



เทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรม



โดย
นางสาวนิตติมา ลักขณานุกรักษ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล แผน ก แบบ ก 2 ระดับปริญญามหาบัณฑิต

ภาควิชาคอมพิวเตอร์

มหาวิทยาลัยศิลปากร

ปีการศึกษา 2565

ลิขสิทธิ์ของมหาวิทยาลัยศิลปากร

เทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรม



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล แผน ก แบบ ก 2 ระดับปริญญามหาบัณฑิต
ภาควิชาคอมพิวเตอร์
มหาวิทยาลัยศิลปากร
ปีการศึกษา 2565
ลิขสิทธิ์ของมหาวิทยาลัยศิลปากร

THE DISTANCE - BASED SELECTION TECHNIQUE FOR CROSSOVER IN GENETIC
ALGORITHM



A Thesis Submitted in Partial Fulfillment of the Requirements
for Master of Science INFORMATION TECHNOLOGY AND DIGITAL INNOVATION

Department of COMPUTER SCIENCE

Silpakorn University

Academic Year 2022

Copyright of Silpakorn University

61318302 : เทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล แผน ก แบบ ก 2 ระดับปริญญาโท
คำสำคัญ : ขั้นตอนเชิงพันธุกรรม, การไขว้เปลี่ยน, การคัดเลือกโดยธรรมชาติ, ฟังก์ชันระยะทาง

นางสาว นิติมา ลักขณานุรักษ์: เทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนใน
ขั้นตอนเชิงพันธุกรรม อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก : ผู้ช่วยศาสตราจารย์ ดร. คทา ประดิษฐ์วงศ์

งานวิจัยนี้มีจุดมุ่งหมายเพื่อศึกษาและพัฒนาเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรม ด้วยเทคนิคการคัดเลือก 3 เทคนิค ได้แก่ เทคนิคการเลือกตามด้วยระยะทางแฮมมิง (HS) เทคนิคการเลือกตามด้วยสหสัมพันธ์โคไซน์ (CS) และเทคนิคการเลือกตามด้วยสหสัมพันธ์เพียร์สัน (PS) ซึ่งจะนำไปทดลองเปรียบเทียบกับเทคนิคการเลือกแบบดั้งเดิม 2 เทคนิค ได้แก่ การคัดเลือกแบบวงล้อรูเล็ต (Roulette Wheel Selection) การคัดเลือกแบบการแข่งขัน (Tournament Selection) ในการทำการทดลองครั้งนี้ เทคนิคการคัดเลือกทั้งหมด จะได้รับการประเมินด้วย 4 ปัญหา ได้แก่ ปัญหาบิตหนึ่งมากที่สุด (One max problem) ปัญหาบิตศูนย์มากที่สุด (Zero max problem) ปัญหาบิตสูงสุดที่เกิดจากการสุ่ม (Random max problem) และปัญหากับดัก (Trap problem) โดยมีเกณฑ์การวัดผลคือจำนวนรุ่นที่พบคำตอบ และค่าความเหมาะสม เมื่อไม่พบคำตอบที่ถูกต้อง จากการทดลองได้แบ่งแนวทางออกเป็น 2 กลุ่มตามลักษณะของปัญหา สำหรับปัญหากับดักเทคนิคการเลือกตามระยะทาง สามารถช่วยให้หลุดจากจุดเหมาะสมวงแคบ (local optima) ซึ่งสามารถช่วยเพิ่มประสิทธิภาพในการหาคำตอบได้ และสำหรับปัญหาอื่นๆ เทคนิคการคัดเลือกแบบประลองเป็นเทคนิคที่เหมาะสมกว่าเทคนิคอื่นๆ

61318302 : Major INFORMATION TECHNOLOGY AND DIGITAL INNOVATION

Keyword : genetic algorithm, crossover, natural selection, distance function

MISS Nitima LUKKANANURUK : The Distance - Based Selection Technique for Crossover in Genetic Algorithm Thesis advisor : Assistant Professor kata praditwong, Ph.D.

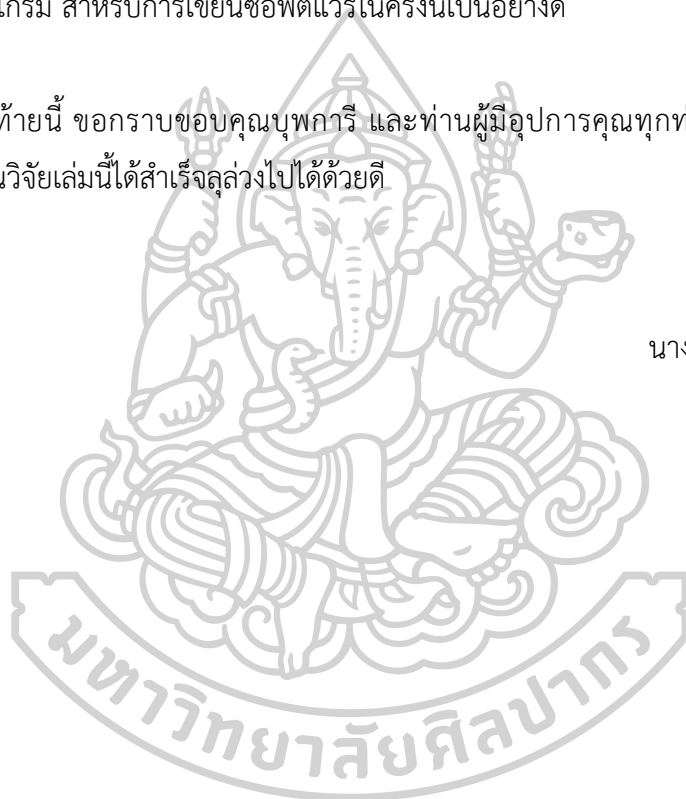
The aim of this research is to study and develop the natural inspired parent selections for the crossover operator in genetic algorithms. There are three distance-based methods of mating selection: the hamming distance-based selection (HS), the cosine coefficient distance-based selection (CS), and the Pearson coefficient distance-based selection (PS). The experiment conducts the comparison of the distance-based selection methods with two traditional selections: the roulette wheel selection (RWS) and the tournament selection (TS). In the experiment, all selection methods are evaluated based on four binary testing problems: one-max, zero-max, random-max, and two trap problems. The measurement criterion is the number of generations when the answer is found and the fitness values when the correct answer is not found. From the experimental results, the suitable approaches are divided into two groups according to the characteristics of the benchmark problems. For the trap problem with many local optima, the distance-based selection methods outperformed the traditional selection. However, for the other benchmark problems, the tournament selection is the better method than others.

กิตติกรรมประกาศ

วิทยานิพนธ์นี้สามารถสำเร็จลุล่วงได้ด้วยดี ด้วยความช่วยเหลือ การให้ความรู้ ให้คำปรึกษา คำแนะนำต่างๆ ที่เป็นประโยชน์ ตลอดจนส่งเสริมผลักดัน แก่ไขข้อบกพร่องต่างๆ ในงานวิจัยนี้ ด้วยความละเอียดถี่ถ้วนจาก ผู้ช่วยศาสตราจารย์ ดร.คทา ประดิษฐ์วงศ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ และ ผู้ช่วยศาสตราจารย์ ดร.สุพจน์ เฮงพระพรหม อาจารย์ประจำสาขาวิชาวิทยาการข้อมูล คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏนครปฐม และขอขอบคุณ นายณัฐ โปธิ์สุวรรณ ที่ได้ให้คำแนะนำการเขียนโค้ดโปรแกรม สำหรับการเขียนซอฟต์แวร์ในครั้งนี้เป็นอย่างดี

สุดท้ายนี้ ขอกราบขอบคุณบุพการี และท่านผู้มีอุปการคุณทุกท่านที่ให้การสนับสนุนเป็นกำลังใจงานวิจัยเล่มนี้ได้สำเร็จลุล่วงไปได้ด้วยดี

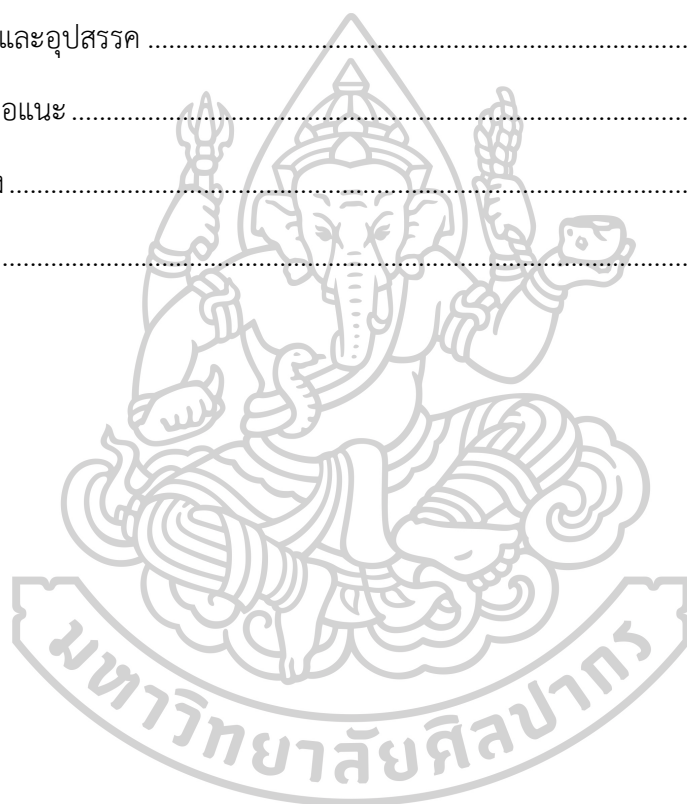
นางสาว นิติมา ลักขณานุรักษ์



สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฌ
สารบัญรูปภาพ.....	ญ
บทที่ 1 บทนำ.....	1
1.1 ความสำคัญและที่มาของปัญหาการวิจัย.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	2
1.3 ขอบเขตของการวิจัย.....	2
1.4 กรอบแนวคิดในการวิจัย.....	2
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	5
2.1 ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm : GA).....	5
2.2 ระยะทางแฮมมิง (Hamming Distance).....	13
2.3 สหสัมพันธโคไซน์ (Cosine Coefficient).....	14
2.4 สหสัมพันธเพียร์สัน (Pearson Coefficient).....	14
2.6 สรุปรงานวิจัยที่เกี่ยวข้อง.....	17
บทที่ 3 วิธีดำเนินการวิจัย.....	20
3.1 ปัญหาที่นำมาใช้ในงานวิจัย.....	20
3.2 การออกแบบการทดลอง.....	22
3.3 ขั้นตอนการทดลอง.....	26

3.4 เครื่องมือและอุปกรณ์ที่ใช้ในการวิจัย	28
บทที่ 4 ผลการวิจัย	29
4.1 ผลการทดสอบประสิทธิภาพเทคนิคการเลือกตามระยะทาง	29
4.2 สรุป	48
บทที่ 5 สรุป อภิปรายผล และข้อเสนอแนะ	49
5.1 สรุปผลและอภิปรายผล	49
5.2 ปัญหาและอุปสรรค	50
5.3 ข้อเสนอแนะ	50
รายการอ้างอิง	51
ประวัติผู้เขียน	53



สารบัญตาราง

	หน้า
ตารางที่ 1 ตัวอย่างการคัดเลือกแบบการแข่งขัน.....	9
ตารางที่ 2 ตัวอย่างการคัดเลือกแบบการแข่งขันโดยไม่ลำเอียง.....	9
ตารางที่ 3 สรุปงานวิจัยที่เกี่ยวข้อง	18
ตารางที่ 4 ค่าความเหมาะสมจากกรณีขนาดของกบดักเท่ากับ 4.....	21
ตารางที่ 5 ค่าความเหมาะสมจากกรณีขนาดของกบดักเท่ากับ 4.....	21
ตารางที่ 6 การตั้งค่าพารามิเตอร์ (Parameter Setting).....	27
ตารางที่ 7 ผลการทดสอบการคัดเลือกด้วยเทคนิคแบบประลอง	30
ตารางที่ 8 ผลการทดสอบการคัดเลือกด้วยเทคนิควงล้อรูเล็ก.....	32
ตารางที่ 9 ผลการทดสอบการคัดเลือกด้วยเทคนิควงล้อรูเล็ก.....	34
ตารางที่ 10 ผลการทดสอบการคัดเลือกด้วยเทคนิคสหสัมพันธ์โคไซน์.....	36
ตารางที่ 11 ผลการทดสอบการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน	38
ตารางที่ 12 จำนวนครั้งที่พบคำตอบของการคัดเลือกด้วยเทคนิคแบบประลอง.....	39
ตารางที่ 13 จำนวนครั้งที่พบคำตอบของการคัดเลือกด้วยเทคนิควงล้อรูเล็ก.....	39
ตารางที่ 14 จำนวนครั้งที่พบคำตอบของการคัดเลือกด้วยเทคนิคระยะทางแฮมมิง	40
ตารางที่ 15 จำนวนครั้งที่พบคำตอบของการคัดเลือกด้วยเทคนิคสหสัมพันธ์โคไซน์	40
ตารางที่ 16 จำนวนครั้งที่พบคำตอบของการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน.....	41
ตารางที่ 17 ผลการทดสอบเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิง พันธุกรรม (เฉพาะปัญหาบิตหนึ่งมากที่สุด ปัญหาบิตศูนย์มากที่สุด และปัญหาบิตสูงสุดจากการสุ่ม)	42
ตารางที่ 18 ผลการทดสอบเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิง พันธุกรรม (เฉพาะปัญหากบดัก ขนาดกบดัก 3 และขนาดกบดัก 5).....	43

สารบัญรูปภาพ

	หน้า
ภาพที่ 1 กรอบแนวคิดในการวิจัย (Conceptual Framework).....	2
ภาพที่ 2 แผนผังแสดงขั้นตอนการทำงานของขั้นตอนวิธีเชิงพันธุกรรม.....	6
ภาพที่ 3 วิธีการแบบวงล้อรูเล็ท	8
ภาพที่ 4 สายโครโมโซมสำหรับการไขว้เปลี่ยนแบบจุดเดียว	10
ภาพที่ 5 สายโครโมโซมที่ผ่านการไขว้เปลี่ยนแบบจุดเดียว	11
ภาพที่ 6 การกลายพันธุ์แบบจุดเดียว	12
ภาพที่ 7 การกลายพันธุ์แบบหลายจุด	12
ภาพที่ 8 การหาค่าระยะทางแฮมมิง	13
ภาพที่ 9 กราฟค่าความเหมาะสมจากกรณีขนาดของกบดักเท่ากับ 4	22
ภาพที่ 10 ตัวอย่างการคำนวณในเทคนิคการคัดเลือกด้วยระยะทางแฮมมิง (HS).....	23
ภาพที่ 11 ตัวอย่างการคำนวณในเทคนิคการคัดเลือกด้วยสหสัมพันธ์โคไซน์ (CS).....	24
ภาพที่ 12 ตัวอย่างการคำนวณในเทคนิคการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน (PS).....	26
ภาพที่ 13 ขั้นตอนการทดลองของเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรม.....	26
ภาพที่ 14 แผนภูมิเส้นเปรียบเทียบในปัญหาบิตหนึ่งมากที่สุด.....	44
ภาพที่ 15 แผนภูมิเส้นเปรียบเทียบในปัญหาบิตศูนย์มากที่สุด	45
ภาพที่ 16 แผนภูมิเส้นเปรียบเทียบในปัญหาบิตสูงสุดที่เกิดจากการสุ่ม	45
ภาพที่ 17 แผนภูมิเส้นเปรียบเทียบในปัญหากับดัก ขนาดกับดัก 3.....	46
ภาพที่ 18 แผนภูมิเส้นเปรียบเทียบในปัญหากับดัก ขนาดกับดัก 5.....	47

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มาของปัญหาการวิจัย

ปัญญาประดิษฐ์ (Artificial intelligence : AI) มีวัตถุประสงค์เพื่อต้องการพัฒนาระบบให้มีความสามารถและความฉลาดที่เท่าเทียมหรือมากกว่ามนุษย์ เช่น ด้านเสียงพูด ด้านการมองเห็น ด้านการเข้าใจในภาษา ด้านการวางแผนและการใช้เหตุผล ซึ่งสาขาที่ถูกลำดับให้มีความสำคัญสำหรับเทคโนโลยีปัญญาประดิษฐ์ ดังเช่น ศาสตร์การเรียนรู้ของเครื่อง (Machine Learning) ถือว่าเป็นกระบวนการที่มีความสำคัญอย่างมากในการเรียนรู้ (Learning) จากสถานการณ์ เพื่อให้เกิดกระบวนการวางแผนและการตัดสินใจได้ตั้งมนุษย์ ซึ่งเป็นศาสตร์ที่ทั่วโลกกำลังจับตามอง โดยที่การคำนวณเชิงวิวัฒนาการ (Evolutionary Computation) เป็นกระบวนการการหาผลลัพธ์ของปัญหาที่มีความซับซ้อน เช่น การหาอัตราส่วนของสูตรสารเคมีที่มีความเหมาะสมที่สุด การคำนวณเส้นทางการขนส่งสินค้าของบริษัทให้เกิดความคุ้มค่ามากที่สุด เป็นต้น ปัญหาเหล่านี้้นำการคำนวณเชิงวิวัฒนาการมาใช้ในหาคำตอบ การคำนวณเชิงวิวัฒนาการมีเทคนิคในการแก้ปัญหาด้วยวิธีต่างๆ หลากหลายวิธีซึ่งวิธีที่ได้รับความนิยมสูงมีการนำไปใช้อย่างแพร่หลาย คือ ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm : GA) เป็นขั้นตอนวิธีการปรับค่าให้เหมาะสม ด้วยกระบวนการตามหลักการของพันธุศาสตร์และการคัดเลือกโดยธรรมชาติ เป็นขั้นตอนวิธีที่ได้รับความนิยมวิธีหนึ่ง สำหรับใช้ในการหาคำตอบที่ดีที่สุดหรือใกล้เคียงที่สุด [1]

ขั้นตอนวิธีเชิงพันธุกรรม เริ่มต้นด้วยการสร้างประชากร (Population) สำหรับใช้ในการคัดเลือกสายพันธุ์ (Selection) โดยค่าความเหมาะสม (Fitness Value) ที่สูง จะมีโอกาสการถูกคัดเลือกมากกว่าค่าความเหมาะสมที่ต่ำ จะทำเป็นรุ่นๆ ในแต่ละรุ่นคำตอบที่ได้จะมีการเปลี่ยนแปลง ซึ่งเกิดการปฏิบัติการทางสายพันธุ์ (Genetic Operation) ได้แก่ การสืบพันธุ์ (Reproduction) การไขว้เปลี่ยน (Crossover) หรือการกลายพันธุ์ (Mutation) จนเกิดเป็นประชากรรุ่นใหม่ต่อไป ที่ให้ค่าความเหมาะสมที่ดียิ่งขึ้น การวิวัฒนาการจะถูกทำซ้ำตามกระบวนการข้างต้นจนกระทั่งได้คำตอบที่มีความเหมาะสมหรือจนกระทั่งพบเงื่อนไขของการสิ้นสุด

ดังนั้นสำหรับงานวิจัยนี้ผู้วิจัยได้มีแนวคิดสำหรับเทคนิคการเลือกตามระยะทางระหว่างสายโครโมโซม โดยใช้สายโครโมโซมที่ถูกคัดเลือกเป็นสายแรก มาคัดเลือกสายโครโมโซมอีกสาย [2] [3] เพื่อนำไปใช้ในกระบวนการไขว้เปลี่ยน สำหรับการหาคำตอบของปัญหา

1.2 วัตถุประสงค์ของการวิจัย

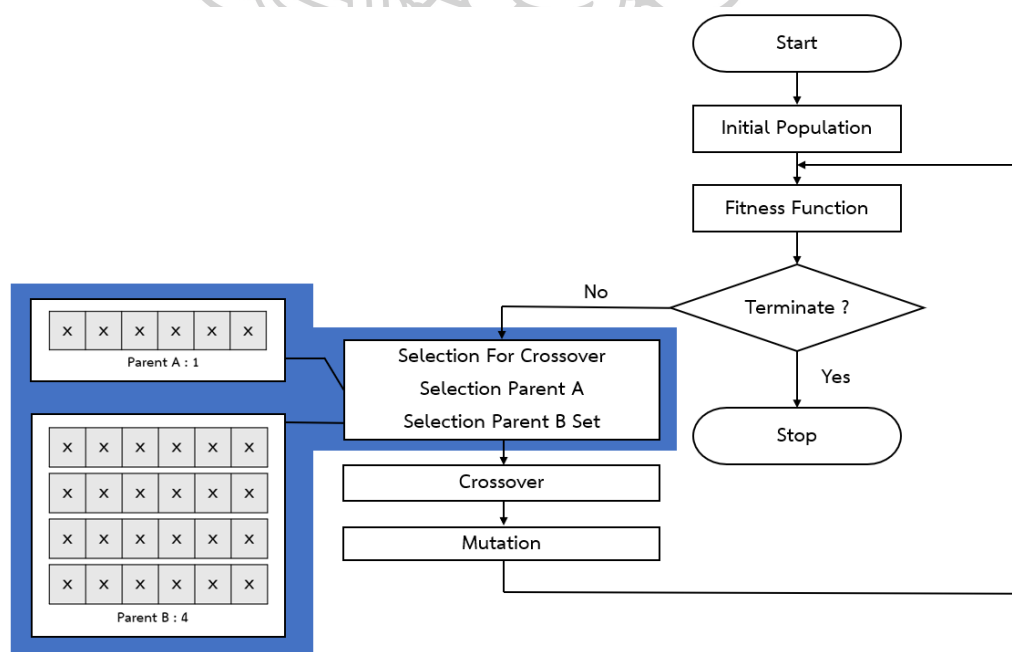
เพื่อศึกษาและพัฒนาเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรมที่มีประสิทธิภาพ

1.3 ขอบเขตของการวิจัย

งานวิจัยนี้ได้พัฒนาเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรม โดยนำไปทดสอบกับปัญหาต่างๆ ดังนี้ ปัญหาบิตหนึ่งมากที่สุด (One Max Problem) ปัญหาบิตศูนย์มากที่สุด (Zero Max Problem) ปัญหาบิตสูงสุดที่เกิดจากการสุ่ม (Random Max Problem) และปัญหากับดัก (Trap Problem) [4] ด้วยการวัดประสิทธิภาพจากคำตอบและระยะเวลาของการค้นหาคำตอบ

1.4 กรอบแนวคิดในการวิจัย

สำหรับการวิจัย เรื่องเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรม มีกรอบแนวคิดในการวิจัยประกอบด้วยขั้นตอนในการวิจัย ดังภาพที่ 1



ภาพที่ 1 กรอบแนวคิดในการวิจัย (Conceptual Framework)

กรอบแนวคิดในการวิจัยเรื่องเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรมที่นำเสนอใน ภาพที่ 1 ซึ่งเป็นการสรุปภาพรวมของการดำเนินงานสำหรับงานวิจัยในครั้งนี้ โดยในงานนี้จะนำเสนอในส่วนของกระบวนการคัดเลือกสำหรับการไขว้เปลี่ยน เมื่อกระบวนการเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรมได้เริ่มต้นขึ้น จะเกิดการสร้างประชากร (Initial Population) ที่เป็นสายโครโมโซมตามค่าพารามิเตอร์ที่กำหนดไว้ นำไปประเมินค่าความเหมาะสม (Fitness Function) ตามปัญหาที่กำหนด ตรวจสอบหาคำตอบจากสายโครโมโซมทั้งหมด โดยกระบวนการจะสิ้นสุดหากพบคำตอบหรือเงื่อนไขของการสิ้นสุด หากไม่พบคำตอบกระบวนการจะดำเนินการต่อ ในส่วนของการคัดเลือก (Selection) ดังนี้

ขั้นตอนที่ 1 เลือกสายโครโมโซม A (Parent A) ด้วยวิธีการแบบวงล้อรูเล็ต (Roulette Wheel Selection)

ขั้นตอนที่ 2 เลือกชุดสายโครโมโซม B (Parent B) ด้วยวิธีการแบบวงล้อรูเล็ต (สำหรับงานวิจัยนี้ กำหนดให้ 1 ชุด มีสายโครโมโซมจำนวน 4 สาย)

ขั้นตอนที่ 3 หาค่าระยะทางระหว่างสายโครโมโซม A กับ ชุดสายโครโมโซม B โดยใช้เทคนิคการเลือกตามด้วยระยะทางแฮมมิง (HS) เทคนิคการเลือกตามด้วยสหสัมพันธ์โคไซน์ (CS) และเทคนิคการเลือกตามด้วยสหสัมพันธ์เพียร์สัน (PS) สำหรับการหาค่าสัมบูรณ์ระหว่างสายโครโมโซม A กับชุดสายโครโมโซม B โดยจะหาทีละ 1 สาย

ขั้นตอนที่ 4 คัดเลือกสายโครโมโซม B ที่มีความแตกต่างกับสายโครโมโซม A มากที่สุด และส่งกลับเพื่อใช้ในการคลอสโอเวอร์

1.5 นิยามศัพท์

1.5.1 สายโครโมโซม (Chromosome String) หมายถึง สายบิตของคุณลักษณะ ซึ่งเป็นตัวแทนของคำตอบที่ต้องการ

1.5.2 ประชากร (Population) หมายถึง กลุ่มของสายโครโมโซม (Group of Chromosomes)

1.5.3 การสืบพันธุ์ (Reproduction) หมายถึง การคัดลอกประชากรที่ดีไปสร้างเป็นประชากรรุ่นถัดไป

1.5.4 การคลอสโอเวอร์ (Crossover) หมายถึง การนำ Parent 2 ตัว มาสร้างประชากรใหม่ เพื่อนำมาใช้เปลี่ยนให้เกิดประชากรรุ่นถัดไป

1.5.5 การกลายพันธุ์ (Mutation) หมายถึง การเปลี่ยนแปลงยีนสายโครโมโซม

1.5.6 ค่าความเหมาะสม (Fitness) หมายถึง ค่าที่แสดงระดับความเหมาะสมกับปัญหาของโครโมโซม

1.5.7 ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm: GA) หมายถึง ขั้นตอนวิธีการหาคำตอบของปัญหาความเหมาะสม ด้วยกระบวนการตามหลักการของพันธุศาสตร์และการคัดเลือกโดยธรรมชาติ

1.5.8 ระยะทางแฮมมิง (Hamming Distance) หมายถึง การหาจำนวนตำแหน่งที่มีความแตกต่างกัน ระหว่าง 2 สายโครโมโซมที่มีความยาวเท่ากัน

1.5.9 สหสัมพันธ์โคไซน์ (cosine correlation) หมายถึง การวัดความคล้ายคลึงระหว่าง 2 เวกเตอร์ โดยการวัดมุมโคไซน์ของเวกเตอร์ทั้งสอง

1.5.10 สหสัมพันธ์เพียร์สัน (Pearson correlation) หมายถึง ค่าที่แสดงถึงความสัมพันธ์ของ 2 ตัวแปร ค่าสัมประสิทธิ์จะมีค่าอยู่ระหว่าง -1.0 ถึง $+1.0$ โดยหากค่าเข้าใกล้ -1.0 แสดงให้เห็นว่าทั้งตัวแปรทั้งสองตัวมีความสัมพันธ์ในเชิงตรงกันข้าม หากเข้าใกล้ $+1.0$ แสดงให้เห็นว่าทั้งตัวแปรทั้งสองตัวมีความสัมพันธ์ไปในทิศทางเดียวกัน แต่หากตัวแปรทั้งสองตัวแปรมีค่าสัมประสิทธิ์เท่ากับ 0 แสดงว่าตัวแปรทั้งสองตัวไม่มีความสัมพันธ์กัน

1.6 ประโยชน์ของการวิจัย

1.6.1 สามารถเพิ่มความสามารถในการหาค่าความเหมาะสมของคำตอบ เพื่อนำไปสู่การประยุกต์ใช้ให้มีประสิทธิภาพที่ดียิ่งขึ้น

1.6.2 สามารถศึกษาความเป็นไปได้ของเทคนิคการเลือกตามระยะทาง ด้วยระยะทางแฮมมิง สหสัมพันธ์โคไซน์ และสหสัมพันธ์เพียร์สัน ในการคัดเลือกสายโครโมโซม B ที่มีความแตกต่างกับสายโครโมโซม A มากที่สุด เพื่อนำมาใช้สำหรับการไขว้เปลี่ยนสำหรับการหาคำตอบของปัญหา

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

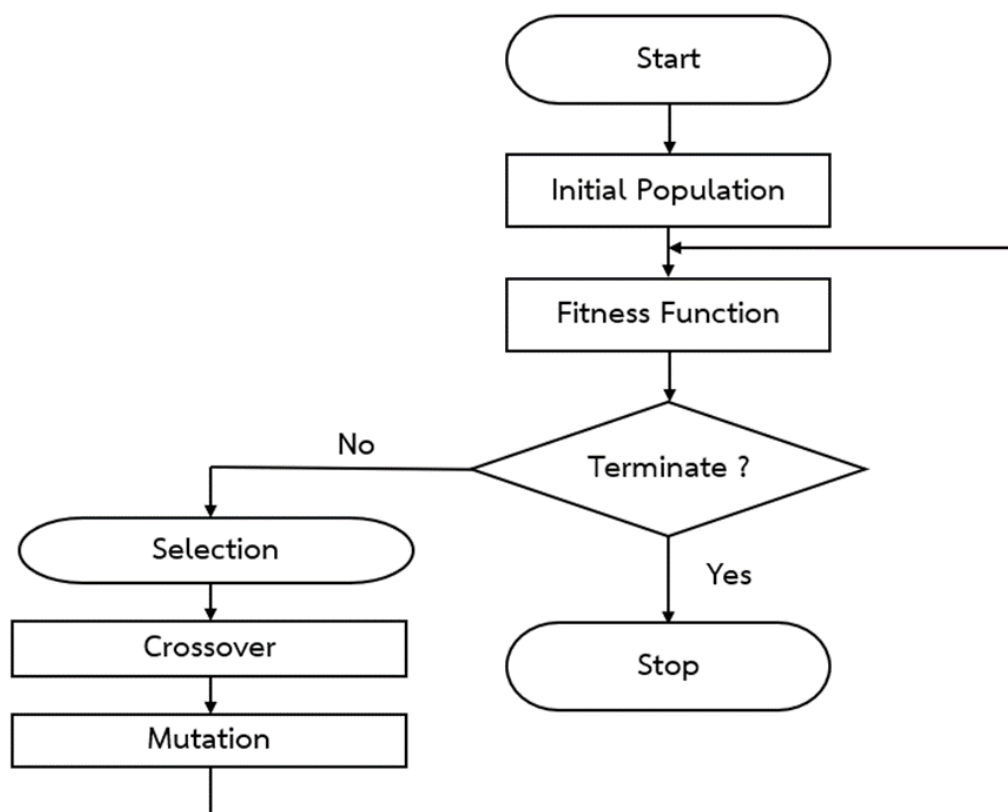
จากงานวิจัยเรื่องเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรม ผู้วิจัยได้นำเสนอเนื้อหาที่เน้นถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยมีรายละเอียดของเนื้อหาประกอบด้วยหัวข้อย่อย 3 หัวข้อ ดังนี้

- 2.1 ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm : GA)
- 2.2 ระยะทางแฮมมิง (Hamming Distance)
- 2.3 สหสัมพันธโคไซน์ (Cosine Coefficient)
- 2.4 สหสัมพันธเพียร์สัน (Pearson Coefficient)
- 2.5 งานวิจัยที่เกี่ยวข้อง
- 2.6 สรุปงานวิจัยที่เกี่ยวข้อง

2.1 ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm : GA)

ขั้นตอนวิธีเชิงพันธุกรรม เป็นเทคนิคทางด้านปัญญาประดิษฐ์ ที่มีความสำคัญสำหรับการใช้ในด้าน การเรียนรู้ ด้านการค้นหา และด้านการเพิ่มประสิทธิภาพ ด้วยการเรียนแบบกระบวนการตาม หลักการของพันธุศาสตร์และการคัดเลือกโดยธรรมชาติ มีกระบวนการปรับปรุงหรือเปลี่ยนแปลงบาง ส่วนประกอบแล้วสร้างขึ้นมาใหม่ โดยมีจุดเด่นในด้านการทนทานต่อความผิดพลาดในการหาผลลัพธ์ จากข้อมูลที่มีความยากและซับซ้อน ซึ่งกระบวนการค้นหาไม่ได้มีรูปแบบที่แน่นอนขึ้นอยู่กับ ลักษณะเฉพาะของข้อมูลหรือรูปแบบ ดังนั้นขั้นตอนวิธีเชิงพันธุกรรมจึงถูกนำไปประยุกต์ใช้ในการ แก้ปัญหาได้อย่างหลากหลาย เช่น การคำนวณหาเส้นทางการเดินทางสำหรับการส่งพัสดุให้มีความ คุ่มค่ามากที่สุด เป็นขั้นตอนวิธีที่ได้รับความนิยมวิธีหนึ่ง สำหรับใช้ในการหาคำตอบที่ดีที่สุดหรือ ใกล้เคียงที่สุด เพื่อแก้ปัญหาและลดความซับซ้อนของปัญหา ซึ่งถือได้ว่าเป็นเครื่องมือเชิงคำนวณอย่าง หนึ่ง

2.1.1 ขั้นตอนการทำงานของขั้นตอนวิธีเชิงพันธุกรรม



ภาพที่ 2 แผนผังแสดงขั้นตอนการทำงานของขั้นตอนวิธีเชิงพันธุกรรม

จากภาพที่ 2 แผนผังแสดงขั้นตอนการทำงานของขั้นตอนวิธีเชิงพันธุกรรม สามารถสรุปเป็นขั้นตอนการทำงานแบบละเอียดได้ดังนี้

2.1.1.1 สร้างประชากร (Initial Population)

ข้อมูลที่ถูกนำมาใช้ในขั้นตอนวิธีเชิงพันธุกรรม มีรูปแบบที่หลากหลาย ดังนั้นต้องเริ่มด้วยกระบวนการแปลงข้อมูลให้อยู่ในของอาร์เรย์ (Array) หรือสายโครโมโซมที่มีขนาดเท่ากัน [5] [6] โดยตัวแปรที่ต้องการค้นหาจะอยู่ในรูปของยีน เมื่อนำหลายยีนมารวมกันจะอยู่ในรูปของสายโครโมโซมซึ่งหมายถึงผลลัพธ์ของปัญหา โดยเป็นรูปแบบที่เหมาะสมสำหรับขั้นตอนวิธีเชิงพันธุกรรม การแปลงข้อมูลสามารถทำได้หลากหลายวิธี เช่น แปลงให้อยู่ในรูปของเลขฐานสอง หรือแปลงให้อยู่ในรูปของค่าจริง สามารถพิจารณาได้จากปัญหาที่ต้องการแก้ เช่น การกำหนดทิศทางการเคลื่อนที่ของหุ่นยนต์

เมื่อข้อมูลที่ถูกนำมาใช้ได้ถูกแปลงให้มีความเหมาะสมสำหรับขั้นตอนวิธีเชิงพันธุกรรมเรียบร้อยแล้ว จากนั้นจะมีการสุ่มโครโมโซมขึ้นมา เพื่อเป็นประชากรตั้งต้น โดยมีพารามิเตอร์ที่กำหนดขึ้น เป็นตัวกำหนดจำนวนของโครโมโซม ส่วนตัวแปรของปัญหาซึ่งแทนด้วยจำนวนยีนเป็นตัวกำหนดส่วนความยาวของโครโมโซม โดยมีผลต่อการหาค่าคำตอบที่ดีที่สุด

2.1.1.2 การประเมินค่าความเหมาะสม (Fitness Function)

ในขั้นตอนการประเมินค่าความเหมาะสม การเลือกฟังก์ชันมาใช้ในการประเมินค่าความเหมาะสมต้องพิจารณาด้วยความระมัดระวัง เนื่องจากฟังก์ชันที่นำมาใช้ จะไม่มีรูปแบบของฟังก์ชันที่ตายตัว โดยขึ้นอยู่กับลักษณะของปัญหาและรูปแบบของคำตอบ

ความสำคัญของค่าความเหมาะสมที่ต้องให้ค่าของความเหมาะสมในแต่ละโครโมโซม เพื่อใช้สำหรับการตัดสินใจคัดเลือกประชากรพ่อและประชากรแม่ในการเป็นประชากรตั้งต้น ในการเกิดของประชากรรุ่นใหม่ หากค่าของความเหมาะสมที่ต่ำ จะไม่ถูกคัดเลือก โดยจะคัดเลือก ค่าของความเหมาะสมที่สูง ให้มีโอกาสในการสืบทอดประชากรรุ่นถัดไป

2.1.1.3 การวิวัฒนาการ (Evolution)

การวิวัฒนาการเริ่มต้นกระบวนการด้วยการคัดเลือกสายพันธุ์ และดำเนินการต่อด้วยปฏิบัติการทางสายพันธุ์

2.1.1.3.1 การคัดเลือกสายพันธุ์ (Selection)

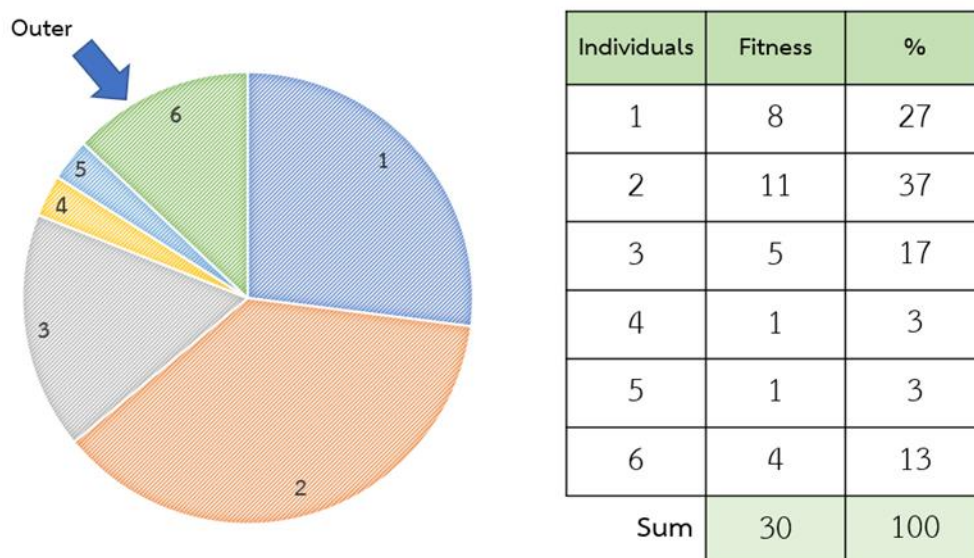
การคัดเลือกสายพันธุ์ โดยโครโมโซมที่มีค่าของความเหมาะสมที่สูง จะมีโอกาสการถูกคัดเลือกที่มากกว่า โครโมโซมที่มีค่าของความเหมาะสมที่ต่ำ เพื่อคัดเลือกโครโมโซมที่มีความเหมาะสมสำหรับการเป็นประชากรตั้งต้น วิธีการคัดเลือกโครโมโซมที่มีความเหมาะสมมีอยู่หลากหลายวิธี เช่น วิธีการแบบวงล้อรูเล็ต (Roulette Wheel Selection) วิธีการแบบประลอง (Tournament Selection)

- วิธีการแบบวงล้อรูเล็ต (Roulette Wheel Selection)

สำหรับวิธีการแบบวงล้อรูเล็ต ทุกโครโมโซมอาจมีโอกาสได้ถูกคัดเลือก โดยโครโมโซมที่มีค่าความเหมาะสมที่มากก็จะมีโอกาสการถูกคัดเลือกได้มาก [7] แต่หากโครโมโซมที่มีค่าความเหมาะสมที่น้อยก็จะมีโอกาสที่ถูกคัดเลือกได้น้อย โดยมีวิธีการกำหนดโอกาสของความน่าจะเป็น

ไว้สำหรับทุกโครโมโซม แต่ปริมาณของโอกาสของความน่าจะเป็นจะขึ้นอยู่กับค่าความเหมาะสมของแต่ละโครโมโซม

เมื่อทุกโครโมโซมถูกคำนวณค่าความน่าจะเป็นให้อยู่ในช่องขนาดของปริมาณความน่าจะเป็นแล้ว หลังจากนั้นจะเริ่มกระบวนการสุ่มเพื่อหาประชากรตั้งต้น โดยโครโมโซมที่มีขนาดช่องที่กว้างกว่าจะมีโอกาสถูกคัดเลือกมากกว่าโครโมโซมที่มีขนาดช่องที่เล็ก โดยการหมุน 1 ครั้ง จะเกิดการคัดเลือกโครโมโซมได้ 1 ตัว โดยจำนวนครั้งของรอบหมุนจะเท่ากับจำนวนของประชากรตั้งต้นที่ต้องการ ดังภาพที่ 3 วิธีการแบบวงล้อรูเล็ต



ภาพที่ 3 วิธีการแบบวงล้อรูเล็ต

จากภาพที่ 3 วิธีการแบบวงล้อรูเล็ต ในครั้งนี้มีจำนวนโครโมโซมทั้งสิ้น 6 ตัว โดยแต่ละโครโมโซมมีพื้นที่ขนาดของช่องที่ไม่เท่ากัน ขึ้นอยู่กับค่าความเหมาะสมของแต่ละโครโมโซม ซึ่งจากกรณีนี้ โครโมโซมที่น่าจะมีโอกาสถูกคัดเลือกมากที่สุด คือ โครโมโซมตัวที่ 2 ด้วยโอกาสการถูกคัดเลือกที่ 37%

- วิธีการแบบประลอง (Tournament Selection)

สำหรับวิธีการแบบประลอง วิธีนี้เริ่มต้นด้วยการเลือกประชากรแบบสุ่ม โดยต้องกำหนดขนาดหรือจำนวนให้น้อยกว่าจำนวนประชากรทั้งหมด และคัดเลือกประชากรที่มีค่าความเหมาะสมที่สุดจากกลุ่มของขนาดหรือจำนวน เพื่อใช้เป็นตัวแทน ประชากรตัวที่ผ่านการเปรียบเทียบแล้ว สามารถใส่คืนเพื่อรอการสุ่มใหม่

ตารางที่ 1 ตัวอย่างการคัดเลือกแบบการแข่งขัน

Population	1	2	3	4	5
Fitness	8	10	4	2	6

จากตารางที่ 1 สามารถอธิบายได้ดังนี้ เริ่มต้นมีประชากรทั้งหมด 5 ตัว ซึ่งแต่ละตัวมีค่าความเหมาะสมที่ต่างกัน มีการกำหนดค่า ขนาดหรือจำนวนเป็น 2 จากนั้นสุ่มประชากรจำนวน 2 ตัว จากประชากรทั้งหมด ซึ่งในกรณีนี้ได้สุ่มประชากรตัวที่ 3 และ 5 และเปรียบเทียบค่าความเหมาะสมระหว่างประชากรตัวที่ 3 และ 5 โดยคัดเลือกประชากรตัวที่ 5 เพราะ ค่าความเหมาะสมของประชากรตัวที่ 5 ซึ่งคือ 6 มากกว่าค่าความเหมาะสมของประชากรตัวที่ 3 ซึ่งคือ 4

ตารางที่ 2 ตัวอย่างการคัดเลือกแบบการแข่งขันโดยไม่ลำเอียง

Time	1	2	3	4	5
Contestant No.1	1	2	3	4	5
Contestant No.2	3	1	2	4	5
Contestant No.3	5	1	4	2	3
Population Selection	1	2	2	2	5

นอกจากนี้ยังมีการคัดเลือกแบบการแข่งขันอีกรูปแบบ คือ การคัดเลือกแบบการแข่งขันโดยไม่ลำเอียง (Unbiased Tournament Selection) โดยอ้างถึง ตารางที่ 1 จากตารางดังกล่าว สามารถอธิบายโดยใช้วิธีการแบบประลองโดยไม่ลำเอียงได้ดังนี้ เริ่มต้นมีประชากรทั้งหมด 5 ตัว ซึ่งแต่ละตัวมีค่าความเหมาะสมที่ต่างกัน มีการกำหนดค่าประชากร เป็น 5 และขนาดของการแข่งขัน (Tournament size) เป็น 3 สร้างลำดับของผู้แข่งขันตัวที่ 1 ถึง 5 จำนวน 3 ลำดับ จากนั้นเปรียบเทียบประชากรในตำแหน่งเดียวกันของทุกลำดับ และเลือกตัวที่มีค่าเหมาะสมที่สุดจนครบ 5 ตัว ตามที่นำเสนอในตารางที่ 2

2.1.1.3.2 ปฏิบัติการทางสายพันธุ์ (Genetic Operation)

- การสืบพันธุ์ (Reproduction)

เป็นวิธีการคัดเลือกสายโครโมโซม เพื่อใช้ยีนจากสายโครโมโซมของรุ่นก่อน ในการเกิดสายโครโมโซมในรุ่นถัดไป ซึ่งการสืบพันธุ์นั้น จะเลือกสายโครโมโซมที่มีค่าความเหมาะสมที่สุด ด้วยการคัดลอกสายโครโมโซมสายนั้นมาใช้ โดยมีจุดประสงค์เพื่อไม่ต้องการให้สายโครโมโซมที่ให้ค่าความเหมาะสมที่สุดสูญหาย

- การไขว้เปลี่ยน (Crossover)

เป็นวิธีการไขว้เปลี่ยนของยีนระหว่างสายโครโมโซม ซึ่งในการไขว้เปลี่ยนนั้นจะเกิดสายโครโมโซมใหม่ จำนวน 2 สายโครโมโซม การไขว้เปลี่ยนสามารถแบ่งออกเป็นวิธีการต่างๆ ได้ดังนี้

- การไขว้เปลี่ยนแบบจุดเดียว (Single Point Crossover)

สำหรับวิธีการนี้ จะสุ่มตำแหน่งของจุดตัดเพียงจุดเดียว [8] ซึ่งจุดตัดมีไว้สำหรับระบุตำแหน่ง สำหรับการไขว้เปลี่ยนของยีน ระหว่างสายโครโมโซมทั้ง 2 สายโครโมโซม

ตัวอย่างการไขว้เปลี่ยนแบบจุดเดียว ซึ่งจากตัวอย่างนี้ใช้เลขฐานสอง โดยมีจำนวนยีนในแต่ละสายโครโมโซมเท่ากับ 8 และเลือกจุดตัดในตำแหน่งที่ 3 ดังภาพที่ 4

Parent 1	0	1	0	0	0	0	1	0
Parent 2	1	0	1	0	1	1	1	1

ภาพที่ 4 สายโครโมโซมสำหรับการไขว้เปลี่ยนแบบจุดเดียว

หลังจากนั้นนำมาไขว้เปลี่ยน จะได้ผลลัพธ์ดังภาพที่ 5 สายโครโมโซมที่ผ่านการไขว้เปลี่ยนแบบจุดเดียว

Offspring1	0	1	0	0	1	1	1	1
Offspring2	1	0	1	0	0	0	1	0

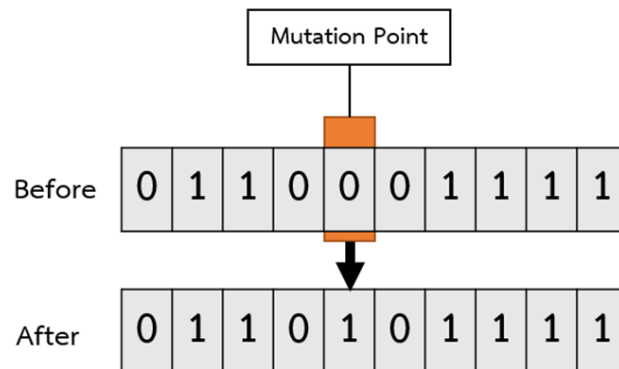
ภาพที่ 5 สายโครโมโซมที่ผ่านการไขว้เปลี่ยนแบบจุดเดียว

- การกลายพันธุ์ (Mutation)

เมื่อได้ผลลัพธ์ของปัญหาที่หาได้ในบางครั้งไม่ตีพอ ซึ่งคำตอบจะอยู่ในรูปโครโมโซมของประชากรรุ่นถัดไป เกิดมีความคล้ายคลึงกับโครโมโซมของประชากรเดิม ด้วยปัญหาความคล้ายคลึงระหว่างประชากรเดิมกับประชากรรุ่นถัดไป ส่งผลให้เกิดกระบวนการกลายพันธุ์เพื่อต้องการให้โครโมโซมประชากรรุ่นถัดไปแตกต่างจากเดิม แต่คงไว้ซึ่งเค้าโครงของโครโมโซมตั้งต้น โดยโครโมโซมที่มีการกลายพันธุ์แล้วอาจได้ค่าความเหมาะสมที่ดีขึ้น แต่ก็อาจเกิดเหตุการณ์โครโมโซมที่มีการกลายพันธุ์ให้ผลลัพธ์ของปัญหาที่แย่ลง

- การกลายพันธุ์แบบจุดเดียว (Single Point Mutation)

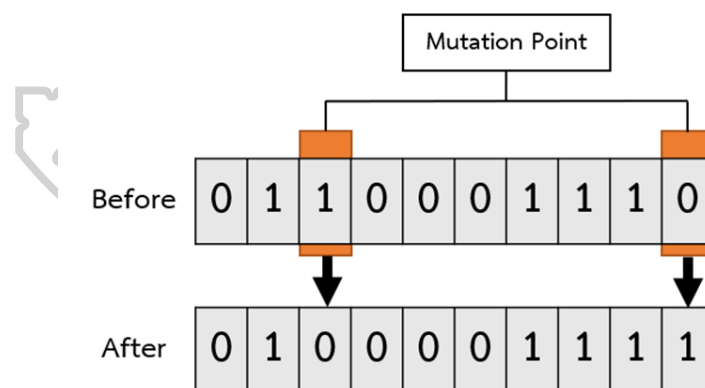
สำหรับวิธีการกลายพันธุ์แบบจุดเดียว วิธีนี้จะถูกใช้เมื่อต้องการเลือกจุดสำหรับการกลายพันธุ์เพียงจุดเดียว [9] ซึ่งตำแหน่งของการกลายพันธุ์ได้มาจากกระบวนการสุ่ม จากภาพที่ 6 การกลายพันธุ์แบบจุดเดียว โดยแสดงตัวอย่างของวิธีการกลายพันธุ์แบบจุดเดียว โดยยีนของโครโมโซมอยู่ในรูปของเลขฐานสอง ซึ่งค่าจะมีแค่ “0” และ “1” เพียงเท่านั้นโดยมีการกลายพันธุ์ในตำแหน่งที่ 5



ภาพที่ 6 การกลายพันธุ์แบบจุดเดียว

○ การกลายพันธุ์แบบหลายจุด (Multi Point Mutation)

สำหรับวิธีการกลายพันธุ์แบบหลายจุด วิธีนี้จะถูกใช้เมื่อต้องการเลือกจุดสำหรับการกลายพันธุ์มากกว่าหนึ่งจุด ซึ่งตำแหน่งของการกลายพันธุ์ได้มาจากกระบวนการสุ่ม จากภาพที่ 7 การกลายพันธุ์แบบหลายจุดโดยแสดงตัวอย่างของวิธีการกลายพันธุ์แบบหลายจุด โดยยีนของสายโครโมโซมอยู่ในรูปของเลขฐานสอง ซึ่งค่าจะมีแค่ “0” และ “1” เพียงเท่านั้นโดยมีจำนวนของการกลายพันธุ์อยู่ที่ 2 ในตำแหน่งที่ 3 และ 10



ภาพที่ 7 การกลายพันธุ์แบบหลายจุด

2.1.1.4 การประเมินผล (Evaluation)

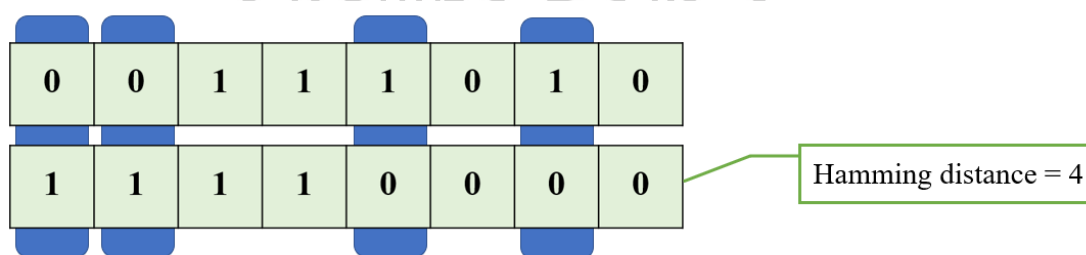
สำหรับกระบวนการนี้ เป็นการประเมินผลโครโมโซมของประชากรรุ่นถัดไปที่เกิดจากประชากรตั้งต้น เป็นกระบวนการหาค่าความเหมาะสมของคำตอบ โดยไม่สามารถระบุเจาะจงค่าหรือวิธีการที่แน่นอนได้ จำเป็นต้องวิเคราะห์ปัญหาเสียก่อน

2.1.1.5 การหยุด (Terminate)

สำหรับกระบวนการหยุด จะเกิดขึ้นเมื่อค้นพบสายโครโมโซมที่เป็นคำตอบแล้วส่งคำตอบออกไปหรือเมื่อครบรุ่นที่ได้กำหนดไว้แล้ว กระบวนการจะหยุดแล้วส่งสายโครโมโซมที่ดีที่สุดกลับไป โดยวิธีขั้นตอนเชิงพันธุกรรมจะสิ้นสุดกระบวนการด้วย การพบสายโครโมโซมที่ให้ค่าความเหมาะสมของคำตอบที่ดีที่สุด ในประชากรรุ่นสุดท้าย [10] ซึ่งเป็นผลลัพธ์ของปัญหาที่ได้จากวิธีการนี้

2.2 ระยะทางแฮมมิง (Hamming Distance)

ระยะทางแฮมมิง เป็นวิธีการที่ใช้สำหรับการหาจำนวนตำแหน่งที่มีสัญลักษณ์หรืออักขระที่แตกต่างกัน ระหว่าง 2 ข้อความที่มีความยาวเท่ากัน [11] โดย ริชาร์ด แฮมมิง (Richard Hamming) เป็นผู้ที่น่าเสนอทฤษฎีนี้ ในคริสต์ทศวรรษ 1950 โดยใช้สำหรับการตรวจสอบค่า ต่อมาระยะทางแฮมมิงถูกนำมาใช้สำหรับด้านการสื่อสาร เพื่อสำหรับการตรวจสอบความผิดพลาด ของจำนวนเลขฐานสองที่คงที่ หรือ ระยะทางสัญญาณ (Signal distance) โดยการสื่อสารในระบบดิจิทัลโดยส่วนมาก อาจเกิดโอกาสที่ข้อมูลที่ส่งผ่านระบบสื่อสารมีการผิดพลาด จึงมีการนำการเข้ารหัสควบคุมความผิดพลาดเชิงดิจิทัล (Error Control Coding : ECC) เข้ามาใช้ เพื่อลดความผิดพลาดของการสื่อสาร แต่มีเกิดโอกาสที่จะผิดพลาดในด้านข้อมูลสูงขึ้น ดังนั้นระยะทางแฮมมิงได้ถูกใช้ในการวิเคราะห์หลายๆด้าน โดยการนำมาต่อยอดจากทฤษฎี เช่น ด้านการออกแบบฮาร์ดแวร์คอมพิวเตอร์ เป็นต้น



ภาพที่ 8 การหาค่าระยะทางแฮมมิง

จากภาพที่ 8 การหาค่าระยะทางแฮมมิง จะแสดงให้เห็นว่า เมื่อมีสายโครโมโซมที่มีความยาวเท่ากัน จะสามารถใช้การหาระยะทางแฮมมิงได้ โดยเปรียบเทียบและหาบิตที่แตกต่างกัน สรุปได้ว่าค่าระยะทางแฮมมิงระหว่างทั้ง 2 สายโครโมโซมนี้ คือ 4

ระยะทางแฮมมิง สามารถแสดงรูปของสมการ ตามสมการที่ 1

$$d(A, B) = \sum_{i=1}^n |A_i - B_i| \quad (1)$$

โดยที่ระยะทางแฮมมิง คือ $d(A, B)$ ความยาวของสายโครโมโซมของทั้งคู่ เท่ากับ n

2.3 สหสัมพันธ์โคไซน์ (Cosine Coefficient)

สหสัมพันธ์โคไซน์ หรือความคล้ายคลึงโคไซน์ (Cosine Similarity) เป็นการวัดมุมของโคไซน์ของเวกเตอร์ทั้งสอง [12] ซึ่งเป็นการวัดความคล้ายคลึงระหว่าง 2 เวกเตอร์ สามารถคำนวณจากสมการที่ 2

$$\cos(\theta) = \frac{\sum_{i=1}^n (A_i \times B_i)}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (2)$$

โดยนำ $A = (a_1, a_2, a_3, \dots, a_n)$ และ $B = (b_1, b_2, b_3, \dots, b_n)$ ทั้ง 2 เวกเตอร์ที่ต้องนำมาเปรียบเทียบ

ค่าสหสัมพันธ์โคไซน์จะมีค่าอยู่ระหว่าง -1 ถึง 1 โดยมีความหมายดังนี้

เข้าใกล้ 1 มีความหมายว่า ทั้ง 2 เวกเตอร์มีความสัมพันธ์กันมากไปในทิศทางเดียวกัน

เข้าใกล้ -1 มีความหมายว่า ทั้ง 2 เวกเตอร์มีความสัมพันธ์กันมากไปในทิศทางตรงข้ามกัน

เข้าใกล้ 0 มีความหมายว่า ทั้ง 2 เวกเตอร์ไม่มีความสัมพันธ์กัน

2.4 สหสัมพันธ์เพียร์สัน (Pearson Coefficient)

สหสัมพันธ์เพียร์สัน เป็นการใช้ชุดข้อมูล 2 ชุด หรือตัวแปร 2 ตัว เพื่อศึกษาความสัมพันธ์เชิงเส้น (linear relationship) โดยต้องเป็นตัวแปรต่อเนื่องอยู่ในมาตราอันตรภาคชั้น (Interval Scale) ขึ้นไป [12] และมีการแจกแจงปกติสองตัวแปร (Bivariate Normal Distribution) สามารถคำนวณจากสมการที่ 3

$$r = \frac{(1/(N-1))(\sum XY - ((\sum X)(\sum Y)/N))}{S_x S_y} \quad (3)$$

กำหนดให้ r เป็น ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน

$\sum X$ เป็น ผลรวมของข้อมูลที่คำนวณได้จากสายโครโมโซมสายที่ 1 (A)

$\sum Y$ เป็น ผลรวมของข้อมูลที่คำนวณได้จากสายโครโมโซมสายที่ 2 (B)

$\sum XY$ เป็น ผลรวมของผลคูณระหว่างสายโครโมโซมสายที่ 1 และ 2

S_x เป็น การหาค่าส่วนเบี่ยงเบนมาตรฐานที่ได้จากสายโครโมโซมสายที่ 1

S_y เป็น การหาค่าส่วนเบี่ยงเบนมาตรฐานที่ได้จากสายโครโมโซมสายที่ 2

N เป็น ขนาดของกลุ่มตัวอย่าง

การเปรียบเทียบ ค่าสหสัมพันธ์แบบเพียร์สันจะมีค่าอยู่ระหว่าง -1 ถึง 1 โดยมีความหมายดังนี้

เข้าใกล้ 1 มีความหมายว่า มีความสัมพันธ์กันมากไปในทิศทางเดียวกัน

เข้าใกล้ 0 มีความหมายว่า ไม่มีความสัมพันธ์กัน

เข้าใกล้ -1 มีความหมายว่า มีความสัมพันธ์กันมากไปในทิศทางตรงข้ามกัน

2.5 งานวิจัยที่เกี่ยวข้อง

จากการศึกษางานวิจัยของ [13] เรื่อง The Gambler's Ruin Problem, Genetic Algorithms, and the Sizing of Populations พบว่า สำหรับปัญหาที่ติด ความยาวของสายโครโมโซมที่มีความเหมาะสม คือ 80 บิต ซึ่งการเข้าสู่ค่าตอบในขั้นตอนเชิงพันธุกรรม จะสอดคล้องกับขนาดของประชากร โดยได้เปรียบเทียบการสุ่มในขั้นตอนวิธีเชิงพันธุกรรม กับปัญหาที่ใช้ในการสุ่มแบบมาตรฐาน

จากการศึกษางานวิจัยของ [2] เรื่อง A Study of Mate Selection Schemes in Genetic Algorithms พบว่า ในงานวิจัยนี้ให้ความสำคัญกับรูปแบบพื้นฐานของทฤษฎีสำหรับการเลือกคู่ที่มีความเหมือนหรือแตกต่าง ด้วยวิธีการคำนวณลักษณะเฉพาะจากปัจจัยต่างๆ ซึ่งพบว่า การจับคู่ที่มีความแตกต่างกันมาก สามารถสร้างความหลากหลายได้มากกว่าการจับคู่ที่มีความใกล้เคียงกัน

จากการศึกษางานวิจัยของ [14] เรื่อง Comparison of Performance between Different Selection Strategies on Simple Genetic Algorithms พบว่า การเปรียบเทียบประสิทธิภาพของ อัลกอริทึมทางพันธุกรรมอย่างง่าย (SGA) ด้วยการคัดเลือกสายพันธุ์ ซึ่งใช้การคัดเลือกแบบวงล้อรูเล็ต และการคัดเลือกแบบการแข่งขัน และใช้การปฏิบัติทางสายพันธุ์ ได้แก่ การไขว้เปลี่ยน และการ

กลายพันธุ์ จากผลลัพธ์แสดงให้เห็นว่า การคัดเลือกแบบการแข่งขันมีประสิทธิภาพที่ดีกว่าการคัดเลือกแบบวงล้อรูเล็ต

จากการศึกษางานวิจัยของ [15] เรื่อง Assortative Mating in Genetic Algorithms for Dynamic Problems พบว่า การจับคู่แบบแบ่งกลุ่มในขั้นตอนวิธีเชิงพันธุกรรมสำหรับการแก้ปัญหาแบบไดนามิก ซึ่งสำรวจอัตราการกลายพันธุ์ พบว่า การกลายพันธุ์ที่แตกต่างส่งผลต่อการแก้ปัญหาในขั้นตอนวิธีเชิงพันธุกรรม ซึ่งวิธีการที่ใช้ในการจับคู่จะส่งผลต่ออัตราการกลายพันธุ์ และประสิทธิภาพของขั้นตอนวิธีเชิงพันธุกรรม โดยอัตราการกลายพันธุ์ที่ต่ำจะให้ผลลัพธ์ที่ดีสำหรับขั้นตอนวิธีเชิงพันธุกรรม

จากการศึกษางานวิจัยของ [3] เรื่อง A Novel Mating Approach for Genetic Algorithms พบว่า ในงานวิจัยนี้ได้นำเสนอแนวทางใหม่ในการจับคู่ระหว่างสายโครโมโซม ซึ่งสามารถประยุกต์ใช้ร่วมกับเกณฑ์ที่ใช้การพิจารณาความคล้ายคลึงแบบดั้งเดิม ใช้สำหรับวัดระยะห่างระหว่างประชากรหรือตามความเหมาะสม โดยงานวิจัยนี้ได้แนะนำพารามิเตอร์ที่เรียกว่า “ดัชนีการจับคู่” ที่สามารถช่วยให้การจับคู่ที่มีความแตกต่างกัน ด้วย 2 แนวทางใหม่ คือ 1) กำหนดการจับคู่ด้วยความคล้ายคลึงกันระหว่างประชากรหรือจากค่าความเหมาะสมของแต่ละสายโครโมโซม ซึ่งจะตรงกันข้ามกับการจับคู่แบบดั้งเดิม 2) ใช้ตัวสายโครโมโซมสายแรกเพื่อคัดเลือกสายโครโมโซมอีกสาย สำหรับการจับคู่จะขึ้นอยู่กับฟังก์ชันความเหมาะสมและกระบวนการทำงาน

จากการศึกษางานวิจัยของ [16] เรื่อง Linkage-Based Distance Metric in the Search Space of Genetic Algorithms พบว่า การวัดระยะทางระหว่างสายโครโมโซมในขั้นตอนวิธีเชิงพันธุกรรมโดยปกติ จะวัดระยะทางเมื่อต้องการนำไปใช้ในการหาพื้นที่ระบบเมตริก เพื่อสามารถนำไปคำนวณได้อย่างมีประสิทธิภาพ สำหรับการวัดระยะทางแฮมมิง สามารถทำให้เห็นถึงปัญหา และช่องว่างจากมุมมองที่มีความแตกต่าง

จากการศึกษางานวิจัยของ [17] เรื่อง A New Adaptive Hungarian Mating Scheme in Genetic Algorithms พบว่า งานวิจัยนี้ได้แนะนำรูปแบบการจับคู่แบบปรับตัวด้วยวิธีการแบบฮังการีด้วยการใช้อัลกอริทึมสำหรับการคัดเลือกรูปแบบการจับคู่ด้วยวิธีการแบบฮังการี โดยประชากรแต่ละคู่ที่ได้รับการคัดเลือกสำหรับใช้ในการแก้ปัญหาต้องลดอัตราการจับคู่ในรุ่นถัดไปให้น้อยลง โดยระยะห่างของประชากรแต่ละรุ่นต้องได้รับการพิจารณาการเพิ่มหรือลด ด้วยคณิตศาสตร์เชิงการจัด (Combinatorics)

จากการศึกษางานวิจัยของ [4] เรื่อง Performance Comparison and Analysis of Crossover Techniques in Genetic Algorithms for the Function of Unitation Problems พบว่า การวัดประสิทธิภาพของวิธีการไขว้เปลี่ยนในรูปแบบต่างๆ สำหรับขั้นตอนวิธีเชิงพันธุกรรม โดยการไขว้เปลี่ยนที่นำมาเปรียบเทียบมีทั้งหมด 4 วิธี คือ การไขว้เปลี่ยนแบบจุดเดียว การไขว้เปลี่ยนแบบสองจุด การไขว้เปลี่ยนแบบเอกรูป และการไขว้เปลี่ยนแบบวงแหวน โดยทดลองกับปัญหาในกลุ่มที่เป็นฟังก์ชันยูนิเทชัน เช่น ปัญหาบิตหนึ่งมากที่สุด ปัญหาบิตศูนย์มากที่สุด ปัญหาบิตสูงสุดที่เกิดจากการสุ่ม และปัญหากับดัก จากการทดลองพบว่า การไขว้เปลี่ยนแบบวงแหวนสามารถหาคำตอบได้อย่างมีประสิทธิภาพ แม้แต่ปัญหากับดัก ที่เป็นปัญหาที่ยากและมีการหลอก ประเด็นที่น่าสนใจที่ค้นพบจากการทดลองนี้ก็คือ ถึงแม้ว่าการไขว้เปลี่ยนแบบวงแหวนจะมีประสิทธิภาพที่ดี แต่เมื่อนำไปใช้กับปัญหาบิตสูงสุดที่เกิดจากการสุ่ม พบว่าวิธีนี้ไม่ได้เป็นวิธีที่ดีที่สุด แต่ในการไขว้เปลี่ยนแบบเอกรูปที่หาคำตอบได้ผลลัพธ์ต่ำกว่า ในปัญหาอื่นๆ กลับได้ผลลัพธ์ที่ดีกว่า

จากการศึกษางานวิจัยของ [8] เรื่อง A New Crossover Mechanism for Genetic Algorithm with Rank-based Selection Method พบว่า การไขว้เปลี่ยน เป็นวิธีที่ดีที่สุดสำหรับการหาคำตอบในขั้นตอนเชิงพันธุกรรม ในงานวิจัยนี้ได้นำเสนอวิธีการไขว้เปลี่ยนแบบใหม่ คือ วิธีการปฏิบัติการทางสายพันธุ์สำหรับการไขว้เปลี่ยนแบบการหาค่าเฉลี่ยแบบไขว้ (Cross Average Crossover Operator) ซึ่งจะเลือกยีนแรกของสายโครโมโซมสายแรก และยีนสุดท้ายของสายโครโมโซมอีกสาย สำหรับการไขว้เปลี่ยน เพื่อสร้างประชากรรุ่นถัดไป (Offspring) และกระบวนการจะทำซ้ำจนกว่าจะถึงยีนสุดท้ายของสายโครโมโซมแรก และถึงยีนแรกของสายโครโมโซมอีกสาย

2.6 สรุปงานวิจัยที่เกี่ยวข้อง

จากการค้นคว้ายังไม่พบการสมการมาใช้กับเทคนิคการคัดเลือกในขั้นตอนวิธีเชิงพันธุกรรม มีเพียงงานวิจัยที่นำ เทคนิค, วิธีการ, และปัญหาต่างๆ ที่ใช้ในขั้นตอนวิธีเชิงพันธุกรรมมาเปรียบเทียบ และหาประสิทธิภาพ ดังสรุปข้อมูลไว้ในตารางที่ 3

ตารางที่ 3 สรุปงานวิจัยที่เกี่ยวข้อง

ลำดับ	งานวิจัย	ลักษณะงาน	เทคนิค
1	The Gambler's Ruin Problem, Genetic Algorithms, and the Sizing of Populations [13]	การเปรียบเทียบประสิทธิภาพ เทคนิคการคัดเลือก และหาขนาดประชากรที่เหมาะสม โดยทดสอบกับปัญหา	Building blocks (BBs), The Gambler's Ruin Problem, One max problem, Trap functions
2	A Study of Mate Selection Schemes in Genetic Algorithms [2]	แนะนำกรอบการทำงานสำหรับการคัดเลือกในขั้นตอนวิธีเชิงพันธุกรรม	Transient Markov Chain Analysis
3	Comparison of Performance between Different Selection Strategies on Simple Genetic Algorithms [14]	การเปรียบเทียบประสิทธิภาพ เทคนิคการคัดเลือก	Roulette Wheel Selection, Tournament Selection, Crossover, Mutation
4	Assortative Mating in Genetic Algorithms for Dynamic Problems [15]	การเปรียบเทียบประสิทธิภาพของการกลายพันธุ์	Mutation
5	A Novel Mating Approach for Genetic Algorithms [3]	การเปรียบเทียบประสิทธิภาพของเทคนิคการคัดเลือก	Random, Fitness-based-best-last
6	Linkage-Based Distance Metric in the Search Space of Genetic Algorithms [16]	นำเสนอเทคนิคการวัดระยะทาง	Hamming Distance, Fitness distance correlation (FDC)

ลำดับ	งานวิจัย	ลักษณะงาน	เทคนิค
7	A New Adaptive Hungarian Mating Scheme in Genetic Algorithms [17]	การเปรียบเทียบประสิทธิภาพของเทคนิคการคัดเลือก	Hungarian Mating Scheme
8	Performance Comparison and Analysis of Crossover Techniques in Genetic Algorithms for the Function of Unitation Problems [4]	การเปรียบเทียบประสิทธิภาพการไขว้เปลี่ยน	Crossover, One Max Problem, Zero Max Problem, Random Max Problem, Trap Problem
9	A New Crossover Mechanism for Genetic Algorithm with Rank-based Selection Method [8]	การเปรียบเทียบประสิทธิภาพการไขว้เปลี่ยน	Crossover

บทที่ 3

วิธีดำเนินการวิจัย

ศึกษารายละเอียดของเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรมและปัญหาต่าง ๆ ที่เกี่ยวข้องในการนำมาใช้ประกอบการออกแบบและพัฒนาประสิทธิภาพของกระบวนการเลือกคู่สำหรับการไขว้เปลี่ยนโดยแบ่งเป็นขั้นตอนได้ดังนี้

3.1 ปัญหาที่นำมาใช้ในงานวิจัย

3.2 การออกแบบการทดลอง

3.3 ขั้นตอนการทดลอง

3.1 ปัญหาที่นำมาใช้ในงานวิจัย

3.1.1 ปัญหาบิตหนึ่งมากที่สุด (One Max Problem)

ปัญหาบิตหนึ่งมากที่สุด คือ ปัญหาที่ใช้วิธีการนับเลขหนึ่งของยีนในโครโมโซม โดยค่าความเหมาะสมที่ดีที่สุด คือ ทุกยีนในทั้งโครโมโซมเป็นเลขหนึ่งทั้งหมด ซึ่งเป็นปัญหาพื้นฐานที่ใช้มักถูกใช้เพื่อทดลองแก้ปัญหามันในวิธีขั้นตอนเชิงพันธุกรรม เป็นปัญหาที่สามารถเข้าใจได้ง่าย และงานวิจัยที่เกี่ยวข้องกับวิธีขั้นตอนเชิงพันธุกรรมจำนวนมาก ใช้ปัญหาบิตหนึ่งมากที่สุดในการทดลอง ส่งผลให้สามารถเปรียบเทียบประสิทธิภาพและอ้างอิง เมื่อนำมาใช้กับวิธีขั้นตอนเชิงพันธุกรรม

3.1.2 ปัญหาบิตศูนย์มากที่สุด (Zero Max Problem)

ปัญหาบิตศูนย์มากที่สุด คือ ปัญหาที่ใช้วิธีการนับเลขศูนย์ของยีนในโครโมโซม โดยค่าความเหมาะสมที่ดีที่สุด คือ ทุกยีนในทั้งโครโมโซมเป็นเลขศูนย์ทั้งหมด เป็นปัญหาที่สามารถเข้าใจได้ง่าย สามารถเปรียบเทียบประสิทธิภาพ เมื่อนำมาใช้กับวิธีขั้นตอนเชิงพันธุกรรม

3.1.3 ปัญหาบิตสูงสุดที่เกิดจากการสุ่ม (Random Max Problem)

ปัญหาบิตสูงสุดที่เกิดจากการสุ่ม คือ ปัญหาที่มีการสุ่มค่าตอบขึ้นมาโดยค่าความเหมาะสมจะคำนวณจากจำนวนของบิตที่ตรงกับคำตอบที่กำหนดขึ้น [9] ดังนั้นค่าความเหมาะสมที่ดีที่สุดคือการที่ทุกบิตในสายโครโมโซมตรงกับคำตอบที่กำหนดขึ้น ซึ่งเป็นปัญหาที่สามารถเข้าใจได้ง่าย สามารถเปรียบเทียบประสิทธิภาพ เมื่อนำมาใช้กับวิธีขั้นตอนเชิงพันธุกรรม

3.1.4 ปัญหาทับดัก (Trap Problem)

ปัญหาทับดัก คือ ปัญหาที่มีความยาก มีจุดประสงค์เพื่อต้องการลอกขั้นตอนการผลิตโครโมโซมคำตอบที่ไม่มี ความใกล้เคียงกับคำตอบ [18] & [19] สำหรับค่าความเหมาะสมที่ดีที่สุดสำหรับคำตอบของปัญหาทับดัก คือ ทุกบิตมีค่าเป็น 1 เริ่มต้นด้วยการกำหนดค่าขนาดของทับดัก (Trap Size) ซึ่งขนาดของปัญหาคือจำนวนของบิต มีการประเมินหาค่าความเหมาะสมในส่วนของบิต และรวมค่าความเหมาะสมของบิต เพื่อนำไปคำนวณเป็นคำตอบของค่าความเหมาะสมของโครโมโซม เช่น กรณีที่ขนาดของทับดัก เท่ากับ 4 และการคำนวณค่าของโครโมโซมครั้งละ 4 บิต

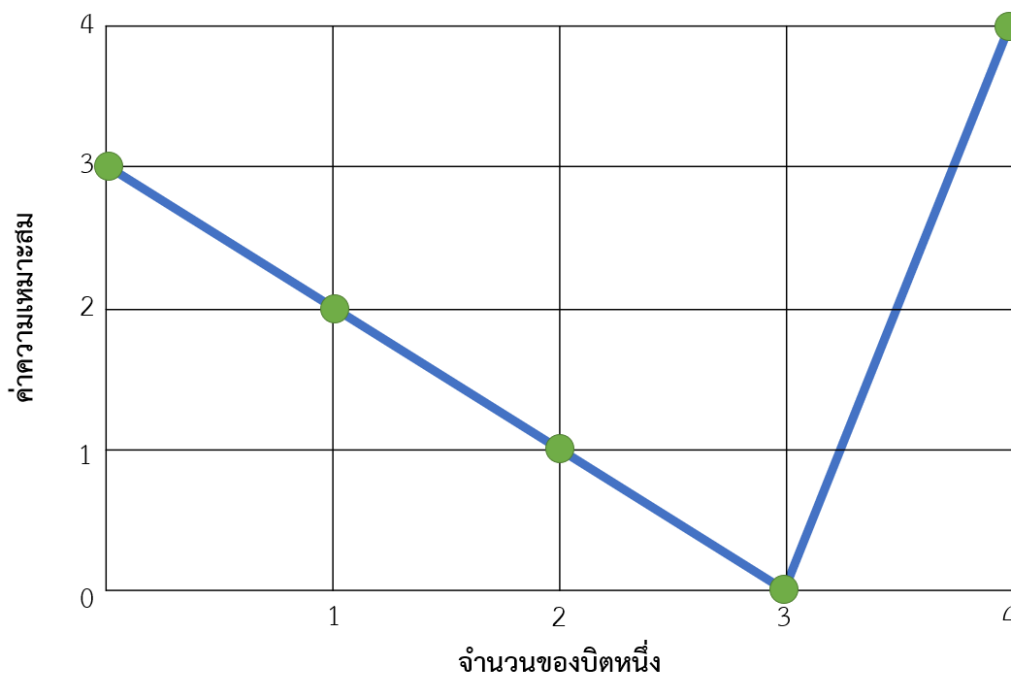
ตารางที่ 4 ค่าความเหมาะสมจากกรณีขนาดของทับดักเท่ากับ 4

ค่าความเหมาะสม	รูปแบบของโครโมโซม	จำนวนของบิตที่เป็น 1
4	1111	4
3	0000	0
2	1000, 0100, 0010, 0001	1
1	1100, 1001, 0011, 0110	2
0	1110, 1101, 1011, 0111	3

ตารางที่ 5 ค่าความเหมาะสมจากกรณีขนาดของทับดักเท่ากับ 4

จำนวนของบิตหนึ่ง	0	1	2	3	4
ค่าความเหมาะสม	3	2	1	0	4

จากตารางที่ 4 และตารางที่ 5 พบว่า จากเงื่อนไขการที่ทุกบิตมีค่าเป็น 1 ในสายโครโมโซมจะได้ค่าความเหมาะสมสูงสุด และเงื่อนไขการที่ทุกบิตมีค่าเป็น 0 ในสายโครโมโซมจะได้ค่าความเหมาะสมรองลงมา แต่หากจำนวนบิตที่มีค่าเป็น 1 ยิ่งมากแต่ไม่ใช่คำตอบจะให้ค่าความเหมาะสมที่ยิ่งน้อยลง โดยแสดงให้เห็นในภาพที่ 9



ภาพที่ 9 กราฟค่าความเหมาะสมจากกรณีขนาดของกับดักเท่ากับ 4

3.2 การออกแบบการทดลอง

ทดสอบเทคนิคการคัดเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรม ด้วยเทคนิคต่างๆ ดังนี้

3.2.1 เทคนิคการคัดเลือกด้วยระยะทางแฮมมิง (HS)

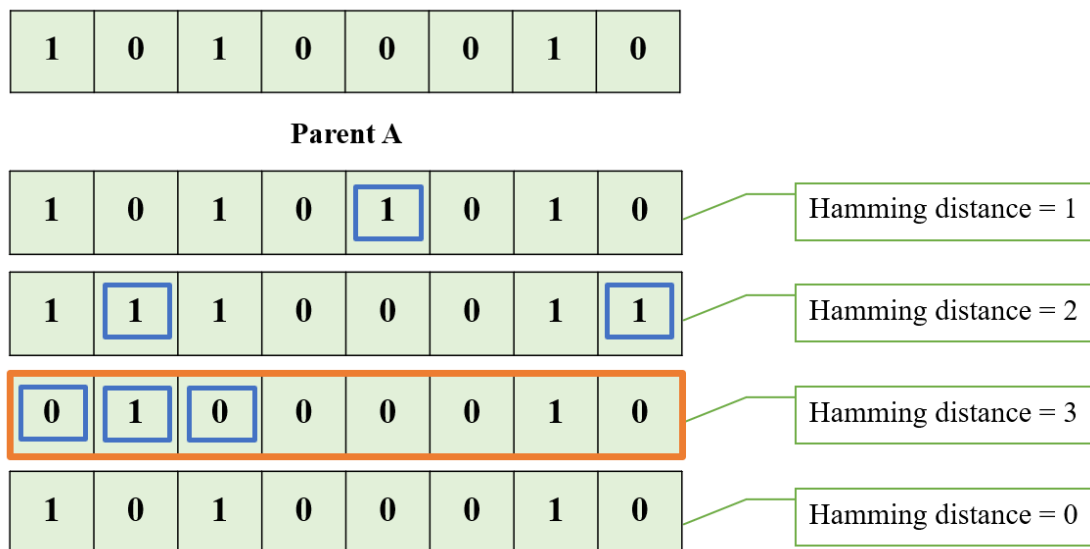
ระยะทางแฮมมิง เป็นเทคนิคที่ใช้สำหรับการหาจำนวนตำแหน่งที่มีความแตกต่าง ระหว่าง 2 สายโครโมโซมที่มีความยาวเท่ากัน สามารถแสดงรูปของสมการ ตามสมการที่ 1

$$d(A, B) = \sum_{i=1}^n |A_i - B_i| \quad (1)$$

กำหนดให้สายโครโมโซมสายแรกเป็น สายโครโมโซม A (Parent A) และสายโครโมโซมอีกสายเป็น สายโครโมโซม B (Parent B) โดยที่ระยะทางแฮมมิง คือ $d(A, B)$ ความยาวของสายโครโมโซมของทั้งคู่ เท่ากับ n

การหาประชากร จะเริ่มต้นด้วยการแบ่งประชากร ออกเป็น 2 สายโครโมโซมซึ่งประกอบด้วยสายโครโมโซม A และสายโครโมโซม B โดยที่สายโครโมโซม A จะใช้วิธีการคัดเลือกแบบวงล้อรูเล็ต

สำหรับสายโครโมโซม B จะสุ่มจากประชากรทั้งหมดด้วยวิธีการแบบวงล้อรูเล็ต ตามขนาดของการแข่งขันที่ได้กำหนดไว้ ซึ่งในงานวิจัยนี้กำหนดให้เป็น 4 นำสายโครโมโซม A มาเปรียบเทียบกับชุดสายโครโมโซม B ทีละสาย แบบบิตต่อบิต เพื่อให้ได้ค่าระยะทางแฮมมิง [11] และเลือกสายโครโมโซม B จากชุดสายโครโมโซม B ที่มีค่าระยะทางแฮมมิงมากที่สุด ตัวอย่างของการคำนวณในเทคนิคการคัดเลือกด้วยระยะทางแฮมมิง (HS) แสดงดังภาพที่ 10



Parent B set with four members.

ภาพที่ 10 ตัวอย่างการคำนวณในเทคนิคการคัดเลือกด้วยระยะทางแฮมมิง (HS)

3.2.2 เทคนิคการคัดเลือกด้วยสหสัมพันธ์โคไซน์ (CS)

สหสัมพันธ์โคไซน์ เป็นเทคนิคที่ใช้สำหรับการวัดความคล้ายคลึงระหว่าง 2 เวกเตอร์ โดยการวัดมุมโคไซน์ของเวกเตอร์ทั้งสอง โดยใช้เทคนิคดังกล่าวในการทดลองครั้งนี้

การหาประชากร จะเริ่มต้นด้วยการแบ่งประชากร ออกเป็น 2 สายโครโมโซมซึ่งประกอบด้วย สายโครโมโซม A และสายโครโมโซม B โดยที่สายโครโมโซม A จะใช้วิธีการคัดเลือกแบบวงล้อรูเล็ต สำหรับสายโครโมโซม B จะสุ่มจากประชากรทั้งหมดด้วยวิธีการแบบวงล้อรูเล็ต ตามขนาดของการแข่งขันที่ได้กำหนดไว้ ซึ่งในงานวิจัยนี้กำหนดให้เป็น 4 โดยนำสายโครโมโซม A มาคิดคำนวณหาค่าสหสัมพันธ์โคไซน์กับชุดสายโครโมโซม B ทีละสาย เพื่อหาค่าสหสัมพันธ์โคไซน์ โดยใช้สมการที่ 2

$$\cos(\theta) = \frac{\sum_{i=1}^n (A_i \times B_i)}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (2)$$

โดยนำ $A = (a_1, a_2, a_3, \dots, a_n)$ และ $B = (b_1, b_2, b_3, \dots, b_n)$ ทั้ง 2 เวกเตอร์ที่ต้องนำมาเปรียบเทียบ

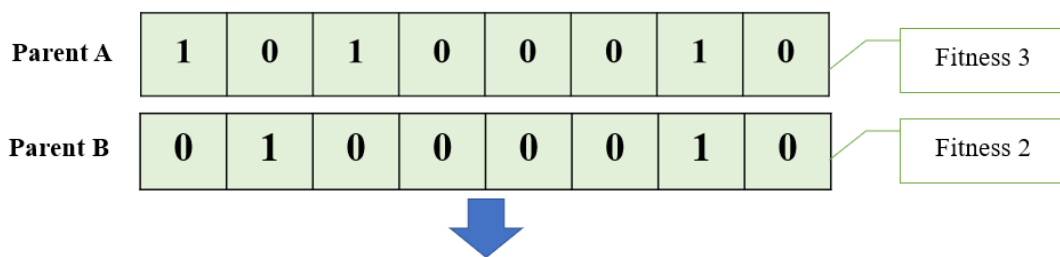
ค่าสหสัมพันธ์โคไซน์จะมีค่าอยู่ระหว่าง -1 ถึง 1 โดยมีความหมายดังนี้

หากค่าเข้าใกล้ 1 มีความหมายว่า ทั้ง 2 เวกเตอร์มีความสัมพันธ์กันมากไปในทิศทางเดียวกัน

หากค่าเข้าใกล้ -1 มีความหมายว่า ทั้ง 2 เวกเตอร์มีความสัมพันธ์กันมากไปในทิศทางตรงข้ามกัน

หากค่าเข้าใกล้ 0 มีความหมายว่า ทั้ง 2 เวกเตอร์ไม่มีความสัมพันธ์กัน

และเลือกสายโครโมโซม B ที่ให้ค่าสหสัมพันธ์โคไซน์น้อยที่สุด ซึ่งแสดงถึงความสัมพันธ์กันน้อยที่สุด หรือสายโมโซมทั้งสองมีความแตกต่างกันมาก ตัวอย่างของวิธีการคำนวณด้วยสหสัมพันธ์โคไซน์ (CS) [20] แสดงดังภาพที่ 11



$$\text{Cosine Coefficient} = \frac{3 \div 2}{\sqrt{(1^2 + 0 + 1^2 + 0 + 0 + 0 + 1^2 + 0)} \times \sqrt{0 + 1^2 + 0 + 0 + 0 + 0 + 1^2 + 0}}$$

$$\text{Cosine Coefficient} = \frac{1.5}{2.45}$$

$$\text{Cosine Coefficient} = 0.61$$

ภาพที่ 11 ตัวอย่างการคำนวณในเทคนิคการคัดเลือกด้วยสหสัมพันธ์โคไซน์ (CS)

3.2.3 เทคนิคการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน (PS)

สหสัมพันธ์เพียร์สัน เป็นเทคนิคที่ใช้สำหรับหาความสัมพันธ์ ระหว่าง 2 ตัวแปร โดยใช้เทคนิคดังกล่าวในการทดลองครั้งนี้

การหาประชากร จะเริ่มต้นด้วยการแบ่งประชากร ออกเป็น 2 สายโครโมโซมซึ่งประกอบด้วยสายโครโมโซม A และสายโครโมโซม B โดยที่สายโครโมโซม A จะใช้วิธีการคัดเลือกแบบวงล้อสุ่มเล็ก สำหรับสายโครโมโซม B จะสุ่มจากประชากรทั้งหมดด้วยวิธีการแบบวงล้อสุ่มเล็ก ตามขนาดของการ

แข่งขันที่ได้กำหนดไว้ ซึ่งในงานวิจัยนี้กำหนดให้เป็น 4 โดยนำสายโครโมโซม A มาคิดคำนวณหาค่าสหสัมพันธ์เพียร์สันกับชุดสายโครโมโซม B ทีละสาย เพื่อหาค่าสหสัมพันธ์เพียร์สัน โดยใช้สมการที่ 3

$$r = \frac{(1/(N-1))(\sum XY - ((\sum X)(\sum Y)/N))}{S_x S_y} \quad (3)$$

กำหนดให้ r เป็น ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน

$\sum X$ เป็น ผลรวมของข้อมูลที่คำนวณได้จากสายโครโมโซมสายที่ 1 (A)

$\sum Y$ เป็น ผลรวมของข้อมูลที่คำนวณได้จากสายโครโมโซมสายที่ 2 (B)

$\sum XY$ เป็น ผลรวมของผลคูณระหว่างสายโครโมโซมสายที่ 1 และ 2

S_x เป็น การหาค่าส่วนเบี่ยงเบนมาตรฐานที่ได้จากสายโครโมโซมสายที่ 1

S_y เป็น การหาค่าส่วนเบี่ยงเบนมาตรฐานที่ได้จากสายโครโมโซมสายที่ 2

N เป็น ขนาดของกลุ่มตัวอย่าง

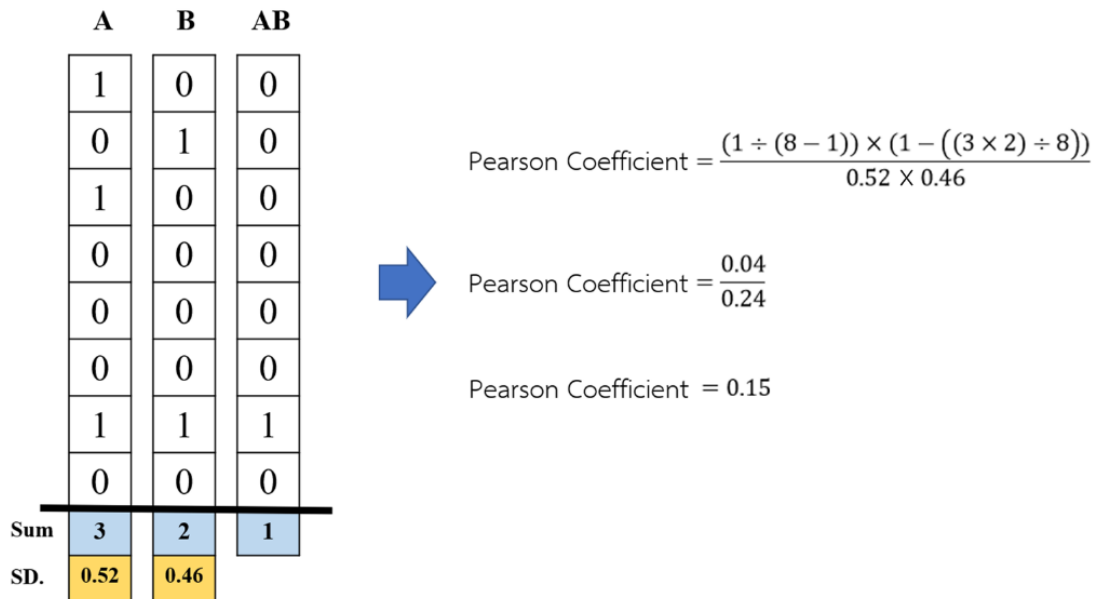
การเปรียบเทียบ ค่าสหสัมพันธ์แบบเพียร์สันจะมีค่าอยู่ระหว่าง -1 ถึง 1 โดยมีความหมายดังนี้

เข้าใกล้ 1 มีความหมายว่า มีความสัมพันธ์กันมากไปในทิศทางเดียวกัน

เข้าใกล้ 0 มีความหมายว่า ไม่มีความสัมพันธ์กัน

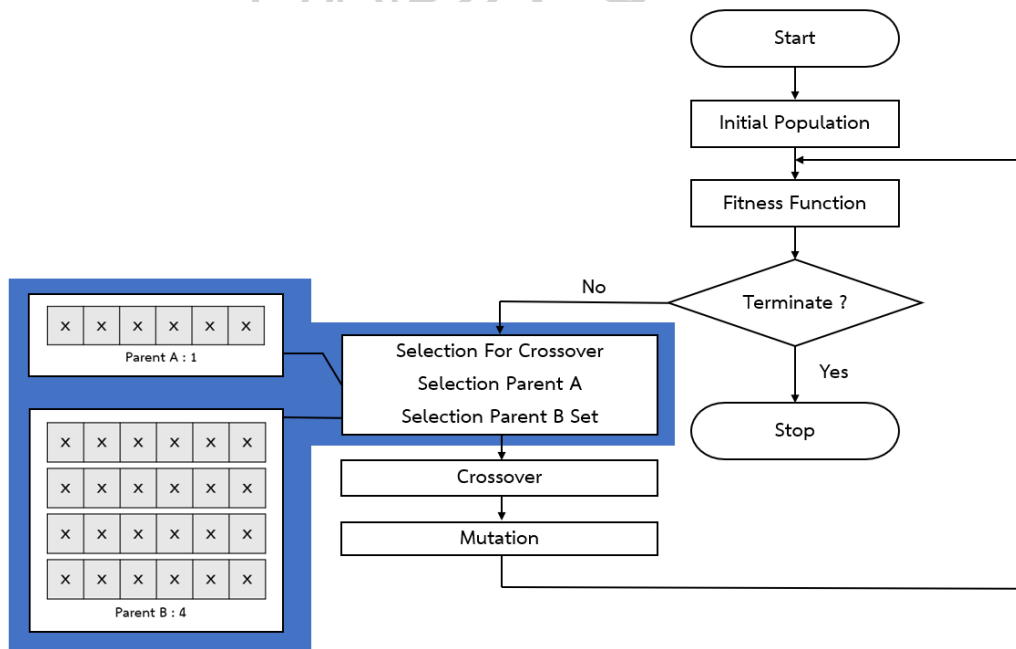
เข้าใกล้ -1 มีความหมายว่า มีความสัมพันธ์กันมากไปในทิศทางตรงข้ามกัน

คำนวณค่าสหสัมพันธ์เพียร์สันของสายโครโมโซม A และ สายโครโมโซม B1 กำหนดให้ได้ค่าเท่ากับ P1 จากนั้น คำนวณค่าสหสัมพันธ์เพียร์สันของสายโครโมโซม A และ สายโครโมโซมที่เหลือคือ B2, B3, B4 ซึ่งจะได้ค่า P2, P3, P4 ตามลำดับ และเลือกสายโครโมโซม B1, B2, B3, B4 ที่ให้ค่าสหสัมพันธ์เพียร์สันหรือ ค่า P น้อยที่สุด ซึ่งแสดงถึงความสัมพันธ์กันน้อยที่สุด หรือสายโครโมโซมที่มีความแตกต่างกันมาก เช่น ถ้าค่า P1 น้อยสุด ก็จะเลือก สายโครโมโซม B1 ตัวอย่างของการคำนวณในเทคนิคการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน (PS) [21] แสดงดังภาพที่ 12



ภาพที่ 12 ตัวอย่างการคำนวณในเทคนิคการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน (PS)

3.3 ขั้นตอนการทดลอง



ภาพที่ 13 ขั้นตอนการทดลองของเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรม

จากภาพที่ 13 (ซึ่งเหมือนกับภาพที่ 1) กระบวนการทำงานจะเริ่มต้นด้วยการสร้างประชากร (Initial Population) ที่เป็นสายโครโมโซมตามค่าพารามิเตอร์ที่กำหนดไว้ นำไปประเมินค่าความ

เหมาะสม (Fitness Function) ตามปัญหาที่กำหนด ตรวจสอบหาคำตอบจากสายโครโมโซมทั้งหมด ถ้าไม่พบคำตอบกระบวนการทำงานจะทำงานในส่วนของการคัดเลือกต่อไป กระบวนการจะสิ้นสุดหากพบคำตอบหรือพบเงื่อนไขสำหรับการยุติทำงาน (Terminate)

3.3.1 เงื่อนไขสำหรับการยุติทำงาน (Terminate)

ในงานวิจัยครั้งนี้ การทำงานของเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรมจะยุติเมื่อค้นพบคำตอบ หรือรันโปรแกรมจนถึงกำหนดจำนวนรุ่นสูงสุด (Number of generations) ที่ได้กำหนดไว้ [10]

ในการทดสอบสำหรับขั้นตอนวิธีเชิงพันธุกรรม จำเป็นต้องมีการกำหนดค่าพารามิเตอร์ต่างๆ โดยทุกปัญหาที่นำมาใช้ จะถูกทดสอบด้วยค่าพารามิเตอร์เดียวกัน เพื่อนำผลลัพธ์ที่ได้มาเปรียบเทียบ ซึ่งค่าพารามิเตอร์ที่ใช้ในงานวิจัยครั้งนี้ [4] จะแสดงใน ตารางที่ 5

ตารางที่ 6 การตั้งค่าพารามิเตอร์ (Parameter Setting)

Parameter	Value
Chromosome length	30,45,60
Population size	30,80
Tournament size	4
Reproduction rate	0.1
Crossover rate	0.8
Mutation rate	0.1
Number of generations	500
Number of runs	100

จากตารางที่ 6 ได้กำหนดไว้สำหรับการทดสอบในปัญหาต่างๆ โดยมีการกำหนดจำนวนสมาชิกประชากร (Population size) เป็น 30 และ 80 สำหรับความยาวของสายโครโมโซม (Chromosome length) เป็น 30,45, และ 60 บิต กำหนดค่าอัตราการสืบพันธุ์ (Reproduction rate) เป็น 0.1 ค่าอัตราการกลายพันธุ์ (Mutation rate) เป็น 0.1 โดยเป็นการกลายพันธุ์แบบจุดเดียว (Single Point Mutation) และค่าอัตราการไขว้เปลี่ยน (Crossover rate) เป็น 0.8 โดยเป็นการไขว้เปลี่ยนแบบจุดเดียว (Single Point Crossover) ในการรันโปรแกรมแต่ละครั้งกำหนดจำนวนรุ่นสูงสุด ที่ใช้ในการวิวัฒนาการคำตอบที่ 500 และทดลองซ้ำ (Number of runs) จำนวน 100 รอบ เพื่อใช้การหาค่าเฉลี่ย (Average) ของรุ่นที่พบคำตอบ (Generation) และค่าความเหมาะสมของคำตอบ

3.4 เครื่องมือและอุปกรณ์ที่ใช้ในการวิจัย

งานวิจัยนี้เป็นงานวิจัยเชิงทดลอง เพื่อเปรียบเทียบประสิทธิภาพ เครื่องมือและอุปกรณ์ที่ใช้ในงานวิจัยมีดังนี้

ฮาร์ดแวร์ที่ใช้ในการวิจัยประกอบด้วย

โน้ตบุ๊ก MSI GP72M 7REX-1046XTH LEOPARD PRO

ซีพียู Intel Core i7-7700HQ

จีพียู GeForce GTX 1050Ti (4GB GDDR5)

ฮาร์ดดิสก์ความจุ 1 TB

หน่วยความจำ 8 GB

ซอฟต์แวร์ที่ใช้ในการวิจัยประกอบด้วย

ระบบปฏิบัติการ Windows 10 Home

Dev-c++ Version 5.11



บทที่ 4

ผลการวิจัย

ในบทนี้ ผู้วิจัยได้นำเสนอผลดำเนินงาน จากการทดลองกับเทคนิคการคัดเลือก 3 เทคนิค ได้แก่ เทคนิคการเลือกตามด้วยระยะทางแฮมมิง (HS) เทคนิคการเลือกตามด้วยสหสัมพันธ์โคไซน์ (CS) และ เทคนิคการเลือกตามด้วยสหสัมพันธ์เพียร์สัน (PS) ซึ่งจะนำไปทดลองเปรียบเทียบกับเทคนิคการคัดเลือกแบบดั้งเดิม 2 เทคนิค ได้แก่ การคัดเลือกแบบวงล้อรูเล็ต การคัดเลือกแบบการแข่งขัน ในการทดลองครั้งนี้ เทคนิคการคัดเลือกทั้งหมด จะได้รับการประเมินด้วย 4 ปัญหา ได้แก่ ปัญหาชนิดหนึ่งมากที่สุด ปัญหาชนิดศูนย์มากที่สุด ปัญหาชนิดสูงสุดที่เกิดจากการสุ่ม ปัญหาจับตัก ขนาดของปัญหาเท่ากับ 3 และ 5 ด้วยการกำหนดค่าพารามิเตอร์ไว้ดังนี้

1. ขนาดของประชากร (Population size) 30
 - ขนาดของปัญหา (Problem size) 30
 - ขนาดของปัญหา (Problem size) 45
 - ขนาดของปัญหา (Problem size) 60
2. ขนาดของประชากร (Population size) 80
 - ขนาดของปัญหา (Problem size) 30
 - ขนาดของปัญหา (Problem size) 45
 - ขนาดของปัญหา (Problem size) 60

สำหรับเกณฑ์การวัดผลในงานวิจัยครั้งนี้ จะวัดผลด้วยจำนวนรุ่นที่พบคำตอบ และค่าความเหมาะสม โดยในงานวิจัยนี้ จำนวนรุ่นที่พบคำตอบ คือ ค่าเฉลี่ยของรุ่นที่พบคำตอบหรือกรณีที่ไม่มีพบคำตอบจะเป็นจำนวนรุ่นสูงสุด และค่าความเหมาะสม คือ ค่าเฉลี่ยของค่าความเหมาะสม จากการทดลองซ้ำ จำนวน 100 รอบ

4.1 ผลการทดสอบประสิทธิภาพเทคนิคการเลือกตามระยะทาง

ผลการทดสอบประสิทธิภาพการคัดเลือกตามระยะทาง ด้วยเทคนิคต่างๆ ด้วยการทดสอบหาประสิทธิภาพของเทคนิค ได้ผลลัพธ์ดังนี้

4.1.1 การคัดเลือกด้วยเทคนิคแบบประลอง

ตารางที่ 7 ผลการทดสอบการคัดเลือกด้วยเทคนิคแบบประลอง

Problem	การคัดเลือกด้วยเทคนิคแบบประลอง			
	Population Size	Problem Size	Generation	Fitness Value
One Max	30	30	44.82	30
	30	45	79.61	45
	30	60	125.62	60
	80	30	20.97	30
	80	45	37.90	45
	80	60	58.90	60
Zero Max	30	30	46.54	30
	30	45	82.62	45
	30	60	123.52	60
	80	30	20.33	30
	80	45	37.03	45
	80	60	59.94	60
Random Max	30	30	42.59	30
	30	45	87.30	45
	30	60	121.45	60
	80	30	19.71	30
	80	45	38.28	45
	80	60	58.22	60
Trap (3)	30	30	500	10.04
	30	45	500	15.34
	30	60	500	20.53
	80	30	500	10.12
	80	45	500	15.46
	80	60	500	20.46
Trap (5)	30	30	500	16.46
	30	45	500	22.68
	30	60	500	29.36
	80	30	500	17.78
	80	45	500	24.16
	80	60	500	30.72

จากตารางที่ 7 พบว่าการคัดเลือกด้วยเทคนิคแบบประลอง ในปัญหาบิตหนึ่งมากที่สุด, ปัญหาบิตศูนย์มากที่สุด, และปัญหาบิตสูงสุดที่เกิดจากการสุ่ม สามารถหาคำตอบได้ แต่ในปัญหากับดักขนาดกับดัก 3 และ 5 ไม่สามารถหาคำตอบได้ โดยพบค่าเฉลี่ยของรุ่นที่พบคำตอบดีที่สุดที่ 19.71 ในปัญหาบิตสูงสุดที่เกิดจากการสุ่ม ด้วยขนาดของประชากร 80 และขนาดของปัญหาเท่ากับ 30



4.1.2 การคัดเลือกด้วยเทคนิควงล้อรูเล็ต

ตารางที่ 8 ผลการทดสอบการคัดเลือกด้วยเทคนิควงล้อรูเล็ต

Problem	การคัดเลือกด้วยเทคนิควงล้อรูเล็ต			
	Population Size	Problem Size	Generation	Fitness Value
One Max	30	30	50.60	30
	30	45	102.19	45
	30	60	153.15	60
	80	30	21.78	30
	80	45	41.78	45
	80	60	70.04	60
Zero Max	30	30	48.03	30
	30	45	98.10	45
	30	60	159.04	60
	80	30	23.17	30
	80	45	43.32	45
	80	60	68.07	60
Random Max	30	30	51.57	30
	30	45	100.98	45
	30	60	150.03	60
	80	30	22.13	30
	80	45	41.65	45
	80	60	67.94	60
Trap (3)	30	30	500	26.30
	30	45	500	38.20
	30	60	500	50.15
	80	30	419.15	28.55
	80	45	490.83	41.33
	80	60	495.44	54.06
Trap (5)	30	30	500	26.03
	30	45	500	38.48
	30	60	500	51.14
	80	30	485.46	27.56
	80	45	500	40.41
	80	60	500	52.76

จากตารางที่ 8 พบว่าการคัดเลือกด้วยเทคนิควงล้อสุ่ม ในปัญหาบิตหนึ่งมากที่สุด, ปัญหาบิตศูนย์มากที่สุด, และปัญหาบิตสูงสุดที่เกิดจากการสุ่ม สามารถหาคำตอบได้ แต่ในปัญหากับดักขนาดกับดัก 3 สามารถหาคำตอบได้ ในขนาดของประชากร 80 และในปัญหากับดัก ขนาดกับดัก 5 สามารถหาคำตอบได้ ในขนาดของประชากร 80 และขนาดของปัญหาเท่ากับ 30 พบค่าเฉลี่ยของรุ่นที่พบคำตอบดีที่สุดที่ 27.56



4.1.3 การคัดเลือกด้วยเทคนิคระยะทางแฮมมิง

ตารางที่ 9 ผลการทดสอบการคัดเลือกด้วยเทคนิควงล้อรูเล็ต

Problem	การคัดเลือกด้วยเทคนิคระยะทางแฮมมิง			
	Population Size	Problem Size	Generation	Fitness Value
One Max	30	30	80.73	30
	30	45	172.82	44.99
	30	60	280.78	59.96
	80	30	41.95	30
	80	45	92.80	45
	80	60	150.01	60
Zero Max	30	30	81.46	30
	30	45	169.69	45
	30	60	271.81	59.97
	80	30	46	30
	80	45	95.81	45
	80	60	143.16	60
Random Max	30	30	76.17	30
	30	45	167.06	45
	30	60	281.93	59.95
	80	30	44.59	30
	80	45	90.09	45
	80	60	151.52	60
Trap (3)	30	30	364.22	28.73
	30	45	491.32	39.54
	30	60	500	50.69
	80	30	104.45	29.99
	80	45	335.72	44.06
	80	60	489.28	56.56
Trap (5)	30	30	494.11	27.35
	30	45	500	39.81
	30	60	500	51.84
	80	30	386.74	29.13
	80	45	475.73	42.55
	80	60	500	55.33

จากตารางที่ 9 พบว่าการคัดเลือกด้วยเทคนิคระยะทางแฮมมิง ในทุกปัญหาสามารถหาคำตอบได้ โดยปัญหาบิตหนึ่งมากที่สุด, ปัญหาบิตศูนย์มากที่สุด, และปัญหาบิตสูงสุดที่เกิดจากการสุ่มขนาดของประชากร 30 และขนาดของปัญหาเท่ากับ 60 ไม่สามารถหาคำตอบได้ทุกครั้ง และปัญหาบิตหนึ่งมากที่สุด ขนาดของประชากร 30 และขนาดของปัญหาเท่ากับ 45 มีโอกาสน้อยมากไม่สามารถหาคำตอบได้ โดยในปัญหากับดัก ขนาดกับดัก 3 ไม่สามารถหาคำตอบได้ ในขนาดของประชากร 30 และขนาดของปัญหาเท่ากับ 60 และในปัญหากับดัก ขนาดกับดัก 5 ไม่สามารถหาคำตอบได้ ในขนาดของประชากร 30 และ 45 และขนาดของปัญหาเท่ากับ 30 และ 60 พบค่าเฉลี่ยของรุ่นที่พบคำตอบที่ดีที่สุดที่ 41.95 ในปัญหาบิตหนึ่งมากที่สุด ด้วยขนาดของประชากร 80 และขนาดของปัญหาเท่ากับ 30



4.1.4 การคัดเลือกด้วยเทคนิคสหสัมพันธ์โคไซน์

ตารางที่ 10 ผลการทดสอบการคัดเลือกด้วยเทคนิคสหสัมพันธ์โคไซน์

Problem	การคัดเลือกด้วยเทคนิคสหสัมพันธ์โคไซน์			
	Population Size	Problem Size	Generation	Fitness Value
One Max	30	30	54.43	30
	30	45	113.55	45
	30	60	159.06	60
	80	30	26.41	30
	80	45	50.23	45
	80	60	74.05	60
Zero Max	30	30	61.67	30
	30	45	121.35	45
	30	60	171.43	60
	80	30	32.31	30
	80	45	55.55	45
	80	60	79.83	60
Random Max	30	30	52.65	30
	30	45	111.04	45
	30	60	161.31	60
	80	30	25.81	30
	80	45	48	45
	80	60	74.01	60
Trap (3)	30	30	500	24.59
	30	45	471.29	34
	30	60	353.91	47.11
	80	30	500	26.71
	80	45	500	38
	80	60	500	48.98
Trap (5)	30	30	500	25.65
	30	45	429.89	41.14
	30	60	473.46	46.83
	80	30	500	26.98
	80	45	500	38.68
	80	60	500	49.94

จากตารางที่ 10 พบว่าการคัดเลือกด้วยเทคนิคสหสัมพันธ์โคไซน์ ในทุกปัญหาสามารถหาคำตอบได้ แต่ในปัญหากับดัก ขนาดกับดัก 3 และ 5 สามารถหาคำตอบได้ ในขนาดของประชากร และขนาดของปัญหา 30, 45 และ 30, 60 พบค่าเฉลี่ยของรุ่นที่พบคำตอบดีที่สุดที่ 25.81 ในปัญหา บิตสูงสุดที่เกิดจากการสุ่ม (Random Max Problem) ด้วยขนาดของประชากร และขนาดของปัญหา ที่ 80, 30



4.1.5 การคัดเลือกด้วยสหสัมพันธ์เพียร์สัน

ตารางที่ 11 ผลการทดสอบการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน

Problem	การคัดเลือกด้วยสหสัมพันธ์เพียร์สัน			
	Population Size	Problem Size	Generation	Fitness Value
One Max	30	30	61.85	30
	30	45	110.93	45
	30	60	168.27	60
	80	30	33.29	30
	80	45	53.87	45
	80	60	83.46	60
Zero Max	30	30	71.98	30
	30	45	128.51	45
	30	60	185.84	60
	80	30	33.06	30
	80	45	59.42	45
	80	60	88.56	60
Random Max	30	30	82.25	30
	30	45	143.61	45
	30	60	200.02	60
	80	30	40.9	30
	80	45	73.67	45
	80	60	103.88	60
Trap (3)	30	30	460.90	24.34
	30	45	313.17	36.98
	30	60	134.06	56.98
	80	30	356.62	28.19
	80	45	493.88	36
	80	60	496.34	43.41
Trap (5)	30	30	480.68	21.75
	30	45	388.74	33.54
	30	60	297.30	52.50
	80	30	465.04	26.44
	80	45	495.42	32.75
	80	60	500	37.12

จากตารางที่ 11 พบว่าการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน ในทุกปัญหาสามารถหาคำตอบได้ ยกเว้นในปัญหากับดัก ขนาดกับดัก 5 ไม่สามารถหาคำตอบได้ ในขนาดของประชากร และขนาดของปัญหา 60 พบค่าเฉลี่ยของรุ่นที่พบคำตอบดีที่สุดที่ 33.06 ในปัญหาบิตศูนย์มากที่สุด ด้วยขนาดของประชากรเท่ากับ 80 และขนาดของปัญหาเท่ากับ 30

จากผลการทดสอบในวิธีการต่างๆ สามารถสรุปเป็นตารางได้ดังนี้

ตารางที่ 12 จำนวนครั้งที่พบคำตอบของการคัดเลือกด้วยเทคนิคแบบประลอง

การคัดเลือกด้วยเทคนิคแบบประลอง						
Population Size	Problem Size	One max	Zero max	Random max	Trap3	Trap5
30	30	100	100	100	0	0
30	45	100	100	100	0	0
30	60	100	100	100	0	0
80	30	100	100	100	0	0
80	45	100	100	100	0	0
80	60	100	100	100	0	0

จากตารางที่ 12 พบว่าการคัดเลือกด้วยเทคนิคแบบประลอง ในปัญหาบิตหนึ่งมากที่สุด, ปัญหาบิตศูนย์มากที่สุด, และปัญหาบิตสูงสุดที่เกิดจากการสุ่ม สามารถพบคำตอบได้ทุกครั้ง แต่ในปัญหากับดัก ขนาดกับดัก 3 และ 5 ไม่สามารถพบคำตอบได้

ตารางที่ 13 จำนวนครั้งที่พบคำตอบของการคัดเลือกด้วยเทคนิควงล้อรูเล็ต

การคัดเลือกด้วยเทคนิควงล้อรูเล็ต						
Population Size	Problem Size	One max	Zero max	Random max	Trap3	Trap5
30	30	100	100	100	0	0
30	45	100	100	100	0	0
30	60	100	100	100	0	0
80	30	100	100	100	22	3
80	45	100	100	100	3	0
80	60	100	100	100	1	0

จากตารางที่ 13 พบว่าการคัดเลือกด้วยเทคนิควงล้อรูเล็ต ในปัญหาบิตหนึ่งมากที่สุด, ปัญหาบิตศูนย์มากที่สุด, และปัญหาบิตสูงสุดที่เกิดจากการสุ่ม สามารถพบคำตอบได้ทุกครั้ง แต่ในปัญหากับดัก ขนาดกับดัก 3 สามารถหาคำตอบได้ ในขนาดของประชากร 80 และในปัญหากับดัก ขนาดกับดัก 5 สามารถหาคำตอบได้ ในขนาดของประชากร 80 และขนาดของปัญหา 30 พบคำตอบได้จำนวน 3 ครั้ง จากจำนวน 100 ครั้งที่หาคำตอบ

ตารางที่ 14 จำนวนครั้งที่พบคำตอบของการคัดเลือกด้วยเทคนิคระยะทางแฮมมิง

การคัดเลือกด้วยเทคนิคระยะทางแฮมมิง						
Population Size	Problem Size	One max	Zero max	Random max	Trap3	Trap5
30	30	100	100	100	61	2
30	45	99	100	100	5	0
30	60	96	97	95	0	0
80	30	100	100	100	99	35
80	45	100	100	100	62	7
80	60	100	100	100	8	0

จากตารางที่ 14 พบว่าการคัดเลือกด้วยเทคนิคระยะทางแฮมมิง (HS) ในปัญหาบิตหนึ่งมากที่สุด, ปัญหาบิตศูนย์มากที่สุด, และปัญหาบิตสูงสุดที่เกิดจากการสุ่ม สามารถพบคำตอบได้เกือบทุกครั้ง ที่ทดสอบ แต่ในปัญหากับดัก ขนาดกับดัก 3 และขนาดกับดัก 5 สามารถพบคำตอบได้ในบางครั้งของการทดสอบ แต่ในขนาดของประชากร 30 และขนาดของปัญหา 60 ไม่สามารถพบคำตอบได้เลย

ตารางที่ 15 จำนวนครั้งที่พบคำตอบของการคัดเลือกด้วยเทคนิคสหสัมพันธ์โคไซน์

การคัดเลือกด้วยเทคนิคสหสัมพันธ์โคไซน์						
Population Size	Problem Size	One max	Zero max	Random max	Trap3	Trap5
30	30	100	100	100	0	0
30	45	100	100	100	11	37
30	60	100	100	100	36	15
80	30	100	100	100	0	0
80	45	100	100	100	0	0
80	60	100	100	100	0	0

จากตารางที่ 15 พบว่าการคัดเลือกด้วยเทคนิคสหสัมพันธ์โคไซน์ (CS) ในปัญหาบิตหนึ่งมากที่สุด, ปัญหาบิตศูนย์มากที่สุด, และปัญหาบิตสูงสุดที่เกิดจากการสุ่ม สามารถพบคำตอบได้ทุกครั้ง แต่ในปัญหากับดัก ขนาดกับดัก 3 และขนาดกับดัก 5 ในขนาดของประชากร 30 และขนาดของปัญหา 45 และขนาดของประชากร 30 และขนาดของปัญหา 60 สามารถพบคำตอบได้ในบางครั้งของการทดสอบ

ตารางที่ 16 จำนวนครั้งที่พบคำตอบของการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน

การคัดเลือกด้วยสหสัมพันธ์เพียร์สัน						
Population Size	Problem Size	One max	Zero max	Random max	Trap3	Trap5
30	30	100	100	100	16	7
30	45	100	100	100	54	37
30	60	100	100	100	88	55
80	30	100	100	100	47	16
80	45	100	100	100	3	1
80	60	100	100	100	2	0

จากตารางที่ 16 พบว่าการคัดเลือกด้วยเทคนิคสหสัมพันธ์เพียร์สัน ในปัญหาบิตหนึ่งมากที่สุด , ปัญหาบิตศูนย์มากที่สุด, และปัญหาบิตสูงสุดที่เกิดจากการสุ่ม สามารถพบคำตอบได้ทุกครั้ง แต่ในปัญหาก็บิตก ขนาดกับบิตก 3 และขนาดกับบิตก 5 สามารถพบคำตอบได้เกือบทุกการตั้งค่าพารามิเตอร์ แต่ในปัญหาก็บิตก ขนาดกับบิตก 5 ขนาดของประชากร 80 และขนาดของปัญหา 60 ไม่สามารถพบคำตอบได้

จากการทดสอบการคัดเลือกด้วยปัญหาต่างๆ ที่นำมาใช้ในงานวิจัยนี้ สามารถสรุปเป็นตารางการเปรียบเทียบ ในตารางที่ 17 และ 18 ดังนี้



ตารางที่ 17 ผลการทดสอบเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรม (เฉพาะปัญหาบิตหนึ่งมากที่สุด ปัญหาบิตศูนย์มากที่สุด และปัญหาบิตสูงสุดจากการสุ่ม)

Problem	Population Size & Problem Size	Average and standard deviation of numbers of generations				
		RWS	TS	HS	CS	PS
One max	30,30	50.60±17.84	44.82±15.05	80.73±39.86	54.43±19.78	61.85±15.87
	30,45	102.19±27.28	79.61±19.03	172.82±79.58	113.55±37.78	110.93±30.18
	30,60	153.15±36.50	125.62±29.45	280.78±96.70	159.06±38.41	168.27±38.12
	80,30	21.78±4.57	20.97±7.13	41.95±14.84	26.41±7.29	33.29±6.56
	80,45	41.78±8.49	37.90±10.37	92.80±26.60	50.23±13.32	53.87±11.77
	80,60	70.04±18.85	58.90±12.20	150.01±37.55	74.05±14.46	83.46±15.37
Zero max	30,30	48.03±18.14	46.54±14.35	81.46±41.67	61.67±18.69	71.98±22.29
	30,45	98.10±26.46	82.62±22.16	169.69±61.77	121.35±36.31	128.51±38.03
	30,60	159.04±40.83	123.52±28.14	271.81±91.83	171.43±46.70	185.84±38.04
	80,30	23.17±6.44	20.33±6.58	46.00±13.84	32.31±8.22	33.06±8.51
	80,45	43.32±10.02	37.03±8.09	95.81±37.72	55.55±13.62	59.42±12.46
	80,60	68.07±14.14	59.94±11.58	143.16±36.19	79.83±20.06	88.56±18.66
Random max	30,30	51.57±19.78	42.59±13.05	76.17±37.73	52.65±19.58	82.25±26.02
	30,45	100.98±37.78	87.30±23.38	167.06±56.38	111.04±33.62	143.61±42.53
	30,60	150.03±38.41	121.45±29.38	281.93±96.45	161.31±37.27	200.02±48.86
	80,30	22.13±7.29	19.71±5.67	44.59±15.65	25.81±6.14	40.90±13.01
	80,45	41.65±13.32	38.28±8.67	90.09±25.92	48.00±10.87	73.67±21.62
	80,60	67.94±14.46	58.22±12.27	151.52±44.60	74.01±18.21	103.88±23.41

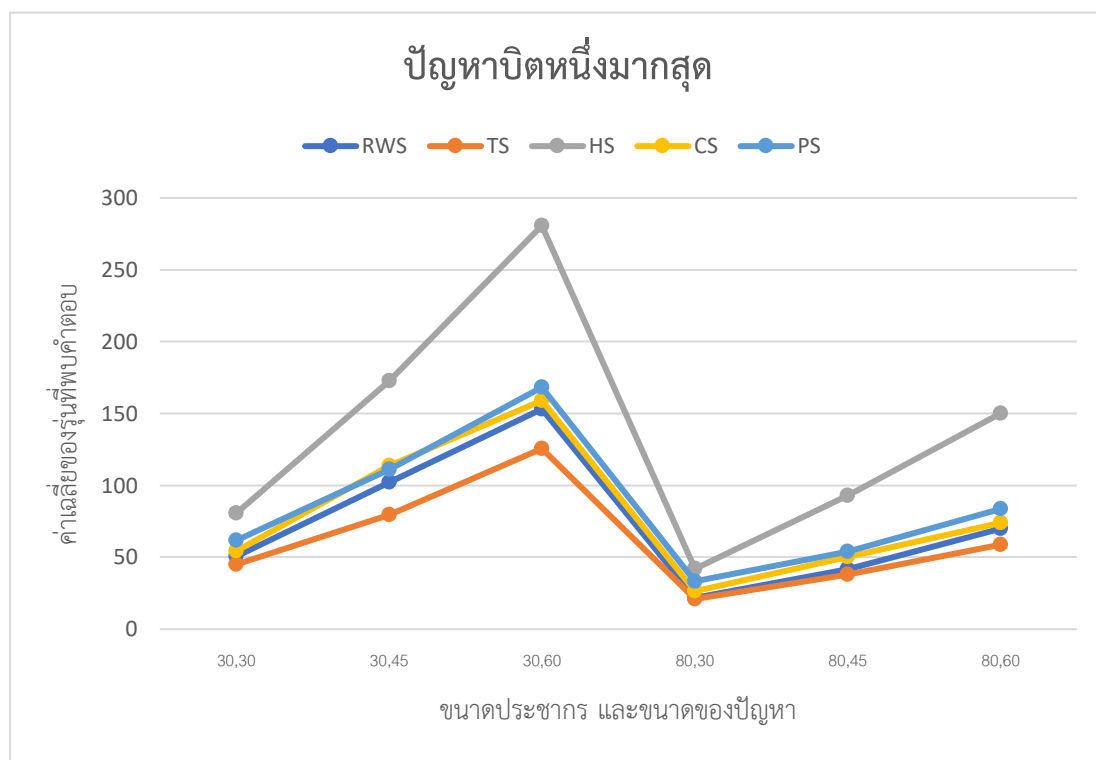
ตารางที่ 18 ผลการทดสอบเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรม (เฉพาะปัญหาที่ติด ขนาดที่ติด 3 และขนาดที่ติด 5)

Problem	Population Size & Problem Size	Average and standard deviation of numbers of generations				
		RWS	TS	HS	CS	PS
Trap(3)	30,30	500.00±19.78	500.00±0.00	364.22±19.78	500.00±0.00	460.90±102.93
	30,45	500.00±37.78	500.00±0.00	491.32±37.78	471.29±96.62	313.17±192.86
	30,60	500.00±38.41	500.00±0.00	500.00±38.41	353.91±199.83	134.06±162.95
	80,30	419.15±7.29	500.00±0.00	104.45±7.29	500.00±0.00	356.62±184.45
	80,45	490.83±13.32	500.00±0.00	335.72±13.32	500.00±0.00	493.88±39.26
	80,60	495.44±14.46	500.00±0.00	489.28±14.46	500.00±0.00	496.34±26.25
Trap(5)	30,30	500.00±0.00	500.00±0.00	494.11±44.16	500.00±0.00	480.68±73.94
	30,45	500.00±0.00	500.00±0.00	500.00±0.00	429.89±115.74	388.74±163.34
	30,60	500.00±0.00	500.00±0.00	500.00±0.00	473.46±79.35	297.30±203.51
	80,30	485.46±83.10	500.00±0.00	386.74±178.53	500.00±0.00	465.04±98.73
	80,45	500.00±0.00	500.00±0.00	475.73±95.18	500.00±0.00	495.42±45.80
	80,60	500.00±0.00	500.00±0.00	500.00±0.00	500.00±0.00	500.00±0.00

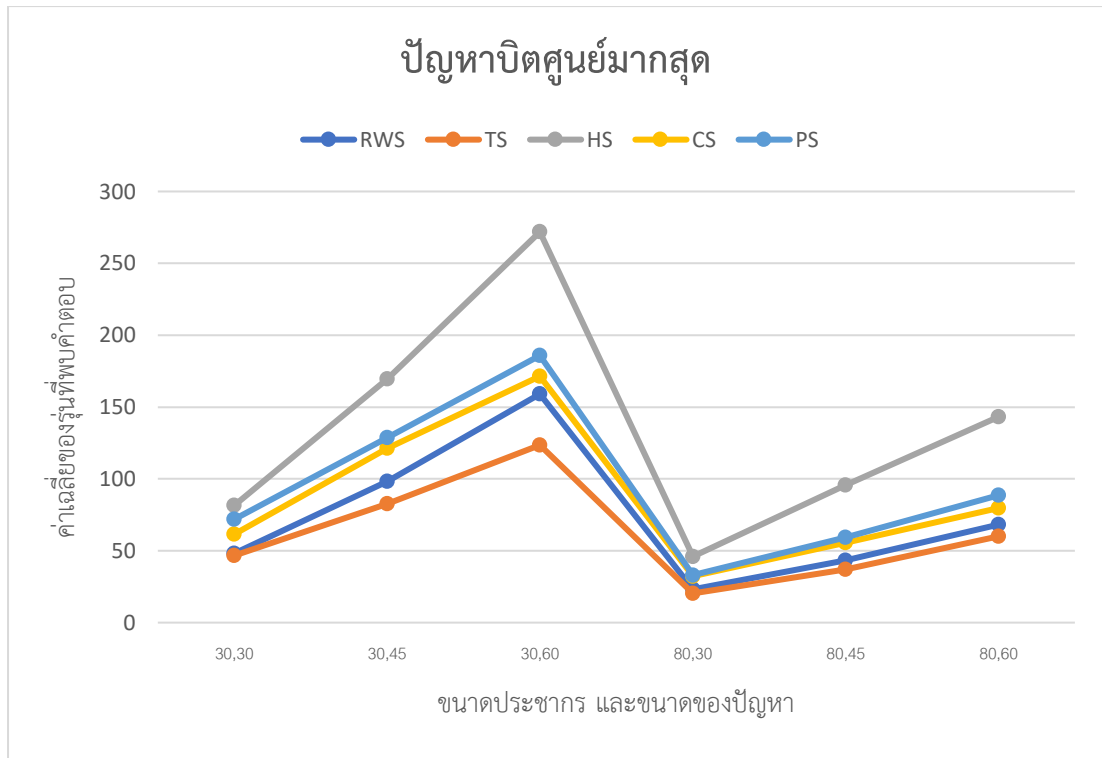
จากผลการทดสอบพบว่า เทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรมสามารถแก้ปัญหาได้ดี หากค่าขนาดของประชากร และขนาดของปัญหาที่ต่ำ และสามารถแก้ปัญหาได้ดีกว่าสำหรับค่าขนาดของประชากร และขนาดของปัญหาที่สูง การนำไปทดสอบกับปัญหาที่ติด ที่มีขนาดที่ติด 5 มีการตั้งค่าพารามิเตอร์ที่ 80, 60 ในการเปรียบเทียบครั้งนี้ ไม่มีวิธีใดที่สามารถแก้ปัญหาได้ และไม่มีวิธีการใดที่สามารถแก้ปัญหาในการรันทั้งหมด ค่าเฉลี่ยของรุ่นที่ดีที่สุดได้แก่ 21.78, 41.95, 20.33, 33.06 และ 25.81 สำหรับการคัดเลือกแบบวงล้อสุ่ม การคัดเลือก

แบบการแข่งขัน เทคนิคการคัดเลือกด้วยระยะทางแฮมมิงg เทคนิคการคัดเลือกด้วยสหสัมพันธ์โคไซน์ และเทคนิคการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน ตามลำดับ

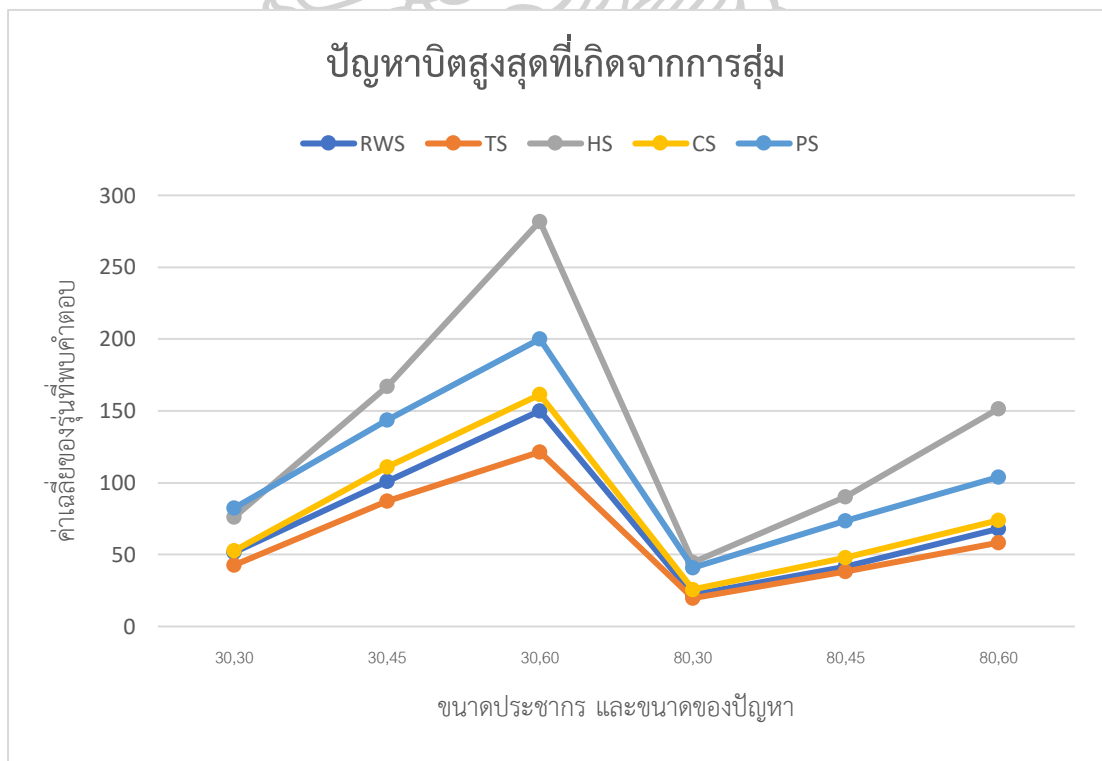
สามารถแสดงผลในรูปแบบแผนภูมิเส้น (line chart) ในภาพที่ 14 – 18 เพื่อให้เห็นผลลัพธ์ที่ชัดเจนยิ่งขึ้น โดยแต่ละแผนภูมิเส้นจะแสดง 1 ปัญหาที่ใช้ในการทดสอบ โดยเปรียบเทียบกับเทคนิค และการตั้งค่าพารามิเตอร์ ที่นำเสนอในงานวิจัยนี้



ภาพที่ 14 แผนภูมิเส้นเปรียบเทียบในปัญหาบิตหนึ่งมากที่สุด

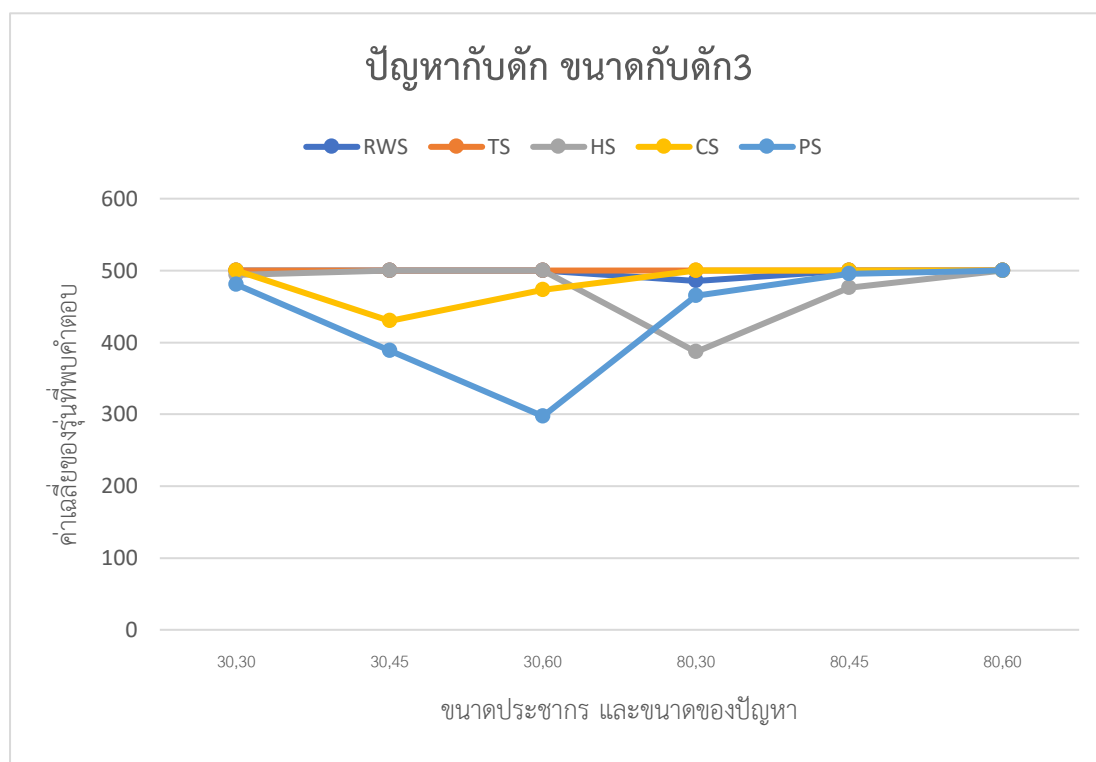


ภาพที่ 15 แผนภูมิเส้นเปรียบเทียบในปัญหาบิตศูนย์มากที่สุด



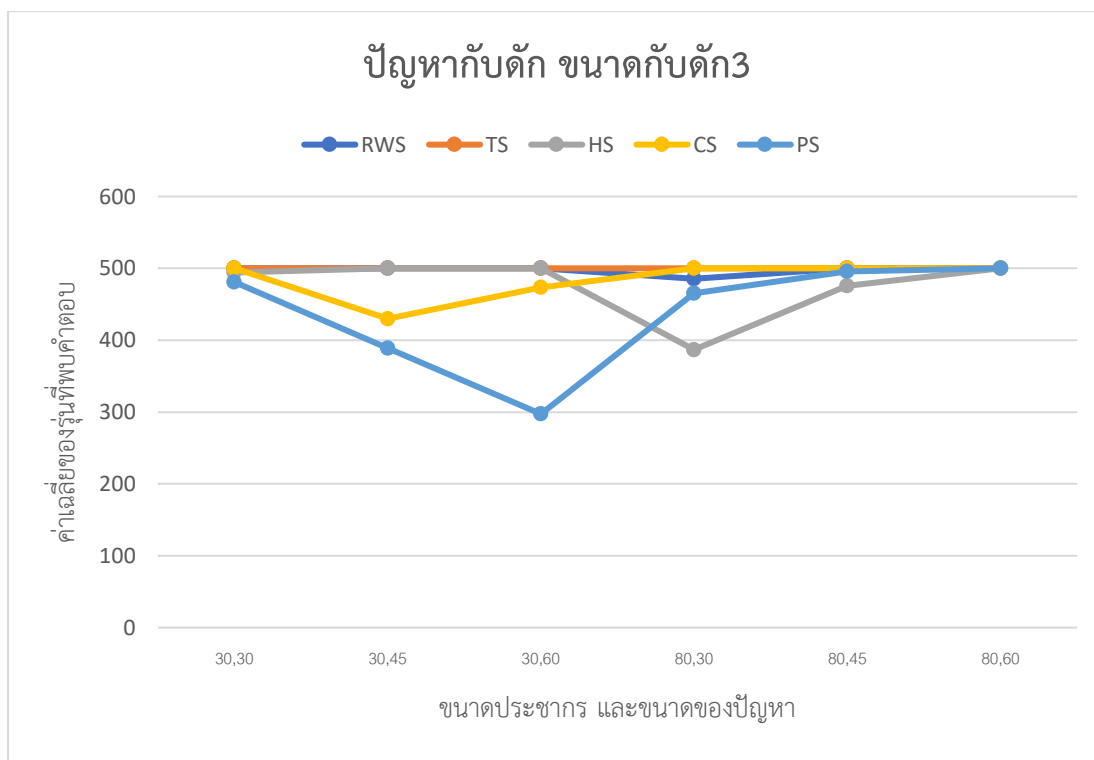
ภาพที่ 16 แผนภูมิเส้นเปรียบเทียบในปัญหาบิตสูงที่สุดที่เกิดจากการสุ่ม

จากภาพที่ 14, 15, และ 16 ในปัญหาบิตหนึ่งมากที่สุด ปัญหาบิตศูนย์มากที่สุด และปัญหาบิตสูงที่สุดที่เกิดจากการสุ่ม การคัดเลือกแบบการแข่งขันมีประสิทธิภาพในการหาคำตอบได้ดีที่สุด และเทคนิคการคัดเลือกด้วยระยะทางแฮมมิงมีประสิทธิภาพในการหาคำตอบได้น้อยที่สุด ตามแผนภูมิเส้นที่ได้นำเสนอ ซึ่งทั้ง 3 ปัญหาดังกล่าว มีประสิทธิภาพในการหาคำตอบได้ใกล้เคียงกันมาก



ภาพที่ 17 แผนภูมิเส้นเปรียบเทียบในปัญหากับดัก ขนาดกับดัก 3

จากภาพที่ 17 ในปัญหากับดัก ขนาดกับดัก 3 เทคนิคการคัดเลือกด้วยระยะทางแฮมมิง มีประสิทธิภาพในการหาคำตอบได้ดีที่สุด แต่ในการตั้งค่าพารามิเตอร์ ขนาดของประชากร 30 ขนาดของปัญหา 45 และขนาดของประชากร 30 ขนาดของปัญหา 60 ไม่มีประสิทธิภาพในการหาคำตอบได้ เทคนิคการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน มีประสิทธิภาพในการหาคำตอบได้ใกล้เคียงกับเทคนิคการคัดเลือกด้วยระยะทางแฮมมิง แต่ในการตั้งค่าพารามิเตอร์ ขนาดของประชากร 80 ขนาดของปัญหา 45 และขนาดของประชากร 80 ขนาดของปัญหา 60 ไม่มีประสิทธิภาพในการหาคำตอบได้ และการคัดเลือกแบบการแข่งขัน ไม่มีประสิทธิภาพในการหาคำตอบในปัญหานี้



ภาพที่ 18 แผนภูมิเส้นเปรียบเทียบในปัญหากับดัก ขนาดกับดัก 5

จากภาพที่ 18 ในปัญหากับดัก ขนาดกับดัก 5 เทคนิคการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน มีประสิทธิภาพในการหาคำตอบได้ดีที่สุด แต่ในการตั้งค่าพารามิเตอร์ ขนาดของประชากร 80 ขนาดของปัญหา 45 และขนาดของประชากร 80 ขนาดของปัญหา 60 ไม่มีประสิทธิภาพในการหาคำตอบได้ และการคัดเลือกแบบการแข่งขัน ไม่มีประสิทธิภาพในการหาคำตอบในปัญหานี้

จากภาพที่ 17 และ 18 ในปัญหากับดัก ขนาดกับดัก 3 และขนาดกับดัก 5 เทคนิคการคัดเลือกด้วยระยะทางแฮมมิง และเทคนิคการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน มีประสิทธิภาพในการหาคำตอบได้ดีใกล้เคียงกัน หากขนาดของประชากร และขนาดของปัญหา ที่ใกล้เคียงในทุกเทคนิควิธีการ จะไม่มีประสิทธิภาพในการหาคำตอบ ในเทคนิคการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน และเทคนิคการคัดเลือกด้วยสหสัมพันธ์โคไซน์ หากขนาดของประชากร น้อยกว่าขนาดของปัญหา จะมีประสิทธิภาพในการหาคำตอบที่ดี แต่เทคนิคการคัดเลือกด้วยระยะทางแฮมมิง หากขนาดของประชากร มากกว่าขนาดของปัญหา จะมีประสิทธิภาพในการหาคำตอบที่ดี และการคัดเลือกแบบการแข่งขัน ไม่มีประสิทธิภาพในการหาคำตอบในปัญหานี้

4.2 สรุป

ผลการเปรียบเทียบการทดสอบเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรม ด้วยวิธีการการคัดเลือกด้วยเทคนิควงล้อรูเล็ก, วิธีการคัดเลือกด้วยเทคนิคแบบประลอง การคัดเลือกด้วยเทคนิคระยะทางแฮมมิง การคัดเลือกด้วยเทคนิคสหสัมพันธ์โคไซน์ และการคัดเลือกด้วยสหสัมพันธ์เพียร์สัน ในปัญหาบิตหนึ่งมากที่สุด ปัญหาบิตศูนย์มากที่สุด ปัญหาบิตสูงสุดที่เกิดจากการสุ่ม และปัญหากับดัก ขนาดกับดัก 3 และขนาดกับดัก 5 พบว่า เทคนิคการเลือกตามระยะทาง มีประสิทธิภาพการหาคำตอบได้ดี ในปัญหากับดัก และเทคนิคการคัดเลือกมาตรฐาน มีประสิทธิภาพการหาคำตอบได้ดี ในปัญหาบิตหนึ่งมากที่สุด ปัญหาบิตศูนย์มากที่สุด และปัญหาบิตสูงสุดที่เกิดจากการสุ่ม



บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

สำหรับบทนี้เป็นเนื้อหาสุดท้ายของงานวิจัย จะประกอบด้วยหัวข้อทั้งหมด 3 หัวข้อ โดยมีรายละเอียดแต่ละหัวข้อดังนี้

5.1 สรุปผลและอภิปรายผล

5.2 ปัญหาและอุปสรรค

5.3 ข้อเสนอแนะ

5.1 สรุปผลและอภิปรายผล

การวิจัยครั้งนี้ได้ศึกษาเพื่อตอบวัตถุประสงค์ของการวิจัย เพื่อศึกษาและพัฒนาเทคนิคการเลือกตามระยะทางสำหรับการไขว้เปลี่ยนในขั้นตอนเชิงพันธุกรรมที่มีประสิทธิภาพ ซึ่งเป็นการศึกษาพัฒนา และทดสอบเทคนิควิธีการคัดเลือกตามระยะทาง ด้วยปัญหาดังนี้ ปัญหาบิตหนึ่งมากที่สุด ปัญหาบิตศูนย์มากที่สุด ปัญหาบิตสูงสุดที่เกิดจากการสุ่ม และปัญหากับดัก โดยนำมาเปรียบเทียบกับ การคัดเลือกด้วยเทคนิคมาตรฐานดังนี้ การคัดเลือกแบบการแข่งขัน และการคัดเลือกแบบวงล้อรูเล็ต ซึ่งจากการทดสอบพบว่า สำหรับปัญหามาตรฐาน ได้แก่ ปัญหาบิตหนึ่งมากที่สุด ปัญหาบิตศูนย์มากที่สุด และปัญหาบิตสูงสุดที่เกิดจากการสุ่ม ใช้การคัดเลือกด้วยเทคนิคแบบประลอง สามารถแก้ปัญหาได้ดีที่สุด ในขณะที่ปัญหาที่มีความซับซ้อนอย่างปัญหากับดัก การคัดเลือกด้วยค่าระยะทางสามารถช่วยให้หลุดจากกระบวนการปรับเปลี่ยนที่ให้ความสมมูล (local optima) ซึ่งสามารถช่วยเพิ่มประสิทธิภาพในการหาคำตอบสำหรับปัญหากับดักได้ดีกว่า

ขั้นตอนเชิงพันธุกรรมโดยปกติ หากใช้การคัดเลือกด้วยเทคนิคมาตรฐานอย่าง การคัดเลือกแบบการแข่งขัน และการคัดเลือกแบบวงล้อรูเล็ต จะช่วยสนับสนุนปัจจัยที่ส่งผลต่อการพบคำตอบ ในกรณีที่มีการนิยามการประเมินค่าความเหมาะสมได้อย่างเหมาะสม ซึ่งหมายความว่า สายโครโมโซมที่ใกล้เคียงคำตอบจะให้ค่าความเหมาะสมที่ดี จึงส่งผลให้กระบวนการเลือกมาเป็นคำตอบ ทำงานได้ดีมากขึ้น ในขณะที่หากไม่สามารถนิยามการประเมินค่าความเหมาะสมที่เหมาะสมได้ การคัดเลือกด้วยเทคนิคมาตรฐาน จึงไม่สามารถหาคู่ที่เหมาะสม เพื่อนำไปพัฒนาการหาคำตอบได้อย่างมีประสิทธิภาพ

ดังนั้นการออกแบบเทคนิคการคัดเลือกตามระยะทางโดยหาความแตกต่าง ตามที่ได้นำเสนอจะสามารถช่วยแก้ปัญหาดังกล่าวได้

5.2 ปัญหาและอุปสรรค

การวิธีการคัดเลือกตามระยะทาง ด้วยวิธีการต่างๆ จำเป็นต้องใช้เวลาในการรันโปรแกรมเป็นจำนวนมากและใช้เวลานาน หากเครื่องคอมพิวเตอร์ที่ใช้มีประสิทธิภาพที่ไม่ดีพอ

5.3 ข้อเสนอแนะ

5.3.1 งานวิจัยฉบับนี้เป็นการเปรียบเทียบกับวิธีการและปัญหาพื้นฐาน หากเป็นการนำไปใช้กับปัญหาจริง และมีการตั้งค่าพารามิเตอร์ที่มีความแตกต่างไปจากนี้ น่าจะสามารถนำมาทดสอบวิเคราะห์ แล้วเปรียบเทียบให้เห็นประสิทธิภาพมากยิ่งขึ้น

5.3.2 การเลือกใช้ปัญหา สำหรับการนำมาทดสอบเพิ่มเติม อาจสามารถเพิ่มประสิทธิภาพของวิธีการคัดเลือกตามระยะทางได้

5.3.3 วิธีการคัดเลือกตามระยะทางอาจจะทดลองปรับเลือกใช้วิธีการอื่นที่มีความหลากหลาย เพื่อทดสอบว่าวิธีการแบบใดจะเหมาะสมและให้ประสิทธิภาพที่ดี

5.3.4 อาจกำหนดขนาดของการแข่งขัน (Tournament Size) สำหรับการนำมาทดสอบเพิ่มเติม อาจสามารถเพิ่มประสิทธิภาพของวิธีการคัดเลือกตามระยะทางได้

5.3.5 การกำหนดให้ทุกสายโครโมโซมที่สุ่มตามขนาดของการแข่งขัน มีสิทธิ์จับคู่กันได้ทุกสาย อาจสามารถเพิ่มประสิทธิภาพของวิธีการคัดเลือกตามระยะทางได้

รายการอ้างอิง

1. Reeves, C., *Genetic Algorithm*. 2010: Hard book of Metaheuristics. 109-139.
2. Huang, C.-F. *A Study of Mate Selection Schemes in Genetic Algorithms-Part I*. in *Massachusetts Institute of Technology*. 2002. Citeseer.
3. Galan, S.F., O.J. Mengshoel, and R. Pinter, *A Novel Mating Approach for Genetic Algorithms*. *Evolutionary Computation*, 2013. 21(2): p. 197-229.
4. Pumsuwan, D. and S. Rimcharoen, *Performance Comparison and Analysis of Crossover Techniques in Genetic Algorithms for the Fuction of Unitation Problems*. *Science and Technology Mahasarakham University*, 2018. 37(4).
5. Mauldin, M.L. *Maintaining Diversity in Genetic Search*. in *AAAI*. 1984.
6. Golberg, D.E., *Genetic algorithms in search, optimization, and machine learning*. Addison wesley, 1989. 1989(102): p. 36.
7. Ruamsub, T., *Plant Leaf Classification Using Leaf Shape and Texture Features with Genetic Algorithm*. 2016, Sripatum University.
8. Orong, M.Y., A.M. Sison, and R.P. Medina. *A new crossover mechanism for genetic algorithm with rank-based selection method*. in *2018 5th International Conference on Business and Industrial Research (ICBIR)*. 2018. IEEE.
9. Boonlong, K., *Increment of Chromosome Length in Single-Objective Genetic Algorithm*. *Burapha Science Journal*, 2013. 18(1): p. 15-25.
10. Onnoi, P., *The Efficiency Improvement of Compact Genetic Algorithm Using Weighted Update by Fitness Values*. 2014, Nakhon Pathom Rajabhat University.
11. W., H.R., *Error detecting and error correcting code*. *Bell System Technical Journal*, 1950. 29(2): p. 147-160.
12. Manpiensuk, C. and S. Hengpraproh, *The Efficiency of the Similarity Function to the k-Nearest Neighbors Algorithm for Data Classification*, in the *8 th NPRU National Academic Conference*. 2016.
13. Harik, G., et al., *The gambler's ruin problem, genetic algorithms, and the sizing of populations*. *Evolutionary computation*, 1999. 7(3): p. 231-253.
14. Zhong, J., et al. *Comparison of performance between different selection*

- strategies on simple genetic algorithms.* in *International conference on computational intelligence for modelling, control and automation and international conference on intelligent agents, web technologies and internet commerce (CIMCA-IAWTIC'06)*. 2005. IEEE.
15. Ochoa, G., et al. *Assortative mating in genetic algorithms for dynamic problems.* in *Workshops on Applications of Evolutionary Computation*. 2005. Springer.
 16. Kim, Y.-H. and Y. Yoon, *Linkage-based distance metric in the search space of genetic algorithms.* *Mathematical Problems in Engineering*, 2015. 2015.
 17. Jung, C., et al., *A New Adaptive Hungarian Mating Scheme in Genetic Algorithms.* *Discrete Dynamics in Nature and Society*, 2016. 2016.
 18. Goldberg, D.E., K. Deb, and J. Horn. *Massive Multimodality, Deception, and Genetic Algorithms.* in *PPSN*. 1992.
 19. Jones, T. and S. Forrest. *Fitness distance correlation as a measure of problem difficulty for genetic algorithms.* in *ICGA*. 1995.
 20. Pranithanwitthaya, S. and S. Hengprapohm, *The Efficiency of the Feature Selection to the Data Classification, in the 4 th ASEAN Undergraduate Conference in Computing (AUC2) 2016.* 2016.
 21. Edelmann, D., T.F. Móri, and G.J. Székely, *On relationships between the Pearson and the distance correlation coefficients.* *Statistics & Probability Letters*, 2021. 169: p. 108960.

ประวัติผู้เขียน

ชื่อ-สกุล	นางสาวนิติมา ลักขณานุรักษ์
วัน เดือน ปี เกิด	8 กันยายน 2536
สถานที่เกิด	นครปฐม
วุฒิการศึกษา	จบการศึกษาระดับประกาศนียบัตรวิชาชีพ (ปวช.) ประเภทวิชาพาณิชยกรรม สาขาวิชาพาณิชยกรรม สาขางานคอมพิวเตอร์ธุรกิจ จบการศึกษาระดับปริญญาตรี คณะวิทยาการจัดการ สาขาวิชาคอมพิวเตอร์ธุรกิจ มหาวิทยาลัยราชภัฏนครปฐม
ที่อยู่ปัจจุบัน	48/2 หมู่ 6 ซอย 6 ตำบลโพรงมะเดื่อ อำเภอเมืองนครปฐม จังหวัดนครปฐม 73000
ผลงานตีพิมพ์	จักรกฤษณ์ หงส์เวียงจันทร์ นิติมา ลักขณานุรักษ์ และไกรรุ่ง เสงพระพรหม. (2562). การเปรียบเทียบประสิทธิภาพการจำแนกกลุ่มข้อมูล โรค ออทิสติกด้วยเทคนิคเหมืองข้อมูล. งานประชุมวิชาการระดับชาติ ครั้งที่ 11 มหาวิทยาลัยราชภัฏนครปฐม, 321-326. เกมส์ เหลืองประเสริฐ นิติมา ลักขณานุรักษ์ และไกรรุ่ง เสงพระพรหม. (2562). การเปรียบเทียบประสิทธิภาพการพยากรณ์ข้อมูลราคาข้าวหอมด้วย เทคนิคการทำให้เรียบแบบเอกซ์โปเนนเชียลขั้นเดียวซึ่งปรับได้และการ วิเคราะห์ถดถอยเชิงเส้น. งานประชุมวิชาการระดับชาติ ครั้งที่ 11 มหาวิทยาลัยราชภัฏนครปฐม, 327-333. มยุรี เกศากิจ นิติมา ลักขณานุรักษ์ และไกรรุ่ง เสงพระพรหม. (2562). การ เปรียบเทียบประสิทธิภาพเทคนิคการพยากรณ์สำหรับข้อมูลราคาน้ำมันด้วย เทคนิคเหมืองข้อมูล. งานประชุมวิชาการระดับชาติ ครั้งที่ 11 มหาวิทยาลัย ราชภัฏนครปฐม, 334-340. อนัทชา พุ่มพวง นิติมา ลักขณานุรักษ์ และไกรรุ่ง เสงพระพรหม. (2562). การเปรียบเทียบประสิทธิภาพการพยากรณ์ข้อมูลนักเรียนโรงเรียนสาธิต มหาวิทยาลัยราชภัฏนครปฐมด้วยเทคนิคการทำให้เรียบแบบเอกซ์โปเนนเชียลขั้นเดียวซึ่งปรับได้และการหาค่าเฉลี่ยแบบเคลื่อนที่. งานประชุม วิชาการระดับชาติ ครั้งที่ 11 มหาวิทยาลัยราชภัฏนครปฐม, 341-346. พิเชฐ แก้ววิชิต อุดมศักดิ์ กาวิลละ นิติมา ลักขณานุรักษ์ ไกรรุ่ง เสงพระพรหม

และสุพจน์ เสงพระพรหม. (2563).การเปรียบเทียบประสิทธิภาพเทคนิคการพยากรณ์ข้อมูลความเครียดทางการเงินด้วยเทคนิคเหมืองข้อมูล. งานประชุมวิชาการระดับชาติ ครั้งที่ 12 มหาวิทยาลัยราชภัฏนครปฐม, 567-562.

นิติมา ลักขณานุรักษ์ จารุกิตต์ มณีสุข วิภาวดี หุ่นทอง ไกรรุ่ง เสงพระพรหม และสุพจน์ เสงพระพรหม. (2563).การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลประเภทของสัตว์โดยใช้เทคนิคเหมืองข้อมูล. งานประชุมวิชาการระดับชาติ ครั้งที่ 12 มหาวิทยาลัยราชภัฏนครปฐม, 573-579.

ถิรพร เหมือนอุย ทวีศักดิ์ เพชรคนชม นิติมา ลักขณานุรักษ์ ไกรรุ่ง เสงพระพรหม และสุพจน์ เสงพระพรหม. (2563).การเปรียบเทียบประสิทธิภาพจำแนกข้อมูลการเล่นเกมด้วยเทคนิคเหมืองข้อมูล. งานประชุมวิชาการระดับชาติ ครั้งที่ 12 มหาวิทยาลัยราชภัฏนครปฐม, 580-585.

Nitima Lukkananuruk, Kata Praditwong, &Supoj Hengpraproh (2022). The Distance - Based Selection Technique for Crossover in Genetic Algorithm. Paper presented at the International conference on 2022 19th International Joint Conference on Computer Science and Software Engineering (JCSSE).

