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i

θ_k แทน เวกเตอร์ค่าพารามิเตอร์ความสามารถ

3.1.2 โมเดลการวัดมิติที่ไม่สามารถทดแทนกันได้ (Non-compensatory model) Sympson (1978 cited in Ackerman, 1994) เสนอโมเดลการวัดที่มีคุณสมบัติเดียวกันกับโมเดลการวัดมิติที่ไม่สามารถทดแทนกันได้ เรียกว่า Multidimensional Non-compensatory Two-Parameter Logistic Model (MNC2PL Model) โดยโมเดลดังกล่าว ได้แสดงพารามิเตอร์ค่าความยากของข้อสอบและพารามิเตอร์ค่าอำนาจจำแนกของข้อสอบในแต่ละมิติ และกำหนดค่าของส่วนประกอบในสมการ ถ้าค่าความน่าจะเป็นในการตอบข้อสอบถูกลดลงส่วนจำนวนมิติจะเพิ่มขึ้น สามารถคำนวณได้ดังสมการ 17

$$P(X_i = 1) = \prod_{k=1}^{m} \frac{e^{a_{ik}(\theta_k - b_{ik})}}{1 + e^{a_{ik}(\theta_k - b_{ik})}} \qquad (17)$$

เมื่อ X_i แทน คะแนน (0, 1) ของข้อสอบข้อที่ i

a_{ik} แทน เวกเตอร์ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i

b_{ik} แทน สเกลาร์ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i

θ_k แทน เวกเตอร์ค่าพารามิเตอร์ความสามารถ

3.2 ประเภทคะแนนที่ได้จากแบบวัดมีหลายค่า (Polytomously scored) โมเดลการตอบสนองข้อสอบแบบพหุมิติแบบที่มีการตรวจให้คะแนน 2 ค่า (Dichotomously scored) ที่นำเสนอในข้างต้นมีข้อจำกัดที่สำคัญ คือ การตรวจให้คะแนนรายข้อ ต้องเป็นแบบ 2 ค่า (Binary) เช่น การตรวจให้คะแนนแบบ 0 และ 1, ถูกหรือผิด, เห็นด้วย หรือไม่เห็นด้วย เป็นต้น ในขณะที่เครื่องมือทางการศึกษาและจิตวิทยาหลายประเภท แต่ละข้อมักมีหลายรายการคำตอบที่กำหนดลำดับหรือน้ำหนักคะแนนต่างกัน เพราะน่าจะให้สารสนเทศและความเที่ยงจากการตอบที่สูงกว่าการตรวจให้คะแนนแบบ 2 ค่า เช่น แบบวัดเจตคติ แบบวัดบุคลิกภาพ แบบวัดความสนใจในอาชีพ เป็นต้น จึงได้มีการพัฒนาโมเดลการตอบสนอง ข้อสอบสำหรับใช้กับการตรวจให้คะแนนรายข้อมากกว่า 2 ค่า ซึ่งปัจจุบันมีรูปแบบของเครื่องมือที่นิยมใช้กัน เช่น มาตรประมาณค่า (Rating scale) การตรวจให้คะแนนความรู้บางส่วน (Partial credit) การตรวจให้คะแนนตามลำดับขั้นของรายการหลายคำตอบ (Ordered-response categories) เป็นต้น (ศิริชัย กาญจนวาสี, 2555, หน้า155) ซึ่งมีโมเดลดังนี้

3.2.1 Multidimensional Random Coefficients Multinomial Logit Model: MRCML เป็นโมเดลที่ได้มีการปรับปรุงเพื่อเพิ่มความสามารถมาจากโมเดลการตอบสนองข้อสอบของราชส์ (Rasch family of items response model) การประยุกต์โมเดลดังกล่าวเริ่มต้นจากนักวิจัยได้ศึกษาโมเดลนี้ เช่น Adams (1997) Wang (1999) Wang, Wilson and Adams (1997) มีแนวคิดพื้นฐานการพัฒนาโมเดลจากการสร้างแบบพิมพ์ (Building-block) เป็นโมเดลสำหรับข้อสอบข้อที่ i ซึ่งเป็นข้อสอบที่สามารถตอบได้หลายคำตอบและสามารถให้คะแนนได้หลายค่า (k) หรืออาจเรียกว่ามิติที่สามารถเป็นไปได้ทั้งหมด ซึ่งมีลักษณะเป็นกลุ่มขนาดใหญ่ โดยกำหนดให้เป็น d (d = 1, ..., d) ส่วนผู้ทดสอบที่ตอบ ข้อสอบกำหนดให้เป็น p (p = 1, ..., p) แล้วความน่าจะเป็นของผู้ทดสอบที่ตอบข้อสอบข้อที่ i $\left(P_{ik}\right)$ ซึ่งมีอยู่ k กลุ่ม เมื่อเปรียบเทียบกับ k-1 กลุ่ม $\left(P_{ik-1}\right)$ แล้วจะได้เป็นฟังก์ชันเชิงเส้นตรงของความสามารถแฝงบนมิติ $\left(\theta_d\right)$และความยากเชิงสัมพัทธ์ของกลุ่ม k (Briggs & Wilson, 2003) ซึ่งคำนวณได้ดังสมการ 18

$$k\left(\delta_{ik}\right):\log\left(\frac{P_{ik}}{P_{ik-1}}\right)=\theta_d-\delta_{ik} \tag{18}$$

MRCML Model ซึ่งคำนวณได้ดังสมการ 19 ได้ดังนี้ คือ (Wilson, Adam & Wang, 1997 Cited in Briggs & Wilson, 2003)

$$P\left(X_{ik}=1;A,B,\xi/\theta\right)=\frac{\exp\left(b_{ik}\theta+a'_{ik}\xi\right)}{\sum_{k=1}^{k_i}\exp\left(b_{ik}\theta+a'_{ik}\xi\right)} \tag{19}$$

เมื่อ a'_{ik} แทน ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i

b_{ik} แทน ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i

θ แทน ค่าพารามิเตอร์ความสามารถของผู้ทดสอบ

3.2.2 Multidimensional Partial Credit Model: MPCM เป็นโมเดลที่ปรับขยายมาจากโมเดลของราชส์แบบ 1 พารามิเตอร์สำหรับแบบทดสอบที่ตรวจให้คะแนนหลายค่า ซึ่งคำนวณได้ดังสมการ 20 (Reckase, 2009, p. 33)

$$P\left(u_{ij}=k/\theta_j\right)=\frac{e^{\left[\sum_{u=0}^{k}\left(\theta_j-\delta_{iu}\right)\right]}}{\sum_{v=0}^{mi}e^{\left[\sum_{u=0}^{v}\left(\theta_j-\delta_{iu}\right)\right]}} \quad (20)$$

เมื่อ δ_{iu} แทน ค่าพารามิเตอร์ค่า threshold สำหรับชุดคะแนนความยากของข้อสอบข้อที่ i บนมิติที่ u

3.2.3 Multidimensional Generalized Partial Credit Model: MGPC เป็นโมเดลที่ปรับขยายมาจากโมเดล Generalized Partial Credit Model (GPC) ซึ่ง ได้รับการออกแบบมาสำหรับอธิบายปฏิสัมพันธ์ของผู้ทดสอบด้วยข้อสอบที่มีคะแนนมากกว่าสองค่า ซึ่งคำนวณได้ดังสมการ 21 (Reckase, 2009, p. 37)

$$P\left(u_{ij}=k/\theta_j\right)=\frac{e^{\left[\sum_{u=1}^{k}Da_i\left(\theta_j-b_i+d_{iu}\right)\right]}}{\sum_{v=1}^{mi}e^{\left[\sum_{u=1}^{v}Da_i\left(\theta_j-b_i+d_{iu}\right)\right]}} \quad (21)$$

เมื่อ k แทน คะแนนในข้อสอบ (0, 1, 2,...,)

b_i แทน ค่าพารามิเตอร์ความยากของข้อสอบข้อที่ i

d_{iu} แทน พารามิเตอร์ของความสัมพันธ์ของข้อสอบข้อที่ i แต่ละข้อที่มีความสัมพันธ์กับมิติที่ u

3.2.4 Multidimensional Graded Response Model: MGRM เป็นโมเดลสำหรับแบบทดสอบที่ตรวจให้คะแนนแบบหลายค่า ซึ่งออกแบบโดย Muraki and Carson (Muraki & Carson, 1993 Cited in Reckase, 2009, p. 39) สามารถคำนวณได้ดังสมการ 22 (Reckase, 2009, p. 39)

$$P\left(u_{ij}=k/\theta_j\right)=\frac{1}{\sqrt{2\pi}}\int_{a_i\left(\theta_j-b_i,k+1\right)}^{a_i\left(\theta_j-b_{ik}\right)}e^{\frac{-t^2}{2}}dt \quad (22)$$

เมื่อ k แทน คะแนนบนข้อสอบ (0, 1, 2,..., m)

a_i แทน เวกเตอร์ค่าพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i

b_{ik} แทน พารามิเตอร์ของความสัมพันธ์ของผู้ทดสอบแต่ละคนกับขั้นตอนของการทำข้อสอบแต่ละข้อ

จากการศึกษาปรากฏว่า การวิเคราะห์ข้อมูลในงานวิจัยนี้เป็นแบบวัด 2 ค่า (Dichotomously) คือ 0 กับ 1 นั่นคือ ตอบผิดมีค่าเป็น 0 และ ตอบถูกมีค่าเป็น 1 แล้วใช้โมเดลการวิเคราะห์ข้อสอบแบบ 3 พารามิเตอร์ นั้นคือ ค่าอำนาจจำแนกของข้อสอบ ค่าความยากของข้อสอบ และ ค่าโอกาสการเดาของข้อสอบ เพราะผู้ทดสอบมีโอกาสการเดาจากการทำข้อสอบ จึงนำมาวิเคราะห์ร่วมด้วย

ตอนที่ 4 การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์

1. การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computerized Adaptive Testing : MCAT) เป็นวิธีการทดสอบที่มีการขยายแนวคิดมาจากการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบเอกมิติ (Unidimensional Computerized Adaptive Testing: CAT) ซึ่งเป็นวิธีการเฉพาะที่ใช้ในการทดสอบความสามารถและคุณลักษณะแฝงอื่น ๆ ในการเลือกข้อสอบให้เหมาะกับระดับความสามารถของผู้เข้าสอบ (Frey & Seitz, 2009) เป็นการลดจำนวนข้อสอบให้ผู้ทดสอบตอบข้อสอบและเพิ่มความแม่นยำของการวัด เมื่อเปรียบเทียบกับการทดสอบแบบดั้งเดิมที่มีจำนวนข้อสอบคงที่ในแบบสอบ (Fixed Item Test: FIT) และเป็นการขยายแนวคิดจากโมเดลการตอบสนองข้อสอบแบบเอกมิติสู่โมเดลการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory Model: MIRT) ซึ่งถือว่าคุณลักษณะแฝงของบุคคลมากกว่า 1 องค์ประกอบ (มิติ หรือ ด้าน) ที่ส่งผลต่อการตอบข้อสอบ และการทดสอบแบบปรับเหมาะกำลังได้รับความนิยมอย่างมากในการนำไปทดสอบความรู้หรือทักษะของนักศึกษาหรือบุคคลทั่วไป จึงเกิดการรวมกันของทั้งสองวิธีนี้เรียกว่า การทดสอบปรับเหมาะแบบพหุมิติ (Multidimensional Adaptive Testing: MAT) (Segall, 1996, 2010, p. 62; Reckase, 2009, p. 95) ต่อมานำคอมพิวเตอร์มาใช้ในกระบวนการทดสอบ จึงเรียกว่า การทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (Multidimensional Computerized Adaptive Testing: MCAT) วิธีการทดสอบดังกล่าวนี้เป็นการทดสอบที่มีประสิทธิภาพในการวัดสูง ในด้านความแม่นยำในการวัด ในด้านการลดความยาวของข้อสอบหรือจำนวนข้อสอบ ในด้านการวัดที่ตรงกับสภาพจริงของลักษณะข้อสอบที่สัมพันธ์กับความสามารถของผู้ทดสอบ

ดังนั้น ในการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ จึงสามารถลดจำนวนข้อสอบลงได้มากกว่าการทดสอบด้วยการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์เอกมิติ ประมาณ 30-50% และลดจานวนข้อสอบได้มากกว่าการทดสอบแบบดั้งเดิม ที่กำหนดจำนวนข้อสอบไว้ประมาณ 70% โดยไม่สูญเสียความแม่นยำ (Frey & Seitz, 2009) และการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติมีประสิทธิภาพ ที่สูงกว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบเอกมิติ ถึง 1.3 เท่า (Frey & Seitz, 2009) อย่างไรก็ตาม ถ้ากระบวนการทดสอบได้ใช้วิธีการเลือกข้อสอบแบบสุ่ม (Random) การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ จะส่งผลให้การทดสอบมีประสิทธิภาพสูงกว่าวิธีการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบเอกมิติ ถึง 3.7 เท่า (Frey & Seitz, 2009)

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ เป็นการขยายความสามารถการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบเอกมิติสู่พหุมิติ ซึ่งการนำเสนอเกี่ยวกับการประมาณค่าความสามารถของผู้ทดสอบกับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ

2. โมเดลการตอบสนองข้อสอบแบบพหุมิติ

การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ สามารถใช้งานร่วมกับโมเดลการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory Model: MIRT) ได้หลากหลายโมเดล (Frey & Seitz, 2009) เช่น โมเดลโลจิสติกแบบพหุมิติ (Multidimensional Logistic Model) โมเดลปกติสะสมแบบพหุมิติ (Multidimensional normal ogive model) โมเดลสำหรับการวัดมิติที่ไม่สามารถทดแทนได้ (Models for Non-compensatory dimensional) โมเดลสำหรับการวัดการเรียนรู้และการเปลี่ยนแปลง (Models for learning and change) โมเดลในการระบุโครงสร้างของระดับคุณลักษณะ (Models with specified trait level structures) และโมเดลสำหรับการจำแนกกลุ่มบุคคล (Models for distinct classes of persons) โดยทั่วไปในการระบุความน่าจะเป็นของโมเดลการตอบสนองข้อสอบแบบพหุมิติ สำหรับการตอบสนองกับข้อสอบที่ถูกต้อง (i = 1) ขึ้นอยู่กับความสามารถแฝง และคุณลักษณะของข้อสอบที่ครอบคลุมด้วยพารามิเตอร์ข้อสอบหนึ่งพารามิเตอร์หรือมากกว่า (Frey & Seitz, 2009) โดย Segall (1996, 2010, p. 62) นำเสนอโมเดลการระบุความน่าจะเป็นในการตอบที่ถูกต้องของโมเดลโลจิสติกแบบพหุมิติ ซึ่งคำนวณได้ดังสมการ 23

$$P_i(\theta) \equiv \Pr ob\left(U_{ij}=1 \middle| \theta_{ij}\right) = c_i + \frac{1-c_i}{1+\exp^{\left[-Da_i'(\theta-b_i1)\right]}} \qquad (23)$$

เมื่อ $-Da_i(\theta - b_i 1)$ แทน $-D\sum_{k-1}^{p} a_{ki}(\theta_k - b_i)$

U_{ij} แทน ตัวแปรสุ่มที่มีสองค่า (0,1) จากข้อสอบข้อที่ i จากผู้ทดสอบคนที่ j

D แทน ค่าคงที่ เท่ากับ 1.702

a'_{ij} แทน เวกเตอร์ 1 x p ของพารามิเตอร์อำนาจจำแนกของข้อสอบข้อที่ i

b_i แทน พารามิเตอร์ความยากของข้อสอบข้อที่ i

c_i แทน พารามิเตอร์โอกาสการเดาของข้อสอบข้อที่ i

1 แทน เวกเตอร์ p x 1 เมื่อ p คือจำนวนมิติ

โมเดลการระบุความน่าจะเป็นในการตอบที่ถูกของโมเดลปกติสะสมแบบพหุมิติ ซึ่งคำนวณได้ดังสมการ 24 (Bock & Schilling, 2003; Reckase, 2009, p. 95)

$$P_i(\theta) = \Pr ob\left(U_{ij}=1\middle|\theta_j\right) = c_j + \left(1-c_j\right)\frac{1}{\sqrt{2\pi}}\int_{-z_j(\theta_j)}^{\infty}\exp\left(-t^2/2\right)df$$

$$= c_i + \left(1-c_j\right)F + z_j\left(q_j\right) \qquad (24)$$

เมื่อ $z_i(\theta)$ แทน $d_i + a_i\theta_j = d_i + a_i\theta_{j1} + a_{i2}\theta_{j2} + ... + a_{ip}\theta_{ip}$

F แทน เมทริกซ์ความแปรปรวนร่วมของมิติขนาด p x p

c_i แทน พารามิเตอร์โอกาสการเดาของข้อสอบข้อที่ i

d_i แทน Easiness Intercept ของข้อสอบข้อที่ i

a_i แทน เวกเตอร์อำนาจจำแนกของข้อสอบข้อที่ i

3. องค์ประกอบของการทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ

การทดสอบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติแบ่งออก 5 ขั้นตอน ซึ่งเป็นขั้นตอนที่ประยุกต์ตามกรอบแนวคิดของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT) (Thompson & Weiss, 2011) ได้แก่ 1) คลังข้อสอบ (Item bank) 2) คัดเลือกข้อสอบเริ่มต้น (Starting point) 3) การคัดเลือกข้อสอบข้อถัดไป (Item selection) 4) การประมาณค่าความสามารถ (Ability estimation) และ 5) เกณฑ์การยุติการทดสอบ (Stopping rule)

3.1 การสร้างคลังข้อสอบ (Item bank) เป็นคลังรวบรวมข้อสอบหรือแบบทดสอบที่มีคุณภาพ สำหรับการทดสอบแบบปรับเหมาะแบบพหุมิติ ผู้ทดสอบจะได้รับข้อสอบที่เหมาะสมกับความสามารถของผู้ทดสอบแต่ละคน โดยข้อสอบจะต้องมีค่าความยากของข้อสอบกระจายเต็มระดับความสามารถของผู้ทดสอบ ค่าอำนาจจำแนกของข้อสอบสูงเพื่อให้ประสิทธิภาพการทดสอบสูงสุด

และค่าโอกาสการเดาของข้อสอบเข้าใกล้หรือเท่ากับศูนย์ ข้อสอบควรมีจำนวนมากเพียงพอและกระจายทุกระดับความสามารถของผู้ทดสอบ ดังนั้น ต้องใช้ข้อสอบจำนวนมากสำหรับสร้างคลังข้อสอบ ทำให้คลังข้อสอบสำหรับการทดสอบแบบปรับเหมาะมีขนาดใหญ่กว่าการทดสอบดั้งเดิม (Wainer, Dorans, Flaugher, Green, & Mislevy, 2000, pp. 105-112)

การสร้างข้อสอบที่มีจำนวนมาก สำหรับขั้นตอนการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ ได้แก่ 1) สร้างข้อสอบใหม่ให้มีจำนวนเพียงพอในแต่ละเนื้อหา เน้นที่ครอบคลุมเนื้อหาหลักและกฎเบื้องต้นในการเขียนข้อสอบ 2) ตรวจทานคุณภาพของข้อสอบเพื่อให้ข้อสอบมีคุณภาพสูงสุดเท่าที่เป็นไปได้ 3) ทดสอบเบื้องต้นเพื่อตรวจสอบข้อบกพร่องของข้อสอบใหม่ 4) เลือกชุดย่อยของข้อสอบใหม่ ตอบข้อสอบและวิเคราะห์ผลสอบตามทฤษฎีการทดสอบแบบมาตรฐานเดิมและทฤษฎีการตอบข้อสอบแบบใหม่ เพื่อคัดข้อสอบไม่เหมาะสมออกไปหรือปรับปรุงข้อสอบใหม่ 5) เปรียบเทียบสมดุลเนื้อหาของคลังข้อสอบด้วยแบบทดสอบตามทฤษฎีการทดสอบแบบมาตรฐานเดิม เพื่ออ้างไปยังสมดุลเนื้อหาของคลังข้อสอบในการทดสอบปรับเหมาะพหุมิติ ประเมินโปรแกรมการทดสอบด้วยการจำลองพฤติกรรมการตอบข้อสอบของผู้ทดสอบระดับความสามารถที่ต่างกัน คะแนนทั้ง 2 การทดสอบนี้ต้องเปรียบเทียบกันได้ และ 6) แสดงข้อสอบตามแนวทางการจัดการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ต่อไป (Wainer, Dorans, Flaugher, Green, & Mislevy, 2000, pp. 105-112)

วิธีเลือกข้อสอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ ให้สามารถเปรียบเทียบได้กับการทดสอบใช้กระดาษและดินสอ มีความจำเป็นต้องจัดสมดุลเนื้อหาเพื่อใช้เป็นตัวแทนของแบบทดสอบถ้าการทดสอบใช้กระดาษและดินสอเป็นพหุมิติ แต่การทดสอบแบบปรับเหมาะข้อตกลงกำหนดว่าคลังข้อสอบต้องเป็นมิติเดียว การแก้ปัญหานี้อาจใช้วิธีการแบ่งคลังข้อสอบใหญ่เป็นคลังข้อสอบย่อย คลังข้อสอบย่อยแต่ละคลังแทนแต่ละมิติ คะแนนแต่ละส่วนมาจากแต่ละมิติหรือแต่ละเนื้อหาผลการทดสอบหาโดยคำนวณรวมน้ำหนักของคะแนนต่าง ๆ ในแต่ละมิติ แต่วิธีนี้อาจยุ่งยากในการทดสอบและการรายงานผลตามมา (Wainer, Dorans, Flaugher, Green, & Mislevy, 2000, pp. 105-112) Green et al. (1984) ให้เหตุผลสำหรับใช้คลังข้อสอบมิติเดียวว่า รูปแบบการทดสอบแบบปรับเหมาะที่เหมาะสมเกิดได้ ก็ต่อเมื่อข้อสอบมีมิติเด่นมิติเดียว (Dominant dimension) ข้อสอบมิติเด่น 2 มิติ หรือข้อสอบพหุมิติมีแนวโน้มว่าให้ค่าอำนาจจำแนกต่ำ และการเลือกข้อสอบอำนาจจำแนกสูงมีแนวโน้มให้ความเป็นมิติเดียวโดยตัวของมันเองอยู่แล้ว

การสร้างคลังข้อสอบมากกว่าหนึ่งคลังข้อสอบหรือแบ่งคลังข้อสอบใหญ่เป็นคลังข้อสอบย่อยสามารถสุ่มเลือกคลังข้อสอบและหมุนเวียนการใช้ชุดแบบทดสอบไปตามคลังข้อสอบต่าง ๆ ได้ (Stocking, 1988, 1994 cited in Wang & Kolen, 2001) การสร้างคลังข้อสอบจำนวนมากต้องเปรียบเทียบระหว่างคลังข้อสอบก่อนนำไปใช้ พิจารณาปัจจัยจากต่าง ๆ ดังนี้ 1) ขนาดของคลัง

ข้อสอบ 2) จำนวนคลังข้อสอบเมื่อใช้พร้อม ๆ กัน 3) การทับซ้อนของคลังข้อสอบ 4) วิธีเข้าสู่แต่ละคลังข้อสอบ 5) การหมุนเวียนใช้คลังข้อสอบและเกณฑ์การออกจากคลังข้อสอบ 6) กฎการนำข้อสอบกลับมาใช้ใหม่ 7) การทดสอบข้อสอบก่อนนำมาใช้ วิธีการคำนวณข้อสอบ และวิธีการควบคุมช่วงความสามารถบนเส้นต่อเนื่อง และ 8) วิธีการเลือกข้อสอบเพื่อให้ครอบคลุมแต่ละเนื้อหาวิชา ควบคุมการแสดงข้อสอบอัตราการทับซ้อนและอื่น ๆ (Wang & Kolen, 2001)

ขนาดคลังข้อสอบมีส่วนช่วยเพิ่มประสิทธิภาพการทดสอบ Weiss (1985) แนะนำว่าการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ จะให้ประสิทธิภาพเมื่อคลังข้อสอบมีจำนวนข้อสอบมากเพียงพอสำหรับจำแนกผู้ทดสอบ การศึกษาของ Xing and Hambleton (2004) ปรากฏว่า คลังข้อสอบขนาดใหญ่ทำให้สารสนเทศแบบทดสอบเพิ่มขึ้นตามไปด้วย สอดคล้องกับความคิดเห็นของ Lord (1980) ซึ่งอ้างถึงการศึกษาในสถานการณ์จำลอง เมื่อใช้ความยาวแบบทดสอบ 25 ข้อ ขนาดคลังข้อสอบ 363 ข้อ และ 183 ข้อ ปรากฏว่า ขนาดคลังข้อสอบ 363 ข้อสามารถให้สารสนเทศได้มากกว่า เพราะคลังข้อสอบขนาดใหญ่มีโอกาสเลือกข้อสอบดีที่สุดจัดให้ผู้ทดสอบได้มากกว่าคลังข้อสอบขนาดเล็ก

การกำหนดขนาดคลังข้อสอบพิจารณาปัจจัยหลายอย่าง (Stocking, 1994 cited in Chalhoub-Deville, Alcaya & Lozier,1996) กล่าวถึงปัจจัยในการออกแบบคลังข้อสอบที่ส่งผลต่อขนาดคลังข้อสอบ เช่น วิธีการเลือกข้อสอบ การควบคุมสมดุลเนื้อหามาตรวัดทางจิตวิทยา กฎการยุติการทับซ้อนข้อสอบ การให้คะแนนแบบทดสอบ ความเป็นคู่ขนานกับแบบทดสอบใช้กระดาษและดินสอ เป็นต้น Way (1997 cited in Bergstrom & Lunz, 1999) อ้างถึงลักษณะแบบทดสอบที่สามารถนำไปประมาณขนาดคลังข้อสอบได้ไว้ว่า 1) ความยาวแบบทดสอบต่อขนาดคลังข้อสอบเท่ากับ 1 ต่อ 6-8 2) ใช้ข้อสอบเพียง 10-15 เปอร์เซ็นต์ของประชากรผู้ทดสอบ 3) อัตราการทับซ้อนข้อสอบที่จัดให้แก่ผู้ทดสอบแต่ละคู่ต้องไม่มากเกิน 15-20 เปอร์เซ็นต์ และ 4) เปอร์เซ็นต์การทับซ้อนข้อสอบสูงสุดของแบบทดสอบแบบปรับเหมาะ 2 ชุดไม่เกิน 40 เปอร์เซ็นต์ ระหว่างผู้ทดสอบความสามารถระดับเดียวกัน

Segall (2005) ได้แนะนำขนาดคลังข้อสอบควรมีขนาดประมาณ 6-8 เท่าของความยาวแบบทดสอบ ส่วน de Ayla (2009, p. 376) เสนอว่า คลังข้อสอบควรมีขนาดประมาณ 8-12 เท่าของจำนวนข้อสอบที่ใช้ในการทดสอบ กล่าวคือ ถ้าในการทดสอบใช้ข้อสอบประมาณ 25 ข้อ คลังข้อสอบควรมีข้อสอบ 200-300 ข้อ และ Weiss (2011) เสนอให้มีข้อสอบไม่น้อยกว่า 200 ข้อ Thompson and Weiss (2011) เสนอให้มีข้อสอบประมาณ 400 ข้อ แม้ว่าวิธีดำเนินการของการทดสอบปรับเหมาะจะดีเพียงใดหากคลังข้อสอบที่พัฒนาขึ้น ไม่มีคุณภาพ การทดสอบครั้งนี้ก็ถือว่าไม่มีคุณภาพเช่นกัน ดังนั้นการออกแบบและพัฒนาคลังข้อสอบสำหรับการทดสอบแบบปรับเหมาะจึงเป็นสิ่งที่สำคัญที่ควรคำนึงถึง (Reckase, 2009, pp. 336-337)

3.2 การคัดเลือกข้อสอบเริ่มต้น (Starting point) เป็นการกำหนดจุดเริ่มต้นของการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ มีผู้วิจัยนำเสนอหลายวิธี โดยทั่วไปจะใช้ข้อสอบข้อแรกที่มีค่าความยากปานกลาง ตั้งแต่ -1 ถึง 1 จัดให้ผู้ทดสอบแต่ละคน เมื่อไม่ทราบค่าสารเทศเบื้องต้นหรือระดับความสามารถของผู้ทดสอบ จึงจำเป็นต้องพิจารณาว่าจะใช้เกณฑ์ใดในการเริ่มต้นการทดสอบ ต่อมา Thompson & Weiss (2011) ได้นำเสนอแนวทางในการกำหนดเกณฑ์เริ่มต้นของการเลือกข้อสอบ ได้แก่ 1) สุ่มเลือกข้อสอบโดยใช้โปรแกรมคอมพิวเตอร์ 2) เลือกข้อสอบที่มีค่าความยากปานกลาง 3) ทดสอบเบื้องต้นก่อนการทดสอบจริงเพื่อนำผลมากำหนดข้อแรก

3.3 การคัดเลือกข้อสอบข้อถัดไป (Item selection) สำหรับการเลือกข้อสอบจากคลังข้อสอบเพื่อให้ผู้ทดสอบตอบ โดยหลักการคัดเลือกข้อสอบมีเป้าหมายเพื่อลดความแปรปรวนของเวกเตอร์ความสามารถแบบพหุมิติชั่วคราว (Frey & Seitz, 2009) วิธีการคัดเลือกข้อสอบที่ใช้กับการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ อาศัยภาวะความน่าจะเป็นสูงสุดหรือต่ำสุดจากการประมาณค่าความสามารถ (θ) และมีวิธีการคัดเลือกข้อสอบอยู่หลากหลายวิธี แต่ละวิธีสามารถใช้ร่วมกับวิธีการประมาณค่าทุกแบบและยังสามารถใช้วิธีการคัดเลือกข้อสอบหลายวิธีร่วมกันได้ (Reckase, 2009, p. 327)

3.4 การประมาณค่าความสามารถ (Ability estimation) การประมาณค่าความสามารถของโมเดลการตอบสนองข้อสอบแบบพหุมิติ (MIRT) มีเกณฑ์การประมาณค่าที่แตกต่างกันหลายแนวทาง เช่น Maximum Likelihood Criterion, Maximum a Posteriori Bayesian Criterion และ Newton-Raphson Method (Reckase, 2009, p. 139) แต่โดยทั่วไปการประมาณค่าความสามารถของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ แบบพหุมิติ (MCAT) ที่นิยมมีอยู่ 2 วิธี คือ วิธีการประมาณค่าแบบภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation: ML) และวิธีการประมาณค่าแบบเบส์ (Bayesian Estimation: BE) (Reckase, 2009, p. 314) จากข้อตกลงเบื้องต้นเกี่ยวกับความเป็นอิสระของข้อสอบและผู้ทดสอบในทฤษฎีการตอบสนองข้อสอบ ถ้าตัวแปรความสามารถมีค่าคงที่การตอบสนองข้อสอบของผู้ทดสอบต่อข้อสอบแต่ละข้อของแบบทดสอบ ถูกคาดหวังว่าเป็นอิสระจากกัน ในเชิงสถิติการประมาณค่าความสามารถด้วยเทคนิคการประมาณค่าแบบภาวะน่าจะเป็นสูงสุด ก่อให้เกิดความไม่ชัดเจนในการประมาณค่าของเวกเตอร์ความสามารถ ณ จุดเริ่มต้นของการทดสอบ แต่วิธีการประมาณค่าแบบเบส์จะหลีกเลี่ยงข้อบกพร่องนี้ แต่ก็ยังไม่มีข้อสรุปเกี่ยวกับปัญหาในการประมาณค่าความสามารถ ในทางปฏิบัติจึงมีการใช้วิธีการประมาณค่าทั้งแบบภาวะน่าจะเป็นสูงสุดและแบบเบส์ (Frey & Seitz, 2009) และมีวิธีการประมาณค่าความสามารถที่สำคัญ ดังนี้

3.2.1 วิธีการประมาณค่าแบบภาวะน่าจะเป็นสูงสุด คำนวณได้ดังสมการ 25 (Segall, 2010, p. 62; Reckase, 2009, p. 140)

$$f\left(\mu_{i1}=\mu_{i1},\mu_{i2}=\mu_{12},...,\mu_{in}=\mu_{in}|\theta\right)=L\left(\mu|\theta\right)=\prod_{i=1}^{n}P\left(u_{ij}|\theta_j\right)^{\mu_{ij}}Q\left(u_{ij|\theta_i}\right)^{1-u_{ij}} \quad (25)$$

เมื่อ $f\left(\mu_{i1}=\mu_{i1},\mu_{i2}=\mu_{12},...,\mu_{in}=\mu_{in}|\theta\right)$ แทน ฟังก์ชันการตอบข้อสอบ ข้อที่จะนำไปใช้ในการคำนวณในขั้นตอนการคัดเลือกข้อสอบและการให้คะแนน

$L\left(\mu|\theta\right)$ แทน ฟังก์ชันความน่าจะเป็น

$P\left(u_{ij}|\theta_j\right)^{\mu_{ij}}$ แทน ความน่าจะเป็นในการตอบข้อสอบถูก

$Q\left(u_{ij}|\theta_j\right)^{1-u_{ij}}$ แทน $1-P\left(u_{ij}|\theta_j\right)^{\mu_{ij}}$

3.2.2 วิธีการประมาณค่าแบบเบส์ (Bayesian Estimation: BE) มีวิธีการที่สามารถคำนวณได้ดังสมการ 26 (Segall, 2010, pp. 60-64; Reckase, 2009, pp. 144-145)

$$f\left(\theta|m\right)=\frac{L\left(u|\theta\right)f\left(\theta\right)}{f\left(u\right)}=\frac{L\left(u|\theta\right)f\left(\theta\right)}{\int_{-\infty}^{\infty}f\left(u|\theta\right)f\left(\theta\right)d\theta} \quad (26)$$

เมื่อ $L\left(u|\theta\right)$ แทน ฟังก์ชันความน่าจะเป็น

$f\left(\theta\right)$ แทน ความน่าจะเป็นที่เกิดขึ้นก่อนของ θ

วิธีการประมาณค่าแบบเบส์ มีข้อตกเบื้องต้นว่า มีการแจกแจงเป็นแบบปกติหลายตัวแปร (Multivariate normal) ที่มีค่าเฉลี่ยเท่ากับ 0 และมีเมทริกซ์ความแปรปรวนลบกับความแปรปรวนร่วมเท่ากับ ϕ แสดงได้ดังสมการ 27

$$f\left(\theta\right)=\left(2p\right)^{-p/2}\left|\varphi\right|^{-1/2}\exp\left[-\frac{1}{2}\left(\theta-\mu\right)'\varphi^{-1}\left(\theta-\mu\right)\right] \quad (27)$$

$f\left(u\right)$ แทน ความน่าจะเป็นที่เกิดขึ้นภายหลังของ u คำนวณได้จาก สมการ 28

$$f\left(u\right)=\int_{-\infty}^{\infty}f\left(u|\theta\right)f\left(\theta\right)d\theta \quad (28)$$

3.2.3 วิธีการประมาณค่าแบบนิวตันราฟ (Newton-Raphson Method) วิธีการนี้ใช้ควบคู่ไปกับการประมาณค่าวิธีแบบภาวะน่าจะเป็นสูงสุด ในกรณีที่ไม่สามารถประมาณค่าความสามารถได้ (Segall, 1996; Chen, 2009, pp. 1-2; Frey & Seitz, 2009) ซึ่งคำนวณได้ดังสมการ 29 (Segall, 1996; Chen, 2009, pp. 1-2)

$$\theta^{(j+1)} = \theta^{(j)} - \delta^{(i)} \tag{29}$$

เมื่อ $\delta^{(j)}$ แทน เวกเตอร์ p x 1 และคำนวณค่าได้ดังสมการ 30

$$\delta^{(j)} = \left[H\left(q^{(j)}\right)\right]^{-1}\left(\left(\frac{\partial}{\partial_q} 1nL\left(u|\theta\right)\right)\right) \tag{30}$$

เมื่อ $H\left(\theta^{(j)}\right)$ แทน เมทริกซ์ p x p เป็นค่าของอนุพันธ์อันดับสองของฟังก์ชันความน่าจะเป็น ดังสมการ 31

$$H(\theta) = \begin{bmatrix} \frac{\partial^2 1nL}{\partial\theta_1^2} & \frac{\partial^2 1nL}{\partial\theta_1\partial\theta_2} & \frac{\partial^2 1nL}{\partial\theta_1\partial\theta_p} \\ \frac{\partial^2 1nL}{\partial\theta_2\partial\theta_1} & \frac{\partial^2 1nL}{\partial\theta_2^2} & \frac{\partial^2 1nL}{\partial\theta_2\partial\theta_p} \\ \frac{\partial^2 1nL}{\partial\theta_p\partial\theta_1} & \frac{\partial^2 1nL}{\partial\theta_q\partial\theta_2} & \frac{\partial^2 1nL}{\partial\theta_p^2} \end{bmatrix} \tag{31}$$

และ $\frac{\partial}{\partial_q} 1nL\left(u|\theta\right)$ เป็นค่าของอนุพันธ์อันดับหนึ่งของฟังก์ชันความน่าจะเป็นของ $\theta^{(j)}$ ซึ่งคำนวณได้ดังสมการ 32

$$\frac{\partial}{\partial_q} 1nL\left(u|\theta\right) = \begin{bmatrix} \frac{\partial}{\partial\theta_1} 1nL\left(m|\theta\right) \\ \frac{\partial}{\partial\theta_2} 1nL\left(u|\theta\right) \\ M \\ \frac{\partial}{\partial\theta_p} 1nL\left(u|\theta\right) \end{bmatrix} \tag{32}$$

3.2.4 วิธีการประมาณค่าแบบกำลังสองน้อยที่สุด (Least Squares Estimation) วิธีการนี้พบได้ไม่บ่อยนักในการประมาณค่าพารามิเตอร์ของทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ แต่ก็เป็นวิธีที่ใช้ในโปรแกรม NOHARM ซึ่งคำนวณได้ดังสมการ 33 (Reckase, 2009, pp. 147-148)

$$SS_\theta = \sum_{i=1}^{n} \left(u_i - P\left(u_i = 1 | \theta, a_i, d_i\right) \right)^2 \quad (33)$$

เมื่อ SS_θ แทน ผลรวมของความแตกต่างยกกำลังของแต่ละเวกเตอร์ θ

u_i แทน คะแนนที่ได้จากข้อสอบข้อที่ i (มีค่าเป็น 0 และ 1)

3.5 เกณฑ์การยุติการทดสอบ (Stopping rule) กฎการยุติการทดสอบของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์โดยทั่วไปใช้กฎการยุติ 3 ลักษณะได้แก่ 1) เกณฑ์การยุติการทดสอบเมื่อผู้ทดสอบใช้ข้อสอบในคลังข้อสอบครบทุกข้อ 2) เกณฑ์การยุติการทดสอบแบบความยาวคงที่ (Fixed length) และ 3) เกณฑ์ยุติการทดสอบแบบความยาวยืดหยุ่น (Variable length) (Reckase, 2009, p. 335) หรือการพิจารณาจากค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า (Standard Error of Estimation: SEE) โดยทั่วไปจะใช้กฎการยุติการทดสอบจากค่าความคลาดเคลื่อนมาตรฐานกำหนดไว้ที่ 0.30 (Thissen, 1990) ส่วนงานวิจัยที่ผ่านมาจะกำหนด SEE ≤ 0.55, 0.5 และ 0.30 (Seo, 2011; Thissen, 1990) สำหรับงานวิจัยการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์แบบพหุมิติ (MCAT) ที่ผ่านมา งานวิจัยส่วนใหญ่ใช้การจำกัดจำนวนข้อสอบ ปรากฏว่าการใช้ข้อสอบจำนวน 30 ข้อ มีค่าความลำเอียงน้อยกว่าการใช้ข้อสอบจำนวน 20 ข้อ แต่ยังระบุไม่ได้ว่าจำนวนข้อสอบเท่าใดจึงจะเหมาะสมสำหรับการทดสอบปรับเหมาะแบบพหุมิติ (Diao & Reckase, 2009)

ตอนที่ 5 ทฤษฎีวงจรการพัฒนาโปรแกรม

การพัฒนาโปรแกรม หมายถึง การสร้างโปรแกรมใหม่หรือการปรับปรุงโปรแกรมเดิมให้มีประสิทธิภาพมากขึ้น โดยการประยุกต์โปรแกรมคอมพิวเตอร์เป็นเครื่องมือหลักในการทำงาน ในงานวิจัยนี้ จึงได้นำวงจรการพัฒนาโปรแกรม (System Development Life Cycle: SDLC) ของ Singh et al., (2015) มาเป็นแนวทางในการพัฒนาโปรแกรม ซึ่งประกอบด้วย 6 ขั้นตอน ได้แก่ 1) การกำหนดความต้องการและการวิเคราะห์ (Requirement gathering and analysis) 2) การออกแบบ (Design) 3) การพัฒนาโปรแกรม (Implementation and coding) 4) การทดสอบ (Testing) 5) การใช้งาน (Deployment) และ 6) การบำรุงรักษา (Maintenance) ดังนี้

1. การกำหนดความต้องการและการวิเคราะห์ (Requirement gathering and analysis) การกำหนดความต้องการ เป็นขั้นตอนเริ่มต้นของ SDLC เพื่อระบุถึงปัญหาที่เกิดขึ้น กำหนดทางเลือกในการแก้ปัญหา และเลือกวิธีการแก้ไขปัญหา โดยการประยุกต์โปรแกรมคอมพิวเตอร์และเลือกใช้เทคโนโลยีสารสนเทศที่เหมาะสมที่สุดด้วยการศึกษาความเป็นไปได้ จากนั้นจึงกำหนดแผนการดำเนินงานโครงการอย่างละเอียดให้ครอบคลุมทุกกิจกรรมของ SDLC ส่วนการวิเคราะห์โปรแกรมเป็นขั้นตอนต่อเนื่องจากการกำหนดความต้องการ วัตถุประสงค์หลักของขั้นตอนนี้คือ การรวบรวมความต้องการใช้งานของผู้ใช้ทุกฝ่ายที่เกี่ยวข้องกับโปรแกรม และนำมาวิเคราะห์สิ่งที่จำเป็นต่อความต้องการของโปรแกรม เพื่อจัดทำข้อกำหนดเกี่ยวกับความต้องการของโปรแกรม ซึ่งเป็นเอกสารสำคัญสำหรับใช้ประกอบการดำเนินการพัฒนาโปรแกรมทุกขั้นตอน การวิเคราะห์โปรแกรมจึงเป็นขั้นตอนที่จำเป็นอย่างยิ่ง สำหรับงานวิจัยด้านการพัฒนาโปรแกรม ผู้ที่เกี่ยวข้องกับการพัฒนาโปรแกรมไม่ว่าจะเป็นกลุ่มของนักพัฒนาโปรแกรมหรือผู้ใช้งานทุกระดับ จึงควรตระหนักถึงความสำคัญและให้ความร่วมมือในการดำเนินงานขั้นตอนนี้ เพื่อให้การดำเนินงานในขั้นตอนถัด ๆ ไปของ SDLC เป็นไปอย่างถูกต้องและมีประสิทธิภาพ 6 ขั้นตอน ดังนี้

1.1 ขั้นตอนการกำหนดความต้องการ (Requirement phase) การพัฒนาโปรแกรมจะมีผู้ใช้ที่เกี่ยวข้องกับโปรแกรมเป็นจำนวนมาก ตั้งแต่ ผู้ดูแลโปรแกรม อาจารย์ นักเรียน และเจ้าหน้าที่ที่เกี่ยวข้อง ดังนั้น นักวิเคราะห์โปรแกรมจำเป็นต้องวางแผนในการรวบรวมข้อมูลความต้องการใช้งานของผู้ใช้ (User requirement gathering) อย่างรัดกุม มิเช่นนั้นแล้วขั้นตอนนี้อาจจะใช้เวลายาวนานที่สุดของ SDLC เนื่องจากในทางปฏิบัติแล้ว กิจกรรมการรวบรวมความต้องการของผู้ใช้โปรแกรมจะเป็นส่วนที่ยุ่งยากและซับซ้อนที่สุด เพราะผู้ใช้แต่ละคนมีเจตคติต่อโปรแกรมที่จะพัฒนาแตกต่างกัน ส่งผลทำให้การถ่ายทอดความคิดเรื่องความต้องการใช้งานโปรแกรมของผู้ใช้แต่ละคนจึงมีรายละเอียดที่แตกต่างกัน ผลจากการศึกษาประเด็นปัญหาที่เกิดขึ้นในโปรแกรมทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์จะประกอบไปด้วย 3 ขั้นตอน ได้แก่ 1) การนิยามปัญหา เป็นการอธิบายถึงปัญหาทุกประเด็นที่เกิดขึ้นโปรแกรม ซึ่งอาจจะมีทั้งปัญหาหลักและปัญหาอื่น ๆ ที่ตามมา 2) การศึกษาความเป็นไปได้ เป็นการพิจารณาถึงความเหมาะสมและประโยชน์ที่จะได้รับเมื่อเปรียบเทียบกับทรัพยากรที่ใช้ไป 3) การจัดทำแผนการดำเนินงานโครงการ เมื่อผู้วิจัยตัดสินใจเลือกวิธีการในการแก้ไขปัญหาแล้ว ขั้นตอนต่อไปจะเป็นการจัดทำแผนการดำเนินงานโครงการ (Software Development plan)

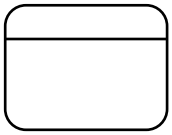
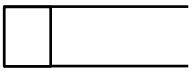
1.2 ขั้นตอนการวิเคราะห์ (Analysis phase) การวิเคราะห์โปรแกรมเป็นขั้นตอนต่อเนื่องจากการสำรวจเบื้องต้น วัตถุประสงค์หลักของขั้นตอนนี้คือ การรวบรวมความต้องการใช้งานของผู้ใช้ทุกฝ่ายที่เกี่ยวข้องกับโปรแกรม ซึ่งเป็นเอกสารสำคัญสำหรับใช้ประกอบการดำเนินการพัฒนาโปรแกรมทุกขั้นตอนไปจนกระทั่งถึงขั้นตอนการตรวจรับงาน การวิเคราะห์โปรแกรมจึงเป็นขั้นตอนที่

จำเป็นอย่างยิ่งสำหรับการวิจัยด้านการพัฒนาโปรแกรม ผู้ที่เกี่ยวข้องกับการพัฒนาโปรแกรมไม่ว่าจะเป็นกลุ่มของนักพัฒนาโปรแกรมหรือผู้ใช้งานทุกระดับ จึงควรตระหนักถึงความสำคัญและให้ความร่วมมือในการดำเนินงาน เพื่อให้การดำเนินงานในขั้นตอนถัด ๆ ไปของ SDLC เป็นไปอย่างถูกต้องและมีประสิทธิภาพ โดยมีขั้นตอนย่อยได้แก่ 1) การรวบรวมความต้องการใช้งานของผู้ใช้ 2) แหล่งข้อมูลของความต้องการใช้งานโปรแกรม 3) เอกสารที่เกี่ยวข้อง และ 4) วิธีการรวบรวมความต้องการของผู้ใช้

1.3 แบบจำลองกระบวนการ (Process modeling) หลังจากที่นักวิเคราะห์โปรแกรมได้จัดทำข้อกำหนดความต้องการของโปรแกรมแล้ว กิจกรรมถัดไปจะเป็นการสร้างแบบจำลอง (Model) เพื่อนำเสนอข้อมูลด้วยภาพหรือสัญลักษณ์ ซึ่งจะทำให้เห็นถึงองค์ประกอบและความสัมพันธ์ขององค์ประกอบเหล่านั้นได้อย่างชัดเจนยิ่งขึ้น ดังคำกล่าวที่ว่า “รูปหนึ่งรูปแทนคำนับพันคำ” ในการวิเคราะห์และออกแบบโปรแกรมจำเป็นต้องมีการสร้างแบบจำลองในหลายมุมมองเพื่อใช้ในการสื่อสารและการสร้างความเข้าใจร่วมกัน ทั้งในทีมนักพัฒนาโปรแกรม เจ้าของโปรแกรม รวมทั้งผู้ใช้ทุกฝ่ายที่เกี่ยวข้อง แบบจำลองแบ่งได้เป็น 2 ประเภทคือ แบบจำลองเชิงกายภาพ (Physical model) สำหรับแสดงรายละเอียดการทำงานของโปรแกรมทั้งในส่วนของกระบวนการทำงาน ข้อมูลนำเข้า ข้อมูลที่เป็นผลลัพธ์และข้อกำหนดต่าง ๆ รวมไปถึงฮาร์ดแวร์ ซอฟต์แวร์ และเทคโนโลยีสารสนเทศที่ใช้ในโปรแกรม จึงกล่าวได้ว่าแบบจำลองเชิงกายภาพจะเน้นการนำเสนอในส่วนวิธีการทำงานอย่างไร (How) ซึ่งการนำเสนอในมุมมองดังกล่าวอาจทำให้รายละเอียดของโปรแกรมบางส่วนขาดหายไป ดังนั้น ในการวิเคราะห์และออกแบบโปรแกรมและประยุกต์งานโปรแกรม โดยแบบจำลองในลักษณะดังกล่าวเรียกว่า แบบจำลองเชิงตรรกะ (Logical model) ในขั้นตอนของการออกแบบโปรแกรมเชิงตรรกะจะใช้แบบจำลองทางตรรกะ เพื่อนำเสนอถึงองค์ประกอบที่ควรจะมีทั้งหมดของโปรแกรมนั้น รวมไปถึงข้อมูลที่ได้จากการทำงานจะถูกส่งไปยังแหล่งข้อมูลใดบ้าง ซึ่งสามารถสร้างได้ ดังนี้

1.3.1 แผนภาพกระแสข้อมูล (Data Flow Diagram: DFD) แผนภาพกระแสข้อมูลเปรียบเสมือนแบบพิมพ์เขียว (Blueprint) ที่แสดงรายละเอียดของการออกแบบบ้านก่อนลงมือก่อสร้างจริง โดยสถาปนิกจะใช้แบบพิมพ์เขียวในการสื่อสารกับเจ้าของบ้านเพื่อให้เกิดความเข้าใจร่วมกันว่า บ้านที่จะสร้างจะมีรูปร่างทรงใด และรายละเอียดของบ้านเป็นอย่างไร เพราะการจำลองข้อมูลเป็นแผนภาพจะทำให้เจ้าของบ้านสามารถจินตนาการได้ง่าย ในขั้นตอนนี้นักวิเคราะห์โปรแกรมและผู้ที่เกี่ยวข้องกับโปรแกรมจำเป็นต้องช่วยกันทบทวนและตรวจสอบจนกระทั่งมั่นใจได้ว่า แผนภาพดังกล่าวครอบคลุมกระบวนการทำงานทั้งหมดที่อยู่ภายใต้ขอบเขตการวิจัยที่ดำเนินการอยู่ เพื่อนำไปใช้ในการออกแบบโปรแกรมต่อไป ซึ่งมีองค์ประกอบและสัญลักษณ์ที่ใช้แผนภาพกระแสข้อมูล

ตารางที่ 2-3 สัญลักษณ์ที่ใช้ในการสร้างแผนภาพกระแสข้อมูล

ภาพสัญลักษณ์	ความหมาย
	ชื่อสัญลักษณ์: กระบวนการ (Process)
	ชื่อสัญลักษณ์: เอนทิตีภายนอก (External entity)
	ชื่อสัญลักษณ์: กระแสข้อมูล (Data flow)
	ชื่อสัญลักษณ์: แหล่งจัดเก็บข้อมูล (Data store)

จากตารางที่ 2-3 สัญลักษณ์ที่ใช้ในการสร้างแผนภาพกระแสข้อมูล สามารถอธิบายคุณลักษณะของสัญลักษณ์ได้ดังนี้

กระบวนการ (Process) เป็นกิจกรรมการทำงานของโปรแกรมหรือกระบวนการทำงานของโปรแกรม มีสัญลักษณ์ที่แสดงถึงกระบวนการทำงานในขั้นตอนนั้น ๆ จำเป็นจะต้องใช้ข้อมูลนำเข้า (Input) และประมวลผลเพื่อให้ได้ข้อมูลที่เป็นผลลัพธ์ (Output) ใดบ้าง โดยข้อมูลนำเข้าอาจมาจากกระบวนการทำงานอื่นอาจเป็นเอนทิตีภายนอก หรือแหล่งจัดเก็บข้อมูล เช่นเดียวกับข้อมูลที่เป็นผลลัพธ์ ซึ่งอาจส่งไปยังกระบวนการทำงานอื่น เอนทิตีภายนอกหรือแหล่งจัดเก็บข้อมูลก็ได้เช่นกัน

เอนทิตีภายนอก (External entity) เป็นแหล่งข้อมูลซึ่งอาจเป็นบุคคล โปรแกรม สารสนเทศ หรือเครื่องมือ/อุปกรณ์ที่สัมพันธ์กับกระบวนการทำงาน โดยแบ่งเป็น 2 แหล่ง ได้แก่ 1) แหล่งข้อมูลนำเข้า (Data source) ที่จำเป็นต้องใช้ประกอบกระบวนการทำงาน และ 2) แหล่งรับข้อมูล (Data destination) ซึ่งเป็นผลลัพธ์ที่ได้จากกระบวนการทำงานนั้น ๆ

กระแสข้อมูล (Data flow) เป็นทิศทางการไหลของข้อมูลนำเข้าหรือส่งออกไปเป็นผลลัพธ์ซึ่งเกี่ยวข้องกับกระบวนการทำงานหนึ่ง ๆ โดยจะเชื่อมโยงเพื่อแสดงความสัมพันธ์ระหว่างกระบวนการทำงานกับเอนทิตีภายนอก หรือกระบวนการทำงานกับแหล่งจัดเก็บข้อมูล

แหล่งจัดเก็บข้อมูล (Data store) เป็นแหล่งจัดเก็บข้อมูลทั้งในส่วนของข้อมูลนำเข้าหรือส่งออกไปเป็นผลลัพธ์ แหล่งจัดเก็บข้อมูลเป็นสัญลักษณ์ที่แสดงถึงแหล่งจัดเก็บข้อมูลทั้งในส่วนของข้อมูลนำเข้าและข้อมูลที่เป็นผลลัพธ์ ในการพัฒนาโปรแกรมจำเป็นต้องมีแหล่งจัดเก็บข้อมูล ไม่เช่นนั้นข้อมูลนำเข้าและข้อมูลที่เป็นผลลัพธ์รวมไปถึงข้อมูลอื่น ๆ ที่จำเป็นสำหรับการประมวลผลจะไม่สามารถจัดเก็บเพื่อใช้งานได้ครั้งต่อไปอีก

1.3.2 กฎในการสร้างแผนภาพกระแสข้อมูล ก่อนการสร้างแผนภาพกระแสข้อมูล นักวิเคราะห์โปรแกรมจำเป็นต้องศึกษากฎที่เป็นพื้นฐานในการสื่อสาร เพื่อให้เกิดความเข้าใจร่วมกัน ทั้งภายในทีมนักพัฒนาโปรแกรม และผู้ที่มีส่วนร่วมในการพัฒนาโปรแกรม โดยสามารถสรุปสาระสำคัญ ได้แก่ 1) แต่ละกระบวนการทำงานจะต้องมีกระแสข้อมูลเข้า (Input) โดยอาจมาจากเอนทิตีภายนอกแหล่งจัดเก็บข้อมูล หรือกระบวนการทำงานอื่น ๆ 2) แต่ละกระบวนการทำงานจะต้องมีกระแสข้อมูลที่เป็นผลลัพธ์ (Output) โดยอาจส่งไปยังเอนทิตีภายนอก แหล่งจัดเก็บข้อมูล หรือกระบวนการทำงานอื่น ๆ 3) เอนทิตีภายนอกจะดำเนินการรับ-ส่งข้อมูลระหว่างกันโดยไม่ผ่านกระบวนการทำงานไม่ได้ 4) แหล่งจัดเก็บข้อมูลจะดำเนินการรับ-ส่งข้อมูลระหว่างกันโดยไม่ผ่านกระบวนการทำงานไม่ได้ 5) เอนทิตี้ภายนอกจะดำเนินการรับ-ส่งข้อมูลกับแหล่งจัดเก็บข้อมูลโดยไม่ผ่านกระบวนการทำงานไม่ได้ และ 6) กระแสข้อมูลจะต้องมีทิศทางการไหลไปยังเป้าหมายเพียงทิศทางเดียวเท่านั้น การใช้ลูกศร 2 ทิศทางกำกับทิศทางการไหลของกระแสข้อมูลจึงไม่ถูกต้อง เช่น ถ้าต้องการค้นหาข้อมูล ทิศทางการไหลของกระแสข้อมูลก็จะออกจากแหล่งจัดเก็บข้อมูล แต่ถ้าต้องการบันทึกข้อมูลทิศทางการไหลของกระแสข้อมูลก็จะตรงไปยังแหล่งจัดเก็บข้อมูล

1.3.3 วิธีการสร้างแผนภาพกระแสข้อมูล แผนภาพกระแสข้อมูลจะมีวิธีการสร้างอย่างเป็นโปรแกรมด้วยหลักการทำงานจากบนลงล่าง (Top-Down approach) โดยสามารถใช้วิธีการแบ่งย่อยแผนภาพ (Decomposition diagram) ได้ความสัมพันธ์ของแผนภาพกระแสข้อมูลแต่ละระดับโดยประกอบด้วย ได้แก่ 1) แผนภาพแสดงภาพรวมของโปรแกรมทั้งหมดเรียกว่า แผนภาพบริบท (Context diagram) เป็นแผนภาพในระดับบนสุด ซึ่งจะแสดงถึงขอบเขตของโปรแกรมว่ามีปฏิสัมพันธ์กับสิ่งแวดล้อมของโปรแกรมอย่างไร โดยจะไม่แสดงรายละเอียดของกระบวนการทำงานภายในโปรแกรม และไม่แสดงแหล่งจัดเก็บข้อมูลต่าง ๆ ของโปรแกรม 2) แผนภาพแสดงองค์ประกอบของกระบวนการทำงานหลักของโปรแกรมเรียกว่า แผนภาพกระแสข้อมูลระดับที่ 0 (Data Flow Diagram Level 0) เป็นแผนภาพในระดับถัดมาจากแผนภาพบริบท โดยแสดงถึงรายละเอียดของกระบวนการทำงานหลักของโปรแกรมว่าจะมีการประสานงานรับ-ส่งข้อมูลและสารสนเทศกับเอนทิตีภายนอกใดบ้าง นอกจากนั้นยังแสดงถึงการเชื่อมโยงการทำงานระหว่างกระบวนการทำงานต่าง ๆ รวมทั้งการจัดเก็บและค้นหาข้อมูลจากแหล่งจัดเก็บข้อมูลต่าง ๆ โดยมีหลักการสร้างแผนภาพระดับที่ 0 และ 3) แผนภาพแสดงรายละเอียดการทำงานของแต่ละกระบวนการทำงานในระดับที่ 0 เรียกว่า แผนภาพกระแสข้อมูลระดับที่ 1 (Data Flow Diagram Level 1) หมายถึง แผนภาพระดับที่ 1 แผนภาพระดับที่ 2 และแผนภาพระดับย่อยลงไปอีกตามความจำเป็นในการวิเคราะห์โปรแกรม ต้องอธิบายถึงรายละเอียดของแผนภาพระดับที่ 1 ซึ่งเป็นแผนภาพย่อยที่ใช้อธิบายรายละเอียดของกระบวนการทำงานหนึ่ง ๆ จากแผนภาพระดับที่ 0 เนื่องจากในกระบวนการทำงานหนึ่ง ๆ จึงจำเป็นต้องมีกิจกรรมย่อยหลายขั้นตอน ดังนั้น การแสดง

รายละเอียดได้มากเท่าใด การพัฒนาโปรแกรมให้ถูกต้องและสอดคล้องกับความต้องการใช้งานก็จะยิ่งมากขึ้นเท่านั้น

1.3.4 แบบจำลองฐานข้อมูล (Data model) เป็นแบบจำลองที่ใช้สำหรับอธิบายถึงโครงสร้าง และความสัมพันธ์ระหว่างข้อมูลภายในฐานข้อมูล จากรูปแบบที่เป็นแนวคิดที่ยากต่อการเข้าใจให้อยู่ในรูปแบบที่สามารถเข้าใจ และจับต้องได้ง่ายขึ้น โดยทั่วไปภายในแบบจำลองของข้อมูลประกอบด้วย 3 ส่วน ได้แก่ 1) ส่วนที่ใช้แทนข้อมูลต่าง ๆ ที่ประกอบกันขึ้นเป็นฐานข้อมูล โดยทั่วไปมักจะแทนด้วยรูปสี่เหลี่ยม 2) ส่วนของกฎต่าง ๆ ที่ใช้ควบคุมความถูกต้องของข้อมูลภายในฐานข้อมูล และ 3) ส่วนของการกระทำต่าง ๆ ที่สามารถใช้งานร่วมกับฐานข้อมูล

2. การออกแบบ (Design) ในวงจรของการพัฒนาโปรแกรม ขั้นตอนของการออกแบบโปรแกรมมีรายละเอียดในการดำเนินงานและผลลัพธ์ที่ได้จากขั้นตอนนี้ค่อนข้างมาก การออกแบบจะแบ่งออกเป็น 2 ส่วน ได้แก่ 1) การออกแบบโปรแกรมเชิงตรรกะ (Logical design) ทั้งในส่วนของผลลัพธ์ส่วนการนำเข้าข้อมูล และส่วนต่อประสานกับผู้ใช้ ที่ได้จากการวิเคราะห์โปรแกรม และ 2) การออกแบบโปรแกรมเชิงกายภาพ (Physical design) เป็นการนำเสนอผลจากออกแบบโปรแกรมเชิงตรรกะมาระบุลักษณะการทำงานของโปรแกรมทางกายภาพ ทั้งในส่วนการกำหนดคุณลักษณะเฉพาะของฮาร์ดแวร์และซอฟต์แวร์ที่ใช้ในโปรแกรม รวมไปถึงการออกแบบโปรแกรมรักษาความปลอดภัย โดยมีส่วนในการออกแบบดังนี้

2.1 การออกแบบผลลัพธ์ของโปรแกรม (Output design) เป็นส่วนที่ควรดำเนินการเป็นลำดับแรกของการออกแบบเชิงตรรกะ เนื่องจากผลลัพธ์ของโปรแกรมที่สอดคล้องกับความต้องการของผู้ใช้โปรแกรม จะส่งผลกระทบไปยังการออกแบบส่วนอื่น ๆ ทั้งการออกแบบส่วนนำเข้าข้อมูล และการออกแบบส่วนประสานกับผู้ใช้ นอกจากนี้ผู้ใช้โปรแกรมจะมีความคาดหวังกับผลลัพธ์ของโปรแกรมที่จะต้องช่วยเพิ่มประสิทธิภาพในการปฏิบัติงาน ดังนั้น ผลลัพธ์ของโปรแกรมจึงมักจะถูกใช้ในการวัดระดับความพึงพอใจ และความสำเร็จในการพัฒนาโปรแกรมนั้น

2.2 ปัจจัยหลักที่เกี่ยวข้องกับการออกแบบผลลัพธ์ของโปรแกรม เพื่อให้การออกแบบผลลัพธ์ของโปรแกรมสอดคล้องกับความต้องการของผู้ใช้โปรแกรมและผลจากการวิเคราะห์โปรแกรมในขั้นตอนที่ผ่านมา นักวิเคราะห์โปรแกรมจำเป็นจะต้องพิจารณาถึงปัจจัยหลักที่เกี่ยวข้องกับผลลัพธ์ของโปรแกรมให้ชัดเจนก่อนออกแบบผลลัพธ์แต่ละรายการ

2.3 การแบ่งพื้นที่ในการแสดงผลลัพธ์ การแสดงผลลัพธ์ไม่ว่าจะผ่านสื่อใด ๆ ก็ตามมักมีการแบ่งพื้นที่ในการนำเสนอผลลัพธ์เป็น 3 ส่วนหลัก ได้แก่ 1) พื้นที่ส่วนหัว (Heading zone) สำหรับแสดงรายละเอียด 2) พื้นที่ส่วนเนื้อหา (Body zone) สำหรับแสดงรายละเอียดเนื้อหาของผลลัพธ์ไม่ว่าจะอยู่ในรูปแบบใดก็ตาม ควรคำนึงถึงการใช้เนื้อหาอย่างสมดุลและมีการเว้นวรรคข้อความอย่าง

เหมาะสม และ 3) พื้นที่ส่วนสรุป (Summary zone) สำหรับแสดงรายละเอียดผลสรุปเนื้อหาของโปรแกรม

2.4 หลักการออกแบบผลลัพธ์ของโปรแกรม การออกแบบผลลัพธ์ของโปรแกรมมีข้อควรระวังในภาพรวม ได้แก่ 1) ควรใช้เทคนิคในการเน้นข้อมูลที่สำคัญ 2) ควรจัดวางรูปแบบข้อมูลผลลัพธ์ให้ถูกต้องตามมาตรฐาน 3) ในกรณีที่เป็นหลักฐานทางการเงิน ควรจะแสดงจำนวณเงินเป็นตัวอักษรด้วย เพื่อใช้ยืนยันความถูกต้องของตัวเลขดังกล่าว 4) การนำเสนอข้อมูลในรูปแบบข้อความ 5) การนำเสนอข้อมูลในรูปแบบตาราง 6) การนำเสนอข้อมูลในรูปแบบกราฟฟิก และ 7) การออกแบบผลลัพธ์จะต้องคำนึงถึงประโยชน์ใช้สอย รวมทั้งการอำนวยความสะดวกในการใช้งานสารสนเทศแก่ผู้ใช้มากที่สุด

2.5 ประเภทของผลลัพธ์ ผลลัพธ์ของโปรแกรมแบ่งได้เป็น 2 ประเภท ได้แก่ 1) ผลลัพธ์สำหรับใช้ติดต่อกับหน่วยงานภายนอก (External output) หมายถึง ผลลัพธ์ของโปรแกรมที่ส่งไปยังผู้ใช้ภายนอก และ 2) ผลลัพธ์สำหรับใช้งานภายในองค์กร (Internal output) หมายถึงผลลัพธ์ของโปรแกรมที่ใช้สนับสนุนการปฏิบัติงานภายในองค์กร รวมไปถึงการบริหารจัดการองค์กรของผู้บริหารระดับต่าง ๆ

3. การพัฒนาโปรแกรม (Implementation and coding) ในขั้นตอนนี้โปรแกรมเมอร์จะเขียนโปรแกรมตามข้อมูลที่ได้จากเอกสารข้อมูลเฉพาะของการออกแบบ ปกติแล้วนักวิเคราะห์โปรแกรมไม่มีหน้าที่เกี่ยวข้องในการเขียนโปรแกรม แต่ถ้าโปรแกรมเมอร์คิดว่าการเขียนอย่างอื่นดีกว่าจะต้องปรึกษานักวิเคราะห์โปรแกรมก่อน เพื่อนักวิเคราะห์จะบอกได้ว่าโปรแกรมที่จะแก้ไขนั้นมีผลกระทบกับโปรแกรมทั้งหมดหรือไม่ การพัฒนาโปรแกรมแบ่งออกเป็น 2 วิธี ดังนี้

3.1 การพัฒนาโปรแกรมเชิงโครงสร้าง (Structured system development) เป็นวิธีการพัฒนาโปรแกรมแบบดั้งเดิม ที่เกี่ยวข้องกับการวิเคราะห์เชิงโครงสร้าง การออกแบบเชิงโครงสร้างและการโปรแกรมเชิงโครงสร้าง โดยมีรายละเอียดดังนี้

3.1.1 เทคนิคการโปรแกรมเชิงโครงสร้าง (Structured programming technique) เป็นเทคนิคในการพัฒนาโปรแกรมที่มีหนึ่งจุดเริ่มต้นและจุดสิ้นสุดโดยจะมีรูปแบบชุดคำสั่งที่ใช้ในการประมวลผล 3 รูปแบบ ดังนี้

3.1.1.1 ชุดคำสั่งเรียงเป็นลำดับ (Sequence)

3.1.1.2 ชุดคำสั่งกำหนดทางเลือกหรือการตัดสินใจ (Decision)

3.1.1.3 ชุดคำสั่งเพื่อการทำซ้ำหรือวนลูป (Repetition/Looping)

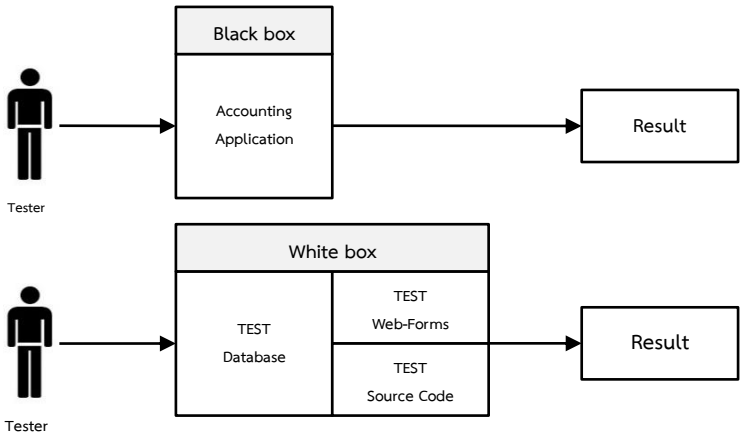
3.1.2 เทคนิคการออกแบบเชิงโครงสร้าง (Structure design technique) เป็นเทคนิคในการนำเสนอโปรแกรมที่พัฒนาผ่านทางผังโครงสร้าง (Structure chart) ซึ่งเป็นแผนภาพที่แสดงให้เห็นถึงโมดูลภาพในโปรแกรมและแสดงวิธีการจัดการกับโมดูลและการส่งผ่านข้อมูลระหว่าง

โมดูล ทำให้นักออกแบบทราบถึงกระบวนการทำงานของโปรแกรม ฟังก์ชันหลักของโปรแกรมข้อมูลที่ใช้ในโปรแกรม และผลลัพธ์ที่ได้จากการทำงานของโปรแกรม ต่อมามีการพัฒนาเทคนิคการออกแบบแฟ้มข้อมูลและฐานข้อมูลขึ้นมาเพื่อใช้ควบคู่กับการออกแบบเชิงโครงสร้าง เช่น ระบบจัดการฐานข้อมูล (DBMS) ที่สามารถออกแบบส่วนติดต่อผู้ใช้และส่วนที่ใช้โต้ตอบกับฐานข้อมูล

3.1.3 การวิเคราะห์เชิงโครงสร้าง (Structure analysis technique) จะวิเคราะห์ข้อมูลจากเอกสาร รายงาน และขั้นตอนการทำงานของโปรแกรมเดิม จากนั้นสร้างแบบจำลองกระบวนการทำงานของโปรแกรม เพื่อแสดงให้เห็นกระบวนการทำงานของโปรแกรมเดิม ข้อมูลที่โปรแกรมต้องการข้อมูลที่ได้จากการทำงานของโปรแกรม ผู้ใช้ที่เกี่ยวข้องกับโปรแกรม และแฟ้มที่ใช้จัดเก็บข้อมูลของโปรแกรมมีการแสดงทิศทางการไหลของข้อมูลในแต่ละกระบวนการ

3.2 การพัฒนาโปรแกรมเชิงวัตถุ (Object-Oriented system development) เป็นวิธีการพัฒนาโปรแกรมที่นักวิเคราะห์โปรแกรมมีมุมมองข้อมูลทุกส่วนของโปรแกรมเป็นเหมือนวัตถุ (Objects) และจัดประเภทของวัตถุตามคุณลักษณะทางนามธรรม (Abstract) ออกเป็นกลุ่ม ๆ เรียกว่า คลาส (Class) ซึ่งจะกำหนดคุณสมบัติของวัตถุนั้นอย่างชัดเจน แต่ละคลาสจะมีสถานะ (Status) และพฤติกรรม (Behavior) ตามบทบาทของงาน โดยมีรายละเอียดหรือคุณสมบัติของวัตถุ (Characteristic) ที่เก็บซ่อน (Encapsulate) ไว้ในคลาส วัตถุแต่ละวัตถุมีความเป็นอิสระต่อกัน สามารถติดสื่อสารกันระหว่างวัตถุด้วยกันในข้อความ (Message) การปรับปรุงหรือเปลี่ยนแปลงโปรแกรม จะทำได้ด้วยการเปลี่ยนแปลงรายละเอียดด้านคุณสมบัติของวัตถุหรือคลาส และฟังก์ชันการทำงานที่เกี่ยวข้องกับวัตถุนั้น ๆ ซึ่งจะไม่ส่งผลกระทบต่อวัตถุอื่น ๆ

4. การทดสอบ (Testing) เมื่อโปรแกรมได้พัฒนาขึ้น ก่อนนำไปใช้งานจริง จำเป็นต้องดำเนินการทดสอบโปรแกรมก่อน การทดสอบเบื้องต้นด้วยการสร้างข้อมูลจำลองขึ้นมาเพื่อตรวจสอบการทำงานของโปรแกรม หากพบข้อผิดพลาดก็ปรับปรุงแก้ไขให้ถูกต้อง การทดสอบโปรแกรมมีการทดสอบ 2 แบบ ได้แก่ 1) Functional testing (Black box testing) และ 2) Structural (White box testing) แสดงในตัวอย่างภาพที่ 2-2



ภาพที่ 2-2 ความแตกต่างของการทดสอบโปรแกรม

4.1 Functional testing (Black box testing) ผู้ทดสอบจะมองทั้งโปรแกรมเป็นเหมือนกล่องดำ (Black box) ไม่สนใจว่าการทำงานภายในเป็นอย่างไร เช่น การทำงานของเครื่องจักร โดยเราแค่ป้อนคำสั่งให้เครื่องจักรแล้ว ดูว่าเครื่องจักรทำงานถูกต้องหรือไม่ ไม่สนใจกระบวนการทำงานว่าข้างในทำอะไรบ้างส่วนการสร้างกรณีทดสอบจะต้องใช้เอกสารกำหนดคุณลักษณะความต้องการใช้งานซอฟต์แวร์ (Requirement specification) แบ่งการประเมินออกเป็น 4 ด้าน ดังนี้

4.1.1 ด้านตรงตามความต้องการของโปรแกรม (Functional requirement test) เป็นการประเมินผลความถูกต้องและประสิทธิภาพของโปรแกรม ตรงตามความต้องการของผู้ใช้โปรแกรมมีความเหมาะสมมากน้อยเพียงใด

4.1.2 ด้านการทำงานของโปรแกรม (Functional test) เป็นการประเมินความถูกต้องและประสิทธิภาพในการทำงานของโปรแกรม สามารถทำงานได้ตามฟังก์ชันของโปรแกรมมีความเหมาะสมมากน้อยเพียงใด

4.1.3 ด้านการใช้งานโปรแกรม (Usability test) เป็นการประเมินลักษณะการออกแบบโปรแกรม มีความง่ายต่อการใช้งานมีความเหมาะสมมากน้อยเพียงใด

4.1.4 ด้านความปลอดภัยของโปรแกรม (Security test) เป็นการประเมินโปรแกรมในด้านการรักษาความปลอดภัยของข้อมูลในโปรแกรม มีความเหมาะสมมากน้อยเพียงใด

4.2 Structural testing (White box testing) การทดสอบแบบ White box นั้นจะต้องรู้ว่าซอฟต์แวร์สร้างขึ้นมาอย่างไร และใช้การทำงานนั้นสร้างกรณีทดสอบขึ้นมา โดยจะต้องกำหนดกรณีทดสอบตามขั้นตอนการทำงานของฟังก์ชันที่สร้างขึ้นมาจริง ๆ เทคนิคในการทดสอบโปรแกรมสามารถแบ่งออกเป็น 2 กลุ่ม ได้แก่ 1) การทดสอบโดยไม่ใช้เครื่องคอมพิวเตอร์ (Manual testing) และ 2) การทดสอบด้วยเครื่องคอมพิวเตอร์ (Automated testing)

4.3 การเปรียบเทียบการทำงานระหว่าง Black box กับ White box การทดสอบแบบ Black box และ White box มีวัตถุประสงค์เพื่อสร้างกรณีทดสอบขึ้นมาทดสอบซอฟต์แวร์ ในส่วนของ Black box จะอาศัยข้อมูลจาก requirement specification เพียงอย่างเดียวในการสร้างกรณีทดสอบ ส่วนการทดสอบแบบ White box นั้น จะอาศัยขั้นตอนการทำงานของซอฟต์แวร์ หรือ source code เพียงอย่างเดียวในการสร้างกรณีทดสอบ ดังนั้น จะต้องใช้การทดสอบทั้ง 2 วิธีร่วมกัน ตัวอย่าง 1) ถ้าใช้การทดสอบแบบ White box เพียงอย่างเดียว ก็จะไม่รู้ว่าเราพัฒนาฟังก์ชันต่าง ๆ ตาม Requirement specification ครบแล้วหรือไม่ และไม่สามารถบอกได้ว่ายังมี Requirement specification ส่วนใดที่ยังไม่ได้พัฒนา หรือ 2) ถ้าใช้การทดสอบแบบ Black box เพียงอย่างเดียว ในกรณีที่มีการสร้างฟังก์ชันที่นอกเหนือไปจากที่กำหนดไว้ใน Requirement specification จะไม่สามารถบอกได้ว่า สร้างฟังก์ชันที่นอกเหนือกับที่กำหนดไว้

5. การใช้งาน (Deployment) เมื่อทดสอบโปรแกรมผ่านมาตรฐานการทดสอบจึงนำไปติดตั้งเพื่อใช้งานบนสถานการณ์จริง จะเปลี่ยนการทำงานจากโปรแกรมเดิมไปเป็นโปรแกรมใหม่ แต่การเปลี่ยนแปลงไปสู่สิ่งใหม่ย่อมมีผลกระทบต่อผู้ใช้งานที่ยังคงมีความคุ้นเคยกับวิธีการดำเนินงานแบบเก่า รวมทั้งข้อจำกัดในเรื่องของความพร้อมในการเปลี่ยนแปลง ดังนั้นจึงควรเลือกแนวทางที่เหมาะสมในการติดตั้ง ซึ่งแบ่งออกเป็น 4 แนวทาง ดังนี้

5.1 การติดตั้งแบบทันทีทันใด (Direct installation) เป็นวิธีการติดตั้งโปรแกรมใหม่ทันที และยกเลิกการใช้งานโปรแกรมเก่าทันที

5.2 การติดตั้งแบบขนาน (Parallel installation) เป็นวิธีการติดตั้งโปรแกรมใหม่ไปพร้อม ๆ กับการใช้งานโปรแกรมเก่าจนกว่าผู้ใช้จะมีความพอใจโปรแกรมใหม่ จึงตัดสินใจหยุดใช้งานโปรแกรมเก่า

5.3 การติดตั้งแบบนำร่อง (Pilot installation) เป็นวิธีการติดตั้งที่มีการใช้งานโปรแกรมแกรมใหม่เพียงหน่วยเดียวเพื่อเป็นการนำร่อง แล้วจึงค่อยปรับเปลี่ยนทั้งหมดเมื่อเห็นว่าโปรแกรมใหม่นั้นลงตัวแล้ว

5.4 การติดตั้งแบบทยอยติดตั้งเป็นระยะ (Phased installation) เป็นวิธีการที่ติดตั้งโปรแกรมใหม่เพียงบางส่วนก่อนระยะหนึ่งควบคู่ไปกับการใช้งานโปรแกรมเก่า แล้วจึงค่อย ๆ ทยอยใช้โปรแกรมงานใหม่เพิ่มขึ้นทีละส่วนจนกระทั่งครบทุกส่วนของโปรแกรมใหม่อย่างเต็มรูปแบบ ในการใช้งานโปรแกรมใหม่ทดแทนโปรแกรมงานเดิม นักวิเคราะห์โปรแกรมจำเป็นต้องมีการแปลงข้อมูลจากโปรแกรมงานเดิมมาให้อยู่ในรูปแบบที่โปรแกรมใหม่ สามารถนำไปใช้งานได้ ในขั้นตอนนี้นักวิเคราะห์โปรแกรมจะต้องดำเนินการจัดทำเอกสารคู่มือโปรแกรม รวมถึงการฝึกอบรมผู้ใช้

6. การบำรุงรักษา (Maintenance) หลังจากที่โปรแกรมที่พัฒนาขึ้นมาใหม่ได้ถูกนำไปใช้งานเป็นที่เรียบร้อย หากพบข้อผิดพลาดหรือข้อบกพร่องจากการทำงานของโปรแกรม นักวิเคราะห์

โปรแกรมจึงจำเป็นต้องดำเนินการติดตามและแก้ไขให้ถูกต้อง รวมถึงกรณีที่ข้อมูลจัดเก็บมีปริมาณมากขึ้น การขยายระบบเครือข่ายเพื่อรองรับเครื่องลูกข่ายที่มีจำนวนมากขึ้น บางกรณีอาจจำเป็นต้องเขียนโปรแกรมเพิ่มเติมหากผู้ใช้มีความต้องการเพิ่มขึ้น ดังนั้น ในขั้นตอนของการกำหนดความต้องการ นักวิเคราะห์โปรแกรมจึงจำเป็นต้องมีการจัดทำเอกสารข้อตกลงร่วมกันทั้งสองฝ่ายถึงขอบเขตในการพัฒนาโปรแกรม และกรณีที่มีการแก้ไขหรือพัฒนาโปรแกรมเพิ่ม

จากการศึกษางานวิจัยดังกล่าว ปรากฏว่า วิธีการควบคุมการใช้ข้อสอบกับวิธีการควบคุมสมดุลเนื้อหาแยกออกจากกัน ดังนั้น ผู้วิจัยมีความสนใจวิธีโหลดบาลานซ์ซึ่งเป็นวิธีใหม่ มาใช้ในการควบคุมการใช้ข้อสอบและการควบคุมสมดุลเนื้อหาพร้อมกันเพื่อให้มีประสิทธิภาพมากขึ้น วิธีโหลดบาลานซ์ซึ่งมีคุณสมบัติการควบคุมสมดุลเนื้อหาโดยใช้วิธีหาค่าเฉลี่ยถ่วงน้ำหนักเพื่อหาตัวแทนด้านเนื้อหา ส่วนการควบคุมการใช้ข้อสอบใช้วิธีเรียงลำดับจากน้อยไปหามากระหว่างความถี่สะสมจากการใช้ข้อสอบและผลต่างของค่าประมาณความสามารถกับค่าสารสนเทศสูงสุด แล้วเลือกข้อสอบข้อแรกเป็นข้อสอบข้อถัดไป แล้วนำวิธีดังกล่าวมาพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ โดยใช้วงจรการพัฒนาโปรแกรม

บทที่ 3
วิธีดำเนินการวิจัย

การวิจัยนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง (LB) 2) พัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยใช้วิธีควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา 2 วิธี คือ 2.1) วิธีโหลดบาลานซ์ซิ่ง (LB) และ 2.2) วิธี Item Eligibility (IE) 3) เปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา 2 วิธี ในด้าน 3.1) จำนวนข้อสอบที่ใช้ในการทดสอบ 3.2) เวลาที่ใช้ในการทดสอบ 3.3) อัตราการแสดงการใช้ข้อสอบสูงสุด 3.4) สัดส่วนข้อสอบที่มีการแสดงมากเกินไป และ 3.5) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และ 4) ศึกษาความสัมพันธ์ระหว่างค่าความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT) การวิจัยแบ่งเป็น 4 ระยะ ดังนี้

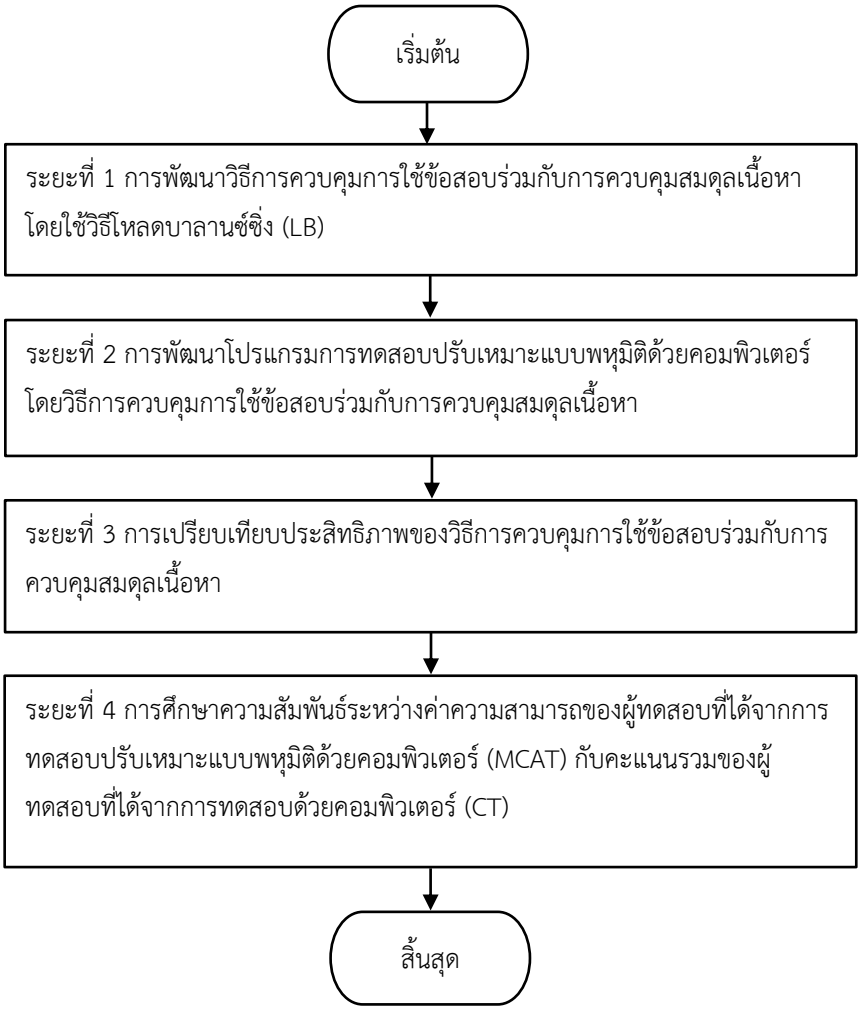
ระยะที่ 1 การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง (LB)

ระยะที่ 2 การพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

ระยะที่ 3 การเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

ระยะที่ 4 การศึกษาความสัมพันธ์ระหว่างค่าความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT)

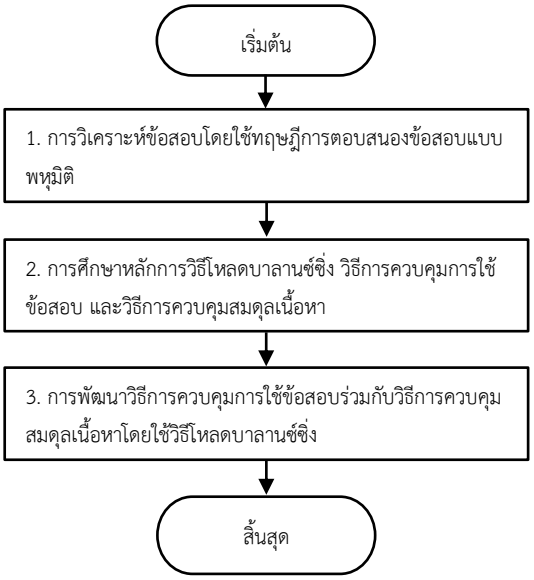
ลำดับการดำเนินการวิจัย แสดงดังภาพที่ 3-1



ภาพที่ 3-1 ลำดับการดำเนินการวิจัยแบ่งตามระยะ

ระยะที่ 1 การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง (LB)

การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง สำหรับการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ แบ่งออกเป็น 3 ขั้นตอน ได้แก่ 1) การวิเคราะห์ข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ 2) การศึกษาหลักการวิธีโหลดบาลานซ์ซิ่ง วิธีการควบคุมการใช้ข้อสอบ และวิธีการควบคุมสมดุลเนื้อหา และ 3) การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับวิธีการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง ขั้นตอนในการพัฒนา แสดงดังภาพที่ 3-2



ภาพที่ 3-2 ขั้นตอนการพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

ขั้นตอนที่ 1 การวิเคราะห์ข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ ซึ่งเป็นขั้นตอนวิเคราะห์ข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ (Multidimensional Item Response Theory: MIRT) แบบ 3 พารามิเตอร์ มีขั้นตอนย่อย 4 ขั้นตอน ดังนี้

1. การรวบรวมข้อสอบ

1.1 เตรียมเอกสารที่ใช้แนบสำหรับออกหนังสือขอความอนุเคราะห์เค้าโครงการวิจัยฉบับย่อ

1.2 ขอความอนุเคราะห์จากวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา เกี่ยวกับการออกหนังสือขอความอนุเคราะห์ข้อมูลผลการทดสอบ V-NET ระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) วิชาเทคโนโลยีสารสนเทศและการสื่อสาร หรือวิชาคอมพิวเตอร์และการสื่อสาร จากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) (สทศ.)

1.3 จัดส่งหนังสือขอความอนุเคราะห์ข้อมูลถึงผู้อำนวยการ สทศ. เลขที่หนังสือ ศธ. 6224/0208 ลงวันที่ 20 มิถุนายน 2561

1.4 ได้ตอบรับหนังสือจาก สทศ. เลขที่หนังสือ สทศ.0001/996 เรื่องเอกสารรับผลการทดสอบ ระหว่างปี 2556-2560 รวม 5 ปี จำนวน 240 ข้อ เมื่อวันที่ 26 มิถุนายน 2561 เป็นรูปแบบไฟล์ Excel

2. การวิเคราะห์คุณภาพข้อสอบโดยใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ

2.1 ตรวจสอบความสมบูรณ์ของข้อสอบ V-NET ทั้งข้อข้อสอบ ตัวเลือก และเฉลย ถ้าข้อสอบข้อใดไม่มีความสมบูรณ์คัดข้อสอบข้อนั้นออก โดยแบ่งข้อสอบเป็น 2 มิติ มิติที่ 1 เป็นหมวด

เนื้อหาหลัก 5 หมวด ตามเนื้อหาวิชาเทคโนโลยีสารสนเทศและการสื่อสาร ได้แก่ 1) เครือข่าย 2) ข้อมูล 3) ซอฟต์แวร์ 4) ฮาร์ดแวร์ 5) กระบวนการและผู้ใช้ และมิติที่ 2 เป็นความรู้ทางไอซีที ประกอบด้วย 6 ด้าน ได้แก่ 1) การเข้าถึงสารสนเทศ 2) การจัดการสารสนเทศ 3) การประเมิน 4) การสร้างสิ่งใหม่จากความเข้าใจ 5) การสื่อสาร และ 6) การใช้ไอซีทีอย่างเหมาะสม ผู้เชี่ยวชาญได้พิจารณาความตรงเชิงเนื้อหาด้วยดัชนีความตรงเชิงเนื้อหา (Content Validity Index: CVI) จำแนกข้อสอบทางด้านหมวดเนื้อหาและทางด้านความรู้ทางไอซีที จำนวน 3 คน ประกอบด้วย

1) รองศาสตราจารย์ ดร.ปณิตา วรรณพิรุณ อาจารย์ประจำคณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

2) ผู้ช่วยศาสตราจารย์ ดร.สยาม แกมขุนทด อาจารย์ประจำคณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

3) ผู้ช่วยศาสตราจารย์ ดร.รุ่งทิวา เสาร์สิงห์ อาจารย์ประจำคณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ

จากภาพที่ 3-3 แสดงตัวอย่างการวิเคราะห์เนื้อหาข้อสอบความรู้ทางไอซีที โดยผู้เชี่ยวชาญพิจารณาในแต่ละข้อ สามารถวัดได้ตรงตามที่ต้องการหรือไม่ เป็นหมวดเนื้อหาด้านฮาร์ดแวร์ จาก 5 ด้าน และด้านความรู้ทางไอซีที มี 6 ด้าน พิจารณาความสอดคล้องของเนื้อหา 4 ตัวเลือก ได้แก่ ไม่สอดคล้อง สอดคล้องบางส่วน ค่อนข้างสอดคล้อง และมีความสอดคล้อง การหาสัดส่วนของข้อสอบที่ผ่านการประเมินจากผู้เชี่ยวชาญ (เฉพาะข้อสอบที่ผู้เชี่ยวชาญตอบ ค่อนข้างสอดคล้อง หรือมีความสอดคล้อง) หารด้วยข้อสอบทั้งหมดในด้านนั้น และการหาค่าเฉลี่ยของข้อสอบทั้งหมด ค่าดัชนีความตรงเชิงเนื้อหา (Average I-CVI) อยู่ระหว่าง 0.90 ถึง 1.00 (Polit & Beck, 2006)

หมวดเนื้อหา	ความรู้ทางไอซีที						ผู้เชี่ยวชาญพิจารณา			
4) ด้านฮาร์ดแวร์	การเข้าถึงสารสนเทศ	การจัดการสารสนเทศ	การประเมิน	การสร้างสิ่งใหม่	การสื่อสาร	การใช้ไอซีทีอย่างเหมาะสม	ไม่สอดคล้อง	สอดคล้องบางส่วน	ค่อนข้างสอดคล้อง	มีความสอดคล้อง
1. คุณพ่อมีปัญหาเปิดเครื่องคอมพิวเตอร์ส่วนบุคคล (PC) ไม่ติด โดยอาการไม่มีไฟฟ้าเข้ามาในตัวเครื่อง ข้อใดต่อไปนี้เป็นกลุ่มของปัญหา 1) ไม่ได้เสียบปลั๊กไฟ, เมนบอร์ดเสีย, ซอฟต์แวร์มีปัญหา 2) สายไฟในเครื่องขาด, ซอฟต์แวร์มีปัญหา, เมนบอร์ดเสีย *3) ไม่ได้เสียบปลั๊กไฟ, พาวเวอร์ซัพพลายเสีย, เมนบอร์ดเสีย 4) สายไฟในเครื่องขาด, ซอฟต์แวร์มีปัญหา, หน่วยความจำเสีย 5) ไม่ได้เสียบปลั๊กไฟ, พาวเวอร์ซัพพลายเสีย, ซอฟต์แวร์มีปัญหา			/							

ภาพที่ 3-3 ตัวอย่างการวิเคราะห์เนื้อหาข้อสอบความรู้ทางไอซีทีโดยผู้เชี่ยวชาญ

ตารางที่ 3-1 จำนวนข้อสอบจำแนกตามหมวดเนื้อหาและความรู้ทางไอซีที

หมวดเนื้อหา/ความรู้ไอซีที	การเข้าถึงสารสนเทศ	การจัดการสารสนเทศ	การประเมิน	การสร้างสิ่งใหม่จากความเข้าใจ	การสื่อสาร	การใช้ไอซีทีอย่างเหมาะสม	รวม
1) เครือข่าย	5	5	5	5	15	8	43
2) ข้อมูล	10	12	10	12	5	12	61
3) ซอฟต์แวร์	10	8	10	11	5	11	55
4) ฮาร์ดแวร์	5	6	6	5	5	14	41
5) กระบวนการและผู้ใช้	10	6	12	0	6	6	40
รวม	40	37	43	33	36	51	240

2.2 จัดเตรียมข้อมูลจากผลการตอบข้อสอบ พร้อมกำหนดเงื่อนไขการวิเคราะห์คุณภาพของข้อสอบตามหลักการทางทฤษฎีการตอบสนองข้อสอบแบบพหุมิติ 3 พารามิเตอร์

2.3 วิเคราะห์คุณภาพข้อสอบ โดยใช้โปรแกรมสำเร็จรูป MULTILOG V.7.03 (Thissen, Chen, & Bock, 2003) เขียนโปรแกรมนำไฟล์เข้าแล้ววิเคราะห์และแสดงผล

3. การคัดเลือกข้อสอบ คัดเลือกข้อสอบจากผลการวิเคราะห์คุณภาพของข้อสอบ โดยคัดเลือกข้อสอบตรงตามเกณฑ์ที่กำหนดไว้ (Urry, 1977) ได้แก่ ค่าอำนาจจำแนกของข้อสอบ (a) มีค่าตั้งแต่ 0.50 ถึง 2.50 ค่าความยากของข้อสอบ (b) มีค่าตั้งแต่ -2.50 ถึง 2.50 และค่าโอกาสการเดาของข้อสอบ (c) มีค่าไม่เกิน 0.30 แสดงในตารางที่ 3-2

ตารางที่ 3-2 จำนวนข้อสอบที่ผ่านการคัดเลือกจำแนกตามเนื้อหาและความรู้ทางไอซีที

หมวดเนื้อหา/ความรู้ไอซีที	การเข้าถึงสารสนเทศ	การจัดการสารสนเทศ	การประเมิน	การสร้างสิ่งใหม่จากความเข้าใจ	การสื่อสาร	การใช้ไอซีทีอย่างเหมาะสม	รวม
1) เครือข่าย	2	5	4	5	12	6	34
2) ข้อมูล	8	7	7	9	5	11	47
3) ซอฟต์แวร์	9	8	7	11	2	11	48
4) ฮาร์ดแวร์	2	6	6	3	5	13	35
5) กระบวนการและผู้ใช้	6	3	12	0	4	3	28
รวม	27	29	36	28	28	44	192

จากตารางที่ 3-2 จำนวนข้อสอบที่ผ่านการคัดเลือก จำนวน 192 ข้อ จำแนกช่วงค่าอำนาจจำแนกของข้อสอบ และค่าความยากของข้อสอบ ดังนี้

1) ค่าอำนาจจำแนกของข้อสอบ

ระหว่าง 0.50 ถึง 1.00 จำนวน 108 ข้อ

ระหว่าง 1.01 ถึง 1.50 จำนวน 48 ข้อ

ระหว่าง 1.51 ถึง 2.00 จำนวน 26 ข้อ

ระหว่าง 2.01 ถึง 2.50 จำนวน 10 ข้อ

จากข้อสอบ 192 ข้อ ข้อสอบส่วนใหญ่มีค่าอำนาจจำแนกค่อนข้างต่ำ ระหว่าง 0.50 ถึง 1.00 จำนวน 108 ข้อ

2) ค่าความยากของข้อสอบ

ระหว่าง -2.50 ถึง -1.50 จำนวน 14 ข้อ

ระหว่าง -1.51 ถึง -0.50 จำนวน 30 ข้อ

ระหว่าง -0.51 ถึง 0.50 จำนวน 48 ข้อ

ระหว่าง 0.51 ถึง 1.50 จำนวน 59 ข้อ

ระหว่าง 1.51 ถึง 2.50 จำนวน 41 ข้อ

จากข้อสอบ 192 ข้อ ข้อสอบส่วนใหญ่มีค่าความยากปานกลางจนถึงค่อนข้างยาก ระหว่าง 0.51 ถึง 1.50 จำนวน 59 ข้อ

4. การจัดทำคลังข้อสอบ เป็นการนำข้อสอบที่มีคุณภาพ จำนวน 192 ข้อ บันทึกลงในคลังข้อสอบที่พัฒนาขึ้น ประกอบด้วย ข้อสอบ ตัวเลือก คำตอบ หมวดเนื้อหา ความรู้ทางไอซีที ค่าอำนาจจำแนกของข้อสอบ (a) ค่าความยากของข้อสอบ (b) ค่าโอกาสการเดาของข้อสอบ (c)

ขั้นตอนที่ 2 การศึกษาหลักการวิธีโหลดบาลานซ์ซิ่ง วิธีการควบคุมการใช้ข้อสอบ และวิธีการควบคุมสมดุลเนื้อหา ซึ่งเป็นขั้นตอนการศึกษาเอกสารและวรรณกรรมที่เกี่ยวข้อง ให้เข้าใจถึงหลักการทำงานของวิธีต่าง ๆ สามารถนำไปพัฒนาและออกแบบวิธีการทำงานได้อย่างถูกต้อง โดยมีหลักการของแต่ละวิธีดังนี้

1. วิธีโหลดบาลานซ์ซิ่ง (LB) เป็นวิธีการทางเครือข่ายคอมพิวเตอร์ (Network) แบ่งหน้าที่จากจำนวนรวมของงานที่คอมพิวเตอร์ต้องทำ ระหว่างคอมพิวเตอร์สองเครื่องหรือมากกว่า เป็นการช่วยแบ่งภาระการทำงานที่เข้ามาจากเครื่องคอมพิวเตอร์ของผู้ใช้ ให้กระจายไปในกลุ่มของเครื่องแม่ข่าย ทำให้ภาระงานไม่เน้นหนักไปที่เครื่องใดเครื่องหนึ่ง และมีภาระงานเท่าเทียมกัน Singh et al. (2015) ได้นำเสนอกระบวนการทำงาน 3 ขั้นตอน ได้แก่ 1) การสำรวจและเคลื่อนย้ายข้อมูล 2) การคำนวณภาระงาน และ 3) การควบคุมการทำงาน ผู้วิจัยนำ 3 ขั้นตอนมาประยุกต์ใช้กับวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาและปรับปรุงอัลกอริทึมวิธีการเดิมในขั้นตอนที่ 2 และขั้นตอนที่ 3 เพื่อประสิทธิภาพการทำงานและแก้ไขปัญหาที่ดีขึ้น ดังนี้

1.1 การสำรวจและเคลื่อนย้ายข้อมูล เป็นขั้นตอนแรกที่อัลกอริทึมสำรวจข้อสอบที่มีอยู่ในคลังข้อสอบทั้งหมด เพื่อนำข้อมูลลักษณะของข้อสอบทั้งหมดมาเตรียมความพร้อมสำหรับการคำนวณ แสดงดังภาพที่ 3-4

```
Migration_Agent ( )
Input: VM_configurations fromChannel_Agent(VM_initial)
Output: Search similar VM from other Datacenters
{
Accept VM_configurations fromchannel_Agent(VM_initial);
Search a Data_center;
Check VM_Load_Table;
If (found)
Return (A);
Else
Return (NA);
On_receiving (self_destroy);
Kill (MA_i);
}
```

ภาพที่ 3-4 อัลกอริทึมการสำรวจและเคลื่อนย้ายข้อมูล

1.2) การคำนวณภาระงาน เป็นวิธีการควบคุมสมดุลเนื้อหาของข้อสอบ ให้ครอบคลุมเนื้อหาทุกด้านเมื่อมีการทดสอบของผู้ทดสอบ โดยการคำนวณจากเปอร์เซ็นต์การใช้ข้อสอบของแต่ละด้าน แล้วนำเปอร์เซ็นต์ที่ได้เปรียบเทียบกับค่าคงที่ นั้นคือ 25 ดังภาพที่ 3-5 ผู้วิจัยนำอัลกอริทึมดังกล่าวมาเขียนโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ ในส่วนของการควบคุมสมดุลเนื้อหา แล้วทดสอบประสิทธิภาพของการควบคุมสมดุลเนื้อหา จำนวน 20 ครั้ง ซึ่งแบ่งเป็น 2 แบบ ได้แก่ 1) แบบตั้งใจทำข้อสอบ จำนวน 10 ครั้ง และ 2) แบบไม่ตั้งใจทำข้อสอบ จำนวน 10 ครั้ง ปรากฏว่า เมื่อมีการยุติการทดสอบ แบบที่ 1 ตั้งใจทำข้อสอบ มีจำนวนข้อสอบที่ใช้ในการทดสอบประมาณ 10 ถึง 15 ข้อ และแบบที่ 2 ไม่ตั้งใจทำข้อสอบ มีจำนวนข้อสอบที่ใช้ในการทดสอบประมาณ 8 ถึง 15 ข้อ ซึ่งทั้ง 2 แบบ มีเนื้อหาบางด้านไม่ถูกนำไปใช้ทดสอบและเน้นหนักไปด้านใดด้านหนึ่ง ทั้งนี้เกิดจากเกณฑ์ยุติเข้าเงื่อนไขของค่าความคลาดเคลื่อนมาตรฐาน (SE) ต่ำกว่า 0.30 ตามทฤษฎี จากปัญหาที่เกิดขึ้น ผู้วิจัยพิจารณาข้อสอบของแต่ละด้าน มีจำนวนข้อสอบไม่เท่ากัน ดังนั้นการนำเงื่อนไขค่าคงที่ 25 จากอัลกอริทึมมาใช้ ทำให้เกิดข้อจำกัดดังกล่าว ผู้วิจัยจึงศึกษางานวิจัยที่เกี่ยวข้องเพื่อแก้ไขปัญหาค่าคงที่ ปรากฏว่า Bortot and Marques Pereira (2017) นำเสนอวิธีการใช้ค่าเฉลี่ยถ่วงน้ำหนัก (Weighted mean) โดยการนำพีชคณิตเป็นสัญลักษณ์พิสูจน์สูตรการแทนค่าเฉลี่ยถ่วงน้ำหนักที่นำมาแก้สมการกับข้อมูลที่มีความสำคัญไม่เท่ากัน ซึ่งสอดคล้องกับการแก้ไขปัญหาจำนวนข้อสอบของแต่ละด้านไม่เท่ากันได้ ผู้วิจัยจึงนำค่าเฉลี่ยถ่วงน้ำหนักมาแทรกอัลกอริทึมเดิมเพื่อแทนค่าคงที่ 25 แล้วทดสอบประสิทธิภาพจำนวน 20 ครั้ง ผลการทดสอบ ปรากฏว่า เมื่อยุติการทดสอบ จำนวนข้อสอบมีความครอบคลุมเนื้อหาทุกด้านและเนื้อหาไม่เน้นหนักไปด้านใดด้านหนึ่ง ดังนั้น การใช้ค่าเฉลี่ยถ่วงน้ำหนัก สามารถนำมาแก้ปัญหาดังกล่าวได้ และจากการศึกษางานวิจัยที่เกี่ยวข้อง ยังไม่พบการนำค่าเฉลี่ยถ่วงน้ำหนักมาใช้ควบคุมสมดุลเนื้อหา

1) ข้อสอบที่ถูกใช้แล้วไม่นำกลับไปใช้ทดสอบซ้ำของผู้ทดสอบแต่ละคน 2) ไม่ให้มีอัตราการแสดงข้อสอบมากเกินไปจนทำให้เกิดการคาดเดาข้อสอบในกลุ่มผู้ทดสอบ และ 3) การควบคุมข้อสอบที่ถูกนำไปใช้ทดสอบน้อยหรือไม่เคยถูกนำไปใช้ทดสอบให้มีโอกาสนำไปใช้ทดสอบ จากภาพที่ 3-6 ผู้วิจัยนำอัลกอริทึมดังกล่าวมาเขียนโปรแกรม ต่อจากขั้นตอนการคำนวณภาระงาน จากการทดสอบปรากฏว่า การควบคุมการใช้ข้อสอบมีประสิทธิภาพในด้านข้อสอบที่ถูกใช้แล้วไม่นำกลับไปใช้ทดสอบซ้ำของผู้ทดสอบแต่ละคน แต่ประสิทธิภาพอัตราการแสดงข้อสอบมากเกินไปจนทำให้เกิดการคาดเดาข้อสอบในกลุ่มผู้ทดสอบ ยังแสดงจำนวนข้อสอบที่นำไปแสดงมากกว่า 0.2 ตามทฤษฎี และการควบคุมข้อสอบที่ถูกนำไปใช้ทดสอบน้อยหรือไม่เคยถูกนำไปใช้ทดสอบให้มีโอกาสนำไปใช้ทดสอบ ยังมีข้อสอบอีกหลายข้อที่มีโอกาสนำไปใช้ทดสอบแต่ไม่ได้ถูกเลือก ผู้วิจัยจึงพิจารณาข้อสอบที่ถูกนำไปใช้มากเกินอัตราที่กำหนดตามทฤษฎี และข้อสอบที่ไม่ถูกเลือกไปใช้ทดสอบ ปรากฏว่าอัลกอริทึมมีการทำงานแบบสุ่มตัวแทนไปใช้ทดสอบ โดยไม่ได้คำนึงถึงอัตราการใช้ข้อสอบ จากปัญหาดังกล่าว ผู้วิจัยได้ศึกษาวิธีที่มาช่วยแก้ไขปัญหาที่เกิดขึ้น ปรากฏว่า งานวิจัยของ Huang and Jo (1995) ได้เสนองานวิจัย เรื่อง “Data frequency and the number” เป็นการหาความถี่ของข้อมูลเพื่อเลือกหุ้น ถ้าข้อมูลมีความถี่มากแสดงว่าหุ้นตัวนั้นน่าลงทุน แต่ถ้าความถี่น้อยแสดงว่าหุ้นตัวนั้นไม่น่าลงทุน จากงานวิจัยดังกล่าว สอดคล้องกับการแก้ไขปัญหาที่ผู้วิจัยพบ นั่นคือ ถ้าความถี่การใช้ข้อสอบมากแสดงว่าข้อสอบนั้นถูกนำไปใช้ทดสอบมาก ถ้าความถี่การใช้ข้อสอบน้อยแสดงว่าข้อสอบนั้นถูกนำไปใช้ทดสอบน้อย ผู้วิจัยจึงแทรกอัลกอริทึมการเรียงความถี่ของการใช้ข้อสอบจากน้อยไปหามากและเรียงผลต่างที่ได้จากค่าประมาณความสามารถกับค่าสารสนเทศสูงสุดที่เข้าใกล้ 0 จากน้อยไปหามากเช่นกัน ดังภาพที่ 3-6

จากการทดสอบ ปรากฏว่า การนำความถี่ของการใช้ข้อมูลมาใช้ในการควบคุมการใช้ข้อสอบ มีประสิทธิภาพทั้ง 3 ด้าน ได้แก่ 1) ข้อสอบที่ถูกใช้แล้วไม่นำกลับไปใช้ทดสอบซ้ำของผู้ทดสอบแต่ละคน 2) ไม่ให้มีอัตราการแสดงข้อสอบมากเกินไปจนทำให้เกิดการคาดเดาข้อสอบในกลุ่มผู้ทดสอบ และ 3) การควบคุมข้อสอบที่ถูกนำไปใช้ทดสอบน้อยหรือไม่เคยถูกนำไปใช้ทดสอบให้มีโอกาสนำไปใช้ทดสอบ

3. วิธีการควบคุมสมดุลเนื้อหา จากการศึกษางานวิจัย ปรากฏว่า มีวิธีการควบคุมสมดุลเนื้อหาที่นิยมใช้ จำนวน 3 วิธี ได้แก่ 1) การควบคุมสมดุลเนื้อหา (Constrained) 2) การควบคุมสมดุลเนื้อหาประยุกต์ (Modified constrained) และ 3) รูปแบบมัลติโนเมียลประยุกต์ (Modified Multinomial Model: MMM) ดังนั้น วิธีโหลดบาลานซ์ซิ่ง จึงเป็นวิธีใหม่ที่มีคุณสมบัติของการควบคุมสมดุลเนื้อหา โดยใช้ค่าเฉลี่ยถ่วงน้ำหนักของแต่ละหมวดเนื้อหาที่คำนวณจากการใช้ข้อสอบของผู้ทดสอบ ณ ขณะนั้น

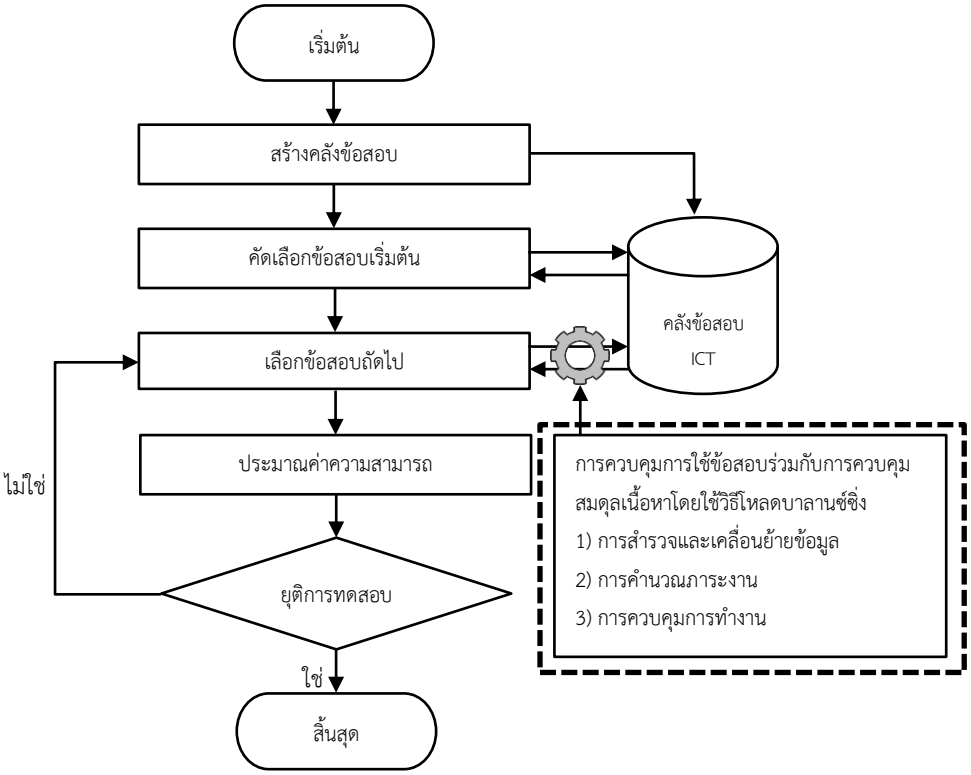
ดังนั้น จากการศึกษาวิธีโหลดบาลานซ์ซิ่ง ในการควบคุมการใช้ข้อสอบและการควบคุมสมดุลเนื้อหาได้เปรียบเทียบประสิทธิภาพการทำงานของวิธี LB กับ วิธี IE ดังตารางที่ 3-3

ตารางที่ 3-3 การเปรียบเทียบวิธี LB กับ วิธี IE

วิธีการ	LB	IE
ควบคุมการใช้ข้อสอบ	1. เรียงลำดับจากน้อยไปหามาก ระหว่างความถี่สะสมของการใช้ข้อสอบกับผลต่างของค่าประมาณความสามารถของผู้ทดสอบ (ML) กับค่าสารสนเทศสูงสุด MIC 2. คัดเลือกข้อสอบลำดับแรกเป็นข้อสอบข้อถัดไป	1. กำหนดให้มีค่าอัตราการใช้ข้อสอบสูงสุดแบบคงที่ เช่น $r^{\max} = 0.2$ 2. ถ้าข้อสอบมีอัตราการใช้ข้อสอบมากกว่า 0.2 ให้ข้อสอบมีสถานะ 0 แต่ถ้าข้อสอบมีอัตราการใช้ข้อสอบต่ำกว่า 0.2 ให้ข้อสอบมีค่าสถานะ 1 แต่ถ้าข้อสอบในคลังมีค่าสถานะ 0 จนไม่สามารถนำมาใช้ทดสอบได้ วิธีการนี้จึงเปลี่ยนสถานะข้อสอบทุกข้อให้เป็น 1 ทั้งหมด 3. นำข้อสอบที่มีสถานะเป็น 1 มาเข้าวิธีการ SH 4. คำนวณส่วนต่างระหว่างค่า ML กับ ค่า MIC แล้วเลือกข้อสอบที่มีค่าเข้าใกล้ 0 มากที่สุดมาเป็นข้อสอบข้อถัดไป
ควบคุมสมดุลเนื้อหา	1. ใช้ค่าเฉลี่ยถ่วงน้ำหนักของแต่ละด้านเนื้อหา ถ้าค่าเฉลี่ยหมวดเนื้อหาใดมากที่สุดหมายถึงจำนวนข้อสอบนั้นถูกนำไปใช้น้อย ให้เป็นตัวแทนหมวดเนื้อหา	ไม่มีการควบคุมสมดุลเนื้อหา

ขั้นตอนที่ 3 การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง นำผลการวิเคราะห์ที่ได้จากการศึกษาในขั้นตอนที่ 2 มาพัฒนาเป็นวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา เป็นอีกหนึ่งทางเลือกที่ช่วยแก้ไขปัญหาดังกล่าว

หลักการทำงานวิธีนี้เข้าไปแทรกระหว่างการเลือกข้อสอบข้อถัดกับคลับข้อสอบ ซึ่งเลือกข้อสอบข้อถัดไปใช้วิธีการหาค่าสารสนเทศสูงสุดใกล้เคียงกับค่าประมาณความสามารถของผู้ทดสอบ ส่วนโหลดบาลานซ์ซิ่ง ทำงานหลังจากเลือกข้อสอบข้อแรก ซึ่งมี 3 ขั้นตอน ได้แก่ 1) การสำรวจและเคลื่อนย้ายข้อมูล 2) การคำนวณภาระงาน และ 3) การควบคุมการทำงาน ดังภาพที่ 3-7



ภาพที่ 3-7 วิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง

จากภาพขั้นตอนวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่งมีรายละเอียดขั้นตอนของอัลกอริทึม ดังนี้

ขั้นตอนที่ 1 การสำรวจและเคลื่อนย้ายข้อมูล เป็นการสำรวจคุณลักษณะของข้อสอบทุกข้อในคลังข้อสอบ ซึ่งมีการทำงาน ดังนี้

1. การค้นหาข้อสอบ เป็นการค้นหาข้อสอบทั้งหมดในคลัง เพื่อหาข้อมูลคุณลักษณะของข้อสอบ

2. การเก็บคุณลักษณะข้อสอบ เป็นการเก็บข้อมูลลักษณะของข้อสอบที่ได้จากการค้นหาข้อสอบ ลงใน Local storage ของอัลกอริทึมโหลดบาลานซ์ซิ่ง มีหมวดเนื้อหาจำนวน 5 ด้าน แสดงตัวอย่างดังตารางที่ 3-4

ตารางที่ 3-4 ตัวอย่างการเก็บลักษณะข้อสอบ

ข้อที่	หมวดเนื้อหา	a	b	c	สถานะ	ความถี่สะสม
1	1	0.55	-0.66	0.25	1	5
2	3	0.59	0.83	0.09	0	4
3	5	1.99	-1.29	0.27	0	0
4	1	1.76	0.18	0.22	0	1
5	2	0.67	0.52	0.11	1	1
6	4	0.61	1.44	0.17	1	0
7	2	0.97	0.79	0.19	1	2
8	4	1.56	-0.67	0.23	1	1
9	3	0.76	0.45	0.14	1	4
10	1	0.82	-0.24	0.05	0	2

หมายเหตุ: หมวดเนื้อหา หมายถึง หมวดเนื้อหาวิชาเทคโนโลยีสารสนเทศมี 5 หมวด
a หมายถึง ค่าอำนาจจำแนกของข้อสอบ
b หมายถึง ค่าความยากของข้อสอบ
c หมายถึง ค่าโอกาสการเดาของข้อสอบ
สถานะ หมายถึง สถานะใช้สอบ นำไปใช้ (1) ยังไม่นำไปใช้ (0)
ความถี่สะสม หมายถึง ความถี่สะสมของข้อสอบในแต่ละข้อได้จากผู้ทดสอบทั้งหมด

ขั้นตอนที่ 2 การคำนวณภาระงาน เป็นวิธีการควบคุมสมดุลเนื้อหาของข้อสอบ มีเป้าหมายเพื่อหาด้านเนื้อหาที่เป็นตัวแทน โดยการคำนวณค่าเฉลี่ยถ่วงน้ำหนักของสถานะการใช้ข้อสอบ ณ ขณะนั้นของผู้ทดสอบ ถ้าสถานะเป็น 1 หมายถึง ข้อสอบข้อนั้นได้ถูกนำไปใช้ทดสอบแล้ว แต่ถ้าสถานะเป็น 0 หมายถึง ข้อสอบข้อนั้นยังไม่ถูกนำไปใช้ทดสอบ ซึ่งขั้นตอนการคำนวณค่าเฉลี่ยถ่วงน้ำหนักสำหรับควบคุมความสมดุลเนื้อหา ดังนี้

1. การคำนวณค่าเฉลี่ยถ่วงน้ำหนัก กำหนดให้เนื้อหามีจำนวน 5 ด้าน ข้อสอบมีจำนวน 192 ข้อ ซึ่งด้านที่ 1 มีจำนวนข้อสอบ 34 ข้อ นำไปใช้สอบ 1 ข้อ ด้านที่ 2 มีจำนวนข้อสอบ 47 ข้อ นำไปใช้สอบ 2 ข้อ ด้านที่ 3 มีจำนวนข้อสอบ 48 ข้อ นำไปใช้สอบ 1 ข้อ ด้านที่ 4 มีจำนวนข้อสอบ 35 ข้อ นำไปใช้สอบ 2 ข้อ และด้านที่ 5 มีจำนวนข้อสอบ 28 ข้อ ยังไม่ถูกนำไปใช้สอบ

1.1 คำนวณจำนวนข้อสอบที่ไม่ถูกนำไปใช้ด้านที่ 1 คำนวณจากสมการ 34

$$Item_{avaliable} = Item_{total} - Item_{used} \quad (34)$$

เมื่อ $Item_{avaliable}$ แทน จำนวนข้อสอบที่ไม่ถูกนำไปใช้

$Item_{total}$ แทน ข้อสอบในด้านที่ 1 ทั้งหมด

$Item_{used}$ แทน ข้อสอบที่ถูกนำไปใช้แล้ว

นั้นคือ 33 = 34 – 1 ดังนั้นจำนวนข้อสอบที่ไม่ถูกนำไปใช้ในด้านที่ 1 คือ 33

1.2 คำนวณเปอร์เซ็นต์ข้อสอบที่ไม่ถูกนำไปใช้ด้านที่ 1 คำนวณจากสมการ 35

$$v_i = \frac{Item_{avaliable}}{Item_{tatal}} \times 100 \quad (35)$$

เมื่อ v_i แทนเปอร์เซ็นต์ของด้านที่ 1 นั่นคือ $\frac{33}{34} \times 100 = 97.06\%$ ดังนั้น เปอร์เซ็นต์ด้านที่ 1 คือ 97.06% คำนวณเปอร์เซ็นต์ขั้นที่ 1.1 และ 1.2 จนครบ 5 ด้าน

1.3 คำนวณค่าเฉลี่ยถ่วงน้ำหนักจากสมการ 36

$$\overline{\mathrm{X}} = \frac{\sum_{i=1}^{n} x_i v_i}{\sum_{i=1}^{n} x_i} \quad (36)$$

เมื่อ $\overline{\mathrm{X}}$ แทน ค่าเฉลี่ยถ่วงน้ำหนัก

n แทน จำนวนด้านของหมวดไอซีที

v_i แทน เปอร์เซ็นต์ข้อสอบที่ไม่ถูกนำไปใช้แต่ละด้าน

x_i แทน จำนวนข้อสอบทั้งหมดของแต่ละด้าน

จากสมการ 36 แทนค่าข้อมูลได้ดังนี้

$$\frac{(34)(97.06)+(47)(95.74)+(48)(95.83)+(35)(97.14)+(28)(0)}{34+47+48+35+28} = 96.88$$

ดังนั้น ค่าเฉลี่ยถ่วงน้ำหนักทั้ง 5 ด้านคือ 96.88

จากตารางที่ 3-5 เป็นการเก็บค่าเฉลี่ยถ่วงน้ำหนักที่ได้จากการคำนวณ ซึ่งสถานะ 0 มีจำนวน 3 ด้าน ได้แก่ 1) ด้านที่ 5 ไม่มีการนำข้อสอบไปใช้ คิดเป็น 100 เปอร์เซ็นต์ 2) ด้านที่ 4 มีการนำข้อสอบไปใช้ คิดเป็น 97.14 เปอร์เซ็นต์ และ 3) ด้านที่ 1 มีการนำข้อสอบไปใช้คิดเป็น 97.06 เปอร์เซ็นต์ ตามลำดับ ซึ่งทั้ง 3 ด้าน เมื่อนำมาเปรียบเทียบกับค่าเฉลี่ยถ่วงน้ำหนัก 96.88 เปอร์เซ็นต์ ปรากฏว่า อัลกอริทึมเลือกข้อสอบด้านที่มีเปอร์เซ็นต์ที่ไม่ถูกนำไปใช้สูงสุดเพื่อเป็นตัวแทน นั้นคือด้านที่ 5 เป็นตัวแทนด้านเนื้อหา

ขั้นตอนที่ 3 การควบคุมการทำงาน เป็นวิธีควบคุมการใช้ข้อสอบโดยนำด้านที่เป็นตัวแทนจากขั้นตอนที่ 2 มาคำนวณเพื่อควบคุมการใช้ข้อสอบ แบ่งการทำงานดังนี้

1. การหาจำนวนข้อสอบที่ยังไม่ถูกนำไปใช้ทดสอบในด้านที่ 5 คำนวณได้จากสมการ 38

$$Item_{avaliable(v)(i)} = Item_{total(v)(i)} - Item_{used(v)(i)} \tag{38}$$

เมื่อ $Item_{avaliable(v)(i)}$ คือข้อสอบที่ยังไม่ใช้ในด้านที่ 5 $Item_{total(v)(i)}$ คือข้อสอบทั้งหมดในด้านที่ 5 และ $Item_{used(v)(i)}$ คือข้อสอบที่ถูกนำไปใช้ในด้านที่ 5 นั้นคือ 28=28-0 ดังนั้น ในด้านที่ 5 มีจำนวนข้อสอบที่สามารถนำไปใช้ได้จำนวน 28 ข้อ

2. การหาค่าสารสนเทศสูงสุดในข้อสอบด้านที่ 5 ดังสมการ 39

$$MItem_{(v)(i)} = b_i + \frac{1}{Da_i} \ln\left[\frac{1+\sqrt{1+8c_i}}{2}\right] \tag{39}$$

เมื่อ $MItem_{(v)(i)}$ แทน ค่าสารสนเทศสูงสุดของข้อสอบข้อที่ i

D แทน ค่าคงที่มีค่าเท่ากับ 1.7

a_i แทน ค่าอำนาจจำแนกของข้อสอบข้อที่ i

b_i แทน ค่าความยากของข้อสอบข้อที่ i

c_i แทน ค่าการเดาของข้อสอบข้อที่ i

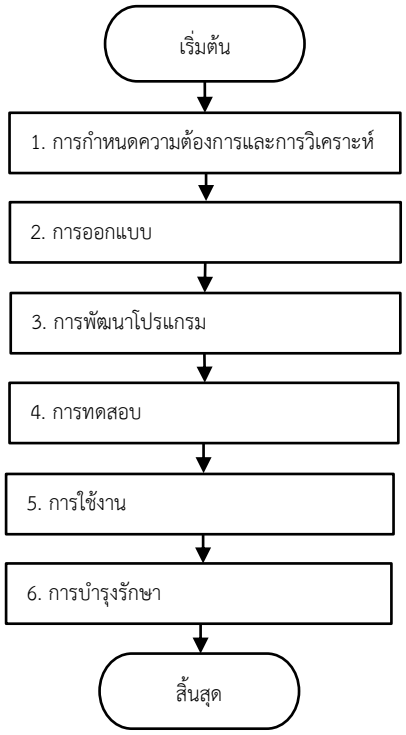
เมื่อหาค่าสารสนเทศสูงสุดในข้อสอบด้านที่ 5 แล้วเก็บข้อมูลลงใน Local storage ของเครื่องลูกข่ายดังตัวอย่างตารางที่ 3-6

3. การคำนวณผลต่างระหว่างค่าประมาณความสามารถของผู้ทดสอบกับค่าสารสนเทศสูงสุด คำนวนได้จากสมการ 40

จากตารางที่ 3-6 แสดงตัวอย่างการคัดเลือกข้อสอบข้อถัดไป โดยการเปรียบเทียบระหว่างค่าสารสนเทศสูงสุดกับค่าประมาณความสามารถของผู้ทดสอบ ณ ขณะนั้น กำหนดให้เป็น 1.01 เมื่อพิจารณาจากตารางช่องค่าสารสนเทศสูงสุดแทนด้วย $MItem_{(v)(i)}$ ข้อ 54 มีค่าสารสนเทศใกล้เคียงกับค่าประมาณความสามารถ คือ 1.05 ดังนั้น อัลกอริทึมเรียงลำดับจากน้อยไปหามากระหว่างส่วนต่าง $TestItem_{(v)(i)}$ กับความถี่สะสมของการใช้ข้อสอบ สรุปว่า ตัวแทนข้อสอบข้อถัดไปจึงเป็นข้อ 54 เนื่องจากมีส่วนต่างและจำนวนการใช้ข้อสอบความถี่สะสมน้อยที่สุด

ระยะที่ 2 การพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

การพัฒนาโปรแกรมฯ สำหรับวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา 2 วิธี ได้แก่ 1) วิธีการโหลดบาลานซ์ซิ่ง (LB) เป็นวิธีใหม่ที่ผู้วิจัยพัฒนาขึ้น 2) วิธี Item Eligibility (IE) เป็นวิธีเก่า การวิจัยนี้อธิบายถึงการพัฒนาโปรแกรมสำหรับวิธีใหม่ โดยใช้ทฤษฎีวงจรการพัฒนาโปรแกรม (SDLC) (Singh et al., 2015) ประกอบด้วย 6 ขั้นตอน ได้แก่ 1) การกำหนดความต้องการและการวิเคราะห์ 2) การออกแบบ 3) การพัฒนาโปรแกรม 4) การทดสอบ 5) การใช้งาน และ 6) การบำรุงรักษา แสดงดังภาพที่ 3-8



ภาพที่ 3-8 วงจรการพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

ขั้นตอนที่ 1 การกำหนดความต้องการและการวิเคราะห์ (Requirement gathering and analysis) ในการพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยใช้วิธีโหลดบาลานซ์ซิ่งในการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา สำหรับทดสอบความรู้ทางไอซีที ผู้วิจัยได้ศึกษาความเป็นไปได้ต่อการพัฒนาโปรแกรมฯ เพื่อทำความเข้าใจส่วนร่วมขององค์ประกอบต่าง ๆ และรายละเอียดของขั้นตอนดำเนินงาน และได้ศึกษาขีดความสามารถขององค์ประกอบของเทคโนโลยีสารสนเทศที่มีผลต่อการพิจารณาพัฒนาโปรแกรม ซึ่งมี 6 องค์ประกอบหลัก ได้แก่ 1) ฮาร์ดแวร์ (Hardware) 2) ซอฟแวร์ (Software) 3) ข้อมูล (Data) 4) เครือข่าย (Network) 5) กระบวนการ (Procedure) และ 6) ผู้ใช้ (People) (Gary & Harry, 2012, p. 7) การกำหนดความต้องการของโปรแกรมและวิเคราะห์โปรแกรม มีดังนี้

1. การกำหนดความต้องการโปรแกรม (Program Technology Requirement)

1.1 ฮาร์ดแวร์ (Hardware) คุณสมบัติเครื่องคอมพิวเตอร์ที่ใช้ในการทดสอบ

1) เป็นเครื่องคอมพิวเตอร์ส่วนบุคคล หรือแบบพกพา (Laptop)

2) มีหน่วยประมวลผลกลาง (CPU) มีความเร็วสัญญาณนาฬิกา 2.4 GHz ขึ้นไป

3) มีหน่วยความจำหลัก (RAM) ชนิด DDR3 หรือดีกว่า มีขนาดไม่น้อยกว่า 4 GB

4) มีหน่วยจัดเก็บข้อมูล (Hard disk) ขนาดความจุไม่น้อยกว่า 1 TB

1.2 ซอฟแวร์ (Software) คุณสมบัติของซอฟแวร์ที่ใช้ในการพัฒนา

1) ภาษาที่ใช้ในการพัฒนาโปรแกรม PHP, AJAX, HTML5, Java Script

2) เครื่องมือในการพัฒนาโปรแกรม Edit Plus, eclipse 2014, Dreamweaver

3) โปรแกรมจัดการฐานข้อมูล MySQL

1.3 ข้อมูล (Data) ขอบเขตของการจัดเก็บข้อมูล

1) ข้อมูลผู้ทดสอบ 2) ข้อมูลคลังข้อสอบ และ 3) ข้อมูลการทดสอบ

1.4 เครือข่าย (Network) การทำงานของโปรแกรมที่ต้องอาศัยเครือข่ายอินเทอร์เน็ต ทำแบบทดสอบผ่านทางอินเทอร์เน็ต (Internet)

1.5 กระบวนการ (Procedure) วิธีการหรือขั้นตอนการเข้าใช้โปรแกรม

1) ผู้ที่ใช้โปรแกรมต้องลงทะเบียนเพื่อเก็บเป็นประวัติ

2) ระหว่างใช้โปรแกรมต้องปฏิบัติตามคู่มือการใช้งาน

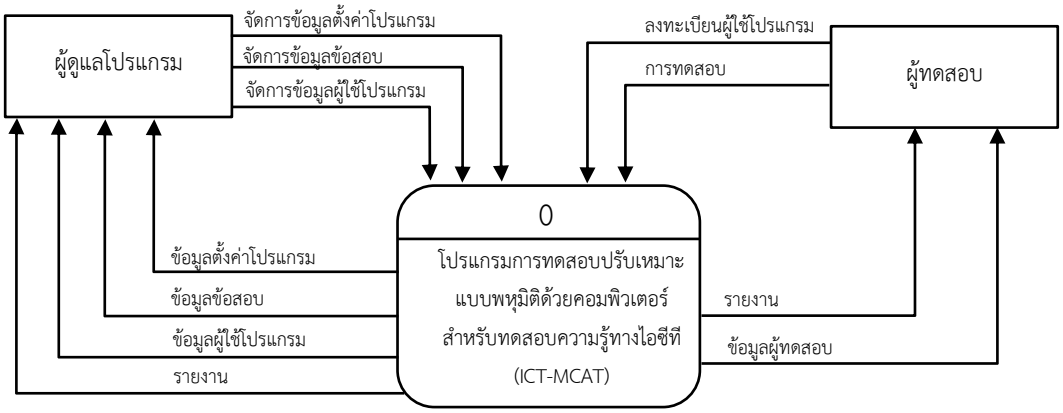
1.6 ผู้ใช้ (People) กำหนดกลุ่มผู้ใช้โปรแกรม

1) ผู้ดูแลโปรแกรม

2) ผู้ทดสอบ

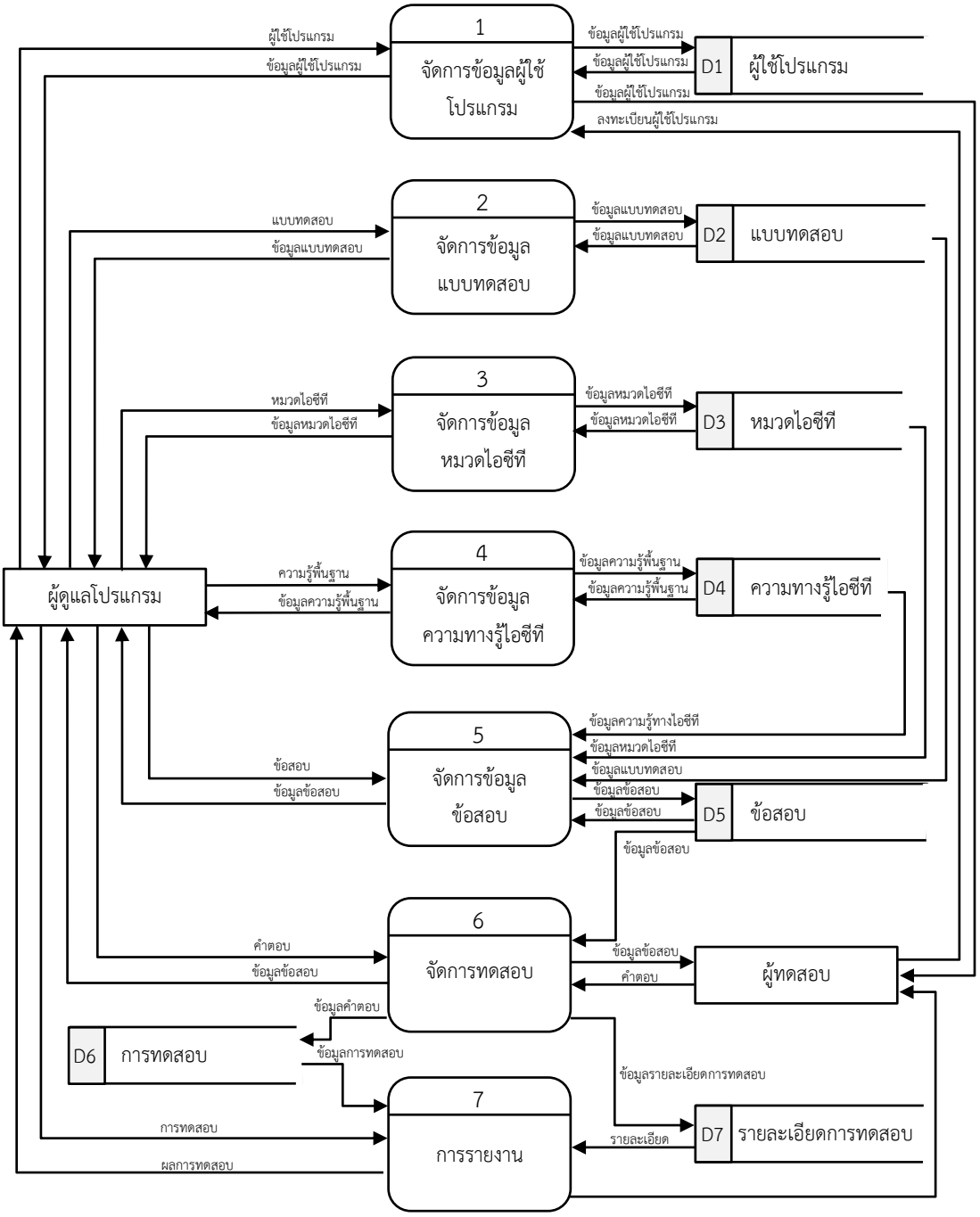
2. การวิเคราะห์กระบวนการทำงานโดยการออกแบบแผนภาพกระแสข้อมูล (Data Flow Diagrams: DFD) แบบจำลองฐานข้อมูล (Data model) และพจนานุกรมข้อมูล (Data dictionary) เพื่อใช้ในการออกแบบและพัฒนาโปรแกรม ดังนี้

2.1 การออกแบบแผนภาพกระแสข้อมูล (Data Flow Diagrams: DFD) แผนภาพบริบท (Context Diagram) แสดงภาพรวมของโปรแกรมทั้งหมด และ DFD Level 0 คือ แผนภาพกระแสข้อมูลในระดับที่แสดงขั้นตอนการทำงานหลักทั้งหมด (Process) ของโปรแกรมในการแสดงทิศทางการไหลของของข้อมูล (Data Flow) และแสดงรายละเอียดของแหล่งจัดเก็บข้อมูล (Data Store) (Gary & Harry, 2012, p. 200; Alan et al., 2013, p. 185)



ภาพที่ 3-9 แผนภาพบริบทการพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

จากภาพที่ 3-9 แสดงภาพรวมของโปรแกรม มีแหล่งข้อมูล 2 แหล่ง ได้แก่ 1) ผู้ดูแลโปรแกรม มีหน้าที่ในการจัดการข้อมูลตั้งค่าโปรแกรม จัดการข้อมูลข้อสอบ จัดการข้อมูลผู้ใช้โปรแกรม รวมทั้งรายงานข้อมูลของผู้ดูแลโปรแกรม และ 2) ผู้ทดสอบ มีหน้าที่ในการลงทะเบียนโปรแกรม การทดสอบความรู้ทางไอซีที และดูรายงานผลการทดสอบ



ภาพที่ 3-10 แผนภาพกระแสข้อมูลการพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

จากภาพที่ 3-10 แสดงกระแสการรับส่งข้อมูลของแต่ละกระบวนการ ประกอบไปด้วย 2 แหล่งข้อมูล 7 กระบวนการ และ 7 แหล่งจัดเก็บข้อมูล โดยอธิบายตามกรอบของกระบวนการดังนี้

กระบวนการที่ 1 การจัดการข้อมูลผู้ใช้โปรแกรม ผู้ดูแลโปรแกรมดำเนินการจัดการข้อมูลผู้ใช้โปรแกรมให้สามารถมีสิทธิ์ในการใช้โปรแกรม เช่น จัดการข้อมูลชื่อ-นามสกุล รหัสบัตรประชาชน เป็นต้น เมื่อดำเนินการจัดการข้อมูลเสร็จ โปรแกรมบันทึกข้อมูลเข้าสู่แหล่งจัดเก็บข้อมูลผู้ใช้โปรแกรม (D1)

กระบวนการที่ 2 การจัดการข้อมูลแบบทดสอบ ผู้ดูแลโปรแกรมดำเนินการจัดการข้อมูลแบบทดสอบเบื้องต้นก่อนการทดสอบ เช่น จำนวนข้อสอบสูงสุด (Item length) และ ค่าความคลาดเคลื่อน (SEE) เมื่อดำเนินการจัดการข้อมูลเสร็จ โปรแกรมบันทึกข้อมูลเข้าสู่แหล่งจัดเก็บข้อมูลแบบทดสอบ (D2)

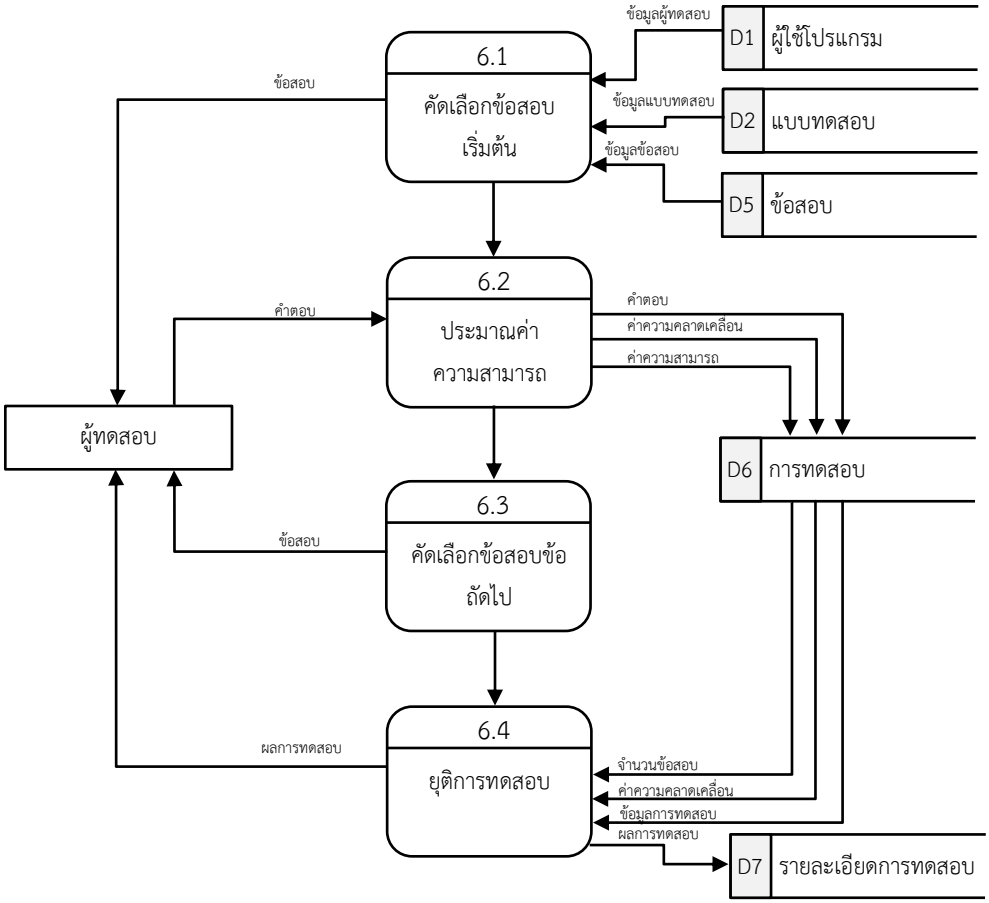
กระบวนการที่ 3 การจัดการข้อมูลหมวดไอซีที ผู้ดูแลโปรแกรมดำเนินการจัดการข้อมูลหมวดไอซีที ประกอบด้วย 5 หมวด ได้แก่ 1) เครือข่าย 2) ข้อมูล 3) ซอฟต์แวร์ 4) ฮาร์ดแวร์ และ 5) กระบวนการและผู้ใช้ เมื่อดำเนินการจัดการข้อมูลเสร็จ โปรแกรมบันทึกข้อมูลเข้าสู่แหล่งจัดเก็บข้อมูลหมวดไอซีที (D3)

กระบวนการที่ 4 การจัดการข้อมูลความทางรู้ไอซีที ผู้ดูแลโปรแกรมดำเนินการจัดการข้อมูลความรู้ทางไอซีที ซึ่งประกอบด้วย 6 ด้าน ได้แก่ 1) การเข้าถึงสารสนเทศ 2) การจัดการสารสนเทศ 3) การประเมิน 4) การสร้างสิ่งใหม่จากความเข้าใจ 5) การสื่อสาร และ 6) การใช้ไอซีทีอย่างเหมาะสม เมื่อดำเนินการจัดการข้อมูลเสร็จ โปรแกรมบันทึกข้อมูลเข้าสู่แหล่งจัดเก็บข้อมูลความรู้ทางไอซีที (D4)

กระบวนการที่ 5 การจัดการข้อมูลข้อสอบ ผู้ดูแลโปรแกรมดำเนินการจัดการข้อมูลข้อสอบที่มีคุณภาพตามเกณฑ์การตรวจสอบคุณภาพทุกข้อ ซึ่งมีค่าอำนาจจำแนกของข้อสอบ (a) อยู่ในช่วง 0.50 ถึง 2.50 ค่าความยากของข้อสอบ (b) อยู่ในช่วง -2.50 ถึง 2.50 และค่าโอกาสในการเดาข้อสอบ (c) มีค่าไม่เกิน 0.30 เมื่อเดำเนินการจัดการข้อมูลเสร็จ โปรแกรมบันทึกข้อมูลเข้าสู่แหล่งจัดเก็บข้อมูลข้อสอบ (D5)

กระบวนการที่ 6 การจัดการทดสอบ โดยการทดสอบได้มาจาก 2 แหล่งข้อมูล ได้แก่ 1) ผู้ดูแลโปรแกรมทำแบบทดสอบความรู้ทางไอซีที และ 2) ผู้ทดสอบทำการทดสอบความรู้ทางไอซีที ซึ่งข้อสอบมาจากข้อมูลข้อสอบ (D5) เมื่อดำเนินการทดสอบเสร็จ โปรแกรมบันทึกข้อมูลเข้าสู่แหล่งจัดเก็บข้อมูลการทดสอบ (D6) และแหล่งจัดเก็บข้อมูลรายละเอียดการทดสอบ (D7)

กระบวนการที่ 7 การรายงาน โปรแกรมนำเสนอรายงานผลการทดสอบผ่านทางจอภาพ แบ่งการนำเสนอออกเป็น 2 ส่วน ได้แก่ รายงานสำหรับผู้ดูแลโปรแกรม และรายงานสำหรับผู้ทดสอบ



ภาพที่ 3-11 แผนภาพกระแสข้อมูลระดับ 1 การจัดการทดสอบ

จากภาพที่ 3-11 แสดงกระแสการรับส่งข้อมูลของกระบวนการที่ 6 การจัดการทดสอบ ประกอบด้วย 4 กระบวนการย่อย ดังนี้

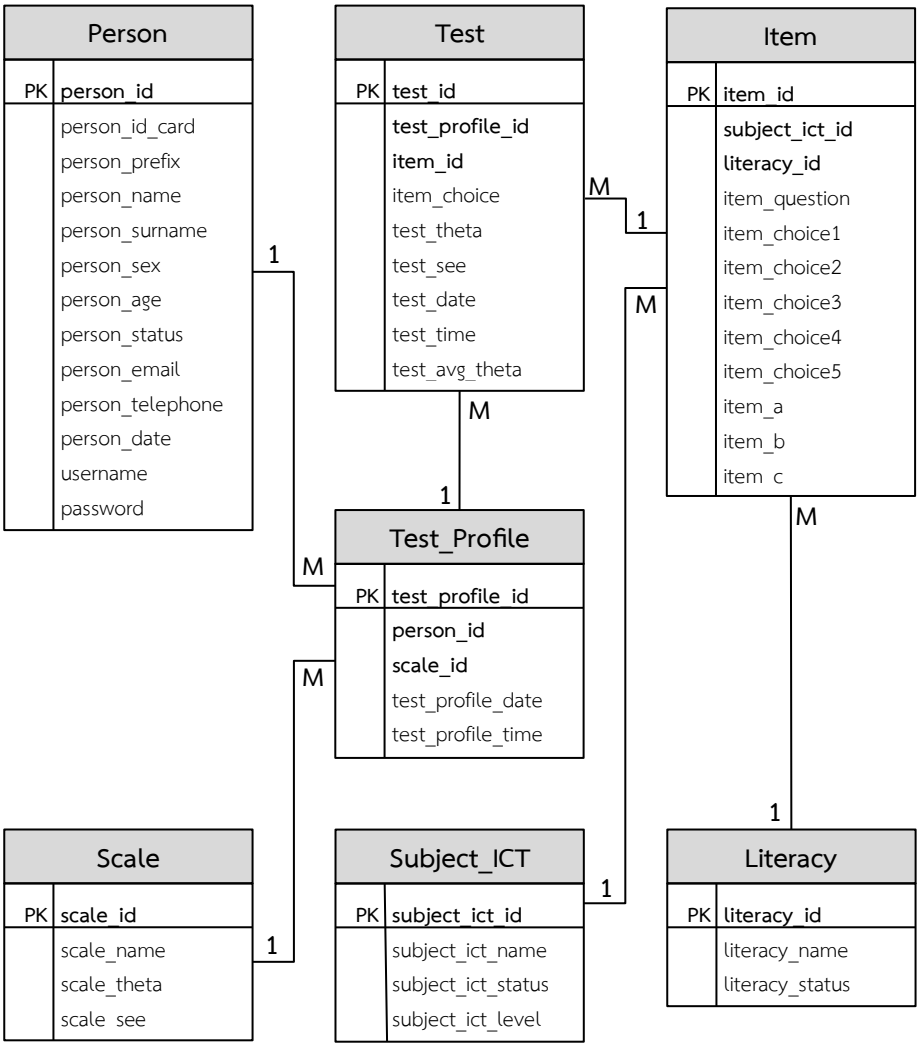
กระบวนการย่อยที่ 1 การคัดเลือกข้อสอบเริ่มต้น เป็นการเลือกข้อสอบข้อแรกให้กับผู้ทดสอบ ด้วยวิธีการสุ่มข้อสอบจากคลังข้อสอบที่มีค่าความยากของข้อสอบ (b) ระหว่าง -1.0 ถึง 1.0

กระบวนการย่อยที่ 2 การประมาณค่าความสามารถ เป็นการประมาณค่าความสามารถของผู้ทดสอบโดยใช้วิธีการประมาณค่าแบบความภาวะน่าจะเป็นสูงสุด (ML) (Reckase, 2009, p. 140)

กระบวนการย่อยที่ 3 การคัดเลือกข้อสอบข้อถัดไป เป็นการคัดเลือกข้อแรกของการเรียงลำดับจากน้อยไปหามากระหว่างค่าถี่สะสมกับผลต่างของค่าประมาณความสามารถของผู้ทดสอบกับค่าสารสนเทศสูงสุด

กระบวนการย่อยที่ 4 การยุติการทดสอบ เป็นกระบวนการที่ทำให้ข้อสอบยุติลง จากผลความคลาดเคลื่อนจากการประมาณค่าความสามารถกำหนดไว้ให้มีค่าน้อยกว่า 0.30 หรือยุติการทดสอบเมื่อข้อสอบหมดคลัง (Thompson & Weiss, 2011)

2.2 แบบจำลองฐานข้อมูล (Data Model) อธิบายความสัมพันธ์ของข้อมูลในโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา สำหรับวัดความรู้ทางไอซีที (Gary & Harry, 2012, p. 223; Alan et al., 2013, pp. 224-230)



ภาพที่ 3-12 แบบจำลองฐานข้อมูลโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

2.3 พจนานุกรมข้อมูล (Data dictionary) แสดงรายละเอียดตารางข้อมูลต่าง ๆ ในฐานข้อมูล (Database) ซึ่งประกอบด้วยแอตทริบิวต์ (Attribute), ชนิดข้อมูล (Data type), ความยาวข้อมูล (Length), คีย์ (Key), รายละเอียด (Description) ทำให้สามารถค้นหารายละเอียดที่ต้องการได้สะดวกมากยิ่งขึ้น พจนานุกรมข้อมูลเป็นการผสมผสานระหว่างรูปแบบของพจนานุกรมโดยทั่วไปและรูปแบบของข้อมูลในโปรแกรมงานคอมพิวเตอร์ เพื่ออธิบายชนิดของข้อมูลแต่ละตัวว่าเป็น ตัวเลข อักขระ ข้อความ หรือวันที่ เป็นต้น (Conceptual Data Model) (Gary & Harry, 2012, p. 223; Alan et al., 2013, p. 185)

ตารางที่ 3-7 รายละเอียดผู้ใช้โปรแกรม

Attribute	Data type	Length	Key	Description
person_id	Number	Auto	PK	รหัสผู้ใช้โปรแกรม
person_id_card	Varchar	30		รหัสบัตร ปชช.
person_prefix	Varchar	10		คำนำหน้า
person_name	Varchar	60		ชื่อ
person_surname	Varchar	60		นามสกุล
person_sex	Varchar	5		เพศ
person_age	Number	5		อายุ
person_status	Varchar	5		สถานะผู้ใช้
person_email	Varchar	60		อีเมล
person_telephone	Varchar	15		โทรศัพท์
person_date	Date			วันที่สมัคร
username	Varchar			ผู้ใช้โปรแกรม
password	Varchar			รหัสผ่าน

ตารางที่ 3-8 รายละเอียดข้อมูลแบบทดสอบ

Attribute	Data type	Length	Key	Description
scale_id	Number	Auto	PK	รหัสแบบทดสอบ
scale_name	Varchar	150		ชื่อแบบทดสอบ
scale_theta	Number			ค่าความสามารถ
scale_see	Number			ค่าคลาดเคลื่อน

ตารางที่ 3-9 รายละเอียดข้อมูลความรู้ทางไอซีที

Attribute	Data type	Length	Key	Description
literacy_id	Number	Auto	PK	รหัสความรู้
literacy_name	Varchar	150		ชื่อความรู้
literacy_status	Number	5		สถานะ

ตารางที่ 3-10 รายละเอียดข้อมูลหมวดไอซีที

Attribute	Data type	Length	Key	Description
subject_ict_id	Number	Auto	PK	รหัสหมวดหมู่
subject_ict_name	Varchar	150		ชื่อหมวดหมู่
subject_ict_status	Number	5		สถานะ
subject_ict_level	Number	5		ระดับ

ตารางที่ 3-11 รายละเอียดข้อมูลการทดสอบ

Attribute	Data type	Length	Key	Description
test_profile_id	Number	Auto	PK	รหัสรายละเอียด
person_id	Number		FK	รหัสผู้ใช้โปรแกรม
scale_id	Number		FK	รหัสแบบทดสอบ
test_profile_date	Date			วันที่ทดสอบ
test_profile_time	time			เวลาทดสอบ

ตารางที่ 3-12 รายละเอียดข้อมูลข้อสอบ

Attribute	Data type	Length	Key	Description
item_id	Number	Auto	PK	รหัสข้อสอบ
subject_ict_id	Number		FK	รหัสหมวดไอซีที
literacy_id	Number		FK	รหัสความรู้
item_question	Varchar			โจทย์
item_choice1	Varchar			ตัวเลือกที่1
item_choice2	Varchar			ตัวเลือกที่2
item_choice3	Varchar			ตัวเลือกที่3
item_choice4	Varchar			ตัวเลือกที่4
item_choice5	Varchar			ตัวเลือกที่5
item_a	Number			อำนาจจำแนก
item_b	Number			ความยาก
item_c	Number			โอกาสในการเดา

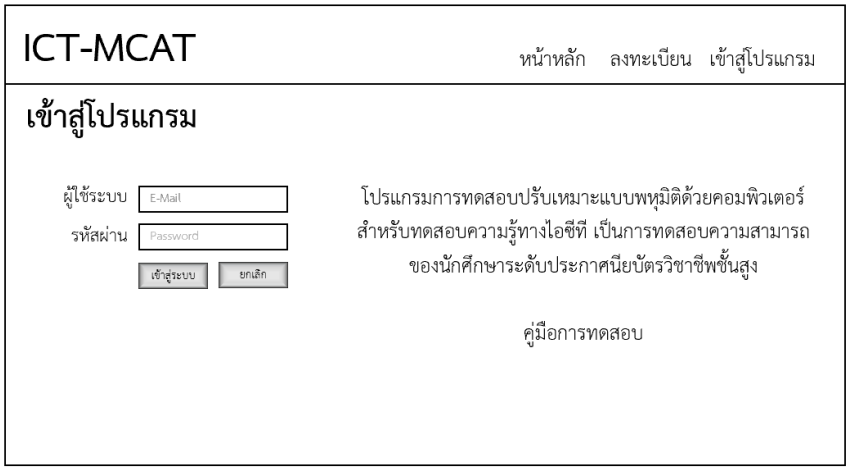
ตารางที่ 3-13 รายละเอียดการทดสอบ

Attribute	Data type	Length	Key	Description
test_id	Number	Auto	PK	รหัสการทดสอบ
test_profile_id	Number		FK	รหัสรายละเอียด
item_id	Number		FK	รหัสข้อสอบ
item_choice	Number			รหัสคำตอบ
test_theta	Number			ค่าความสามารถ
test_see	Number			ค่าคลาดเคลื่อน
test_date	date			วันที่ทดสอบ
test_time	time			เวลาทดสอบ
test_avg_theta	Number			ผลรวม

ขั้นตอนที่ 2 การออกแบบ (Design) แบ่งการออกแบบหน้าจอเป็น 2 ส่วน ได้แก่ 1) ส่วนของผู้ดูแลโปรแกรม และ 2) ส่วนของผู้ทดสอบ (Gary & Harry, 2012, pp. 342-360; Alan et al., 2013, pp. 311-345) ดังนี้

ส่วนที่ 1 ผู้ดูแลโปรแกรม มีรูปแบบการทำงาน ดังนี้

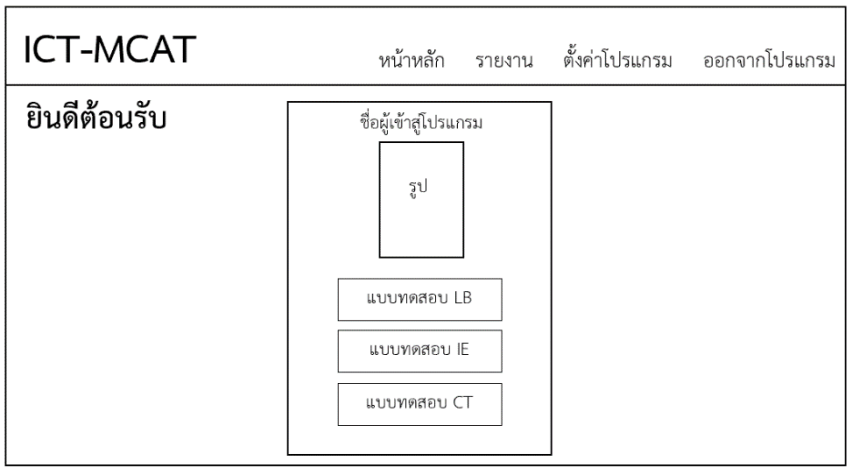
1. การเข้าสู่โปรแกรม



ภาพที่ 3-13 การเข้าสู่โปรแกรม

จากภาพที่ 3-13 การเข้าสู่โปรแกรม ผู้ดูแลโปรแกรมต้องกรอกข้อมูลอีเมลและรหัสผ่านเพื่อเข้าสู่โปรแกรม

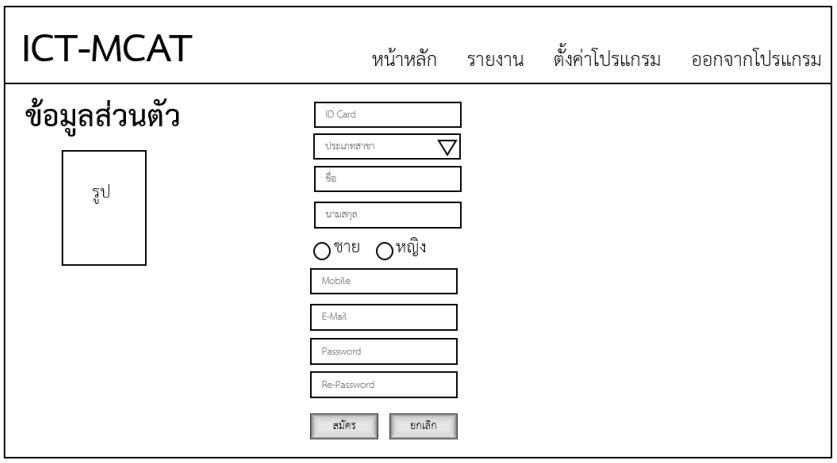
2. หน้าแรกหลังจากเข้าสู่โปรแกรม



ภาพที่ 3-14 หน้าแรกหลังจากเข้าสู่โปรแกรม

จากภาพที่ 3-14 หลังจากเข้าสู่โปรแกรมแบ่งออกเป็น 2 ส่วน ได้แก่ 1) ส่วนของเมนู และ 2) ส่วนของผู้ดูแลโปรแกรม

3. การจัดการข้อมูลผู้ใช้โปรแกรม



ภาพที่ 3-15 การจัดการข้อมูลผู้ใช้โปรแกรม

จากภาพที่ 3-15 ในส่วนของข้อมูลส่วนตัว ผู้ดูแลโปรแกรมสามารถแก้ไขข้อมูลส่วนตัวดังนี้

1) รหัสบัตรประชาชน ใส่หมายเลขบัตรประชาชนจำนวน 13 หลัก และโปรแกรมใช้อัลกอริทึมคำนวนเพื่อตรวจสอบความถูกต้องของรหัสประชาชน

2) ประเภทสาขาวิชา มีให้เลือก 3 ประเภทสาขาวิชา ได้แก่ 2.1) บริหารธุรกิจ 2.2) อุตสาหกรรม และ 2.3) เทคโนโลยีสารสนเทศ

3) ชื่อ ของผู้ดูแลโปรแกรม

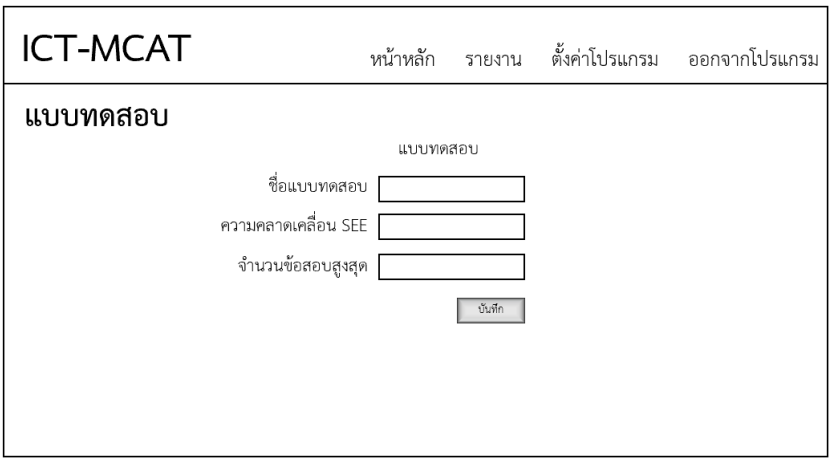
4) นามสกุล ของผู้ดูแลโปรแกรม

5) เพศ มีให้เลือก เพศชาย และเพศหญิง

6) เบอร์โทรศัพท์ ของผู้ดูแลโปรแกรม

7) อีเมลเพื่อใช้ในการ Login เข้าสู่โปรแกรม

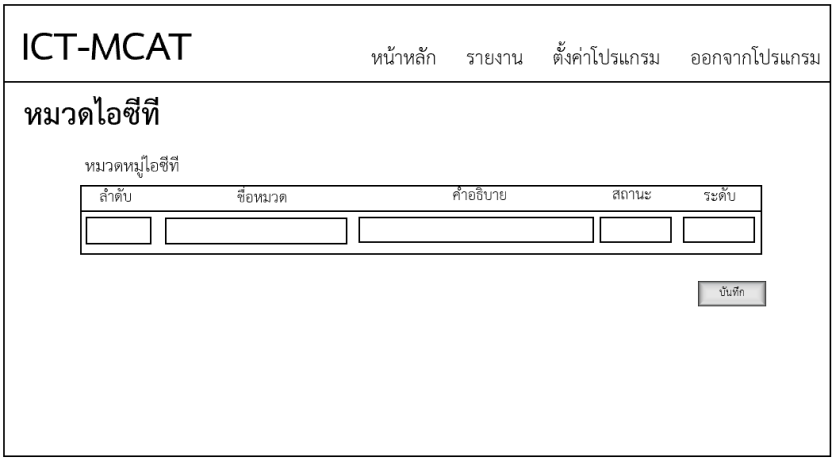
4. การจัดการข้อมูลแบบทดสอบ



ภาพที่ 3-16 การจัดการข้อมูลแบบทดสอบ

จากภาพที่ 3-16 การจัดการข้อมูลแบบทดสอบ ประกอบด้วยชื่อแบบทดสอบ ค่าคลาดเคลื่อน (SEE) และจำนวนข้อสอบสูงสุดที่ใช้สอบ

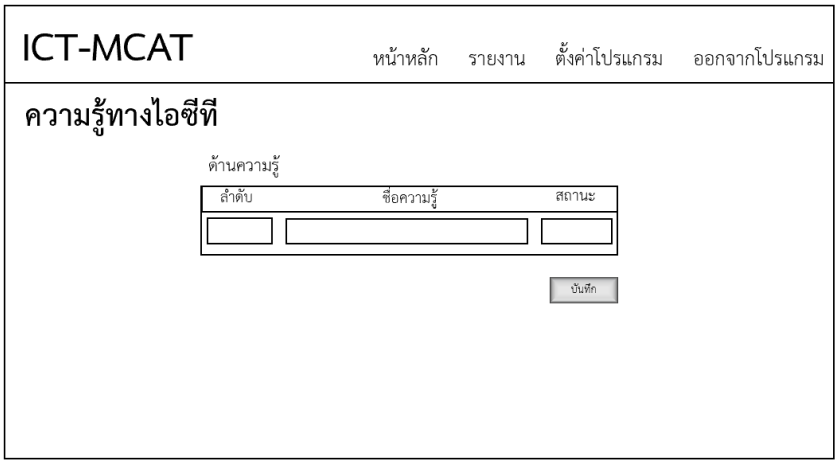
5. การจัดการข้อมูลหมวดไอซีที



ภาพที่ 3-17 การจัดการข้อมูลหมวดไอซีที

จากภาพที่ 3-17 เป็นการจัดการข้อมูลหมวดไอซีที ประกอบด้วย ชื่อหมวดไอซีที คำอธิบายของหมวด สถานะใช้งานของหมวดไอซีที

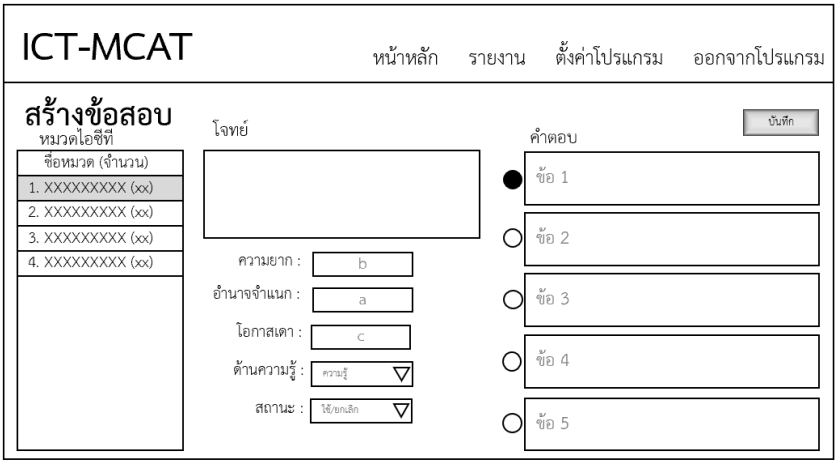
6. การจัดการข้อมูลความรู้ทางไอซีที



ภาพที่ 3-18 การจัดการข้อมูลความรู้ทางไอซีที

จากภาพที่ 3-18 เป็นการจัดการความรู้ทางไอซีที ประกอบด้วย ชื่อความรู้ทางไอซีที และสถานะใช้งานของความรู้ทางไอซีที

7. การจัดการข้อมูลข้อสอบ



ภาพที่ 3-19 การจัดการข้อมูลข้อสอบ

จากภาพที่ 3-19 เป็นการจัดการข้อมูลข้อสอบที่ผ่านการตรวจสอบคุณภาพรายข้อ ประกอบด้วย

1) โจทย์ เป็นการใส่ข้อมูลโจทย์ข้อสอบ

2) คำตอบ เป็นการใส่ข้อมูลคำตอบแบบ 5 ตัวเลือก

3) รูปประกอบ เป็นการใส่รูปภาพประกอบ ได้ทั้งรูปภาพข้อสอบและรูปภาพคำตอบ

4) อำนาจจำแนก (a) เป็นการใส่ตัวเลขของค่าอำนาจจำแนกของข้อสอบ

5) ความยาก (b) เป็นการใส่ตัวเลขของค่าความยากของข้อสอบ

6) โอกาสเดา (c) เป็นการใส่ตัวเลขของค่าโอกาสการเดาของข้อสอบ

7) สารสนเทศสูงสุด (MI) เป็นการใส่ค่าจากการคำนวณสารสนเทศสูงสุด

8) ด้านความรู้ เป็นการใส่ข้อมูลด้านความรู้ทางไอซีที 6 ด้าน

9) หมวดไอซีที เป็นการใส่ข้อมูลด้านเนื้อหา 5 หมวด

10) สถานะ เป็นการเปิดและปิดของข้อสอบ

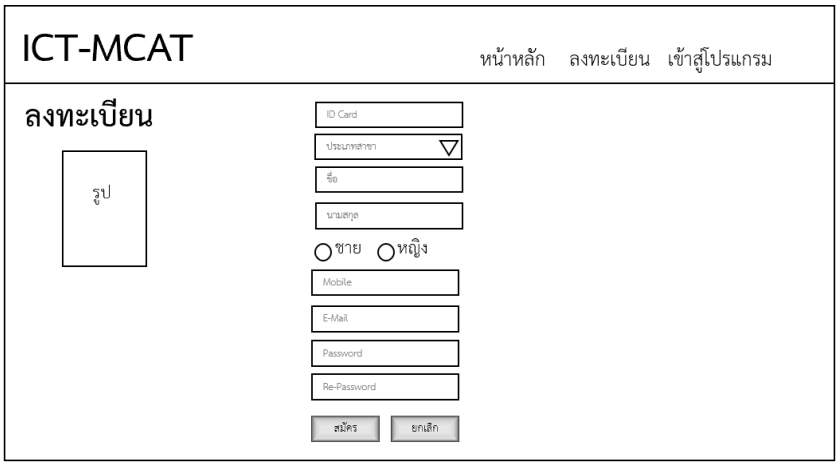
8. รายงาน

ภาพที่ 3-20 รายงานผลการทดสอบ

จากภาพที่ 3-20 เป็นรายงานที่ผู้ดูแลโปรแกรมใช้ในการตรวจสอบความสมบูรณ์ของข้อมูล เช่น รายงานความสมบูรณ์ของผู้ทดสอบ รายงานความสามารถรายบุคคล รายงานอัตราการใช้ข้อสอบ และรายงานความสมดุลเนื้อหา เป็นต้น

ส่วนที่ 2 ผู้ทดสอบ มีรูปแบบการทำงาน ดังนี้

1. การจัดการข้อมูลลงทะเบียนผู้ทดสอบ



ภาพที่ 3-21 การจัดการข้อมูลลงทะเบียนผู้ทดสอบ

จากภาพที่ 3-21 ผู้ทดสอบลงทะเบียน โดยมีรายละเอียดข้อมูลผู้ทดสอบ ดังนี้

1) รหัสบัตรประชาชน ใส่หมายเลขบัตรประชาชนจำนวน 13 หลัก และโปรแกรมใช้อัลกอริทึมคำนวนเพื่อตรวจสอบความถูกต้องของรหัสประชาชน

2) ประเภทสาขาวิชา มีให้เลือก 3 ประเภทสาขาวิชา ได้แก่ 2.1) บริหารธุรกิจ 2.2) อุตสาหกรรม และ 2.3) เทคโนโลยีสารสนเทศ

3) ชื่อ ของผู้ทดสอบ

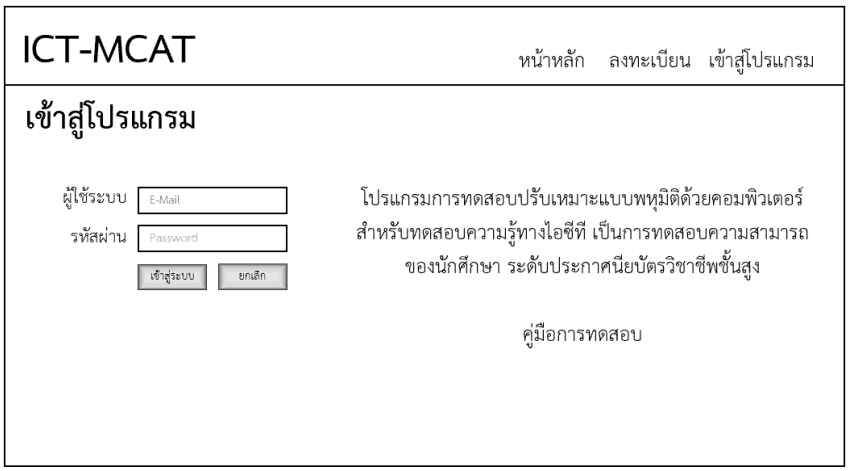
4) นามสกุล ของผู้ทดสอบ

5) เพศ มีให้เลือก เพศชาย และเพศหญิง

6) เบอร์โทรศัพท์ ของผู้ทดสอบ

7) อีเมลเพื่อใช้ในการ Login เข้าสู่โปรแกรม

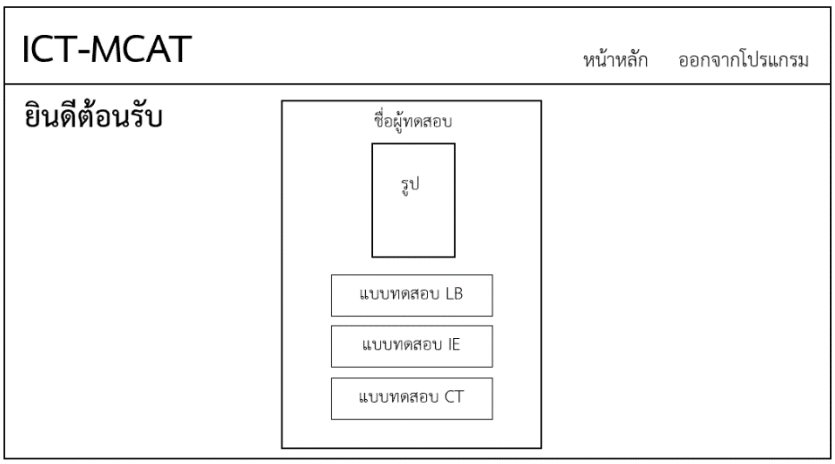
2. การเข้าสู่โปรแกรม



ภาพที่ 3-22 การเข้าสู่โปรแกรมผู้ทดสอบ

จากภาพที่ 3-22 เป็นหน้าการเข้าสู่โปรแกรมของผู้ทดสอบ หลังจากลงทะเบียนเสร็จเรียบร้อย

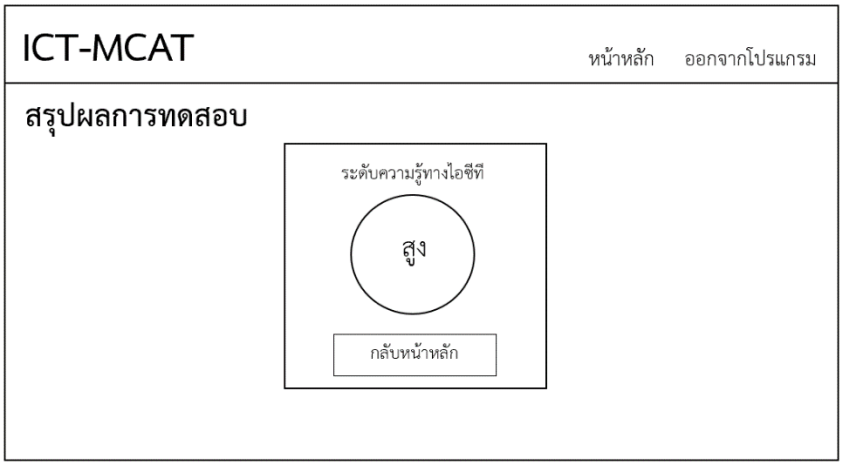
3. การจัดการข้อมูลการทดสอบ



ภาพที่ 3-23 การจัดการข้อมูลการทดสอบ

จากภาพที่ 3-23 เป็นหน้าจอแสดงการเลือกแบบทดสอบ โดยมี 3 แบบ ได้แก่ 1) แบบ LB เป็นวิธีการใหม่ที่ผู้วิจัยได้พัฒนาขึ้น 2) แบบ IE เป็นวิธีการเดิม และ 3) แบบ CT เป็นการทดสอบด้วยคอมพิวเตอร์ ซึ่งแบบที่ 1 และ 2 เป็นการทดสอบปรับเหมาะแบบพหุมิติด้วยยคอมพิวเตอร์

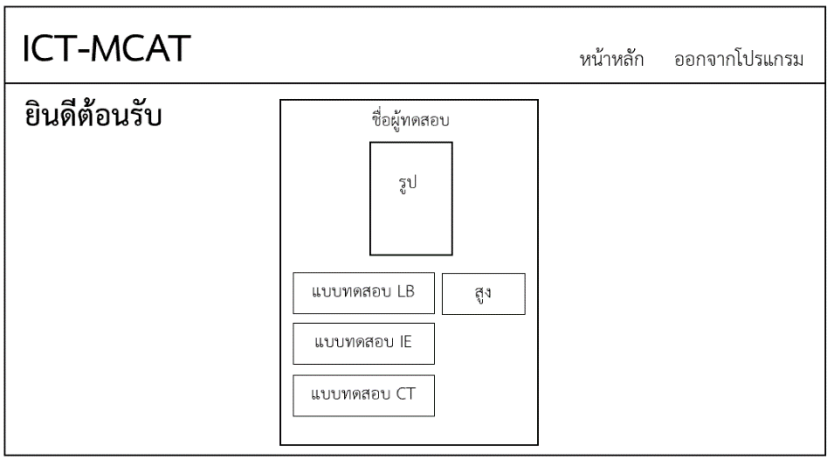
4. รายงานผลการทดสอบเมื่อทดสอบเสร็จ



ภาพที่ 3-24 รายงานผลการทดสอบเมื่อทดสอบเสร็จ

จากภาพที่ 3-24 เมื่อผู้ทดสอบได้ทดสอบแบบใดแบบหนึ่งเสร็จ โปรแกรมแสดงผลการทดสอบแบบภาพรวมของระดับความรู้ทางไอซีที แบ่งออกเป็น 7 ระดับ ได้แก่ 1) สูงมาก 2) สูง 3) ค่อนข้องสูง 4) ปานกลาง 5) ค่อนข้างต่ำ 6) ต่ำ และ 7) ต่ำมาก

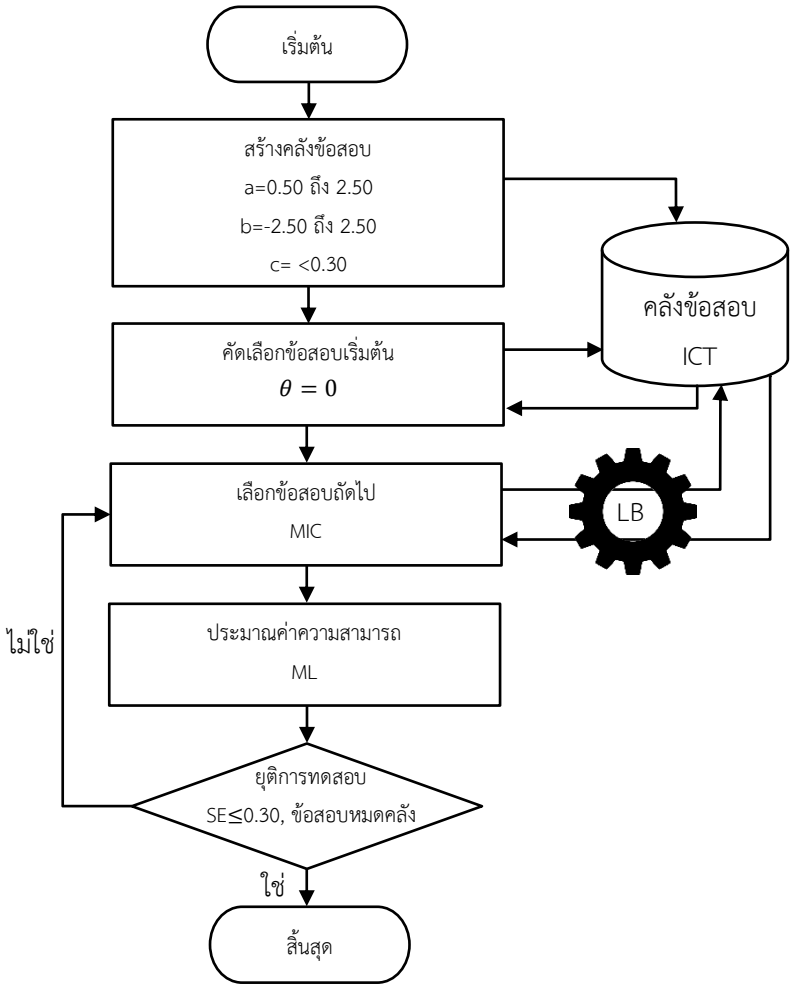
5. รายงานผลการทดสอบแต่ละแบบ



ภาพที่ 3-25 รายงานผลการทดสอบแต่ละแบบ

จากภาพที่ 3-25 ผลการทดสอบของแต่ละแบบที่ผู้ทดสอบได้เข้าทำ โดยผลการทดสอบแสดงระดับความสามารถหลังทำแบบทดสอบเสร็จ

ขั้นตอนที่ 3 การพัฒนาโปรแกรม (Implementation and coding) เป็นกระบวนการนำเอกสารที่วิเคราะห์มาเขียน Source code ให้เป็นไปตามที่ออกแบบเอาไว้ โดยแบ่งออกเป็นส่วน ๆ (Gary & Harry, 2012, pp. 291-292; Alan et al., 2013, pp. 444-445) ใช้ Code ในการพัฒนา คือ PHP, AJAX, HTML5, Java Script ประกอบกัน ใช้ฐานข้อมูล MySQL และโปรแกรม Eclipse 2014 เป็นเครื่องมือ (Tools) ในการพัฒนา Code ดังกล่าว แบ่งออก 5 ขั้นตอนตามกรอบแนวคิดของการทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ (CAT) (Thompson & Weiss, 2011) แสดงดังภาพที่ 3-26



ภาพที่ 3-26 ขั้นตอนการพัฒนาโปรแกรมตามแนวคิดการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์

จากภาพที่ 3-26 อธิบายถึงขั้นตอนการพัฒนาโปรแกรมตามแนวคิดการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ ซึ่งมีขั้นตอนดังนี้

1) การสร้างคลังข้อสอบ พัฒนาหน้าจอโปรแกรมเพื่อจัดเก็บข้อสอบ มีการเชื่อมโยงฐานข้อมูลชื่อตาราง Subject_ICT, Literacy และ Item เป็นการนำเนื้อหาวิชาเทคโนโลยีสารสนเทศและการสื่อสาร ระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) ที่ได้มาจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ข้อสอบระหว่างปี พ.ศ.2556-2560 ผ่านการวิเคราะห์คุณภาพจำนวน 192 ข้อ

2) การคัดเลือกข้อสอบเริ่มต้น พัฒนาหน้าจอโปรแกรมที่มีอัลกอริทึมการคัดเลือกข้อสอบเริ่มต้น มีค่าความยาก (b) ระหว่าง -1.0 ถึง 1.0 เพื่อนำมาทดสอบเป็นข้อแรกของการเริ่มต้นการทดสอบ

3) การเลือกข้อสอบข้อถัดไป พัฒนาหน้าจอโปรแกรมที่มีอัลกอริทึมค่าสารสนเทศสูงสุดเพิ่มเติมวิธีการควบคุมการใช้ข้อสอบและควบคุมสมดุลเนื้อหา 2 วิธี ได้แก่ 1) วิธี LB และ 2) วิธี IE โดยวิธี LB มีอัลกอรึทึม 3 ขั้นตอน ได้แก่ 1) การสำรวจและเคลื่อนย้ายข้อมูล 2) การคำนวณภาระงาน และ 3) การควบคุมการทำงาน

4) การประมาณค่าความสามารถ ใช้ผลการตอบข้อสอบของผู้ทดสอบสำหรับการประมาณค่าความสามารถโดยใช้วิธีการประมาณค่าแบบ ML

5) เกณฑ์ยุติการทดสอบ พัฒนา Code ในการยุติการทดสอบโดยใช้ผลความคลาดเคลื่อนจากการประมาณค่าความสามารถซึ่งกำหนดไว้ให้มีค่าน้อยกว่า 0.30 เป็นเกณฑ์ในการยุติการทดสอบหรือยุติการทดสอบเมื่อข้อสอบหมดคลัง

ขั้นตอนที่ 4 การทดสอบ (Testing) ในระยะการทดสอบโปรแกรมแบ่งออกเป็น 2 ส่วนดังนี้

ส่วนที่ 1 การทดสอบโปรแกรม (Program testing) การทดสอบโปรแกรมโดยใช้ Black box testing เป็นแบบประเมินโปรแกรมหลังจากการพัฒนา เพื่อตรวจสอบคุณภาพของโปรแกรมให้เป็นไปตามกรอบการวางแผนพัฒนาโปรแกรม (Gary & Harry, 2012, pp. 525-526; Alan et al., 2013, p. 453) โดยการทดสอบแบ่งออกเป็น 4 ด้าน ได้แก่ 1) ความตรงตามความต้องการของโปรแกรม (Functional requirement test) เป็นการประเมินผลความถูกต้องและประสิทธิภาพของโปรแกรม ตรงตามความต้องการของผู้ใช้โปรแกรมมีความเหมาะสมมากน้อยเพียงใด 2) การทำงานของโปรแกรม (Functional test) เป็นการประเมินความถูกต้องและประสิทธิภาพในการทำงานของโปรแกรม สามารถทำงานได้ตามฟังก์ชันของโปรแกรมมีความเหมาะสมมากน้อยเพียงใด 3) การใช้งานโปรแกรม (Usability test) เป็นการประเมินลักษณะการออกแบบโปรแกรม มีความง่ายต่อการใช้งานมีความเหมาะสมมากน้อยเพียงใด และ 4) ความปลอดภัยของโปรแกรม (Security test) เป็นการประเมินโปรแกรมในด้านการรักษาความปลอดภัยของข้อมูลในโปรแกรม มีความเหมาะสมมากน้อย

เพียงใด โดยใช้แบบประเมินผลที่พัฒนาขึ้น ซึ่งมีลักษณะเป็นมาตรประมาณค่า 5 ระดับ เกณฑ์การให้คะแนน มีดังนี้

1) เกณฑ์การให้คะแนนระดับประสิทธิภาพ

1 หมายถึง น้อยที่สุด

2 หมายถึง น้อย

3 หมายถึง ปานกลาง

4 หมายถึง มาก

5 หมายถึง มากที่สุด

2) เกณฑ์การแปลความหมายค่าเฉลี่ย ดังนี้

ค่าเฉลี่ยตั้งแต่ 4.51 ถึง 5.00 หมายถึง โปรแกรมมีประสิทธิภาพอยู่ในระดับมากที่สุด

ค่าเฉลี่ยตั้งแต่ 3.51 ถึง 4.50 หมายถึง โปรแกรมมีประสิทธิภาพอยู่ในระดับมาก

ค่าเฉลี่ยตั้งแต่ 2.51 ถึง 3.50 หมายถึง โปรแกรมมีประสิทธิภาพอยู่ในระดับปานกลาง

ค่าเฉลี่ยตั้งแต่ 1.51 ถึง 2.50 หมายถึง โปรแกรมมีประสิทธิภาพอยู่ในระดับน้อย

ค่าเฉลี่ยตั้งแต่ 1.00 ถึง 1.50 หมายถึง โปรแกรมมีประสิทธิภาพอยู่ในระดับน้อยที่สุด

การประเมินโปรแกรม แบ่งออกเป็น 2 ส่วน คือ การประเมินโดยผู้เชี่ยวชาญและการประเมินโดยผู้ใช้งาน

1. การประเมินโดยผู้เชี่ยวชาญ (Expert testing) เป็นผู้เชี่ยวชาญในสายงานเทคโนโลยีสารสนเทศ ต้องมีประสบการณ์ด้านเทคโนโลยีสารสนเทศที่เกี่ยวข้องและมีประวัติการทำงานด้านวิเคราะห์และออกแบบโปรแกรมพร้อมทั้งทักษะการจัดการเทคโนโลยีสารสนเทศ ดังนี้

1.1 รองศาสตราจารย์ ดร.ปณิตา วรรณพิรุณ อาจารย์ประจำคณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

1.2 ผู้ช่วยศาสตราจารย์ ดร.สยาม แกมขุนทด อาจารย์ประจำคณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

1.3 ผู้ช่วยศาสตราจารย์ ดร.รุ่งทิวา เสาร์สิงห์ อาจารย์ประจำคณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ

1.4 ดร.มิ่ง เทพครเมือง ผู้ช่วยฝ่ายวิจัย นิเทศ และประสบการณ์วิชาชีพ โรงเรียนสาธิตมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร (ฝ่ายประถม)

1.5 ดร.ไพเราะ ราชสมบูรณ์ อาจารย์พิเศษคณะบริหารธุรกิจ มหาวิทยาลัยเกริก และประธานบริษัท ซี เอส เอ็น แอ็ดวานซ์ จำกัด

เมื่อผ่านการตรวจประสิทธิภาพโปรแกรมจากผู้เชี่ยวชาญจึงนำโปรแกรมไปทดสอบกับนักศึกษา 40 คนที่ลงทะเบียนเรียนวิชาเทคโนโลยีสารสนเทศ เพื่อทดสอบปัญหาของโปรแกรม (BUG)

แล้วนำปัญหาจากการทดสอบมาปรับปรุงโปรแกรม แล้วขอจริธรรมงานวิจัย เลขที่ 074/2561 ลง วันที่ 30 ตุลาคม 2561

2. การประเมินโดยผู้ใช้งาน (User testing) การประเมินผลโดยผู้ใช้งาน ด้านการใช้งานโปรแกรม เป็นนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง จำนวน 400 คน ที่ลงทะเบียนเรียนวิชาเทคโนโลยีสารสนเทศและการสื่อสาร ประเมินผลภายหลังจากผู้ใช้งานได้ทดสอบแบบ LB และ IE เสร็จเรียบร้อย ผ่านทางเว็บไซต์ http://ict-mcat.com

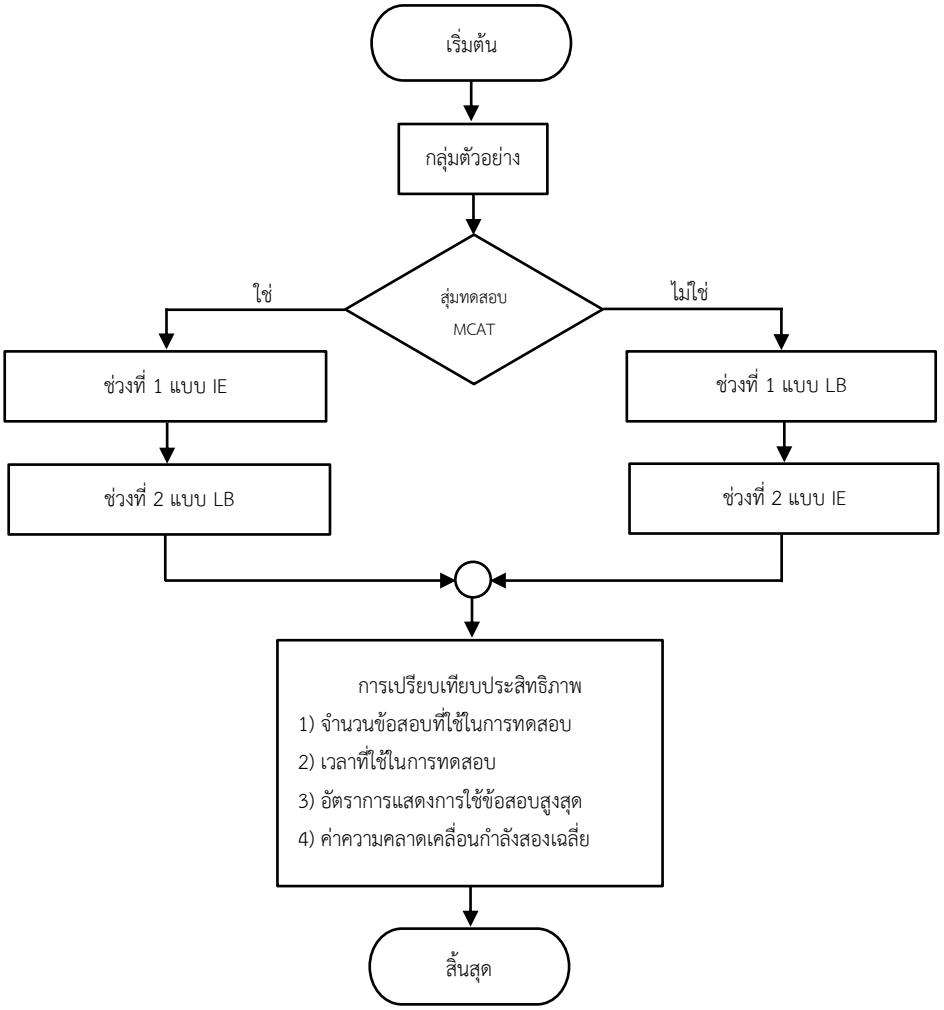
ส่วนที่ 2 การพัฒนาเอกสารผู้ใช้งาน (User document developing) เป็นการจัดทำเอกสารหรือคู่มือการใช้งานโปรแกรม ได้อธิบายถึงวิธีการใช้งานอย่างละเอียด โดยแสดงตัวอย่างหน้าจอการใช้งานพร้อมทั้งผลลัพธ์ที่เกิดขึ้นจากโปรแกรมประกอบคำอธิบาย ผู้ดูแลโปรแกรมและผู้ทดสอบ เพื่อให้สามารถใช้คู่มือเป็นแนวทางในการใช้งานโปรแกรมได้อย่างถูกต้องและรวดเร็ว (Gary & Harry, 2012, p. 528; Alan et al., 2013, p. 464)

ขั้นตอนที่ 5 การใช้งาน (Deployment) เป็นขั้นตอนนำไปใช้งานจริง รวมถึงการนำเสนองานของการใช้งานจริง (Gary & Harry, 2012, p. 572; Alan et al., 2013, p. 479) โดยผู้ทดสอบทำแบบทดสอบผ่านทางอินเทอร์เน็ต (Internet) ซึ่งโปรแกรมและฐานข้อมูลได้จัดเก็บอยู่บนการประมวลผลแบบแบ่งปันทรัพยากรผ่านเครือข่าย (Cloud computing) ผ่านทางเว็บไซต์ http://ict-mcat.com

ขั้นตอนที่ 6 การบำรุงรักษา (Maintenance) เป็นขั้นตอนการดูแลโปรแกรมเพื่อให้โปรแกรมมีประสิทธิภาพในการทำงาน (Gary & Harry, 2012, p. 574; Alan et al., 2013, p. 495) โดยแบ่งการบำรุงรักษา 1) เพื่อแก้ไขข้อผิดพลาดของโปรแกรม เป็นการแก้ไข ปรับปรุงข้อผิดพลาดที่เกิดขึ้นระหว่างการใช้โปรแกรมแกรม 2) เพื่อให้โปรแกรมสามารถรองรับความต้องการที่เพิ่มขึ้น เป็นการพัฒนาเพิ่มเติมความต้องการของผู้ใช้โปรแกรม เพื่อให้ครอบคลุมการใช้งานที่จำเป็น 3) เพื่อบำรุงรักษาโปรแกรมให้ทำงานได้อย่างมีประสิทธิภาพ เป็นการเพิ่มศักยภาพของโปรแกรมให้ทำงานได้อย่างมีประสิทธิภาพ และ 4) เพื่อบำรุงรักษาโปรแกรมป้องกันข้อผิดพลาดที่จะเกิดขึ้น

ระยะที่ 3 การเปรียบเทียบประสิทธิภาพวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

ในขั้นตอนนี้เป็นการเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา สำหรับการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ 2 วิธี คือ 1) วิธีโหลดบาลานซ์ซิ่ง (LB) และ 2) วิธี Item Eligibility (IE) โดยเปรียบเทียบประสิทธิภาพ 4 ด้าน ได้แก่ 1) จำนวนข้อสอบที่ใช้ในการทดสอบ 2) เวลาที่ใช้ในการทดสอบ 3) อัตราการแสดงการใช้ข้อสอบสูงสุด และ 4) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ขั้นตอนแสดงดังภาพที่ 3-27



ภาพที่ 3-27 ขั้นตอนการเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับวิธีการควบคุมสมดุลเนื้อหา

จากภาพที่ 3-27 แสดงขั้นตอนการเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา สำหรับการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ ระหว่าง 2 วิธี ได้แก่ 1) วิธี LB และ 2) วิธี IE ผู้ทดสอบแต่ละคนต้องทำแบบทดสอบ 2 วิธี โดยก่อนเริ่มทดสอบ ผู้ทดสอบสุ่มหยิบบัตร วิธี LB หรือ วิธี IE ได้การทดสอบแบบใดก่อน หลังจากนั้นนำผลที่ได้รับการทดสอบทั้ง 2 วิธี มาเปรียบเทียบผลการทดสอบความรู้ทางไอซีที โดยมีวิธีดำเนินการวิจัยดังนี้

1. กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการวิจัยนี้ เป็นนักศึกษาอาชีวศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) ชั้นปีที่ 2 ทุกสาขา ที่ลงทะเบียนเรียนวิชาเทคโนโลยีสารสนเทศ จากวิทยาลัยเทคโนโลยีภาคตะวันออก (อี.เทค) สังกัดสำนักงานคณะกรรมการการอาชีวศึกษา โดยใช้วิธีการสุ่มอย่างมีระบบ (Systematic random sampling) ตามทะเบียนรายชื่อนักศึกษา จำนวน 400 คน ในภาคเรียนที่ 2 ปีการศึกษา 2561 ทดสอบระหว่าง วันที่ 6-9 พฤศจิกายน 2561

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย เป็นโปรแกรมทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์รายวิชาเทคโนโลยีสารสนเทศ ระดับประกาศนียบัตรวิชาชีพชั้นสูง จำนวนข้อสอบ 192 ข้อ มีความยาวข้อสอบไม่เกิน 30 ข้อ ซึ่งเป็นโปรแกรมในรูปแบบเว็บแอพพลิเคชั่น ผ่านทางเว็บไซต์ http://ict-mcat.com ผู้ใช้สามารถเลือกการทดสอบได้ 2 วิธีคือ 1) วิธี LB และ 2) วิธี IE

3. การเก็บรวบรวมข้อมูล

การวิจัยนี้แบ่งการเก็บรวมรวมข้อมูลเป็น 2 ส่วน ดังนี้

ส่วนที่ 1 การเตรียมเอกสาร แบ่งออกเป็น 3 ขั้นตอนดังนี้

ขั้นตอนที่ 1 เตรียมเอกสารที่ใช้แนบสำหรับออกหนังสือขอความอนุเคราะห์ เช่น เค้าโครงการวิจัยฉบับย่อ คู่มือการใช้โปรแกรม จริยธรรมงานวิจัย เป็นต้น

ขั้นตอนที่ 2 ขอความอนุเคราะห์จากวิทยาลัยวิยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา เกี่ยวกับการออกหนังสือขอความอนุเคราะห์เก็บรวบรวมข้อมูลถึงผู้บริหารวิทยาลัยเทคโนโลยีภาคตะวันออก (อี.เทค)

ขั้นตอนที่ 3 จัดส่งหนังสือขอความอนุเคราะห์เก็บรวบรวมข้อมูลถึงผู้บริหารวิทยาลัยของกลุ่มตัวอย่าง เลขที่ ศธ.6224/0409 วันที่ 24 ตุลาคม 2561 พร้อมแนบใบยินยอมเข้าร่วมการวิจัย ผู้อำนวยการวิทยาลัยได้ลงนามยินยอมเข้าร่วมการวิจัย ลงวันที่ 6 พฤศจิกายน 2561

ส่วนที่ 2 การดำเนินการทดสอบ แบ่งออกเป็น 3 ขั้นตอนดังนี้

ขั้นตอนที่ 1 การเตรียมคลังข้อสอบ การวัดความรู้ทางไอซีทีมี 2 มิติในการจัดทำคลังข้อสอบ มิติที่ 1 เป็นเนื้อหาวิชาที่ใช้ทดสอบ ส่วนมิติที่ 2 เป็นความรู้ทางไอซีทีที่ต้องการทดสอบ

ขั้นตอนที่ 2 อุปกรณ์ที่ใช้ทดสอบ ผู้ทดสอบต้องเตรียมอุปกรณ์ที่สามารถเข้าสู่โปรแกรมทดสอบปรับเหมาะแบบพหุมิติ โดยอุปกรณ์ที่สามารถเข้าถึง เช่น คอมพิวเตอร์ตั้งโต๊ะ สมาร์ทโฟน แท็บเล็ต เป็นต้น แล้วเชื่อมต่อเครือข่ายอินเทอร์เน็ต เปิดเว็บบราวเซอร์ พิมพ์ URL ในช่อง ที่อยู่ (address) http://ict-mcat.com เพื่อเข้าสู่โปรแกรมทดสอบ

ขั้นตอนที่ 3 การดำเนินการทดสอบ แบ่งออกเป็น 3 ขั้นตอนย่อยดังนี้

1. อธิบายวัตถุประสงค์ของการทดสอบและอธิบายการใช้โปรแกรม เพื่อให้กลุ่มตัวอย่างเห็นความสำคัญของการทำข้อสอบในแต่ละข้อ ทำด้วยความเต็มใจและตั้งใจ

2. สุ่มหยิบบัตรทดสอบ เนื่องจากผู้ทดสอบต้องดำเนินการทดสอบ 2 วิธี คือ วิธี LB กับวิธี IE แบ่งออกเป็น 2 เงื่อนไข คือ เงื่อนไขที่ 1 ถ้าคำตอบจากการสุ่มหยิบบัตรคือ “LB” ให้ผู้ทดสอบทำแบบทดสอบวิธี LB ก่อน แล้วตามด้วยทดสอบวิธี IE ส่วนเงื่อนไขที่ 2 ถ้าคำตอบจากการสุ่มหยิบบัตรคือ “IE” ให้ผู้ทดสอบทำแบบทดสอบวิธี IE ก่อน แล้วตามด้วยทดสอบวิธี LB ตัวอย่างบัตรทดสอบดังภาพที่ 3-28

ภาพที่ 3-28 บัตรทดสอบ ICT-MCAT แบบ LB กับ IE

3. ดำเนินการทดสอบด้วยโปรแกรมที่พัฒนาขึ้น ซึ่งผู้ทดสอบดำเนินการทดสอบแบบออนไลน์ แสดงดังตารางที่ 3-14

ตารางที่ 3-14 ลำดับการทดสอบวิธี LB กับวิธี IE

	สอบช่วงที่ 1	พัก	สอบช่วงที่ 2
สุ่มกลุ่มทดสอบ “LB”	แบบ LB	30 นาที	แบบ IE
สุ่มกลุ่มทดสอบ “IE”	แบบ IE		แบบ LB

4. การวิเคราะห์ข้อมูล

การเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา สำหรับการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ 2 วิธี คือ LB กับ IE ซึ่งวิเคราะห์ด้วยสถิติ ดังนี้

4.1 จำนวนข้อสอบที่ใช้ในการทดสอบ เป็นการเปรียบเทียบประสิทธิภาพด้านจำนวนข้อสอบที่ใช้ในการทดสอบ ซึ่งเป็นข้อมูลแบบไม่ต่อเนื่อง จึงใช้สถิตินอนพาราเมตริก ด้วยการทดสอบ Mann-Whitney test

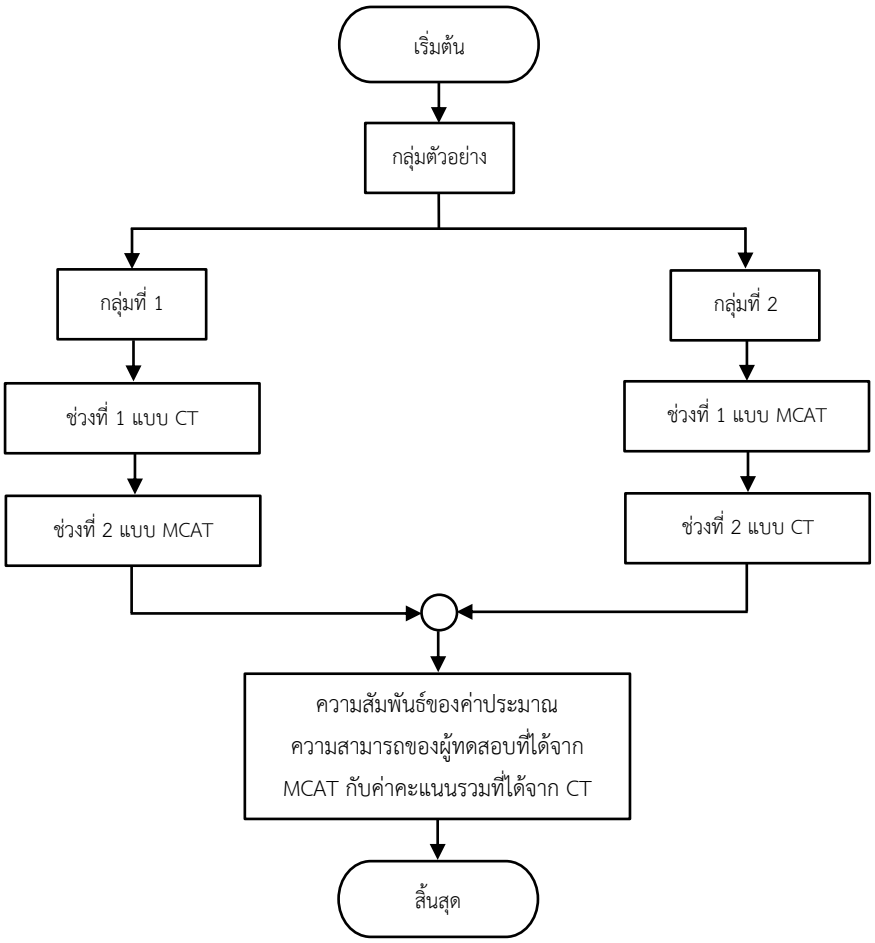
4.2 เวลาที่ใช้ในการทดสอบ เป็นการเปรียบเทียบประสิทธิภาพด้านเวลาในการทดสอบ โดยใช้ค่าเฉลี่ยของเวลามีหน่วยเป็นวินาที ทดสอบโดยใช้ *t*-test

4.3 อัตราการแสดงการใช้ข้อสอบสูงสุด ข้อสอบใด ๆ ที่มีอัตราการแสดงไม่เกิน 0.2 ถ้าข้อสอบที่มีการใช้จำนวนมากแสดงถึงข้อสอบในคลังข้อสอบมีโอกาสนำไปใช้มาก สะท้อนถึงประสิทธิภาพวิธีเลือกข้อสอบไม่เหมาะสม สามารถคำนวณอัตราการแสดงการใช้ข้อสอบสูงสุดจากสมการ 6

4.4 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) เป็นค่าความแม่นยำของค่าประมาณความสามารถของผู้ทดสอบ ซึ่งมีค่าเท่ากับผลต่างกำลังสองเฉลี่ย ค่าประมาณความสามารถมีค่าเข้าใกล้ 0 แสดงว่า ค่าประมาณความสามารถของผู้ทดสอบมีค่าใกล้เคียงกับค่าความสามารถที่แท้จริงของผู้ทดสอบ

ระยะที่ 4 การศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT)

การศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT) ขั้นตอนการดำเนินการแสดงดังภาพที่ 3-29



ภาพที่ 3-29 ขั้นตอนการศึกษาความสัมพันธ์ที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับการทดสอบด้วยคอมพิวเตอร์ (CT)

จากภาพที่ 3-29 แสดงขั้นตอนการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้ทดสอบที่ได้จาก MCAT กับค่าคะแนนรวมของผู้ทดสอบที่ได้จาก CT ผลการศึกษานี้แสดงให้เห็นว่า การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์มีการวัดผลที่สอดคล้องเป็นไปในทิศทางเดียวกับการทดสอบด้วยคอมพิวเตอร์ ด้วยเหตุนี้ ผู้ทดสอบจึงต้องทดสอบทั้ง 2 แบบ โดยใช้โปรแกรมการทดสอบปรับเหมาะแบบพหุมิติที่พัฒนาขึ้นในระยะที่ 2 เป็นเครื่องมือที่ใช้ในการวิจัย การดำเนินการวิจัยในระยะนี้ แบ่งกลุ่มตัวอย่างออกเป็น 2 กลุ่ม โดยกลุ่มที่ 1 ดำเนินการทดสอบแบบ CT ก่อนแล้วจึงทดสอบแบบ MCAT ส่วนในกลุ่มที่ 2 ดำเนินการทดสอบแบบ MCAT ก่อนแล้วจึงทดสอบแบบ CT หลังจากนั้นนำผลการทดสอบทั้ง 2 แบบ มาวิเคราะห์ความสัมพันธ์ โดยมีวิธีดำเนินการวิจัยดังนี้

1. กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการวิจัยนี้ เป็นนักศึกษาอาชีวศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) ชั้นปีที่ 2 สาขาวิชาเทคโนโลยีสารสนเทศ ในวิทยาลัยเทคโนโลยีภาคตะวันออก (อีเทค) สังกัดสำนักงานคณะกรรมการการอาชีวศึกษา โดยใช้วิธีการเลือกตัวอย่างแบบตามสะดวกจำนวน 40 คน ในภาคเรียนที่ 2 ปีการศึกษา 2561

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย เป็นโปรแกรมทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) จำนวน 192 ข้อ ความยาวของข้อสอบ 30 ข้อ และโปรแกรมทดสอบด้วยคอมพิวเตอร์ (CT) มีจำนวนข้อสอบ 40 ข้อ ตามแนวข้อสอบของ สทศ. รายวิชาเทคโนโลยีสารสนเทศและการสื่อสาร ระดับประกาศนียบัตรวิชาชีพชั้นสูง ซึ่งเป็นโปรแกรมในรูปแบบเว็บแอพพลิเคชั่น ผ่านทางเว็บไซต์ http://ict-mcat.com โดยผู้ใช้สามารถเลือกการทดสอบได้ 2 วิธี

3. การเก็บรวบรวมข้อมูล

การวิจัยนี้แบ่งการเก็บรวมรวมข้อมูลเป็น 2 ส่วน ดังนี้

ส่วนที่ 1 การเตรียมเอกสาร แบ่งออกเป็น 3 ขั้นตอนดังนี้

ขั้นตอนที่ 1 เตรียมเอกสารที่ใช้แนบสำหรับออกหนังสือขอความอนุเคราะห์ เช่น เค้าโครงการวิจัยฉบับย่อ คู่มือการใช้โปรแกรม จริยธรรมงานวิจัย เป็นต้น

ขั้นตอนที่ 2 ขอความอนุเคราะห์จากวิทยาลัยวิยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา เกี่ยวกับการออกหนังสือขอความอนุเคราะห์เก็บรวบรวมข้อมูลถึงผู้บริหารวิทยาลัยเทคโนโลยีภาคตะวันออก (อี.เทค)

ขั้นตอนที่ 3 จัดส่งหนังสือขอความอนุเคราะห์เก็บรวบรวมข้อมูลถึงผู้บริหารวิทยาลัยของกลุ่มตัวอย่าง เลขที่ ศธ 6224/0409 วันที่ 24 ตุลาคม 2561 แนบพร้อมใบยินยอมเข้าร่วมการวิจัย ผู้อำนวยการวิทยาลัยได้ลงนามยินยอมเข้ารวมการวิจัย ลงวันที่ 6 พฤศจิกายน 2561

ส่วนที่ 2 การดำเนินการทดสอบ แบ่งออกเป็น 3 ขั้นตอนดังนี้

ขั้นตอนที่ 1 การเตรียมคลังข้อสอบ การวัดความรู้ทางไอซีทีมี 2 มิติในการจัดทำคลังข้อสอบ มิติที่ 1 เป็นเนื้อหาวิชาที่ใช้ทดสอบ ส่วนมิติที่ 2 เป็นความรู้ทางไอซีทีที่ต้องการทดสอบ

ขั้นตอนที่ 2 อุปกรณ์ที่ใช้ทดสอบ ผู้ทดสอบต้องเตรียมอุปกรณ์ที่สามารถเข้าสู่โปรแกรมทดสอบปรับเหมาะแบบพหุมิติ โดยอุปกรณ์ที่สามารถเข้าถึง เช่น คอมพิวเตอร์ตั้งโต๊ะ สมาร์ทโฟน แท็บเล็ต เป็นต้น แล้วเชื่อมต่อเครือข่ายอินเทอร์เน็ต เปิดเว็บบราวเซอร์ พิมพ์ URL ในช่อง ที่อยู่ (Address) http://ict-mcat.com เพื่อเข้าสู่โปรแกรมทดสอบ

ขั้นตอนที่ 3 การดำเนินการทดสอบ แบ่งออกเป็น 3 ขั้นตอนย่อยดังนี้

1. อธิบายวัตถุประสงค์ของการทดสอบและอธิบายการใช้โปรแกรม เพื่อให้กลุ่มตัวอย่างเห็นความสำคัญของการทำข้อสอบในแต่ละข้อ ทำด้วยความเต็มใจและตั้งใจ

2. สุ่มหยิบบัตรทดสอบ เนื่องจากผู้ทดสอบต้องดำเนินการทดสอบ 2 วิธี คือ วิธี MCAT กับวิธี CT จากภาพ 3-29 แบ่งออกเป็น 2 เงื่อนไข คือ เงื่อนไขที่ 1 ถ้าคำตอบจากการสุ่มหยิบบัตรคือ “LB” ให้ผู้ทดสอบทำแบบทดสอบวิธี MCAT แบบ LB ก่อน แล้วตามด้วยทดสอบวิธี CT ส่วนเงื่อนไขที่ 2 ถ้าคำตอบจากการสุ่มหยิบบัตรคือ “CT” ให้ผู้ทดสอบทำแบบทดสอบวิธี CT ก่อน แล้วตามด้วยทดสอบวิธี MCAT แบบ LB ตัวอย่างบัตรทดสอบดังภาพที่ 3-30

ภาพที่ 3-30 บัตรทดสอบ ICT-MCAT แบบ LB กับ CT

3. ดำเนินการทดสอบด้วยโปรแกรม MCAT ที่พัฒนาขึ้นกับโปรแกรม CT ซึ่งผู้ทดสอบดำเนินการทดสอบแบบออนไลน์ แสดงดังตารางที่ 3-15

ตารางที่ 3-15 ลำดับการทดสอบโปรแกรม MCAT กับโปรแกรม CT

	สอบช่วงที่ 1	พัก	สอบช่วงที่ 2
สุ่มกลุ่มทดสอบ “CT”	แบบ CT	30 นาที	แบบ MCAT
สุ่มกลุ่มทดสอบ “LB”	แบบ MCAT		แบบ CT

4. การวิเคราะห์ข้อมูล

การศึกษาความสัมพันธ์ระหว่างค่าความสามารถของผู้ทดสอบที่ได้จาก MCAT กับค่าคะแนนรวมของผู้ทดสอบที่ได้จาก CT ใช้สถิติเชิงบรรยาย ค่าเฉลี่ย (Mean) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ทั้งนี้ค่าประมาณความสามารถของผู้ทดสอบและคะแนนรวมของผู้ทดสอบเป็นข้อมูลแบบต่อเนื่อง จึงใช้การวิเคราะห์สหสัมพันธ์แบบเพียร์สัน (Pearson’s product moment correlation)

บทที่ 4
ผลการวิจัย

การวิจัยนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง (LB) 2) พัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา 2 วิธี คือ 2.1) วิธีโหลดบาลานซ์ซิ่ง (LB) และ 2.2) วิธี Item Eligibility (IE) รูปแบบเว็บแอพพลิเคชั่น 3) เปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา 2 วิธี ในด้าน 3.1) จำนวนข้อสอบที่ใช้ในการทดสอบ 3.2) เวลาที่ใช้ในการทดสอบ 3.3) อัตราการแสดงการใช้ข้อสอบสูงสุด 3.4) สัดส่วนข้อสอบที่มีการแสดงมากเกินไป และ 3.5) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และ 4) ศึกษาความสัมพันธ์ระหว่างค่าความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT) โดยใช้ข้อสอบ V-NET วิชาเทคโนโลยีสารสนเทศและการสื่อสาร ที่ได้มาจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ระหว่างปีการศึกษา 2555-2559 จำนวน 192 ข้อ ที่ผ่านการวิเคราะห์คุณภาพข้อสอบ ผลการวิจัยแบ่งเป็น 4 ตอน ดังนี้

ตอนที่ 1 ผลการพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง (LB)

ตอนที่ 2 ผลการพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

ตอนที่ 3 ผลการเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

ตอนที่ 4 ผลการศึกษาความสัมพันธ์ระหว่างค่าความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT)

สัญลักษณ์ที่ใช้ในการเสนอผลการวิเคราะห์ข้อมูล

ในการนำเสนอผลการวิเคราะห์ข้อมูลในการวิจัยนี้ ได้กำหนดสัญลักษณ์ที่ใช้แทนค่าตัวแปรและค่าสถิติต่าง ๆ ไว้ดังนี้

n	แทน	ขนาดกลุ่มตัวอย่าง
Mean	แทน	ค่าเฉลี่ยเลขคณิต
p-value	แทน	ระดับนัยสำคัญ
SD	แทน	ส่วนเบี่ยงเบนมาตรฐาน
Mdn	แทน	ค่ามัธยฐาน
U	แทน	สถิตินอนพาราเมตริก Mann-Whitney test
Mean Rank	แทน	ค่าเฉลี่ยของอันดับในแต่ละวิธี
Sum of Ranks	แทน	ค่าผลรวมของอันดับในแต่ละวิธี
Min	แทน	ค่าต่ำสุด
Max	แทน	ค่าสูงสุด
t	แทน	สถิติทดสอบ *t*-test
MIER	แทน	อัตราการใช้ข้อสอบสูงสุด
χ^2	แทน	ค่าไค-สแควร์
ko	แทน	อัตราการใช้ข้อสอบ
OEMP	แทน	สัดส่วนข้อสอบที่มีการแสดงมากเกินไป
RMSE	แทน	ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย
Bias	แทน	ค่าความลำเอียง
Total *RMSE*	แทน	ผลรวมค่าความคลาดเคลื่อนกำลังสองเฉลี่ย
Total Bias	แทน	ผลรวมค่าความลำเอียง
*	แทน	มีนัยสำคัญทางสถิติที่ระดับ .05
**	แทน	มีนัยสำคัญทางสถิติที่ระดับ .01
θ	แทน	ค่าความสามารถที่แท้จริงของผู้ทดสอบ
$\hat{\theta}$	แทน	ค่าประมาณความสามารถของผู้ทดสอบ
LB	แทน	วิธีโหลดบาลานซ์ซิ่ง
IE	แทน	วิธี Item Eligibility
MCAT	แทน	การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์
CT	แทน	การทดสอบด้วยคอมพิวเตอร์

ตอนที่ 1 ผลการพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา โดยใช้วิธีโหลดบาลานซ์ซิ่ง (LB)

ผลการพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา โดยใช้วิธีโหลดบาลานซ์ซิ่ง มีรายละเอียดขั้นตอนของอัลกอริทึมที่ประยุกต์ 3 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 การสำรวจและเคลื่อนย้ายข้อมูล เป็นการสำรวจลักษณะของข้อสอบทุกข้อในคลังข้อสอบ ซึ่งแบ่งการทำงาน 2 ส่วน ดังนี้

1. การค้นหาข้อสอบ เป็นการค้นหาข้อสอบทั้งหมดในคลัง เพื่อหาข้อมูลลักษณะของข้อสอบ

2. การเก็บลักษณะข้อสอบ เป็นการเก็บข้อมูลลักษณะของข้อสอบที่ได้จากการค้นหาข้อสอบ ลงใน Local storage ของอัลกอริทึมโหลดบาลานซ์ซิ่ง

ขั้นตอนที่ 2 การคำนวณภาระงาน เป็นวิธีการควบคุมสมดุลเนื้อหาของข้อสอบ โดยการคำนวณค่าเฉลี่ยถ่วงน้ำหนักของสถานะการใช้ข้อสอบ ณ ขณะนั้นของผู้ทดสอบ ซึ่งแบ่งการทำงาน 2 ส่วน ดังนี้

1. การคำนวณค่าเฉลี่ยถ่วงน้ำหนัก ซึ่งมีขั้นตอนย่อย 3 ขั้นตอน คือ 1.1) การคำนวณจำนวนข้อสอบที่ไม่ถูกนำไปใช้ทดสอบ 1.2) การคำนวณเปอร์เซ็นต์การใช้ข้อสอบ และ 1.3) การคำนวณค่าเฉลี่ยถ่วงน้ำหนัก

2. การเก็บข้อมูลค่าเฉลี่ยถ่วงน้ำหนัก เป็นการนำผลจากการคำนวณมาเก็บในตาราง และปรับปรุงสถานะการใช้งาน แล้วเลือกเปอร์เซ็นต์สูงสุดของด้านให้เป็นตัวแทน

ขั้นตอนที่ 3 การควบคุมการทำงาน เป็นวิธีควบคุมการใช้ข้อสอบโดยนำด้านที่เป็นตัวแทนจากการคำนวณภาระงาน มาคำนวณเพื่อควบคุมการใช้ข้อสอบ แบ่งการทำงานเป็น 3 ส่วน ดังนี้

1. การหาจำนวนข้อสอบที่ยังไม่ถูกนำไปใช้ทดสอบในด้านที่เป็นตัวแทน เพื่อไม่ให้ข้อสอบที่ถูกนำไปใช้ทดสอบแล้ว นำกลับมาทดสอบซ้ำ

2. การหาค่าสารสนเทศสูงสุดในข้อสอบด้านที่เป็นตัวแทน เฉพาะข้อสอบที่ไม่ถูกนำไปใช้ทดสอบ ทั้งนี้เพื่อความรวดเร็วในการคำนวณ

3. การคำนวณผลต่างระหว่างค่าประมาณความสามารถของผู้ทดสอบกับค่าสารสนเทศสูงสุด แล้วเรียงลำดับจากน้อยไปหามากระหว่างค่าความถี่สะสมของข้อสอบกับผลต่างที่คำนวณได้ อัลกอริทึมคัดเลือกข้อสอบข้อแรกจากการเรียงลำดับเป็นข้อสอบข้อถัดไป

จากภาพที่ 4-11 เป็นหน้าจอแสดงการเลือกแบบทดสอบ โดยมี 3 แบบ ได้แก่ 1) แบบ LB เป็นวิธีการใหม่ที่ผู้วิจัยได้พัฒนาขึ้น 2) แบบ IE เป็นวิธีการเดิม และ 3) แบบ CT เป็นการทดสอบด้วยคอมพิวเตอร์ ซึ่งแบบที่ 1 และ 2 เป็นการทดสอบปรับเหมาะแบบพหุมิติด้วยยคอมพิวเตอร์

3. รายงานผลการทดสอบเมื่อทดสอบเสร็จ

การรายงานผลการทดสอบของผู้ทดสอบรายบุคคล แบ่งเกณฑ์การประเมินออกเป็น 2 ส่วน 1) เกณฑ์การประเมินระดับค่าความสามารถของผู้ทดสอบ และ 2) เกณฑ์การประเมินระดับคะแนนของผู้ทดสอบ แสดงดังตารางที่ 4-1 และ ตารางที่ 4-2

ตารางที่ 4-1 เกณฑ์การประเมินระดับค่าความสามารถของผู้ทดสอบ วิธี LB และวิธี IE

ค่าประมาณความสามารถของผู้สอบ	ระดับความสามารถ
มากกว่า 2.000	สูงมาก
1.001 ถึง 2.000	สูง
0.501 ถึง 1.000	ค่อนข้างสูง
-0.499 ถึง 0.500	ปานกลาง
-1.499 ถึง -0.500	ค่อนข้างต่ำ
-2.000 ถึง -1.500	ต่ำ
ต่ำกว่า -2.000	ต่ำมาก

ตารางที่ 4-2 เกณฑ์การประเมินระดับคะแนนของผู้ทดสอบ วิธี CT

จำนวนข้อสอบที่ตอบถูก	ระดับความสามารถ
ตั้งแต่ 32 ข้อ ขึ้นไป	สูงมาก
28-31 ข้อ	สูง
24-27 ข้อ	ค่อนข้างสูง
20-23 ข้อ	ปานกลาง
16-19 ข้อ	ค่อนข้างต่ำ
12-15 ข้อ	ต่ำ
ต่ำกว่า 12 ข้อ	ต่ำมาก

ประเด็นที่ 2 ผลการประเมินประสิทธิภาพโปรแกรมปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ ผลจากการประเมินประสิทธิภาพการทดสอบโปรแกรม โดยใช้เครื่องมือ Black box testing เป็นแบบประเมินที่มีลักษณะมาตรประมาณค่า 5 ระดับ การประเมิณมี 2 ส่วน ดังนี้

ส่วนที่ 1 ประเมินโดยผู้เชี่ยวชาญจำนวน 5 คน เพื่อตรวจสอบคุณภาพของโปรแกรมให้เป็นไปตามกรอบการพัฒนาโปรแกรม โดยการทดสอบแบ่งออกเป็น 4 ด้าน ได้แก่ 1) ด้านตรงตามความต้องการของโปรแกรมมีจำนวนข้อสอบ 8 ข้อ 2) ด้านการทำงานของโปรแกรม มีจำนวนข้อสอบ 8 ข้อ 3) ด้านการใช้งานโปรแกรมมีจำนวนข้อสอบ 7 ข้อ และ 4) ด้านความปลอดภัยของโปรแกรมมีจำนวนข้อสอบ 3 ข้อ วันที่ 27-29 ตุลาคม พ.ศ. 2561 แสดงภาพรวมของผลการประเมินประสิทธิแสดงดังตารางที่ 4-3

ตารางที่ 4-3 ผลการประเมินประสิทธิภาพโปรแกรมในภาพรวม

ด้าน	*Mean*	*SD*	ระดับ
1. ตรงตามความต้องการของโปรแกรม	4.45	0.50	มาก
2. การทำงานของโปรแกรม	4.48	0.60	มาก
3. การใช้งานโปรแกรม	4.63	0.55	มากที่สุด
4. ความปลอดภัยของโปรแกรม	4.67	0.49	มากที่สุด
รวม	4.53	0.55	มากที่สุด

จากตารางที่ 4-3 ผลการประเมินประสิทธิภาพของโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์เพื่อทดสอบความรู้ทางไอซีที ปรากฏว่า ภาพรวมของโปรแกรมมีประสิทธิภาพในระดับมากที่สุด (*Mean* = 4.53, *SD* = 0.55) เมื่อพิจารณาเป็นรายด้าน ด้านความปลอดภัยของโปรแกรมอยู่ในระดับมากที่สุด (*Mean* = 4.67, *SD* = 0.49) รองลงมาด้านการใช้งานโปรแกรมอยู่ในระดับมากที่สุด (*Mean* = 4.63, *SD* = 0.55) และ ด้านการทำงานของโปรแกรมอยู่ในระดับมาก (*Mean* = 4.48, *SD* = 0.60) ตามลำดับ แสดงให้เห็นว่า โปรแกรมมีประสิทธิภาพมากที่สุดในการนำไปใช้ทดสอบกับนักศึกษา

ผลการประเมินประสิทธิภาพโปรแกรมรายด้านแสดงดังตารางที่ 4-4 ถึง 4-7

ตารางที่ 4-4 ผลการประเมินประสิทธิภาพด้านตรงตามความต้องการของโปรแกรม

รายการประเมิน	*Mean*	*SD*	ระดับ
1. การจัดการข้อมูลผู้ใช้ระบบและผู้ทดสอบ	4.40	0.55	มาก
2. การจัดการข้อมูลความรู้ทางไอซีที	4.80	0.45	มากที่สุด
3. การจัดการข้อมูลหมวดไอซีที	4.40	0.55	มาก
4. การสร้างข้อสอบ	4.20	0.45	มาก
5. การสร้างชุดข้อสอบ	4.40	0.55	มาก
6. การเลือกและทำแบบทดสอบ	4.80	0.45	มากที่สุด
7. การรายงานผลการทดสอบ	4.40	0.55	มาก
8. การแสดงประวัติผลการทดสอบ	4.20	0.45	มาก
รวม	4.45	0.46	มาก

จากตารางที่ 4-4 ปรากฏว่า การจัดการขอมูลความรู้ทางไอซีที และการเลือกและทำแบบทดสอบ อยู่ในระดับมากที่สุด ส่วนรายการอื่น ๆ อยู่ในระดับมาก

ตารางที่ 4-5 ผลการประเมินประสิทธิภาพด้านการทำงานของโปรแกรม

รายการประเมิน	*Mean*	*SD*	ระดับ
1. สามารถจัดการข้อมูลผู้ใช้ระบบและผู้ทดสอบได้	4.80	0.45	มากที่สุด
2. สามารถจัดการข้อมูลความรู้ทางไอซีทีได้	4.80	0.45	มากที่สุด
3. สามารถจัดการข้อมูลหมวดไอซีทีได้	4.20	0.45	มาก
4. สามารถสร้างข้อสอบได้	4.20	0.84	มาก
5. สามารถสร้างชุดข้อสอบได้	4.20	0.84	มาก
6. สามารถเลือกและทำแบบทดสอบได้	4.60	0.55	มากที่สุด
7. สามารถรายงานผลการทดสอบได้	4.40	0.55	มาก
8. สามารถแสดงประวัติผลการทดสอบได้	4.60	0.55	มากที่สุด
รวม	4.48	0.60	มาก

จากตารางที่ 4-5 ปรากฏว่า การทำงานของโปรแกรมสามารถจัดการข้อมูลผู้ใช้ระบบและผู้ทดสอบได้ สามารถจัดการข้อมูลความรู้ทางไอซีทีได้ สามารถเลือกและทำแบบทดสอบได้ และสามารถแสดงประวัติผลการทดสอบได้ อยู่ในระดับมากที่สุด ส่วนรายการอื่น ๆ อยู่ในระดับมาก

ตารางที่ 4-6 ผลการประเมินประสิทธิภาพด้านการใช้งานโปรแกรม

รายการประเมิน	*Mean*	*SD*	ระดับ
1. ความง่ายต่อการใช้งานของโปรแกรม	4.80	0.45	มากที่สุด
2. ความเหมาะสมในการเลือกใช้ขนาดของตัวอักษร	4.80	0.45	มากที่สุด
3. ความเหมาะสมในการใช้สีของตัวอักษรและภาพ	4.60	0.55	มากที่สุด
4. ความเหมาะสมในการใช้ข้อความ สัญลักษณ์หรือภาพ	4.20	0.84	มาก
5. ความเป็นมาตรฐานเดียวกันในการออกแบบหน้าจอภาพ	4.80	0.45	มากที่สุด
6. ความเหมาะสมในการปฏิสัมพันธ์โต้ตอบกับผู้ใช้	4.80	0.45	มากที่สุด
7. ผู้ใช้คุ้นเคยกับคำศัพท์และสามารถปฏิบัติตามได้โดยง่าย	4.40	0.55	มาก
รวม	4.63	0.55	มากที่สุด

จากตารางที่ 4-6 ปรากฏว่า การใช้งานโปรแกรมมีความง่ายต่อการใช้งานของโปรแกรม มีความเหมาะสมในการเลือกใช้ขนาดของตัวอักษร มีความเหมาะสมในการใช้สีของตัวอักษรและภาพ มีความเป็นมาตรฐานเดียวกันในการออกแบบหน้าจอภาพ และมีความเหมาะสมในการปฏิสัมพันธ์โต้ตอบกับผู้ใช้ อยู่ในระดับมากที่สุด ส่วนรายการอื่น ๆ อยู่ในระดับมาก

ตารางที่ 4-7 ผลการประเมินประสิทธิภาพด้านความปลอดภัยของโปรแกรม

รายการประเมิน	*Mean*	*SD*	ระดับ
1. การกำหนดรหัสผู้ใช้และรหัสผ่านในการตรวจสอบผู้เข้าใช้โปรแกรม	4.60	0.55	มากที่สุด
2. การตรวจสอบสิทธิ์ก่อนการใช้งานของผู้ใช้โปรแกรมในระดับต่าง ๆ เช่น สิทธิ์ผู้ดูแลโปรแกรม สิทธิ์ผู้ทดสอบ	4.60	0.55	มากที่สุด
3. การควบคุมให้ใช้งานตามสิทธิ์ผู้ใช้ได้อย่างถูกต้อง เช่น การใช้งานของผู้ดูแลโปรแกรม การใช้งานของผู้ทดสอบ	4.80	0.45	มากที่สุด
รวม	4.67	0.49	มากที่สุด

จากตารางที่ 4-7 ปรากฏว่า การกำหนดรหัสผู้ใช้และรหัสผ่านในการตรวจสอบผู้เข้าใช้โปรแกรม การตรวจสอบสิทธิ์ก่อนการใช้งานของผู้ใช้โปรแกรมในระดับต่าง ๆ เช่น สิทธิ์ผู้ดูแลโปรแกรม สิทธิ์ผู้ทดสอบ และการควบคุมให้ใช้งานตามสิทธิ์ผู้ใช้ได้อย่างถูกต้อง เช่น การใช้งานของผู้ดูแลโปรแกรม การใช้งานของผู้ทดสอบ อยู่ในระดับมากที่สุด

นอกจากนี้ ผู้เชี่ยวชาญมีข้อเสนอแนะเพิ่มเติมดังนี้

1. ควรแสดงผลการทดสอบเป็นภาพรวมของแต่ละแบบในหน้าแรก

2. การใช้สีข้อความควรเป็นสีดำ และพื้นขาวจะได้อ่านสบายตา

ส่วนที่ 2 ประเมินประสิทธิภาพของโปรแกรมโดยผู้ทดสอบจำนวน 400 คน เป็นการประเมินความคิดเห็นในด้านการใช้งานโปรแกรม ผลการประเมินแสดงดังตารางที่ 4-8

ตารางที่ 4-8 ผลการประเมินประสิทธิภาพด้านการใช้งานโปรแกรมของผู้ทดสอบ

รายการประเมิน	*Mean*	*SD*	ระดับ
1. การเข้าถึงโปรแกรมทำได้ง่ายและสะดวก	4.18	0.78	มาก
2. โปรแกรมมีเมนูที่เข้าถึงส่วนต่าง ๆ ได้ง่ายและสะดวก	4.10	0.72	มาก
3. การสอบโดยใช้โปรแกรมมีความสะดวกกว่าเมื่อเทียบกับการสอบบนกระดาษ	4.13	0.75	มาก
4. เมื่อทดสอบเสร็จโปรแกรมมีการแสดงผลการสอบทันที	4.23	0.78	มาก
5. การแสดงผลมีรายละเอียดครบถ้วน	4.02	0.79	มาก
6. การออกแบบโปรแกรมมีความเหมาะสมกับการทดสอบ	4.07	0.77	มาก
7. การจัดรูปแบบหน้าจอง่ายต่อการใช้งาน	4.13	0.76	มาก
8. การแสดงผลข้อมูลเป็นไปอย่างรวดเร็ว	4.25	0.78	มาก
9. โปรแกรมมีการป้องกันการทำงานผิดพลาดของผู้ใช้งาน	4.11	0.80	มาก
รวม	4.13	0.77	มาก

จากตารางที่ 4-8 ผลการประเมินประสิทธิภาพด้านการใช้งานโปรแกรมของผู้ทดสอบปรากฏว่า โปรแกรมมีประสิทธิภาพด้านการใช้งานอยู่ในระดับมาก (*Mean* = 4.13, *SD* = 0.77)

นอกจากนี้ ผู้ทดสอบมีข้อเสนอแนะเพิ่มเติมดังนี้

1. ปุ่มกดถัดไปเล็กเมื่อทดสอบผ่านสมาร์ทโฟน

2. การแสดงผลข้อความเล็กมากเมื่อทดสอบผ่านสมาร์ทโฟน

ตอนที่ 3 ผลการเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

ในการเปรียบเทียบประสิทธิภาพวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา กลุ่มตัวอย่างเป็นนักศึกษาจากวิทยาลัยเทคโนโลยีภาคตะวันออก (อี.เทค) จำนวน 400 คน คลังข้อสอบความรู้ทางไอซีที ประกอบด้วย 2 ด้าน (มิติ) ได้แก่ 1) หมวดเนื้อหา และ 2) ความรู้ทางไอซีที มีจำนวน 192 ข้อ กำหนดความยาวข้อสอบ 30 ข้อ ส่วนการเปรียบเทียบประสิทธิภาพมี 2 วิธี ได้แก่ 1) วิธี LB เป็นวิธีผู้วิจัยพัฒนาขึ้น และ 2) วิธี IE เป็นวิธีเดิม เปรียบเทียบประสิทธิภาพ 4 ด้าน ได้แก่ 1) จำนวนข้อสอบที่ใช้ในการทดสอบ 2) เวลาที่ใช้ในการทดสอบ 3) อัตราการแสดงการใช้ข้อสอบสูงสุด และ 4) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย

1. การเปรียบเทียบประสิทธิภาพวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาด้านจำนวนข้อสอบที่ใช้ในการทดสอบ แสดงค่าสถิติพื้นฐานดังตารางที่ 4-9 และผลการเปรียบเทียบประสิทธิภาพด้านจำนวนขอสอบที่ใช้ในการทดสอบ แสดงดังตารางที่ 4-10

ตารางที่ 4-9 ค่าสถิติพื้นฐานด้านจำนวนข้อสอบที่ใช้ในการทดสอบ

วิธีการควบคุมการใช้ข้อสอบ	*n*	Min	Max	*Mdn*
LB	400	10	15	13
IE	400	12	16	14

จากตารางที่ 4-9 วิธีการควบคุมการใช้ข้อสอบ LB มีมัธยฐานเท่ากับ 13 ส่วนวิธีการควบคุมการใช้ข้อสอบ IE มีมัธยฐานเท่ากับ 14

ตารางที่ 4-10 ผลการเปรียบเทียบประสิทธิภาพด้านจำนวนข้อสอบที่ใช้ในการทดสอบ

วิธี	*Mdn*	*Mean* Rank	Sum of Ranks	*U*	*p*-value
LB	13	329.70	131879.00	51679.00**	< .01
IE	14	471.30	188521.00		

จากตารางที่ 4-10 ปรากฏว่า วิธี LB (*Mdn* = 13) มีจำนวนข้อสอบที่ใช้ในการทดสอบน้อยกว่า วิธี IE (*Mdn* = 14) ที่ระดับนัยสำคัญทางสถิติ .01 ซึ่งเป็นไปตามสมมติฐานการวิจัยข้อที่ 1 แสดง

ให้เห็นว่า การใช้วิธี LB มีประสิทธิภาพต่อการลดจำนวนข้อสอบที่ใช้ในการทดสอบ เมื่อมีวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

2. การเปรียบเทียบประสิทธิภาพวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาด้านเวลาที่ใช้ในการทดสอบ ค่าสถิติพื้นฐานแสดงดังตารางที่ 4-11 และการเปรียบเทียบประสิทธิภาพด้านเวลาที่ใช้ในการทดสอบแสดงดังตารางที่ 4-12

ตารางที่ 4-11 ค่าสถิติพื้นฐานด้านเวลาที่ใช้ในการทดสอบ

วิธีการควบคุมการใช้ข้อสอบ	*n*	*Mean* (วินาที)	*SD*
LB	400	243.74	88.52
IE	400	257.18	82.67

จากตารางที่ 4-11 วิธีการควบคุมการใช้ข้อสอบ LB มีค่าเฉลี่ยเท่ากับ 243.74 วินาที ส่วนวิธีการควบคุมการใช้ข้อสอบ IE มีค่าเฉลี่ยเท่ากับ 257.18 วินาที

ตารางที่ 4-12 ผลการเปรียบเทียบประสิทธิภาพด้านเวลาที่ใช้ในการทดสอบ

วิธีการควบคุมการใช้ข้อสอบ	*Mean* (วินาที)	*SD*	*t*	*p*-value
LB	243.74	88.52	2.22*	< .05
IE	257.18	82.67		

จากตารางที่ 4-12 ปรากฏว่า ค่าสถิติทดสอบ t = 2.22 แสดงว่า วิธี LB (*Mean* = 243.74, *SD* = 88.52) ใช้เวลาในการทดสอบต่ำสุดเมื่อเทียบกับ วิธี IE (*Mean* = 257.18, *SD* = 82.67) ที่ระดับนัยสำคัญทางสถิติ .05 ซึ่งเป็นไปตามสมมติฐานการวิจัยข้อที่ 2 แสดงให้เห็นว่า การใช้วิธี LB มีประสิทธิภาพต่อการลดเวลาที่ใช้ในการทดสอบ เมื่อมีวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

3. การเปรียบเทียบประสิทธิภาพวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา ด้านอัตราการแสดงการใช้ข้อสอบสูงสุด แสดงดังตารางที่ 4-13

ตารางที่ 4-13 ผลอัตราการแสดงการใช้ข้อสอบในคลังข้อสอบความรู้ทางไอซีที

อัตราการแสดงการใช้ข้อสอบ	วิธีการ	
	LB	IE
$0.4 < ko \leq 0.5$	0	0
$0.3 < ko \leq 0.4$	0	0
$0.2 < ko \leq 0.3$	0	0
$0.1 < ko \leq 0.2$	3	45
$0 < ko \leq 0.1$	180	138
$ko = 0$	9	9
MIER	0.12	0.19
% (ko=0)	0.02	0.02
χ^2	10.16	10.36

จากตารางที่ 4-13 ปรากฏว่า อัตราการใช้ข้อสอบสูงสุดของวิธี LB (MIER = 0.12) น้อยกว่า วิธี IE (MIER = 0.19) ซึ่งเป็นไปตามสมมติฐานการวิจัยข้อที่ 3 ส่วนเปอร์เซ็นต์ข้อสอบที่ไม่ถูกนำไปใช้ระหว่าง วิธี LB และวิธี IE มีจำนวนเท่ากัน (%(ko=0) = 0.02)

4. การเปรียบเทียบประสิทธิภาพวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา ด้านค่าความคลาดเคลื่อนกำลังสองเฉลี่ย แสดงดังตารางที่ 4-14

ตารางที่ 4-14 ผลการวิเคราะห์ความแตกต่างของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) จำแนกตามวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

ค่าสถิติ	วิธีการ	
	LB	IE
RMSE	1.07, 1.11, 1.58, 0.94, 1.88	1.15, 1.12, 1.57, 0.95, 1.71
Bias	0.19, 0.13, 0.59, 0.14, 0.79	0.32, 0.15, 0.57, 0.10, 0.68
Total *RMSE*	6.57	6.48
Total Bias	1.85	1.81

จากตารางที่ 4-14 ผลการวิเคราะห์ความแตกต่างของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย จำแนกตามวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา ปรากฏว่า วิธี IE มีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุด (Total *RMSE* = 6.48, Total Bias = 1.81) ส่วน วิธี LB มีค่า

ความคลาดเคลื่อนกำลังสองเฉลี่ย (Total *RMSE* = 6.57, Total Bias = 1.85) ซึ่งไม่เป็นไปตามสมมติฐานการวิจัยข้อที่ 4 เพราะวิธี IE มีความแม่นยำในการควบคุมการคัดเลือกข้อสอบที่มีค่าความสามารถใกล้เคียงกับผู้ทดสอบมากกว่าวิธี LB ผลจากวิธี LB ควบคุมการใช้ข้อสอบไม่เป็นไปตามสมมติฐาน ส่วนหนึ่งเป็นเพราะจำนวนข้อสอบในคลังมีปริมาณน้อย ประกอบกับมีการควบคุมสมดุลเนื้อหาร่วมด้วย จึงทำให้การกระจายตัวของการประมาณค่าความสามารถไม่ครอบคลุมการควบคุมการใช้ข้อสอบ

จากผลการเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาตามสมติฐานการวิจัยข้อที่ 1 ถึง 4 ระหว่างวิธี LB กับ วิธี IE ดังตารางที่ 4-15

ตารางที่ 4-15 สรุปผลการเปรียบเทียบประสิทธิภาพระหว่างวิธี LB กับวิธี IE

สมมติฐานการวิจัย	สรุปผลการวิจัย
1. จำนวนข้อสอบที่ใช้ในการทดสอบของวิธี LB น้อยกว่า วิธี IE	/
2. เวลาที่ใช้ในการทดสอบของวิธี LB น้อยกว่า วิธี IE	/
3. อัตราการแสดงการใช้ข้อสอบสูงสุดของวิธี LB น้อยกว่า วิธี IE	/
4. ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธี LB น้อยกว่า วิธี IE	×

หมายเหตุ: / หมายถึง สอดคล้องกับสมติฐานการวิจัย

× หมายถึง ไม่สอดคล้องกับสมมติฐานการวิจัย

จากตารางที่ 4-15 ผลการศึกษาแสดงให้เห็นว่า วิธี LB มีประสิทธิภาพสูงในด้านจำนวนข้อสอบที่ใช้ในการทดสอบ ด้านเวลาที่ใช้ในการทดสอบ ด้านอัตราการแสดงการใช้ข้อสอบสูงสุด แม้ว่าประสิทธิภาพวิธี LB ในด้านค่าความคลาดเคลื่อนกำลังสองเฉลี่ยหรือความแม่นยำในการควบคุมการใช้ข้อสอบจะต่ำกว่าวิธี IE แต่วิธี LB เป็นอีกทางเลือกหนึ่งสำหรับการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาของการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์

ตอนที่ 4 ผลการศึกษาความสัมพันธ์ระหว่างค่าประมาณความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT)

ค่าประมาณความสามารถของผู้สอบที่ได้จาก MCAT กับค่าคะแนนรวมที่ได้จาก CT มีความสัมพันธ์กันทางบวก ดังตารางที่ 4-16 จำนวนข้อสอบที่ใช้ในการทดสอบ เวลาที่ใช้ในการทดสอบ และผลการทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์กับการทดสอบด้วยคอมพิวเตอร์

ตารางที่ 4-16 ผลการทดสอบ MCAT กับ CT

คนที่	จำนวนข้อสอบที่ใช้ (ข้อ)		เวลาที่ใช้ในการทดสอบ (วินาที)		ผลการทดสอบ	
	MCAT	CT	MCAT	CT	MCAT ($\hat{\theta}$)	CT (คะแนนรวม)
Mean	14.68	40	258.87	1,174.95	0.04	20.58
SD	1.54	0	128.55	422.55	0.41	4.37

หมายเหตุ: MCAT หมายถึง การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์
CT หมายถึง การทดสอบด้วยคอมพิวเตอร์

จากตารางที่ 4-16 แสดงจำนวนข้อสอบที่ใช้ในการทดสอบ เวลาที่ใช้ในการทดสอบ และผลการทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์กับการทดสอบด้วยคอมพิวเตอร์ ปรากฏว่า การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์มีจำนวนข้อสอบที่ใช้ทดสอบเฉลี่ย 14.68 ข้อ ในขณะที่การทดสอบด้วยคอมพิวเตอร์มีจำนวนข้อสอบ 40 ข้อ แสดงให้เห็นว่า การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ที่ใช้วิธีการควบคุมการใช้ข้อสอบและการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาล่านซ์ซิ่ง มีจำนวนข้อสอบคิดเป็นร้อยละ 36.70 ของการทดสอบด้วยคอมพิวเตอร์

การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ใช้เวลาในการทดสอบเฉลี่ย 258.87 วินาที ขณะที่การทดสอบด้วยคอมพิวเตอร์ใช้เวลาในการทดสอบเฉลี่ย 1,174.95 วินาที แสดงให้เห็นว่า การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ที่ใช้วิธีการควบคุมการใช้ข้อสอบและการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาล่านซ์ซิ่ง ใช้เวลาในการทดสอบคิดเป็นร้อยละ 22.03 ของการทดสอบด้วยคอมพิวเตอร์

ตารางที่ 4-17 ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันระหว่างค่าประมาณความสามารถของผู้ทดสอบกับคะแนนรวมของผู้ทดสอบ

	MCAT	CT	*p*-value
MCAT	1	.59**	< .01
CT		1	

จากตารางที่ 4-17 แสดงค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันระหว่างค่าประมาณความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์กับคะแนนรวมของผู้ทดสอบจากการทดสอบคอมพิวเตอร์ ปรากฏว่า ค่าประมาณความสามารถของผู้ทดสอบกับคะแนนรวมของผู้ทดสอบ มีความสัมพันธ์กันทางบวก ที่ระดับนัยสำคัญทางสถิติ .01 แสดงว่า ผลการทดสอบแบบ MCAT มีความสอดคล้องกับผลการทดสอบแบบ CT นั่นคือ ผู้ทดสอบที่ได้รับค่าประมาณความสามารถจากการทดสอบแบบ MCAT สูงหรือต่ำ จะได้คะแนนรวมจากการทดสอบ CT สูงหรือต่ำด้วย ซึ่งเป็นไปตามสมมติฐานการวิจัยข้อ 5

บทที่ 5
สรุปผลและอภิปรายผล

การควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา เป็นวิธีการเพิ่มประสิทธิภาพในการคัดเลือกข้อสอบข้อถัดไปและเลือกใช้ข้อสอบครอบคลุมเนื้อหาทุกด้าน การวิจัยนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง (LB) 2) พัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา สำหรับทดสอบความรู้ทางไอซีที (ICT-MCAT) มี 2 วิธี คือ 2.1) วิธีโหลดบาลานซ์ซิ่ง (LB) และ 2.2) วิธี Item Eligibility (IE) ในรูปแบบเว็บแอพพลิเคชั่น 3) เปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา 2 วิธี ในด้าน 3.1) จำนวนข้อสอบที่ใช้ในการทดสอบ 3.2) เวลาที่ใช้ในการทดสอบ 3.3) อัตราการแสดงการใช้ข้อสอบสูงสุด และ 3.4) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ใช้กลุ่มตัวอย่างเป็นนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง ปีการศึกษา 2561 จำนวน 400 คน ใช้การสุ่มอย่างเป็นระบบ วิเคราะห์ข้อมูลด้วยค่าความคลาดเคลื่อนกำลังสองเฉลี่ย การทดสอบแมน-วิทนีย์ และการทดสอบค่าที และ 4) ศึกษาความสัมพันธ์ระหว่างค่าความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT) โดยใช้กลุ่มตัวอย่าง จำนวน 40 คน ทดสอบสหสัมพันธ์แบบเพียร์สัน

สรุปผลการวิจัย

1. ผลการพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง (LB) ประกอบด้วย 3 ขั้นตอนหลัก ดังนี้

ขั้นตอนที่ 1 การสำรวจและเคลื่อนย้ายข้อมูล เป็นการสำรวจลักษณะของข้อสอบทุกข้อในคลังข้อสอบ ซึ่งแบ่งการทำงาน 2 ส่วน ได้แก่ 1) การค้นหาข้อสอบ เป็นการค้นหาข้อสอบทั้งหมดในคลัง เพื่อหาข้อมูลลักษณะของข้อสอบ และ 2) การเก็บลักษณะข้อสอบ เป็นการเก็บข้อมูลลักษณะของข้อสอบที่ได้จากการค้นหาข้อสอบ ลงใน Local storage ของอัลกอริทึมโหลดบาลานซ์ซิ่ง

ขั้นตอนที่ 2 การคำนวณภาระงาน เป็นวิธีการควบคุมสมดุลเนื้อหาของข้อสอบ โดยการคำนวณค่าเฉลี่ยถ่วงน้ำหนักของสถานะการใช้ข้อสอบ ณ ขณะนั้นของผู้ทดสอบ ซึ่งแบ่งการทำงาน 2 ส่วน ได้แก่ 1) การคำนวณค่าเฉลี่ยถ่วงน้ำหนัก ซึ่งมีขั้นตอนย่อย 3 ขั้นตอน คือ 1.1) การคำนวณจำนวนข้อสอบที่ไม่ถูกนำไปใช้ทดสอบ 1.2) การคำนวณเปอร์เซ็นต์การใช้ข้อสอบ และ 1.3) การ

คำนวณค่าเฉลี่ยถ่วงน้ำหนัก และ 2) การเก็บข้อมูลค่าเฉลี่ยถ่วงน้ำหนัก เป็นการนำผลจากการคำนวณมาเก็บในตาราง และปรับปรุงสถานะการใช้งาน แล้วเลือกเปอร์เซ็นต์สูงสุดของด้านให้เป็นตัวแทน

ขั้นตอนที่ 3 การควบคุมการทำงาน เป็นวิธีควบคุมการใช้ข้อสอบโดยนำด้านที่เป็นตัวแทนจากการคำนวณภาระงาน มาคำนวณเพื่อควบคุมการใช้ข้อสอบ แบ่งการทำงานเป็น 3 ส่วน ได้แก่ 1) การหาจำนวนข้อสอบที่ยังไม่ถูกนำไปใช้ทดสอบในด้านที่เป็นตัวแทน เพื่อไม่ให้ข้อสอบที่ถูกนำไปใช้ทดสอบแล้ว นำกลับมาทดสอบซ้ำ 2) การหาค่าสารสนเทศสูงสุดในข้อสอบด้านที่เป็นตัวแทน เฉพาะข้อสอบที่ไม่ถูกนำไปใช้ทดสอบ ทั้งนี้เพื่อความรวดเร็วในการคำนวณ และ 3) การคำนวณผลต่างระหว่างค่าประมาณความสามารถของผู้ทดสอบกับค่าสารสนเทศสูงสุด แล้วทำการเรียงลำดับจากน้อยไปหามากระหว่างค่าความถี่สะสมของข้อสอบกับผลต่างที่คำนวณได้ อัลกอริทึมคัดเลือกข้อสอบข้อแรกจากการเรียงลำดับเป็นข้อสอบข้อถัดไป

2. ผลการพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา สรุปผลการพัฒนาโปรแกรม ดังนี้

2.1 ผลการพัฒนาโปรแกรมปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ แบ่งออกเป็น 2 ส่วน คือ ส่วนที่ 1 ผู้ดูแลโปรแกรม มีรูปแบบการทำงานหลัก 8 หน้าจอ ได้แก่ 1) การเข้าสู่โปรแกรมผู้ดูแลโปรแกรม 2) หน้าแรกหลังจากเข้าสู่โปรแกรมของผู้ดูแลโปรแกรม 3) การจัดการข้อมูลผู้ใช้โปรแกรม 4) การจัดการข้อมูลแบบทดสอบ 5) การจัดการข้อมูลหมวดไอซีที 6) การจัดการข้อมูลความรู้ทางไอซีที 7) การจัดการข้อมูลคำถาม และ 8) รายงาน ส่วนที่ 2 ผู้ทดสอบ มีรูปแบบการทำงานหลัก 4 หน้าจอ ได้แก่ 1) การจัดการข้อมูลลงทะเบียนผู้ทดสอบ 2) การเข้าสู่โปรแกรม 3) การจัดการข้อมูลการทดสอบ และ 4) รายงานผลการทดสอบแต่ละแบบ

2.2 ผลการประเมินประสิทธิภาพโปรแกรมปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ แบ่งการประเมินออกเป็น 2 ส่วน คือ ส่วนที่ 1 ประเมินโดยผู้เชี่ยวชาญจำนวน 5 ท่าน ปรากฏว่าโปรแกรมมีประสิทธิภาพโดยรวมในระดับมากที่สุด เมื่อพิจารณาเป็นรายด้าน ด้านความปลอดภัยของโปรแกรมมีประสิทธิภาพในระดับมากที่สุด รองลงมาด้านการใช้งานโปรแกรมมีประสิทธิภาพในระดับมากที่สุด และด้านการทำงานของโปรแกรมมีประสิทธิภาพในระดับมาก ตามลำดับ แสดงให้เห็นว่าโปรแกรมมีประสิทธิมากที่สุดในการนำไปใช้ทดสอบกับนักศึกษา และส่วนที่ 2 ประเมินประสิทธิภาพของโปรแกรมโดยผู้ทดสอบจำนวน 400 คน ปรากฏว่า โปรแกรมมีประสิทธิภาพด้านการใช้งานอยู่ในระดับมาก

3. ผลการเปรียบเทียบประสิทธิภาพของวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา จากกลุ่มตัวอย่าง 400 คน สรุปผลการเปรียบเทียบ ดังนี้

3.1 ผลการเปรียบเทียบประสิทธิภาพด้านจำนวนข้อสอบที่ใช้ในการทดสอบ ปรากฏว่า

วิธี LB มีจำนวนขอสอบที่ใช้ในการทดสอบน้อยกว่า วิธี IE แสดงให้เห็นว่า วิธี LB สามารถลดจำนวนข้อสอบที่ใช้ในการทดสอบ ที่ระดับนัยสำคัญทางสถิติ .01 เป็นไปตามสมมติฐานการวิจัยข้อที่ 1

3.2 ผลการเปรียบเทียบประสิทธิภาพด้านเวลาที่ใช้ในการทดสอบ ปรากฏว่า วิธี LB มีเวลาในการทำแบบทดสอบน้อยกว่า วิธี IE แสดงให้เห็นว่า วิธี LB สามารถลดเวลาที่ใช้ในการทดสอบที่ระดับนัยสำคัญทางสถิติ .05 เป็นไปตามสมมติฐานการวิจัยข้อที่ 2

3.3 ผลการเปรียบเทียบประสิทธิภาพด้านอัตราการแสดงการใช้ข้อสอบสูงสุด ปรากฏว่า วิธี LB มีอัตราการแสดงการใช้ข้อสอบสูงสุดน้อยกว่า วิธี IE แสดงให้เห็นกว่า วิธี LB สามารถควบคุมการใช้ข้อสอบดีกว่า วิธี IE เป็นไปตามสมมติฐานการวิจัยข้อที่ 3

3.4 ผลการเปรียบเทียบความแตกต่างของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ปรากฏว่า วิธี LB มีค่าคลาดเคลื่อนกำลังสองเฉลี่ยมากกว่า วิธี IE แสดงให้เห็นว่า วิธี LB มีความแม่นยำในการประมาณความสามารถน้อยกว่า วิธี IE ไม่เป็นไปตามสมมติฐานการวิจัยข้อที่ 4

4. ผลการศึกษาความสัมพันธ์ค่าประมาณความสามารถที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับค่าคะแนนรวมที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT) จากกลุ่มตัวอย่าง 40 คน ปรากฏว่า ค่าประมาณความสามารถของผู้ทดสอบกับคะแนนรวมของผู้ทดสอบ มีความสัมพันธ์ทางบวก ที่ระดับนัยสำคัญทางสถิติ .01 เป็นไปตามสมมติฐานการวิจัยข้อ 6

อภิปรายผล

การอภิปรายผลการวิจัย แบ่งออกเป็น 4 ประเด็น ได้แก่ 1) การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ชิ่ง 2) การพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา 3) การเปรียบเทียบประสิทธิภาพวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา และ 4)การศึกษาความสัมพันธ์ของคะแนนที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์กับการทดสอบด้วยคอมพิวเตอร์ รายละเอียดมีดังนี้

1. การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ชิ่ง (LB) จำเป็นต้องเข้าใจถึงรูปแบบของข้อสอบที่บรรจุอยู่ในคลังข้อสอบ มีรูปแบบเป็นเอกมิติหรือพหุมิติ ทั้งนี้จึงสามารถเลือกใช้วิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาได้ถูกต้อง จากการศึกษาพบกว่า รูปแบบข้อสอบความรู้ทางไอซีที มีจำนวน 2 มิติ ได้แก่ 1) หมวดเนื้อหา มีจำนวน 5 หมวด และ 2) ความรู้ทางไอซีที มีจำนวน 6 ด้าน ทั้ง 2 มิติมีความสัมพันธ์กันในข้อสอบแต่ละข้อ การวิเคราะห์คุณภาพของข้อสอบความรู้ทางไอซีที จำเป็นต้องเลือกการวิเคราะห์แบบ 3 พารามิเตอร์ จึงใช้ MIRT เพราะเป็นการวิเคราะห์ในรูปแบบพหุมิติ

จากการศึกษาการควบคุมการใช้ข้อสอบ พัฒนาขึ้นมาเพื่อแก้ปัญหาการคัดเลือกข้อสอบข้อถัดไป เนื่องจากวิธีการคัดเลือกข้อสอบข้อถัดไปส่วนใหญ่จะคัดเลือกข้อสอบที่มีค่าอำนาจจำแนกของข้อสอบสูง มาเป็นข้อสอบข้อถัดไป ดังนั้น ข้อสอบเหล่านี้จึงถูกเลือกใช้ในการทดสอบบ่อยครั้ง เป็นผลให้ผู้ทดสอบในรุ่นถัดไป สามารถคาดการณ์ข้อสอบได้ล่วงหน้า จากงานวิจัยของ Georgiadou and Triantafillou (2007) ได้แบ่งวิธีการควบคุมการใช้ข้อสอบออกเป็น 5 กลุ่ม ได้แก่ 1) วิธีการคัดเลือกแบบสุ่ม 2) วิธีการคัดเลือกแบบมีเงื่อนไข 3) วิธีคัดเลือกแบบตามระดับชั้น 4) วิธีคัดเลือกแบบผสม และ 5) วิธีคัดเลือกแบบปรับเหมาะหลายขั้นตอน ส่วนการวิจัยนี้ ได้พัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับวิธีการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง อยู่ในกลุ่มวิธีการคัดเลือกแบบมีเงื่อนไข โดยวิธีการนี้แบ่งการทำงานออกเป็น 3 ขั้นตอน (Singh et al., 2015) ได้แก่ 1) การสำรวจและเคลื่อนย้ายข้อมูล เป็นการค้นหาข้อสอบที่อยู่ในคลังข้อสอบแล้วบันทึกลักษณะของข้อสอบเพื่อเตรียมการคำนวน 2) การคำนวณภาระงาน เป็นการควบคุมสมดุลเนื้อหา และ 3) การควบคุมการทำงาน เป็นการควบคุมการใช้ข้อสอบ

การคำนวณภาระงาน เป็นวิธีการควบคุมสมดุลเนื้อหา ช่วยควบคลุมขอบเขตเนื้อหาที่นำไปใช้ทดสอบเฉลี่ยเท่า ๆ กัน เพื่อป้องกันไม่ให้ผู้ทดสอบได้รับข้อสอบที่มีเนื้อหาเน้นหนักไปด้านใดด้านหนึ่งหรืออาจไม่ได้ทดสอบเนื้อหาบ้างด้าน ในช่วงเริ่มต้นพัฒนาควบคุมสมดุลเนื้อหา โดยการใช้เกณฑ์กลางของอัลกอริทึมโหลดบาล่านซ์ซิ่งที่กำหนดมา 25% จากการทดสอบ ปรากฏว่า เมื่อโปรแกรมยุติการทดสอบ เนื้อหาบางด้านไม่ถูกนำมาทดสอบ สาเหตุมาจากเกณฑ์ยุติเข้าเงื่อนไขของค่าความคลาดเคลื่อนมาตรฐาน (SE) ต่ำกว่า 0.3 จากการพิจารณาปัญหา ปรากฏว่า ข้อสอบในแต่ละด้านมีจำนวนไม่เท่ากัน จึงเป็นสาเหตุหนึ่งที่ทำให้เกิดปัญหาดังกล่าว ดังนั้นการแก้ไขปัญหานี้จึงอาศัยแนวคิดของ Bortot and Marques Pereira (2017) โดยนำค่าเฉลี่ยถ่วงน้ำหนักมาเป็นเกณฑ์กลางของการควบคุมสมดุลเนื้อหา เพราะค่าเฉลี่ยถ่วงน้ำหนักจะแปรผันตามการทดสอบทุกครั้งที่มีการเลือกข้อสอบข้อถัดไป ผลปรากฏว่า วิธีการใช้ค่าเฉลี่ยถ่วงน้ำหนักสามารถช่วยแก้ไขปัญหาข้อสอบของแต่ละด้านมีจำนวนไม่เท่ากันได้ และผู้ทดสอบได้รับข้อสอบครบทุกด้าน

การควบคุมการทำงาน เป็นวิธีการควบคุมการใช้ข้อสอบ เพื่อแก้ไขปัญหาข้อสอบที่ถูกใช้แล้วไม่นำกลับไปใช้ทดสอบซ้ำของผู้ทดสอบแต่ละคน ไม่ให้มีอัตราการแสดงข้อสอบมากเกินไปจนทำให้เกิดการคาดเดาข้อสอบในกลุ่มผู้ทดสอบ และการควบคุมข้อสอบที่ถูกนำไปใช้ทดสอบน้อยหรือไม่เคยถูกนำไปใช้ทดสอบให้มีโอกาสนำไปใช้ทดสอบ ในการวิจัยนี้ ได้พัฒนาวิธีการควบคุมการใช้ข้อสอบโดยอาศัยแนวคิดของ Huang and Jo (1995) ในการใช้ความถี่สะสมเป็นส่วนในการตัดสินใจ เป็นการเรียงลำดับจากน้อยไปหามาก ระหว่างผลต่างของค่าประมาณความสามารถกับค่าสารสนเทศสูงสุดและความถี่สะสมของการใช้ข้อสอบ อัลกอริทึมคัดเลือกข้อสอบลำดับแรกที่มีค่าน้อยสุดมาเป็นข้อสอบข้อถัดไป ผลปรากฏว่า วิธีการนี้สามารถช่วยแก้ไขปัญหาข้อสอบที่ถูกใช้แล้วไม่นำกลับไปใช้

ทดสอบซ้ำของผู้ทดสอบแต่ละคน แก้ไขปัญหาอัตราการแสดงข้อสอบมากเกินไป หรือเกิน 0.2 และสามารถควบคุมข้อสอบที่ถูกนำไปใช้ทดสอบน้อยหรือไม่เคยถูกนำไปใช้ทดสอบให้มีโอกาสนำไปใช้ทดสอบ (Georgiadou & Triantafillou, 2007)

ประเด็นหนึ่งที่สำคัญในการพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาล่านซ์ซิ่ง คือ จำนวนข้อสอบที่ถูกคัดเลือกเป็นข้อสอบข้อถัดไป มีผลทำให้ข้อสอบที่มีอัตราการใช้ข้อสอบสูงกว่า 0.2 มีแนวโน้มลดจำนวนลง (บ่งชี้ว่า จำนวนข้อสอบที่ถูกใช้ซ้ำบ่อยลดลง ซึ่งเป็นข้อดี) แต่ในขณะเดียวกัน ค่า RMSE มีแนวโน้มเพิ่มขึ้น (บ่งชี้ว่า ประสิทธิภาพด้านการประมาณค่าความสามารถของผู้ทดสอบลดลง ซึ่งเป็นข้อเสีย) ผลมาจากการควบคุมเนื้อหา เพื่อหาด้านที่เป็นตัวแทนในการทดสอบ เพราะข้อสอบในด้านที่เป็นตัวแทนอาจมีค่าความยากของข้อสอบต่างจากค่าประมาณความสามารถของผู้ทดสอบ ดังนั้น การแก้ไขปัญหานี้ควรให้แต่ละด้านมีจำนวนข้อสอบที่มีค่าความอยากกระจายครอบคลุมระดับค่าประมาณความสามารถ

2. การพัฒนาโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ในการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา มีการพัฒนา 2 วิธี ได้แก่ 1) วิธี IE และ 2) วิธี LB แบ่งขั้นตอนการพัฒนาโปรแกรม 6 ขั้นตอน ประกอบด้วย 1) การกำหนดความต้องการและการวิเคราะห์ 2) การออกแบบ 3) การพัฒนาโปรแกรม 4) การทดสอบ 5) การใช้งาน และ 6) การบำรุงรักษา ซึ่งสอดคล้องกับกรอบแนวคิดวงจรการพัฒนาโปรแกรมของ Deepti Singh et al. (2015) โดยวิธีการทดสอบปรับเหมาะแบบพหุมิติมี 5 ขั้นตอน (Thompson & Weiss, 2011) 1) การสร้างคลังข้อสอบมีจำนวนข้อสอบความรู้ทางไอซีที 6 ด้าน และ 5 หมวดเนื้อหา ที่ผ่านเกณฑ์จำนวน 192 ข้อ 2) คัดเลือกข้อสอบเริ่มต้น มีค่าความยาก (p) ระหว่าง -1 ถึง 1 3) การเลือกข้อสอบข้อถัดไป ใช้วิธีค่าสารสนเทศสูงสุด (MIC) (Reckase, 2009, p. 52) 4) การประมาณค่าความสามารถใช้เทคนิคการประมาณค่าแบบ (ML) (Reckase, 2009, p. 140) และ 5) เกณฑ์ยุติการทดสอบ โดยใช้ผลความคลาดเคลื่อน (SEE) มีค่าน้อยกว่า 0.30 เป็นเกณฑ์ในการยุติการทดสอบ หรือยุติการทดสอบเมื่อข้อสอบหมดคลังสอดคล้องกับงานวิจัยของ Thompson and Weiss (2011)

จากการพัฒนาโปรแกรมปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ แบ่งออกเป็น 2 ส่วน คือ ส่วนที่ 1 ผู้ดูแลโปรแกรม มีรูปแบบการทำงานหลัก 8 หน้าจอ ส่วนที่ 2 ผู้ทดสอบ มีรูปแบบการทำงานหลัก 4 หน้าจอ ผลการประเมินประสิทธิภาพของโปรแกรมฯ จากผู้เชี่ยวชาญ ปรากฏว่าประสิทธิภาพของโปรแกรมในภาพรวมมีประสิทธิภาพอยู่ในระดับมากที่สุด ดังนั้น ผู้เชี่ยวชาญประเมินว่า โปรแกรมฯ มีประสิทธิภาพช่วยลดอาการเมื่อยล้าของผู้ทดสอบ และประหยัดเวลาในการทดสอบได้อย่างมาก ซึ่งสอดคล้องกับ Gibbons et al. (2008)

3. การเปรียบเทียบประสิทธิภาพวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา ระหว่าง วิธี IE กับ วิธี LB แบ่งการอภิปรายผลการวิจัยเป็น 4 ด้าน ดังนี้

การเปรียบเทียบประสิทธิภาพด้านจำนวนข้อสอบที่ใช้ในการทดสอบ เมื่อพิจารณาค่า Median ปรากฏว่า วิธี LB มีประสิทธิภาพสูงสุดเมื่อเทียบกับวิธี IE ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อที่ 1 แสดงให้เห็นว่า วิธี LB มีประสิทธิภาพสูงในการควบคุมการใช้ข้อสอบที่มีจำนวนน้อยที่สุด (Chang & Ying, 1999) จากการพิจารณาข้อมูลที่ผู้ทดสอบได้ทำการทดสอบ มีผลมาจากผู้ทดสอบได้รับข้อสอบที่มีระดับความยากมากขึ้นและห่างจากค่าประมาณความสามารถ เมื่อพิจารณาจากค่า RMSE จึงทำให้ค่า SEE ระหว่างการทำสอบต่ำกว่า 0.30 ซึ่งเป็นเกณฑ์การยุติการทำข้อสอบ จึงทำให้จำนวนข้อสอบที่ใช้ในการทดสอบลดลง และประเด็นที่ทำให้ระดับความยากของข้อสอบห่างจากค่าประมาณความสามารถของผู้ทดสอบ นั้นคือ วิธีการควบคุมสมดุลเนื้อหา เพราะอัลกอริทึมคัดเลือกเนื้อหาด้านที่เป็นตัวแทนจากการผ่านเงื่อนไขค่าเฉลี่ยถ่วงน้ำหนัก ข้อสอบด้านที่เป็นตัวแทนในด้านเนื้อหามีการแจกแจงค่าความยากไม่ครอบคลุมค่าประมาณความสามารถของผู้ทดสอบ จึงทำให้มีผลต่อการลดจำนวนข้อสอบ

การเปรียบเทียบประสิทธิภาพด้านเวลาที่ใช้ในการทดสอบ เมื่อพิจารณาค่า Mean ปรากฏว่า วิธี LB มีประสิทธิภาพสูงสุดเมื่อเทียบกับวิธี IE ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อที่ 2 แสดงให้เห็นว่า วิธี LB มีประสิทธิภาพสูงในการควบคุมการใช้ข้อสอบที่มีเวลาน้อยที่สุด (Chang & Ying, 1999) จากการพิจารณาขั้นตอนการทดสอบ โดยใช้การสลับแบบทดสอบ เช่น กลุ่มที่ 1 ทดสอบวิธี IE ก่อนแล้วตามด้วยทดสอบวิธี LB ส่วนกลุ่มที่ 2 ทดสอบวิธี LB ก่อน แล้วตามด้วยทดสอบวิธี IE เพื่อป้องกัน Order Effect ของกลุ่มตัวอย่างต่อการทดสอบ บ่งชี้ว่าถ้าผู้ทดสอบได้รับจำนวนข้อสอบที่ลดลงจะใช้เวลาลดลง เป็นข้อดี แต่ถ้าผู้ทดสอบได้รับจำนวนข้อสอบที่ลดลง ในขณะที่ทำแบบทดสอบใช้เวลาในการพิจารณา หรือวิเคราะห์ หรือคำนวณ ก็จะทำให้เวลาเพิ่มขึ้น เป็นข้อเสีย ข้อสังเกตุ ถ้ากำหนดเวลาของข้อสอบในการทำแต่ละข้อ ก็จะเป็นการเร่งให้ผู้สอบรีบทำข้อสอบข้อนั้น ผลที่ได้อาจเป็นข้อดีของเวลาที่ลดลง แต่ข้อเสียคือค่าความสามารถที่ได้นั้น มาจากการคิดหรือการเดา

การเปรียบเทียบประสิทธิภาพด้านอัตราการแสดงการใช้ข้อสอบสูงสุด เมื่อพิจารณาค่า MIER ปรากฏว่า วิธี LB มีประสิทธิภาพสูงสุดเมื่อเทียบกับวิธี IE ซึ่งเป็นไปตามสมมติฐานของการวิจัยข้อที่ 2 แสดงให้เห็นว่า วิธี LB มีประสิทธิภาพในการควบคุมการใช้ข้อสอบที่มีอัตราการใช้ไม่เกิน 0.2 ซึ่งสอดคล้องกับงานวิจัยของ Huebner et al. (2016) ว่าการควบคุมการใช้ข้อสอบให้มีประสิทธิภาพควรมีอัตราการใช้ไม่เกิน 0.2 โดยวิธี LB ทำการเรียงลำดับจากน้อยไปหามากระหว่างผลต่างค่าประมาณความสามารถกับค่าสารสนเทศสูงสุดและความถี่สะสม อัลกอริทึมคัดเลือกข้อสอบลำดับแรกจากการเรียงลำดับ ดังนั้น วิธีการนี้สามารถแก้ไขปัญหาอัตราการใช้ข้อสอบเกิน 0.2 ให้น้อยลงได้

การเปรียบเทียบประสิทธิภาพด้านค่าความคลาดเคลื่อนกำลังสองเฉลี่ย เมื่อพิจารณาค่า RMSE ปรากฏว่า วิธี LB มีประสิทธิภาพต่ำสุดใน 3 ระดับ เมื่อเทียบกับวิธี IE ซึ่งไม่เป็นตามสมมติฐานของการวิจัยข้อที่ 5 แสดงให้เห็นว่า วิธี LB มีประสิทธิภาพในการเลือกข้อสอบที่มีความยากใกล้เคียง

กับค่าประมาณความสามารถต่ำ (Leung et al., 2002) สาเหตุที่ วิธี LB ประสิทธิภาพต่ำสุดเพราะมีข้อจำกัดของการเลือกข้อสอบ เมื่อใช้ค่าเฉลี่ยถ่วงน้ำหนักในการเลือกด้านเนื้อหาแล้วนำความถี่ของการใช้ข้อสอบมาร่วมในควบคุมการเลือกใช้ข้อสอบ มีผลทำให้ข้อสอบมีอัตราการใช้ข้อสอบสูงกว่า 0.2 มีแนวโน้มลดลง บ่งชี้ว่า จำนวนข้อสอบถูกใช้ซ้ำบ่อยลดลง ซึ่งเป็นข้อดี แต่ในขณะเดียวกันค่า RMSE มีแนวโน้มเพิ่มมากขึ้น บ่งชี้ว่า ประสิทธิภาพด้านการประมาณค่าความสามารถของผู้ทดสอบต่อการเลือกใช้ข้อสอบลดลง เป็นข้อเสีย ซึ่งเป็นผลมาจากจำนวนข้อสอบในคลังมีปริมาณไม่เพียงพอต่อการทดสอบ นั้นคือข้อสอบมีการกระจายตัวของค่าความสามารถไม่เพียงพอ จึงส่งผลทำให้การเรียงความถี่ของข้อสอบที่ถูกนำไปใช้มีค่าความยากห่างจากค่าประมาณความสามารถ

4. การศึกษาความสัมพันธ์ระหว่างค่าความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ (MCAT) กับคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ (CT) ผลการศึกษาความสัมพันธ์ระหว่างค่าความสามารถของผู้ทดสอบที่ได้จากการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ที่พัฒนาขึ้นกับคะแนนรวมของผู้ทดสอบที่ได้จากการทดสอบด้วยคอมพิวเตอร์ที่ใช้ข้อสอบคงที่ จำนวน 40 ข้อ ปรากฏว่า เป็นไปตามสมติฐานของการวิจัยข้อ 5 นั้นคือ ผลการทดสอบทั้งสองแบบมีความสัมพันธ์ทางบวก มีขนาดปานกลาง ซึ่งเป็นผลมาจากวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา ทำให้ประสิทธิภาพในการคัดเลือกข้อสอบลดลง เนื่องจากวิธีการดังกล่าว เข้าไปแทรกแซงกระบวนการคัดเลือกข้อสอบข้อถัดไป (Thompson & Weiss, 2011) ซึ่งสอดคล้องกับผลการวิจัยของ Lilley et al. (2004) ที่ได้พัฒนาโปรแกรมการทดสอบปรับเหมาะด้วยคอมพิวเตอร์ สำหรับวัดความรู้ทางภาษาอังกฤษของนักศึกษานานาชาติ ปรากฏว่า การทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์กับการทดสอบด้วยคอมพิวเตอร์มีความสัมพันธ์ทางบวก ขนาดปานกลาง

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. ผู้ที่สนใจวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง สามารถนำไปประยุกต์ในการพัฒนาการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์โดยการแทรกระหว่างการคัดเลือกข้อสอบข้อถัดไปกับคลังข้อสอบ

2. ผู้ที่สนใจพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง สามารถนำผลการพัฒนาเป็นแนวทางในการพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาวิธีการใหม่ เช่น วิธี Round robin หรือ วิธี Sticky

3. ผู้ที่สนใจวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง ควรมีขนาดคลังข้อสอบที่ครอบคลุมความสามารถของผู้ทดสอบ

4. อาจารย์ นักศึกษา หรือบุคคลทั่วไป สามารถวัดความสามารถทางไอซีที ผ่านทางโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์

ข้อเสนอแนะสำหรับการวิจัยต่อไป

1. การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง เป็นการทำงานต่อจากการคัดเลือกข้อสอบข้อถัดไป ดังนั้น ควรพัฒนากระบวนการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ขั้นตอนอื่น เช่น การสร้างคลังข้อสอบ การคัดเลือกข้อสอบข้อแรก การเลือกข้อสอบข้อถัดไป การประมาณค่าความสามารถของผู้สอบ และเกณฑ์ยุติการทดสอบ

2. วิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา มุ่งเน้นเข้าควบคุมการใช้ข้อสอบที่มีอัตราการใช้ข้อสูงให้ลดลง และข้อสอบที่ถูกใช้น้อยได้ถูกเลือกใช้ร่วมด้วย แต่ไม่ครอบคลุมความแม่นยำในการเลือกข้อสอบตามค่าความสามารถให้ตรงกับผู้ทดสอบกรณีที่จำนวนข้อสอบในคลังมีปริมาณน้อย ดังนั้น จึงควรพัฒนาวิธีการดังกล่าวให้ครอบคลุม

3. โปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ที่พัฒนาขึ้น ใช้กับข้อสอบที่มีลักษณะเป็นแบบการตรวจให้คะแนน 2 ค่า คือ ตอบถูกได้ 1 ตอบผิดได้ 0 และยังไม่รองรับการตอบที่ให้คะแนนมากกว่า 2 ค่า ดังนั้น จึงควรพัฒนาโปรแกรมให้รองรับการทดสอบมากกว่า 2 ค่า

บรรณานุกรม

วิจารณ์ พานิช. (2555). *วิถีสร้างการเรียนรู้เพื่อศิษย์*. กรุงเทพฯ: มูลนิธิสดศรี สฤษดิ์วงศ์.

ศิริชัย กาญจนวาสี. (2555). *ทฤษฎีการตอบสนองแนวใหม่* (พิมพ์ครั้งที่ 4). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

Adhikari, M., & Amgoth, T. (2018). Heuristic-based load-balancing algorithm for IaaS cloud. *Future Generation Computer Systems, 81*(1), 156-165.

Amar, M., Anurag, K., Rakesh, K., Rupesh, K., & Prashant, Y. (2011). *SLa driven load balancing for web applications in cloud computing environment.* Paper presented at the Information and Knowledge Management.

Alan, D., Barbara, H. W., & Roberta, M. R. (2013). *Systems analysis and design*. New Jersey: John Wiley & Sons.

Armstrong, R. D., Jones, D. H., Koppel, N. B., & Pashley, P. J. (2004). Computerized Adaptive Testing With Multiple-Form Structures. *Applied Psychological Measurement, 28*(3), 147-164.

Barrada, J. R., Olea, J., & Abad, F. J. (2008). Rotating item banks versus restriction of maximum exposure rates in computerized adaptive testing. *The Spanish Journal of Psychology*, *11*(2), 618-625.

Belov, D. I., & Armstrong, R. D. (2009). Direct and inverse problems of item pool design for computerized adaptive testing. *Educational and Psychological Measurement, 69*(4), 533-547.

Berenbrink, P. , Friedetzky, T., Mallmann-Trenn, F., Meshkinfamfard, S., & Wastell, C. (2018). Threshold load balancing with weighted tasks. *Journal of Parallel and Distributed Computing, 113*(3), 218-226.

Berlinska, J., & Drozdowski, M. (2018). Comparing load-balancing algorithms for MapReduce under Zipfian data skews. *Parallel Computing, 72(1)*, 14-28.

Bortot, S., & Marques Pereira, R. A. (2017). Algebraic representations of the weighted mean. *Fuzzy Sets and Systems, 308*(1), 85-105.

Chang, H. H., & Ying, Z. (1999). a-Stratified Multistage Computerized Adaptive Testing. *Applied Psychological Measurement, 23*(3), 211-222.

Chang, H. H., Qian, J., & Ying, Z. (2001). a-Stratified multistage computerized adaptive testing with b blocking. *Applied Psychological Measurement*, *25*(4), 333-341.

Chang, H. H., & van der Linden, W. J. (2003). Optimal stratification of item pools in a-stratified computerized adaptive testing. *Applied Psychological Measurement*, *27*(4), 262-274.

Chen, P. H. (2015). A sampling and classification item selection approach with content balancing. *Behavior Research Methods, 47*(1), 98-106.

Chen, S. Y., Ankenmann, R. D., & Spray, J. A. (2003). The Relationship Between Item Exposure and Test Overlap in Computerized Adaptive Testing. *Journal of Educational Measurement, 40*(2), 129-145.

Chen, T. B., Zheng, Y. M., Lei, M., Huang, Z. C., Wu, H. T., Chen, H., & Tian, Q. Z. (2005). Assessment of heavy metal pollution in surface soils of urban parks in Beijing, China. *Chemosphere, 60*(4), 542-551.

Cheng, Y., & Chang, H. H. (2009). The maximum priority index method for severely constrained item selection in computerized adaptive testing. *British Journal of Mathematical and Statistical Psychology, 62*(2), 369-383.

Chien, N. K., Son, N. H., & Loc, H. D. (2016). *Load balancing algorithm based on estimating finish time of services in cloud computing.* Paper presented at the Advanced Communication Technology (ICACT), *18*(1), 228-232

Davey, T., & Parshall, C. G. (1995). New algorithms for item selection and exposure control with computerized adaptive testing.

Demirci, İ., & Korçak, Ö. (2018). Cell breathing algorithms for load balancing in Wi-Fi cellular heterogeneous networks. *Computer Networks, 134*(1), 140-151.

Deng, H., Ansley, T., & Chang, H. H. (2010). Stratified and maximum information item selection procedures in computer adaptive testing. *Journal of Educational Measurement, 47*(2), 202-226.

Drasgow, F., Levin, M. V., Williams, B., McCusker, C., Thomasson, G. L., & Lim, R. G. (1990). Evaluation of Optimal Appropriateness Measurement for Use in Practical Settings. *AIR Force Human Resources Laboratory*.

Faynberg, I., Lu, H. L., & Skuler, D. (2016). *Cloud computing: Business trends and technologies*. John Wiley & Sons.

Frey, A., & Seitz, N. N. (2009). Multidimensional adaptive testing in educational and psychological measurement: Current state and future challenges. *Studies in Educational Evaluation, 35*(2), 89-94.

Gary, B. S., & Harry, J. R. (2012). *Systems Analysis and Design*. Boston: Nicole Pinard.

Georgiadou, E., & Triantafillou, A. (2007). A Review of Item Exposure Control Strategies for Computerized Adaptive Testing Developed from 1983 to 2005. *Journal of Technology, Learning, and Assessment, 5*(8), 1-39.

Ghomi, E. J., Rahmani, A. M., & Qader, N. N. (2017). Load-balancing algorithms in cloud computing: a survey. *Journal of Network and Computer Applications*, *88*(2), 50-71.

Ghoneem, M., & Kulkarni, L. (2017). *An Adaptive MapReduce Scheduler for Scalable Heterogeneous Systems.* Paper presented at the Proceedings of the International Conference on Data Engineering and Communication Technology.

Gibbons, R. D., Weiss, D. J., Kupfer, D. J., Frank, E., Fagiolini, A., Grochocinski, V. J., & Immekus, J. C. (2008). Using computerized adaptive testing to reduce the burden of mental health assessment. *Psychiatric Services, 59*(4), 361-368.

Hou, L., Chen, Y., & Cao, Q. (2014). Nonlinear vibration phenomenon of an aircraft rub-impact rotor system due to hovering flight. *Communications in Nonlinear Science and Numerical Simulation, 19*(1), 286-297.

Hou, X., TK, A. K., Thomas, J. P., & Varadharajan, V. (2014). *Dynamic workload balancing for Hadoop MapReduce.* Paper presented at the Big Data and Cloud Computing (BdCloud), *14*(1), 56-62.

Huang, R. D., & Jo, H. (1995). Data frequency and the number of factors in stock returns. *Journal of Banking & Finance, 19*(6), 987-1003.

Huebner, A. R., Wang, C., Quinlan, K., & Seubert, L. (2016). Item exposure control for multidimensional computer adaptive testing under maximum likelihood and expected a posteriori estimation. *Behavior Research Methods, 48*(4), 1443-1453.

Keshvadi, S., & Faghih, B. (2016). A multi-agent based load balancing system in IaaS cloud environment. *Int Rob Auto J, 1*(1), 151-163.

Kingsbury, G. G., & Zara, A. R. (1989). Procedures for Selecting Items for Computerized Adaptive Tests. *Applied Measurement in Education, 2*(4), 359-375.

Kumar, M., & Sharma, S. C. (2017). Dynamic load balancing algorithm for balancing the workload among virtual machine in cloud computing. *Procedia Computer Science, 115*(1), 322-329.

Lee, G. (2014). *Cloud networking: understanding cloud-based data center networks*. Amsterdam: Elsevie Inc.

Lee, Y. H., Ip, E. H., & Fuh, C. D. (2008). A strategy for controlling item exposure in multidimensional computerized adaptive testing. *Educational and Psychological Measurement, 68*(2), 215-232.

Leroux, A. J., Lopez, M., Hembry, I., & Dodd, B. G. (2013). A comparison of exposure control procedures in CATs using the 3PL model. *Educational and Psychological Measurement, 73*(5), 857-874.

Leung, C. K., Chang, H. H., & Hau, K. T. (2000). *Content balancing in stratified computerized adaptive testing designs*: ERIC Clearinghouse.

Leung, C. K., Chang, H. H., & Hau, K. T. (2002). Item selection in computerized adaptive testing: Improving the a-stratified design with the Sympson-Hetter algorithm. *Applied Psychological Measurement, 26*(4), 376-392.

Lilley, M., Barker, T., & Britton, C. (2004). The development and evaluation of a software prototype for computer-adaptive testing. *Computers & Education, 43*(2), 109-123.

Literacy, N. I. (2014). *National Assessment Program ICT Literacy Assessment Framework 2014*. Sydney: ACARA.

Luecht, R. M., & Sireci, S. G. (2011). *A Review of Models for Computer-Based Testing.* New York: College Board.

Lunz, M. E., & Stahl, J. A. (1998). Patterns of item exposure using a randomized CAT algorithm. *In annual meeting of the National Council on Measurement in Education,* San Diego, California.

McBride, J. R., & Martin, J. T. (1983). Reliability and validity of adaptive ability tests in a military setting. In *New horizons in testing* (pp. 223-236).

Murphy, D. L., Dodd, B. G., & Vaughn, B. K. (2010). A comparison of item selection techniques for testlets. *Applied Psychological Measurement, 34*(6), 424-437.

Ozturk N. B., & Dogan N. (2015). Investigating Item Exposure Control Methods in Computerized Adaptive Testing. *Educational Sciences: Theory and Practice, 15*(1), 85-98.

Panel, I. L. (2002). Digital transformation: A framework for ICT literacy. *Educational Testing Service*. New Jersey: ETS.

Parshall, C., Hogarty, K., & Kromrey, J. (1999). Item exposure in adaptive tests: An empirical investigation of control strategies. In *annual meeting of the Psychometric Society,* Lawrence, Kansas.

Polit, D. F., & Beck, C. T. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in nursing & health*, *29*(5), 489-497.

Reckase, M. D. (2009). *Multidimensional Item Response Theory*. New York: Springer.

Revuelta, J., & Ponsoda, V. (1998). A Comparison of Item Exposure Control Methods in Computerized Adaptive Testing. *Journal of Educational Measurement, 35*(4), 311-327.

Service, E. T. (2007). *Digital Transformation A Framework for ICT Literacy*. New Jersey: ETS.

Singh, A., Juneja, D., & Malhotra, M. (2015). Autonomous Agent Based Load Balancing Algorithm in Cloud Computing. *Procedia Computer Science, 45(1)*, 832-841.

Singh, D., Thakur, A., & Chaudhary, A. (2015). A comparative study between waterfall and incremental software development life cycle model. *International Journal of Emerging Trends in Science and Technology*, *2*(4), 220-280.

Stocking, M. L., & Lewis, C. (2000). Methods of controlling the exposure of items in CAT. In *Computerized adaptive testing: Theory and practice* (pp. 163-182).

Sympson, J. B., & Hetter, R. D. (1985). Controlling item-exposure rates in computerized adaptive testing. In *Proceedings of the 27th annual meeting of the Military Testing Association* (pp. 973-977).

Thakur, A., & Goraya, M. S. (2017). A taxonomic survey on load balancing in cloud. *Journal of Network and Computer Applications, 98(1),* 43-57.

Thissen, D., Chen, W. H., & Bock, R. D. (2003). Multilog (version 7)[Computer software]. *Lincolnwood, IL: Scientific Software International*.

Thompson, N. A., & Weiss, D. J. (2011). A Framework for the Development of Computerized Adaptive Tests. *Practical Assessment, Research & Evaluation, 16(1),* 1-9.

Urry, V. W. (1977). Tailored testing: A successful application of latent trait theory. *Journal of Educational Measurement*, *14*(2), 181-196.

Van Der Linden, W. J. (2005). A comparison of item-selection methods for adaptive tests with content constraints. *Journal of Educational Measurement*, *42*(3), 283-302.

Van Der Linden, W. J., & Chang, H. H. (2003). Implementing content constraints in alpha-stratified adaptive testing using a shadow test approach. *Applied Psychological Measurement*, *27*(2), 107-120.

Van Der Linden, W. J., & Veldkamp, B. P. (2004). Constraining item exposure in computerized adaptive testing with shadow tests. *Journal of Educational and Behavioral Statistics*, *29*(3), 273-291.

Vanitha, M., & Marikkannu, P. (2017). Effective resource utilization in cloud environment through a dynamic well-organized load balancing algorithm for virtual machines. *Computers & Electrical Engineering, 57(1)*, 199-208.

Weiss, D. J. (1974). Strategies of adaptive ability measurement. (Research Report 74-5). Psychometric Methods Program, University of Minnesota, Minneapolis.

Yao, L. (2012). Multidimensional CAT Item Selection Methods for Domain Scores and Composite Scores: Theory and Applications. *Psychometrika, 77*(3), 495-523.

Yao, L. (2014). Multidimensional CAT Item Selection Methods for Domain Scores and Composite Scores With Item Exposure Control and Content Constraints. *Journal of Educational Measurement, 51*(1), 18-38.

Zheng, Y., Chang, C. H., & Chang, H. H. (2013). Content-balancing strategy in bifactor computerized adaptive patient-reported outcome measurement. *Quality of Life Research, 22*(3), 491-499.

ภาคผนวก

ภาคผนวก ก
อัลกอริทึมโหลดบาลานซ์ซิ่ง

1. อัลกอริทึมโหลดบาลานซ์ซิ่งแบบเดิม

จากงานวิจัยของ Singh et al. (2015) ได้นำเสนอขั้นตอนของอัลกอริทึมในการทำโหลดบาลานซ์ซิ่งบนเครือข่ายแบบก้อนเมฆ มี 3 ขั้นตอน ดังนี้

1.1 การสำรวจและเคลื่อนย้ายข้อมูล (Migration Agent: MA)

```
Migration_Agent ( )
Input: VM_configurations_fromChannel_Agent(VM_initial)
Output: Search similar VM from other Datacenters
{
Accept VM_configurations_fromchannel_Agent(VM_initial);
Search a Data_center;
Check VM_Load_Table;
If (found)
Return (A);
Else
Return (NA);
On_receiving (self_destroy);
Kill (MA_i);
}
```

ภาพที่ ก-1 การสำรวจและเคลื่อนข้ายข้อมูล

1.2 การคำนวณภาระงาน (Load Agent: LA)

Load_ Agent( ):
Input: Receive request from user;
Output: Allocate_ resources_with_A2LB;
Case I:
{
If (VM_Load_Table==empty( ))
Then allocate_requested_resources;
Maintain_VM_Load_Table;

$$\mu_{available} = \mu_{total} - \mu_{used}$$

$$\upsilon(\%) = \frac{\mu_{available}}{\mu_{total}} \times 100$$

If ($\mathcal{V}$ >25) then
{
allocation_status:=Normal;
}
Else
{
allocation_status:=Critical;
initiateChannel_Agent($VM_{initial}$);
}
Case II:
If (VM_Load_Table≠ empty)
Scan VM_Load_Table;
If(Load_Status(VM_i) ==Critical)
{
Call Channel_Agent($VM_{Load_Balance}$);
Receive <DC_{id},VM_{id}forload_ transfer>;
Transfer_request to DC_{id};
}
Else
Allocate_request to VM_{id};
Update VM_Load_Table;
}

ภาพที่ ก-2 การคำนวณภาระงาน

1.3 การควบคุมการทำงาน (Channel Agent: CA)

```
Channel_Agent(VM_initial):
Input:ReceiveVMiconfiguration from Load_Agent;
Output:Response_Analysis_Table, VM_id;
On_ReceivingVMi fromLoad_Agent( );
Initiate_Migration_agent( );
Receive acknowledgement from Migration_Agents( );
Maintain_Response_Table;
If (response == NA)
{
Send self_destroy(MAi);
}
Else
{
Receive V (MAi);
Maintain Response_Analysis_Table;
Periodically update Response_Analysis_Table;
}

Channel_Agent(VM_Load_Balance):
On incoming request;
{
Scan Response_Analysis_Table;
Prepare list of matching VMi;
L1: for i=1 to n
Large=0;
If (V (VMi) >large)
Large= V (VMi);
V old= V i;
Update message to MAi;
If (V old= = V i )
Return (<DCb VMi>) to Load_Agent;
Else
Goto L1;
}
```

ภาพที่ ก-3 การควบคุมการทำงาน

2. อัลกอริทึมโหลดบาลานซ์ซิ่งแบบปรับใหม่

ผู้วิจัยได้พัฒนาอัลกอริทึมจากงานวิจัยของ Singh et al. (2015) ในขั้นตอนที่ 2 และขั้นตอนที่ 3 ให้สอดคล้องกับการทำงานควบคุมการใช้ข้อสอบร่วมกับควบคุมสมดุลเนื้อหาดังนี้

2.1 การสำรวจและเคลื่อนย้ายข้อมูล (Migration Agent: MA)

Migration_Agent ( )

Input: $Item_{configurations}$ fromChannel_Agent($Item_{initial}$)

Output: Search similar Item from other Datacenters

{

Accept $Item_{configurations}$ fromchannel_Agent($Item_{initial}$);

Search a Data_center;

Check Item_Load_Table;

If (found)

Return (A);

Else

Return (NA);

On_receiving (self_destroy);

Kill (MA_i);

}

2.2 การคำนวณภาระงาน (Load Agent: LA) //ควบคุมความสมดุลเนื้อหา

Load_ Agent( )

Input: Receive request from user;

Output: Allocate_ resources_with_A2LB;

Case I: {

If (Item_Load_Table==empty( ))

Then allocate_requested_resources;

Maintain_Item_Load_Table;

$Item_{(available)} = Item_{total} - Item_{used}$ //จำนวนข้อสอบที่ยังไม่ใช้

$v_i = \frac{Item_{available}}{Item_{total}} \times 100$ //คำนวณเปอร์เซ็นต์ข้อสอบที่ไม่ได้ใช้

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i v_i}{\sum_{i=1}^{n} v_i}$$ //ค่าเฉลี่ยถ่วงน้ำหนัก

If (v_i> $\bar{X}$) then {

allocation_status:= 0 ;

initiateChannel_Agent($Item_{initial}$);

}

Else {

allocation_status:= 1;

}

Case II: {

If (Item_Load_Table≠ empty) //ถ้าข้อมูลในตารางการสำรวจไม่ว่าง

Scan Item_Load_Table;

If(Load_Status($Item_i$) ==Critical) {

Call Channel_Agent($Item_{Load_Balance}$);

Receive <DC_{id},$Item_{id}$forload_ transfer>;

Transfer_request to DC_{id};

}

Else

Allocate_request to $Item_{id}$;

Update Item_Load_Table;

}

2.3 การควบคุมการทำงาน (Channel Agent: CA) //ควบคุมการใช้ข้อสอบ

Channel_Agent($Item_{initial}$):

Input: ReceiveVM_iconfiguration from Load_Agent;

Output: Response_Analysis_Table, VM_id;

On_Receiving$Item_i$ fromLoad_Agent( );

Initiate_Migration_agent( );

Receive acknowledgement from Migration_Agents( );

Maintain_Response_Table;

If (response == NA) { Send self_destroy(MA_i); }

Else { Receive v_i *(MA_i);*

Maintain Response_Analysis_Table;

Periodically update Response_Analysis_Table; }

Channel_Agent(VMLoad_Balance):

On incoming request; {

Scan Response_Analysis_Table;

Prepare list of matching $Item_i$;

L1: for i=1 to n

Large=0;

If (v_i ($Item_i$) >large)

Large= (v ($Item_i$);

$$Item_{(v)(i)} = (Item_{(v)(i)} - Item_{used(v)(i)})$$

$$MItem_{(v)(i)} = MIC\big(Min(Item_{(v)(i)}) \; to \; Max(Item_{(v)(i)})\big)$$

$$TestItem_{(v)(i)} = ABS(ML_{(\theta)} - MItem_{(v)(i)})$$

$v_{i_{old}} = v_i$;

Update message to MA_i;

If ($v_{i_{old}} == v_i$)

Return (<DC_i, $Item_i$>) to Load_Agent;

Else Goto L1;

}

ภาคผนวก ข
รายชื่อผู้เชี่ยวชาญ

รายชื่อผู้เชี่ยวชาญ

1. ราชื่อผู้เชี่ยวชาญสำหรับการประเมินความสอดคล้องเนื้อหาความรู้ทางไอซีที มีดังนี้

1) รองศาสตราจารย์ ดร.ปณิตา วรรณพิรุณ อาจารย์ประจำคณะครุศาสตร์อุตสาหกรรมมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

2) ผู้ช่วยศาสตราจารย์ ดร.สยาม แกมขุนทด อาจารย์ประจำคณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

3) ผู้ช่วยศาสตราจารย์ ดร.รุ่งทิวา เสาร์สิงห์ อาจารย์ประจำคณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ

2. รายชื่อผู้เชี่ยวชาญสำหรับการประเมินประสิทธิภาพโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ มีดังนี้

1) รองศาสตราจารย์ ดร.ปณิตา วรรณพิรุณ อาจารย์ประจำคณะครุศาสตร์อุตสาหกรรมมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

2) ผู้ช่วยศาสตราจารย์ ดร.สยาม แกมขุนทด อาจารย์ประจำคณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

3) ผู้ช่วยศาสตราจารย์ ดร.รุ่งทิวา เสาร์สิงห์ อาจารย์ประจำคณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ

4) ดร.มิ่ง เทพครเมือง ผู้ช่วยฝ่ายวิจัย นิเทศ และประสบการณ์วิชาชีพ โรงเรียนสาธิตมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร (ฝ่ายประถม)

5) ดร.ไพเราะ ราชสมบูรณ์ อาจารย์พิเศษคณะบริหารธุรกิจ มหาวิทยาลัยเกริก และประธานบริษัท ซี เอส เอ็น แอ็ดวานซ์ จำกัด

ที่ ศธ ๖๒๒๔/ว๐๔๐๕

วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา
มหาวิทยาลัยบูรพา
ต.แสนสุข อ.เมือง จ.ชลบุรี ๒๐๑๓๑

๒๒ ตุลาคม ๒๕๖๑

เรื่อง ขอความอนุเคราะห์ตรวจสอบความตรงเชิงเนื้อหาของเครื่องมือวิจัย

เรียน รองศาสตราจารย์ ดร.ปณิตา วรรณพิรุณ

สิ่งที่ส่งมาด้วย เค้าโครงดุษฎีนิพนธ์ และแบบประเมิน จำนวน ๑ ชุด

ด้วย นายธีร์นวัช สุขวิลัยหิรัญ รหัสประจำตัวนิสิต ๕๕๘๑๐๒๕๖ นิสิตหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา ได้รับอนุมัติให้ทำดุษฎีนิพนธ์เรื่อง "การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง สำหรับทดสอบความรู้ทางไอซีที: การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์" ซึ่งอยู่ในความควบคุมดูแลของ รองศาสตราจารย์ ดร.เสรี ชัดแช้ม อาจารย์ที่ปรึกษาหลัก ขณะนี้อยู่ในขั้นตอนการสร้างเครื่องมือวิจัย ในการนี้ วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา ได้พิจารณาแล้วเห็นว่าท่านเป็นผู้เชี่ยวชาญในเรื่องดังกล่าวเป็นอย่างดี จึงขอความอนุเคราะห์จากท่านตรวจสอบความตรงเชิงเนื้อหาของเครื่องมือวิจัยแก่นิสิตในครั้งนี้

จึงเรียนมาเพื่อโปรดพิจารณา วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา หวังเป็นอย่างยิ่งว่าคงจะได้รับความอนุเคราะห์จากท่านด้วยดี และขอขอบคุณมา ณ โอกาสนี้

ขอแสดงความนับถือ

(ผู้ช่วยศาสตราจารย์ ดร.ภัทราวดี มากมี)
คณบดีวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา

วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา
โทร. ๐ ๓๘๑๐ ๒๐๗๗-๘
โทร/ โทรสาร ๐ ๓๘๓๙ ๓๔๘๔
http ://www.rmcs.buu.ac.th

ตารางที่ ข-1 สรุปผลการประเมินประสิทธิภาพโปรแกรมของผู้เชี่ยวชาญ

ด้านตรงตามความต้องการของโปรแกรม							
ข้อที่	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	*Mean*	*SD*
1	4	5	5	4	4	4.40	0.55
2	5	5	5	4	5	4.80	0.45
3	4	5	4	5	4	4.40	0.55
4	4	5	4	4	4	4.20	0.45
5	4	4	4	5	5	4.40	0.55
6	5	4	5	5	5	4.80	0.45
7	4	5	4	5	4	4.40	0.55
8	4	4	5	4	4	4.20	0.45
ด้านการทำงานของโปรแกรม							
ข้อที่	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	*Mean*	*SD*
1	5	5	4	5	5	4.80	0.45
2	5	4	5	5	5	4.80	0.45
3	4	4	4	4	5	4.20	0.45
4	3	4	5	4	5	4.20	0.84
5	3	5	4	5	4	4.20	0.84
6	5	5	5	4	4	4.60	0.55
7	4	4	5	5	4	4.40	0.55
8	5	5	5	4	4	4.60	0.55
ด้านการใช้งานโปรแกรม							
ข้อที่	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	*Mean*	*SD*
1	5	5	4	5	5	4.80	0.45
2	5	5	5	5	4	4.80	0.45
3	4	5	5	4	5	4.60	0.55
4	3	4	5	5	4	4.20	0.84
5	5	5	4	5	5	4.80	0.45
6	5	5	5	4	5	4.80	0.45
7	4	4	4	5	5	4.40	0.55
ด้านความปลอดภัยของโปรแกรม							
ข้อที่	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5	*Mean*	*SD*
1	5	5	4	4	5	4.60	0.55
2	4	5	5	5	4	4.60	0.55
3	5	5	4	5	5	4.80	0.45

ภาคผนวก ค
แบบประเมินประสิทธิภาพโปรแกรม (Black box testing)

แบบประเมินประสิทธิภาพ (Black Box Testing)
โปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์

<u>คำชี้แจง</u>

แบบประเมินประสิทธิภาพต่อการใช้งานโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ฉบับนี้ มีวัตถุประสงค์เพื่อสอบถามความคิดเห็นของท่าน เพื่อวิเคราะห์เนื้อหาที่ใช้ในการประเมินประสิทธิภาพต่อการใช้งานโปรแกรมฯ และประโยชน์ในการพัฒนาโปรแกรมให้มีความสมบูรณ์ แบบประเมินฉบับนี้แบ่งออกเป็น 4 ตอน ดังนี้

ตอนที่ 1 การประเมินประสิทธิภาพด้านตรงตามความต้องการของโปรแกรม
ตอนที่ 2 การประเมินประสิทธิภาพด้านการทำงานของโปรแกรม
ตอนที่ 3 การประเมินประสิทธิภาพด้านการใช้งานโปรแกรม
ตอนที่ 4 การประเมินประสิทธิภาพด้านความปลอดภัยของโปรแกรม
โดยประเมินความเหมาะสมเป็นแบบมาตรประมาณค่า 5 ระดับ

ผู้วิจัยขอกราบขอบพระคุณผู้เชี่ยวชาญเป็นอย่างสูงที่ท่านกรุณาสละเวลาเพื่อให้ความเห็นในการพัฒนาโปรแกรมการโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์

ขอแสดงความนับถือ
นายธีร์นวัช สุขวิลัยหิรัญ
นิสิตระดับปริญญาเอก
สาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา
วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

คำชี้แจง

แบบประเมินประสิทธิภาพโปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์เป็นมาตรวัดแบบประเมินค่า โปรดทำเครื่องหมาย ✓ ลงในช่องที่ตรงกับระดับความคิดเห็นของท่าน

5 หมายถึง โปรแกรมที่พัฒนาขึ้นมีประสิทธิภาพในระดับมากที่สุด
4 หมายถึง โปรแกรมที่พัฒนาขึ้นมีประสิทธิภาพในระดับมาก
3 หมายถึง โปรแกรมที่พัฒนาขึ้นมีประสิทธิภาพในระดับพอใช้
2 หมายถึง โปรแกรมที่พัฒนาขึ้นต้องปรับปรุงแก้ไข
1 หมายถึง โปรแกรมที่พัฒนาขึ้นไม่สามารถนำไปใช้งานได้

ตอนที่ 1 การประเมินประสิทธิภาพด้านตรงตามความต้องการของโปรแกรม

เป็นการประเมินผลความถูกต้องและประสิทธิภาพของโปรแกรมว่า ตรงตามความต้องการของผู้ใช้โปรแกรมมากน้อยเพียงใด

รายการประเมิน	ระดับประสิทธิภาพ				
	5	4	3	2	1
1. การจัดการข้อมูลผู้ใช้ระบบและผู้ทดสอบ					
2. การจัดการข้อมูลความรู้ทางไอซีที					
3. การจัดการข้อมูลหมวดไอซีที					
4. การสร้างข้อสอบ					
5. การสร้างชุดข้อสอบ					
6. การเลือกและทำแบบทดสอบ					
7. การรายงานผลการทดสอบ					
8. การแสดงประวัติการผลทดสอบ					

ตอนที่ 2 การประเมินประสิทธิภาพด้านการทำงานของโปรแกรม

เป็นการประเมินความถูกต้องและประสิทธิภาพในการทำงานของโปรแกรมว่า สามารถทำงานได้ตามฟังก์ชันงานของโปรแกรมมากน้อยเพียงใด

รายการประเมิน	ระดับประสิทธิภาพ				
	5	4	3	2	1
1. สามารถจัดการข้อมูลผู้ใช้ระบบและผู้ทดสอบได้					
2. สามารถจัดการข้อมูลความรู้ทางไอซีทีได้					
3. สามารถจัดการข้อมูลหมวดไอซีทีได้					
4. สามารถสร้างข้อสอบได้					
5. สามารถสร้างชุดข้อสอบได้					
6. สามารถเลือกและทำแบบทดสอบได้					
7. สามารถรายงานผลการทดสอบได้					
8. สามารถแสดงประวัติผลการทดสอบได้					

ตอนที่ 3 การประเมินประสิทธิภาพด้านการใช้งานโปรแกรม

เป็นการประเมินลักษณะการออกแบบโปรแกรมว่า มีความง่ายต่อการใช้งานมากน้อยเพียงใด

รายการประเมิน	ระดับประสิทธิภาพ				
	5	4	3	2	1
1. ความง่ายต่อการใช้งานของโปรแกรม					
2. ความเหมาะสมในการเลือกใช้ขนาดของตัวอักษร					
3. ความเหมาะสมในการใช้สีของตัวอักษรและภาพ					
4. ความเหมาะสมในการใช้ข้อความ สัญลักษณ์หรือภาพ					
5. ความเป็นมาตรฐานเดียวกันในการออกแบบหน้าจอภาพ					
6. ความเหมาะสมในการปฏิสัมพันธ์โต้ตอบกับผู้ใช้					
7. ผู้ใช้คุ้นเคยกับคำศัพท์และสามารถปฏิบัติตามได้โดยง่าย					

ตอนที่ 4 การประเมินประสิทธิภาพด้านความปลอดภัยของโปรแกรม

เป็นการประเมินโปรแกรมในด้านการรักษาความปลอดภัยของข้อมูลในโปรแกรมว่า มีมากน้อยเพียงใด

รายการประเมิน	ระดับประสิทธิภาพ				
	5	4	3	2	1
1. การกำหนดรหัสผู้ใช้และรหัสผ่านในการตรวจสอบผู้เข้าใช้โปรแกรม					
2. การตรวจสอบสิทธิ์ก่อนการใช้งานของผู้ใช้โปรแกรมในระดับต่าง ๆ เช่น สิทธิ์ผู้ดูแลโปรแกรม สิทธิ์ผู้ทดสอบ					
3. การควบคุมให้ใช้งานตามสิทธิ์ผู้ใช้ได้อย่างถูกต้อง เช่น การใช้งานของผู้ดูแลโปรแกรม การใช้งานของผู้ทดสอบ					

ข้อเสนอแนะเพื่อการพัฒนาโปรแกรม

.....

.....

.....

.....

.....

.....

.....

ลงชื่อ ..... ผู้ประเมิน

(.....)

ภาคผนวก ง
คู่มือการใช้โปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์

คู่มือ (ผู้ดูแลโปรแกรม)
โปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์
สำหรับทดสอบความรู้ทางไอซีที (ICT-MCAT)

สาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา
วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา
กันยายน 2561
ลิขสิทธิ์เป็นของมหาวิทยาลัยบูรพา

คำนำ

คู่มือการใช้โปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ สำหรับทดสอบความรู้ทางไอซีที (ICT-MCAT) มีการทำงานของโปรแกรม แบ่งออกเป็น 3 ส่วน ได้แก่ 1) ส่วนของการตั้งค่าโปรแกรม เป็นการตั้งค่าการเริ่มต้นการใช้งาน ประกอบด้วย แบบทดสอบ ด้านความรู้ หมวดไอซีที สร้างข้อสอบ และ สร้างข้อสอบ 2) ส่วนของการทดสอบ เป็นการเลือกแบบทดสอบเพื่อทดสอบความรู้ทางไอซีที ประกอบด้วย วิธีการทดสอบแบบ (Load Balancing: LB) วิธีการทดสอบแบบ (Item Eligibility: IE) และวิธีการทดสอบแบบ (Computer test: CT) และ 3) ส่วนของรายงาน เป็นการรายงานสำหรับผู้ดูแลโปรแกรมและสำหรับผู้ทดสอบ

คู่มือเล่มนี้ นอกจากใช้ประกอบการทำงานของโปรแกรม ยังเหมาะสมกับผู้ที่สนใจพัฒนาโปรแกรมเกี่ยวกับการวัดความรู้ในด้านอื่น ๆ ที่สนใจ และหากคู่มือการใช้โปรแกรมเล่มนี้มีข้อผิดพลาดประการใด ผู้วิจัยต้องขออภัยมา ณ โอกาสนี้

ธีร์นวัช สุขวิลัยหิรัญ

กันยายน 2561

สารบัญ

	หน้า
โปรแกรม ICT-MCAT	1
ส่วนของการตั้งค่าโปรแกรม	3
ส่วนของการทดสอบ	8
ส่วนของรายงาน	11

คู่มือ (ผู้ทดสอบ)
โปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์
สำหรับทดสอบความรู้ทางไอซีที (ICT-MCAT)

สาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา
วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา
กันยายน 2561
ลิขสิทธิ์เป็นของมหาวิทยาลัยบูรพา

คำนำ

คู่มือการใช้โปรแกรมการทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์ สำหรับทดสอบความรู้ทางไอซีที (ICT-MCAT) มีการทำงานของโปรแกรม คือ 1) การจัดการข้อมูลลงทะเบียน 2) การเข้าสู่โปรแกรม 3) การจัดการข้อมูลการทดสอบ และ 4) รายงานผลการทดสอบ

คู่มือเล่มนี้ เป็นขั้นตอนการทำงานของผู้ทดสอบ และหากคู่มือการใช้โปรแกรมเล่มนี้มีข้อผิดพลาดประการใด ผู้วิจัยต้องขออภัยมา ณ โอกาสนี้

ธีร์นวัช สุขวิลัยหิรัญ

กันยายน 2561

สารบัญ

	หน้า
โปรแกรม ICT-MCAT	1
การจัดการข้อมูลลงทะเบียน	3
การจัดการข้อมูลการทดสอบ	4
รายงานผลการทดสอบ	7

รายงานผลการทดสอบ

การรายงานผลการทดสอบของผู้ทดสอบรายบุคคล แบ่งเกณฑ์การประเมินออกเป็น 2 ส่วน ได้แก่ 1) เกณฑ์การประเมินระดับค่าความสามารถของผู้ทดสอบ และ 2) เกณฑ์การประเมินระดับคะแนนของผู้ทดสอบ แสดงดังตารางที่ 1 และ ตารางที่ 2

ตารางที่ 1 เกณฑ์การประเมินระดับค่าความสามารถ สำหรับการทดสอบแบบ LB หรือ IE

ค่าประมาณความสามารถของผู้สอบ	ระดับความสามารถ
มากกว่า 2.000	สูงมาก
1.001 ถึง 2.000	สูง
0.501 ถึง 1.000	ค่อนข้างสูง
-0.499 ถึง 0.500	ปานกลาง
-1.499 ถึง -0.500	ค่อนข้างต่ำ
-2.000 ถึง -1.500	ต่ำ
ต่ำกว่า -2.000	ต่ำมาก

ตารางที่ 2 เกณฑ์การประเมินระดับคะแนนของผู้ทดสอบ สำหรับการทดสอบแบบ CT

จำนวนข้อสอบที่ตอบถูก	ระดับความสามารถ
ตั้งแต่ 32 ข้อ ขึ้นไป	สูงมาก
28-31 ข้อ	สูง
24-27 ข้อ	ค่อนข้างสูง
20-23 ข้อ	ปานกลาง
16-19 ข้อ	ค่อนข้างต่ำ
12-15 ข้อ	ต่ำ
ต่ำกว่า 12 ข้อ	ต่ำมาก

ภาคผนวก จ
จริยธรรมการวิจัย

ที่ ๐๗๔/๒๕๖๑

เอกสารรับรองผลการพิจารณาจริยธรรมการวิจัยในมนุษย์
วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา
มหาวิทยาลัยบูรพา

๑. **ชื่อเรื่องดุษฎีนิพนธ์**

ชื่อเรื่อง: การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซึ่งสำหรับทดสอบความรู้ทางไอซีที: การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์

TITLE: DEVELOPMENT OF THE ITEM EXPOSURE CONTROL WITH CONTENT BALANCE PROCEDURE USING LOAD BALANCING FOR ICT LITERACY: MULTIDIMENSIONAL COMPUTERIZED ADAPTIVE TESTING

๒. **ชื่อนิสิต:** นายธีร์นวัช สุขวิลัยหิรัญ
หลักสูตร ปรัชญาดุษฎีบัณฑิต (Ph.D.) **สาขาวิชา** การวัดและเทคโนโลยีทางวิทยาการปัญญา
รหัส ๕๕๘๑๐๒๕๖

๓. **ผลการพิจารณาของคณะกรรมการพิจารณาจริยธรรมการวิจัยในมนุษย์**

คณะกรรมการพิจารณาจริยธรรมการวิจัยในมนุษย์ วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา ได้พิจารณาแล้วเห็นว่า เค้าโครงดุษฎีนิพนธ์ดังกล่าวเป็นไปตามหลักการของจริยธรรมการวิจัยในมนุษย์ โดยที่ผู้วิจัยเคารพสิทธิและศักดิ์ศรีในความเป็นมนุษย์ ไม่มีการล่วงละเมิดสิทธิ สวัสดิภาพ และไม่ก่อให้เกิดภยันตรายแก่ตัวอย่างการวิจัย กลุ่มตัวอย่าง และผู้เข้าร่วมในโครงการวิจัย

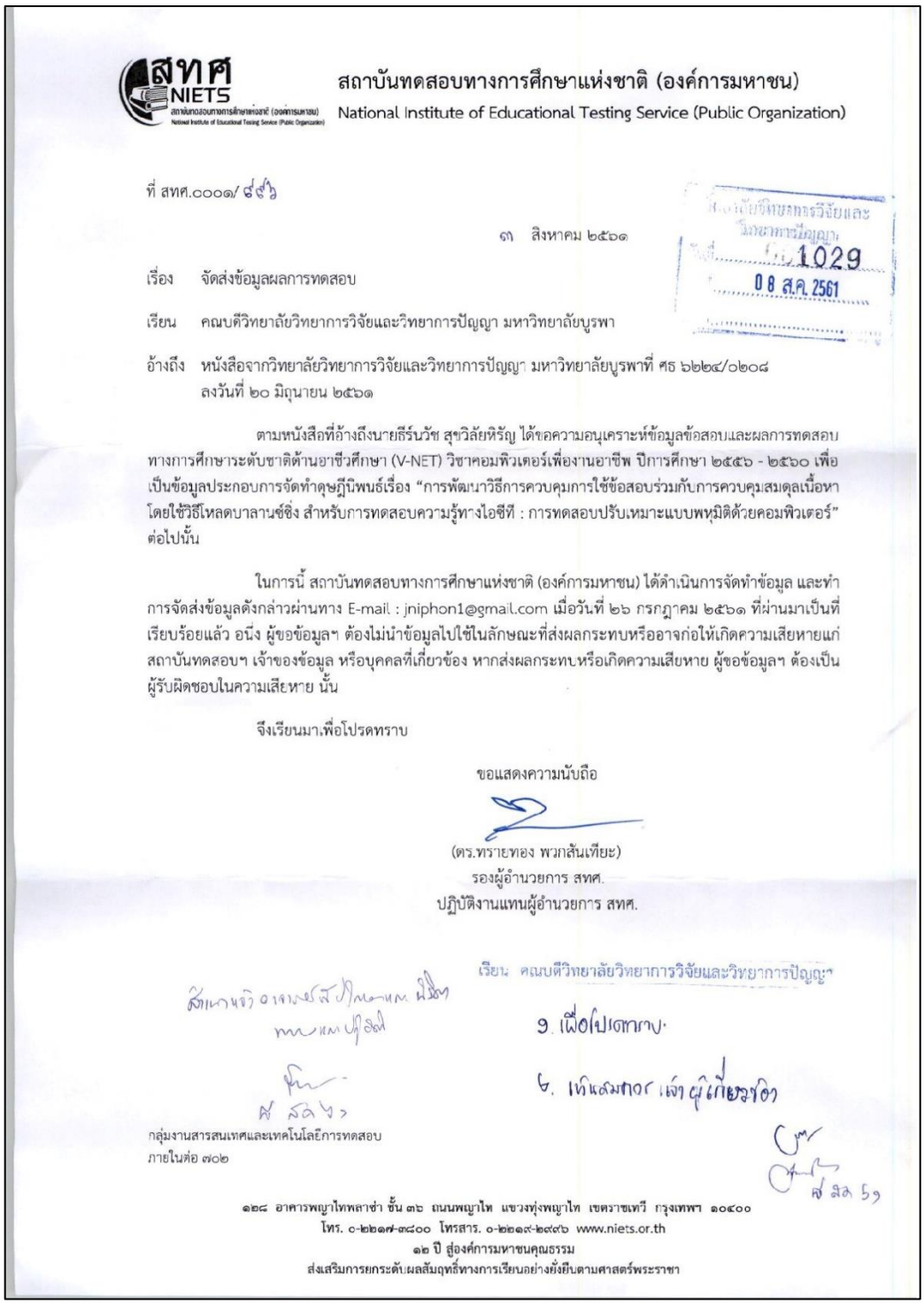
จึงเห็นสมควรให้ดำเนินการวิจัยในขอบข่ายของเค้าโครงดุษฎีนิพนธ์ที่เสนอได้ ตั้งแต่วันที่ออกเอกสารรับรองผลการพิจารณาจริยธรรมการวิจัยในมนุษย์ฉบับนี้ จนถึงวันที่ ๓๐ ธันวาคม พ.ศ. ๒๕๖๑

ออกให้ ณ วันที่ ๓๐ ตุลาคม พ.ศ. ๒๕๖๑

ลงนาม

(ผู้ช่วยศาสตราจารย์ ดร.ภัทราวดี มากมี)
คณบดีวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา
มหาวิทยาลัยบูรพา

ภาคผนวก ฉ
คลังข้อสอบความรู้ทางไอซีที



สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)
National Institute of Educational Testing Service (Public Organization)

ที่ สทศ.๐๐๐๑/ ๘๙๖

๓ สิงหาคม ๒๕๖๑

มหาวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา
รับที่ 1029
๐๘ ส.ค. 2561

เรื่อง จัดส่งข้อมูลผลการทดสอบ

เรียน คณบดีวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

อ้างถึง หนังสือจากวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพาที่ ศธ ๖๒๒๔/๐๒๐๘ ลงวันที่ ๒๐ มิถุนายน ๒๕๖๑

ตามหนังสือที่อ้างถึงนายธีร์นวัช สุขวิลัยหิรัญ ได้ขอความอนุเคราะห์ข้อมูลข้อสอบและผลการทดสอบทางการศึกษาระดับชาติด้านอาชีวศึกษา (V-NET) วิชาคอมพิวเตอร์เพื่องานอาชีพ ปีการศึกษา ๒๕๕๖ - ๒๕๖๐ เพื่อเป็นข้อมูลประกอบการจัดทำดุษฎีนิพนธ์เรื่อง "การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ชิ่ง สำหรับการทดสอบความรู้ทางไอซีที : การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์" ต่อไปนั้น

ในการนี้ สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ได้ดำเนินการจัดทำข้อมูล และทำการจัดส่งข้อมูลดังกล่าวผ่านทาง E-mail : jniphon1@gmail.com เมื่อวันที่ ๒๖ กรกฎาคม ๒๕๖๑ ที่ผ่านมาเป็นที่เรียบร้อยแล้ว อนึง ผู้ขอข้อมูลฯ ต้องไม่นำข้อมูลไปใช้ในลักษณะที่ส่งผลกระทบหรืออาจก่อให้เกิดความเสียหายแก่สถาบันทดสอบฯ เจ้าของข้อมูล หรือบุคคลที่เกี่ยวข้อง หากส่งผลกระทบหรือเกิดความเสียหาย ผู้ขอข้อมูลฯ ต้องเป็นผู้รับผิดชอบในความเสียหาย นั้น

จึงเรียนมาเพื่อโปรดทราบ

ขอแสดงความนับถือ

(ดร.ทรายทอง พวกสันเทียะ)
รองผู้อำนวยการ สทศ.
ปฏิบัติงานแทนผู้อำนวยการ สทศ.

สำเนาหนังสือ อาจารย์ ปิยนาถ นิยมค้า

ส ลล ๖๑

กลุ่มงานสารสนเทศและเทคโนโลยีการทดสอบ
ภายในต่อ ๗๐๒

เรียน คณบดีวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา

๑. เพื่อโปรดทราบ

๒. เห็นสมควรเก็บรวบรวม

ส ลล ๖๑

๑๒๘ อาคารพญาไทพลาซ่า ชั้น ๓๖ ถนนพญาไท แขวงทุ่งพญาไท เขตราชเทวี กรุงเทพฯ ๑๐๔๐๐
โทร. ๐-๒๒๑๗-๓๘๐๐ โทรสาร. ๐-๒๒๑๙-๒๙๙๖ www.niets.or.th
๑๒ ปี สู่องค์การมหาชนคุณธรรม
ส่งเสริมการยกระดับผลสัมฤทธิ์ทางการเรียนอย่างยั่งยืนตามศาสตร์พระราชา

ตารางที่ ฉ-1 คลังข้อสอบความรู้ทางไอซีที ที่ผ่านการวิเคราะห์คุณภาพตามเกณฑ์

ข้อที่	พารามิเตอร์		
	a	b	c
1	0.55	-0.66	0.25
2	0.59	0.83	0.09
3	1.99	-1.29	0.27
4	1.76	0.18	0.22
5	0.67	0.52	0.11
6	0.61	1.44	0.17
7	0.97	0.79	0.19
8	1.56	-0.67	0.23
9	0.76	0.45	0.14
10	0.82	-0.24	0.05
11	0.92	-0.10	0.07
12	0.87	0.66	0.03
13	0.79	-0.61	0.13
14	1.74	0.51	0.21
15	0.89	1.46	0.10
16	0.88	0.81	0.12
17	0.74	1.26	0.18
18	0.73	1.11	0.04
19	1.14	0.16	0.03
20	0.68	-1.14	0.12
⋮	⋮	⋮	⋮
192	0.52	-0.74	0.01

ภาคผนวก ช
ผลการทดสอบวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหา

ตารางที่ ช-1 ผลการทดสอบรายบุคคล ด้านจำนวนข้อสอบที่ใช้ในการทดสอบ และ
ด้านเวลาที่ใช้ในการทดสอบ ระหว่างวิธี LB กับ วิธี IE
ตารางที่ ช-2 ผลอัตราการแสดงการใช้ข้อสอบสูงสุด และสัดส่วนข้อสอบที่มีการแสดง
มากเกินไป ระหว่างวิธี LB กับ วิธี IE
ตารางที่ ช-3 ผลค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ระหว่างวิธี LB กับ วิธี IE

ที่ ศธ ๖๒๒๔/๐๔๐๗

วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา
มหาวิทยาลัยบูรพา
ต.แสนสุข อ.เมือง จ.ชลบุรี ๒๐๑๓๑

๒๔ ตุลาคม ๒๕๖๑

เรื่อง ขอความอนุเคราะห์เก็บรวบรวมข้อมูลเพื่อการวิจัย

เรียน ผู้อำนวยการวิทยาลัยเทคโนโลยีภาคตะวันออก (อี.เทค)

สิ่งที่ส่งมาด้วย เค้าโครงดุษฎีนิพนธ์ และเครื่องมือ จำนวน ๑ ชุด

ด้วย นายธีร์นวัช สุขวิลัยหิรัญ รหัสประจำตัวนิสิต ๕๕๘๑๐๒๕๖ นิสิตหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาการวัดและเทคโนโลยีทางวิทยาการปัญญา ได้รับอนุมัติให้ทำดุษฎีนิพนธ์เรื่อง "การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ซิ่ง สำหรับทดสอบความรู้ทางไอซีที: การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์" ซึ่งอยู่ในความควบคุมดูแลของ รองศาสตราจารย์ ดร.เสรี ชัดแช้ม อาจารย์ที่ปรึกษาหลัก ในการนี้ ผู้วิจัยมีความประสงค์ขอความอนุเคราะห์เก็บรวบรวมข้อมูลจากนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) จำนวน ๓ ประเภทวิชา ได้แก่ ๑) บริหารธุรกิจ ๒) เทคโนโลยีสารสนเทศ และ ๓) ช่างอุตสาหกรรม ที่เคยลงทะเบียนเรียนวิชาเทคโนโลยีสารสนเทศเบื้องต้น หรือวิชาเทคโนโลยีสารสนเทศและการสื่อสาร หรือวิชาคอมพิวเตอร์ โดยจะขอเก็บรวบรวมข้อมูล ระหว่างวันที่ ๑-๒ พฤศจิกายน พ.ศ. ๒๕๖๑ เพื่อนำไปใช้ในงานดุษฎีนิพนธ์ของนิสิตครั้งนี้

จึงเรียนมาเพื่อโปรดพิจารณา วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา หวังเป็นอย่างยิ่งว่าคงจะได้รับความอนุเคราะห์จากท่านด้วยดี และขอขอบคุณมา ณ โอกาสนี้

ขอแสดงความนับถือ

(ผู้ช่วยศาสตราจารย์ ดร.ภัทราวดี มากมี)
คณบดีวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา

วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา
โทร. ๐ ๓๘๑๐ ๒๐๗๗-๘
โทร/ โทรสาร ๐ ๓๘๓๙ ๓๔๘๔
http ://www.rmcs.buu.ac.th

ใบยินยอมเข้าร่วมการวิจัย

หัวข้อดุษฎีนิพนธ์เรื่อง การพัฒนาวิธีการควบคุมการใช้ข้อสอบร่วมกับการควบคุมสมดุลเนื้อหาโดยใช้วิธีโหลดบาลานซ์ชิ่ง สำหรับทดสอบความรู้ทางไอซีที: การทดสอบปรับเหมาะแบบพหุมิติด้วยคอมพิวเตอร์

วันให้คำยินยอม วันที่ 6 เดือน พฤศจิกายน พ.ศ. 2561

ก่อนที่จะลงนามในใบยินยอมเข้าร่วมการวิจัยนี้ ข้าพเจ้าได้รับการอธิบายจากผู้วิจัยถึงวัตถุประสงค์ของการวิจัย วิธีการวิจัย ประโยชน์ที่จะเกิดขึ้นจากการวิจัยอย่างละเอียดและมีความเข้าใจดีแล้ว ข้าพเจ้ายินดีเข้าร่วมโครงการวิจัยนี้ด้วยความสมัครใจ และข้าพเจ้ามีสิทธิที่จะบอกเลิกการเข้าร่วมในโครงการวิจัยนี้เมื่อใดก็ได้ และการบอกเลิกการเข้าร่วมการวิจัยนี้ จะไม่มีผลกระทบใด ๆ ต่อข้าพเจ้า

ผู้วิจัยรับรองว่าจะตอบคำถามต่าง ๆ ที่ข้าพเจ้าสงสัยด้วยความเต็มใจ ไม่ปิดบัง ซ่อนเร้น จนข้าพเจ้าพอใจ ข้อมูลเฉพาะเกี่ยวกับตัวข้าพเจ้าจะถูกเก็บเป็นความลับและจะเปิดเผยในภาพรวมที่เป็นการสรุปผลการวิจัย

ข้าพเจ้าได้อ่านข้อความข้างต้นแล้ว และมีความเข้าใจดีทุกประการ และได้ลงนามในใบยินยอมนี้ด้วยความเต็มใจ

ลงนาม.....ผู้ยินยอม
(ดร.ประเสริฐ กลิ่นชู)
ผู้อำนวยการ วิทยาลัยเทคโนโลยีภาคตะวันออก (อ.เทค)

ลงนาม.....พยาน
(อ.สุทิน พยนต์เลิศ)
หัวหน้าภาคคอมพิวเตอร์ธุรกิจ

ลงนาม.....ผู้ทำวิจัย
(นายธีร์นวัช สุขวิลัยหิรัญ)

ตารางที่ ช-1 ผลการทดสอบรายบุคคล ด้านจำนวนข้อสอบที่ใช้ในการทดสอบ และด้านเวลาที่ใช้ในการทดสอบ ระหว่างวิธี LB กับ วิธี IE

คนที่	ค่าประมาณความสามารถ		จำนวนข้อสอบที่ใช้		เวลาที่ใช้ (วินาที)	
	LB	IE	LB	IE	LB	IE
1	0.08	0.11	14	14	244.80	282.00
2	0.08	-0.19	13	15	415.20	280.80
3	0.14	-0.09	13	14	264.00	211.20
4	0.11	0.05	14	13	358.20	313.20
5	0.23	-0.26	14	16	291.00	160.80
6	0.08	0.11	13	15	325.20	351.00
7	0.08	0.22	14	14	343.80	393.00
8	0.00	0.23	14	15	114.00	271.20
9	0.12	0.08	12	14	336.00	339.00
10	0.08	0.14	14	15	273.00	444.00
11	0.18	0.14	13	14	412.80	192.00
12	0.08	0.03	12	14	307.80	346.20
13	0.13	0.22	14	15	244.80	337.20
14	0.23	0.01	12	13	364.80	393.00
15	0.14	0.22	13	15	324.00	450.00
16	0.12	0.08	12	14	333.00	279.00
17	0.13	0.22	14	14	300.00	454.80
18	0.14	0.11	13	15	325.80	397.20
19	0.18	0.06	13	13	282.00	264.00
20	0.02	0.12	12	16	454.80	420.00
21	0.15	0.13	13	14	390.00	325.20
22	0.08	0.00	14	13	444.00	393.00
⋮	⋮	⋮	⋮	⋮	⋮	⋮
400	0.14	0.15	14	16	208.20	256.80
Median			13	14	-	-
Mean			-	-	243.74	257.18

ตารางที่ ช-2 ผลอัตราการแสดงการใช้ข้อสอบสูงสุด และสัดส่วนข้อสอบที่มีการแสดงมากเกินไป ระหว่างวิธี LB กับ วิธี IE

ข้อที่	พารามิเตอร์			จำนวนข้อสอบใช้		อัตราการใช้ข้อสอบ	
	a	b	c	IE	LB	IE	LB
1	0.55	-0.66	0.25	33	45	0.08	0.11
2	0.59	0.83	0.09	30	26	0.08	0.07
3	1.99	-1.29	0.27	19	20	0.05	0.05
4	1.76	0.18	0.22	35	33	0.09	0.08
5	0.67	0.52	0.11	28	24	0.07	0.06
6	0.61	1.44	0.17	25	25	0.06	0.06
7	0.97	0.79	0.19	33	44	0.08	0.11
8	1.56	-0.67	0.23	34	52	0.09	0.13
9	0.76	0.45	0.14	37	35	0.09	0.09
10	0.82	-0.24	0.05	35	34	0.09	0.09
11	0.92	-0.10	0.07	36	47	0.09	0.12
12	0.87	0.66	0.03	29	27	0.07	0.07
13	0.79	-0.61	0.13	37	30	0.09	0.08
14	1.74	0.51	0.21	28	23	0.07	0.06
15	0.89	1.46	0.10	21	19	0.05	0.05
16	0.88	0.81	0.12	30	19	0.08	0.05
17	0.74	1.26	0.18	20	14	0.05	0.04
18	0.73	1.11	0.04	21	21	0.05	0.05
19	1.14	0.16	0.03	34	24	0.09	0.06
20	0.68	-1.14	0.12	25	19	0.06	0.05
21	1.81	0.19	0.21	33	37	0.08	0.09
22	0.56	-1.15	0.08	26	19	0.07	0.05
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
192	0.52	-0.74	0.01	27	34	0.07	0.09
Median				13	14	-	-
MIER				-	-	0.12	0.19
Ko=0				-	-	0.02	0.02

ตารางที่ ช-3 ผลค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ระหว่างวิธี LB กับ วิธี IE

คนที่	LB: ค่าประมาณความสามารถ ($\hat{\theta}$) กับ ค่าความสามารถที่แท้จริง (θ)					IE: ค่าประมาณความสามารถ ($\hat{\theta}$) กับ ค่าความสามารถที่แท้จริง (θ)				
	หมวด1	หมวด2	หมวด3	หมวด 4	หมวด5	หมวด1	หมวด2	หมวด3	หมวด4	หมวด5
1	-0.61	0.41	1.77	0.42	0.17	-0.54	0.34	1.47	0.23	0.80
	(2.85)	(0.34)	(-0.30)	(0.39)	(0.99)	(2.85)	(0.34)	(-0.30)	(0.39)	(0.99)
2	-0.35	0.36	1.36	0.44	0.20	-0.38	0.36	1.44	0.24	0.88
	(-0.15)	(-0.27)	(-2.52)	(0.19)	(0.50)	(-0.15)	(-0.27)	(-2.52)	(0.19)	(0.50)
3	-0.47	0.41	1.40	0.31	0.17	-0.56	0.35	1.60	0.25	0.18
	(0.65)	(1.57)	(-1.47)	(-0.72)	(-1.10)	(0.65)	(1.57)	(-1.47)	(-0.72)	(-1.10)
4	-0.50	0.33	1.61	0.28	0.15	-0.61	0.47	0.96	0.19	0.77
	(0.15)	(-0.05)	(2.07)	(-0.39)	(1.65)	(0.15)	(-0.05)	(2.07)	(-0.39)	(1.65)
5	-0.38	0.36	1.31	0.36	0.20	-0.56	0.35	1.60	0.25	0.88
	(-0.65)	(-0.32)	(0.36)	(-0.57)	(-1.76)	(-0.65)	(-0.32)	(0.36)	(-0.57)	(-1.76)
6	-0.51	0.37	1.41	0.39	0.15	-0.47	0.30	1.36	0.19	0.67
	(-0.82)	(-0.56)	(0.56)	(0.17)	(0.33)	(-0.82)	(-0.56)	(0.56)	(0.17)	(0.33)
7	-0.40	0.38	1.36	0.43	0.20	-0.51	0.36	1.13	0.19	0.78
	(-1.01)	(-0.44)	(-1.27)	(-0.66)	(0.67)	(-1.01)	(-0.44)	(-1.27)	(-0.66)	(0.67)
8	-0.31	0.35	1.22	0.40	0.20	-0.55	0.32	1.68	0.22	0.80
	(1.01)	(-0.12)	(-0.41)	(0.27)	(0.15)	(1.01)	(-0.12)	(-0.41)	(0.27)	(0.15)
9	-0.33	0.34	1.28	0.41	0.20	-0.56	0.35	1.77	0.26	0.55
	(-0.56)	(-2.04)	(-2.02)	(-0.85)	(-0.96)	(-0.56)	(-2.04)	(-2.02)	(-0.85)	(-0.96)
10	-0.35	0.36	1.36	0.44	0.20	-0.36	0.34	1.44	0.25	0.88
	(-0.97)	(0.87)	(-1.60)	(0.59)	(0.25)	(-0.97)	(0.87)	(-1.60)	(0.59)	(0.25)
11	-0.37	0.37	1.46	0.47	0.56	-0.38	0.36	1.44	0.24	0.88
	(-0.17)	(-0.17)	(-1.54)	(-0.17)	(-0.21)	(-0.17)	(-0.17)	(-1.54)	(-0.17)	(-0.21)
12	-0.36	0.38	1.36	0.48	0.48	-0.56	0.35	1.60	0.25	0.88
	(-0.68)	(1.64)	(-1.91)	(-2.06)	(0.20)	(-0.68)	(1.64)	(-1.91)	(-2.06)	(0.20)
13	-0.38	0.38	1.48	0.46	0.70	-0.40	0.35	1.50	0.22	0.63
	(-0.11)	(1.13)	(-1.40)	(0.19)	(-0.32)	(-0.11)	(1.13)	(-1.40)	(0.19)	(-0.32)
14	-0.41	0.38	1.52	0.44	0.79	-0.41	0.38	1.50	0.22	0.77
	(-1.87)	(-1.16)	(-1.25)	(0.95)	(0.51)	(-1.87)	(-1.16)	(-1.25)	(0.95)	(0.51)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
400	0.26	-0.06	0.33	-0.05	0.40	0.24	0.02	0.33	-0.07	0.22
	(-0.25)	(-0.35)	(-1.13)	(-1.08)	(0.80)	(-0.25)	(-0.35)	(-1.13)	(-1.08)	(0.80)
RMSE	1.07	1.11	1.58	0.94	1.88	1.15	1.12	1.57	0.95	1.71
Bias	0.19	0.13	0.59	0.14	0.79	0.32	0.15	0.57	0.10	0.68

ภาคผนวก ซ

ผลการทดสอบ MCAT กับ CT

ตารางที่ ซ-1 สรุปผลการทดสอบ MCAT กับ CT

คนที่	จำนวนข้อสอบที่ใช้ในการทดสอบ (ข้อ)		เวลาที่ใช้ในการทดสอบ (นาที: วินาที)		ผลการทดสอบ	
	MCAT	CT	MCAT	CT	MCAT ($\hat{\theta}$)	CT (คะแนนรวม)
1	14	40	196.80	1,093.80	0.24	27
2	17	40	146.40	1,408.20	0.68	26
3	15	40	312.00	1,024.20	0.12	23
4	11	40	305.40	1,415.40	0.19	25
5	13	40	211.20	1,350.00	0.28	26
6	16	40	187.80	1,828.20	0.74	22
7	13	40	318.00	1,380.00	0.14	23
8	14	40	249.60	1,144.80	0.15	21
9	13	40	69.00	76.20	-1.54	16
10	16	40	155.40	721.20	-0.14	17
11	15	40	325.80	2,059.20	0.22	25
12	14	40	190.80	1,636.80	-0.19	22
13	16	40	328.80	1,233.60	-0.10	17
14	16	40	75.00	855.00	-0.05	14
15	13	40	316.20	387.60	0.03	22
16	12	40	243.00	2,127.00	0.22	26
17	14	40	362.40	991.20	0.26	24
18	16	40	256.80	983.40	0.13	20
19	15	40	208.20	1,512.00	-0.03	16
20	14	40	214.80	972.00	0.13	24
⋮	⋮	⋮	⋮	⋮	⋮	⋮
40	15	40	192.00	1,503.60	0.20	25
Mean	14.68	40	258.87	1,174.95	0.04	20.58
SD	1.54	0	128.55	422.55	0.41	4.37