เปรียบเทียบประสิทธิภาพวิธีการไขว้เปลี่ยนใน LZWGA

สมรักษ์ นุ่มนาค* และ วรเศรษฐ สุวรรณิก*

บทคัดย่อ

การเข้ารหัสแบบ LZW ในขั้นตอนวิธีเชิงพันธุกรรม (LZWGA) เป็นอัลกอริทึมที่พัฒนามาจากขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm) LZWGA เข้ารหัสโครโมโซมในรูปแบบที่สามารถคลายได้ด้วยอัลกอริทึม Lempel-Ziv-Welch (LZW) การเข้ารหัสแบบนี้ช่วยลดขนาดของโครโมโซมลงและทำให้อัลกอริทึมสามารถแก้ปัญหาขนาดใหญ่มากได้ งานวิจัยนี้ เปรียบเทียบประสิทธิภาพวิธีการไขว้เปลี่ยนแบบต่าง ๆ ใน LZWGA โดยทดลองกับปัญหา OneMax และปัญหา Royal Road ขนาดหนึ่งล้านบิต ผลการทดลองแสดงให้เห็นว่าวิธีการไขว้เปลี่ยนแบบยูนิฟอร์มให้ประสิทธิภาพที่ดีที่สุด โดยหลังจากการวิวัฒนาการดำเนินไป 100 รุ่น วิธีการไขว้เปลี่ยนแบบยูนิฟอร์มให้ค่าความเหมาะสมเฉลี่ยดีกว่าวิธีการไขว้เปลี่ยนแบบจุดเดียว 1.59 เท่า และ 2.0 เท่า และดีกว่าวิธีการไขว้เปลี่ยนแบบสองจุด 1.6 เท่า และ 2.2 เท่า ในปัญหา OneMax และ Royal Road ตามลำดับ

คำสำคัญ: ขั้นตอนวิธีเชิงพันธุกรรม การบีบอัดข้อมูล ปัญหา OneMax ปัญหา Royal Road วิธีการไขว้เปลี่ยน

1. บทนำ

ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm – GA) เป็นวิธีการค้นหาที่เลียนแบบวิวัฒนาการของสิ่งมีชีวิต [1] การค้นหากระทำกับกลุ่มของโครโมโซม โดยโครโมโซมของขั้นตอนวิธีเชิงพันธุกรรมมีลักษณะเป็นสายบิต (binary string) เมื่อขนาดของโครโมโซมมีขนาดใหญ่ขึ้น ปริภูมิการค้นหาก็จะมีขนาดใหญ่ขึ้นตามไปด้วย

LZWGA ใช้วิธีการเข้ารหัสโครโมโซมแบบบีบอัด ทำให้โครโมโซมมีขนาดเล็กลง และการวิวัฒนาการกระทำกับโครโมโซมที่ถูกบีบอัด ในขณะที่ค่าความเหมาะสมของโครโมโซมจะวัดกับโครโมโซมที่ถูกคลาย [2] โครโมโซมที่ถูกบีบอัดเป็นอาร์เรย์ของจำนวนเต็ม เช่น ถ้าโครโมโซมของ LZWGA สำหรับปัญหา OneMax ถูกเข้ารหัสด้วยเลข {1, 0, 2, 4} เมื่อนำโครโมโซมนี้มาคลายจะได้โครโมโซมของ GA ปกติคือ 1010101

การเปรียบเทียบประสิทธิภาพระหว่าง LZWGA, Compressed GA และ GA พบว่า LZWGA ใช้จำนวนรุ่นในการแก้ปัญหา OneMax ขนาด 256 บิต น้อยกว่า GA 205 เท่า และ LZWGA ใช้จำนวนรุ่นในการแก้ปัญหา OneMax ขนาด 2048 บิต น้อยกว่า Compressed GA 14 เท่า [2]

วิธีการไขว้เปลี่ยนเป็นส่วนสำคัญในการสร้างประชากรรุ่นถัดไป ซึ่งจุดประสงค์ของวิธีการไขว้เปลี่ยนคือรวมส่วนที่ดีของโครโมโซมพ่อและแม่เข้าไว้ด้วยกัน โดยวิธีการไขว้เปลี่ยนที่นิยมใช้กันทั่วไปได้แก่ วิธีการไขว้เปลี่ยนแบบจุดเดียว วิธีการไขว้เปลี่ยนแบบสองจุด และวิธีการไขว้เปลี่ยนแบบยูนิฟอร์ม

งานวิจัยนี้เปรียบเทียบประสิทธิภาพระหว่าง วิธีการไขว้เปลี่ยนที่นิยมใช้กัน ที่เหมาะสมต่อ LZWGA โดยตัดการกลายพันธุ์ออก ผลการทดลองแสดงในส่วนที่ 5

2. วิธีการไขว้เปลี่ยน

การทำงานของ GA สามารถเขียนในรูปชุดคำสั่งเทียมได้ดังนี้

1) สร้างประชากรรุ่นแรก

2) วัดค่าความเหมาะสมของประชากรทุกตัว

3) สร้างประชากรรุ่นใหม่โดยใช้ตัวดำเนินการทางพันธุกรรมกับโครโมโซมรุ่นก่อนหน้า

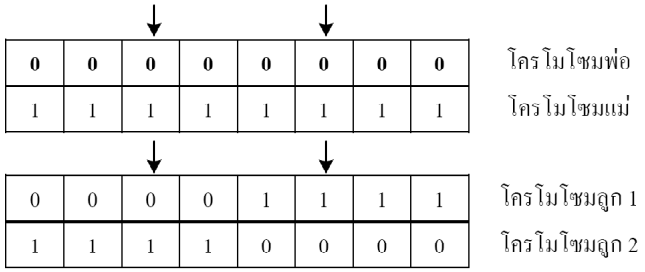
4) ทำซ้ำขั้นตอนที่ 2 และ 3 จนกระทั่งพบคำตอบที่ต้องการหรือถึงจำนวนรุ่นที่กำหนด

การสร้างประชากรในรุ่นถัดไปทำได้โดยใช้ตัวดำเนินการทางพันธุกรรม (genetic operators) กระทำกับโครโมโซมที่ถูกเลือก โดยวิธีการไขว้เปลี่ยนเป็นตัวดำเนินงานทางพันธุกรรมแบบหนึ่ง ซึ่งวิธีการไขว้เปลี่ยน เป็นการนำเอาโครโมโซมสองตัวมาผสมกันเพื่อให้ได้โครโมโซมตัวใหม่ วิธีการไขว้เปลี่ยนที่นิยมใช้กันคือ วิธีการไขว้เปลี่ยนแบบจุดเดียว วิธีการไขว้เปลี่ยนแบบสองจุด และวิธีการไขว้เปลี่ยนแบบยูนิฟอร์ม

* *ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์*

2.1 วิธีการไขว้เปลี่ยนแบบจุดเดียว (One-Point Crossover)

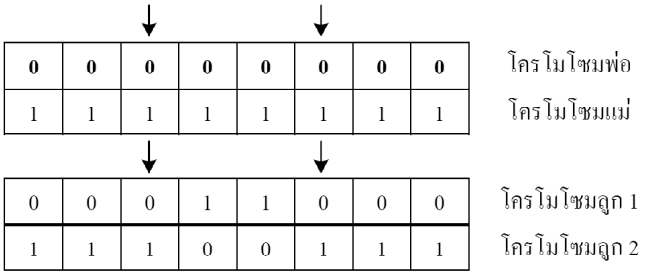
วิธีการไขว้เปลี่ยนแบบจุดเดียว ใช้วิธีสุ่มจุดตัดขึ้นมาหนึ่งจุด เพื่อใช้เป็นจุดที่ใช้ในวิธีการไขว้เปลี่ยน โครโมโซมใหม่ที่ได้จะเกิดจากบางส่วนก่อนจุดตัดในโครโมโซมพ่อผสมกับบางส่วนหลังจุดตัดในอีกโครโมโซมแม่ แสดงดังภาพที่ 1 เช่น สุ่มจุดตัดตำแหน่งที่ 4 โดยโครโมโซมลูก 1 จะเกิดจากตำแหน่งที่ 1 ถึง 4 ของโครโมโซมพ่อรวมกับตำแหน่งที่ 5 ถึง 8 ของโครโมโซมแม่ในขณะที่โครโมโซมลูก 2 เกิดจากตำแหน่งที่ 4 ถึง 8 ของโครโมโซมแม่รวมกับตำแหน่งที่ 1 ถึง 4 ของโครโมโซมพ่อ



ภาพที่ 1 วิธีการไขว้เปลี่ยนแบบจุดเดียว

2.2 วิธีการไขว้เปลี่ยนแบบสองจุด (Two-Point Crossover)

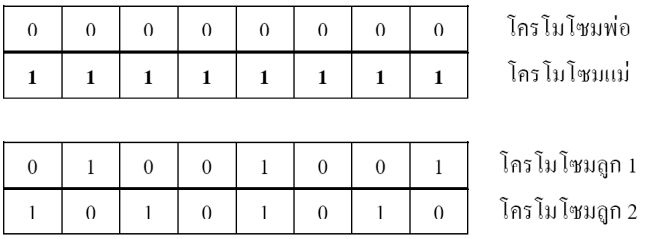
วิธีการไขว้เปลี่ยนแบบสองจุด ใช้วิธีสุ่มจุดตัดขึ้นมาสองจุด เพื่อใช้เป็นจุดในวิธีการไขว้เปลี่ยน โดยโครโมโซมใหม่ที่ได้จะเกิดจากบางส่วนระหว่างจุดตัดในโครโมโซมพ่อผสมกับบางส่วนก่อนและหลังจุดตัดในอีกโครโมโซมแม่ แสดงดังภาพที่ 2 เช่น สุ่มจุดตัดตำแหน่งที่ 3 และ 6 โดยโครโมโซมลูก 1 จะเกิดจากตำแหน่งที่ 1 ถึง 3 ของโครโมโซมพ่อรวมกับตำแหน่งที่ 4 ถึง 5 ของโครโมโซมแม่รวมกับตำแหน่งที่ 6 ถึง 8 ของโครโมโซมพ่อ ในขณะที่โครโมโซมลูก 2 เกิดจากตำแหน่งที่ 1 ถึง 3 ของโครโมโซมแม่รวมกับตำแหน่งที่ 4 ถึง 5 ของโครโมโซมพ่อรวมกับตำแหน่งที่ 6 ถึง 8 ของโครโมโซมแม่



ภาพที่ 2 วิธีการไขว้เปลี่ยนแบบสองจุด

2.3 วิธีการไขว้เปลี่ยนแบบยูนิฟอร์ม (Uniform Crossover)

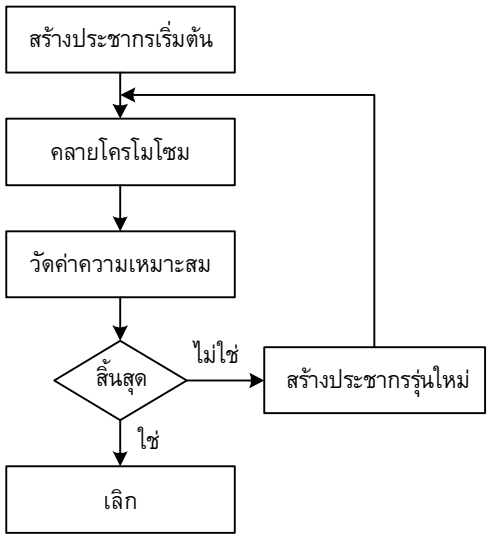
วิธีการไขว้เปลี่ยนแบบยูนิฟอร์ม ใช้วิธีสุ่มเลือกจากโครโมโซมพ่อและแม่ในทุกตำแหน่ง ซึ่งแต่ละตำแหน่งของโครโมโซมใหม่จะมีค่าระหว่างสองค่าคือค่าของโครโมโซมพ่อและแม่ในตำแหน่งตำแหน่งเดียวกัน แสดงดังภาพที่ 3 เช่น โครโมโซมลูกตำแหน่งที่ 1 สุ่มได้โครโมโซมแม่ ตำแหน่งที่ 2 สุ่มได้โครโมโซมพ่อ ตำแหน่งที่ 3 สุ่มได้โครโมโซมแม่ เป็นต้น



ภาพที่ 3 วิธีการไขว้เปลี่ยนแบบยูนิฟอร์ม

3. LZWGA

การทำงานของ LZWGA คล้ายกับการทำงานของ GA แต่ LZWGA จะมีกระบวนการคลายโครโมโซมเพิ่มเข้ามา โดยจะคลายโครโมโซมก่อนแล้วจึงวัดค่าความเหมาะสม ดังแสดงในภาพที่ 4



ภาพที่ 4 กระบวนการทำงานของ LZWGA

3.1 ประชากร

โครโมโซมของ GA จะเป็นเลขฐานสอง ส่วนโครโมโซมที่ถูกบีบอัดของ LZWGA จะเป็นอาร์เรย์ของเลขจำนวนเต็มในช่วง 0 – n โดยค่า n คือขนาดของโครโมโซมที่ถูกบีบอัด

3.2 การสร้างประชากรรุ่นแรก

โครโมโซมในรุ่นแรกของ GA สร้างขึ้นจากการสุ่มโดยไม่มีเงื่อนไข ในขณะที่โครโมโซมในรุ่นแรกของ LZWGA สร้างขึ้นจากการสุ่มโดยมีเงื่อนไขบางอย่างเพื่อให้สามารถคลายได้ เงื่อนไขดังกล่าวคือตำแหน่งแรกของโครโมโซมที่ถูกบีบอัดจะมีค่าได้ระหว่าง 0 กับ 1 ส่วนตำแหน่งถัดไปจะมีค่าได้ไม่เกินตำแหน่งนั้น เช่น โครโมโซมที่ถูกบีบตำแหน่งที่ 1 มีค่าอยู่ใน {0,1} โครโมโซมตำแหน่งที่ 2 มีค่าอยู่ใน {0,1,2} ซึ่งสามารถสรุปได้ตามสมการที่ 1

$$a_1 = r \bmod(i+1), \in \{1 \ldots n\} \qquad (1)$$

โดย

a_i = ตำแหน่งที่ i ของโครโมโซมที่บีบอัดอยู่

r = เลขสุ่ม

n = ขนาดของโครโมโซม

4. การทดลอง

ผู้วิจัยทำการทดลองกับปัญหา OneMax ขนาดหนึ่งล้านบิต และปัญหา Royal Road ขนาดหนึ่งล้านบิต โดยตัดการกลายพันธุ์ออกเพื่อต้องการดูผลจากวิธีการไขว้เปลี่ยนเพียงอย่างเดียว เนื่องจาก LZWGA เป็นอัลกอริทึมที่มีการสุ่ม จะทดลองซ้ำ 30 ครั้ง และรายงานค่าเฉลี่ย

4.1 ปัญหา OneMax

ปัญหา OneMax คือ ปัญหาการนับจำนวนเลขหนึ่งในโครโมโซม โดยค่าความเหมาะสมที่ได้คือ จำนวนเลขหนึ่งทั้งหมดที่พบในโครโมโซม ดังนั้นโครโมโซมที่ดีที่สุดคือโครโมโซมที่ทุกบิตมีค่าเป็นหนึ่ง

ปัญหา OneMax เป็นปัญหาที่ทำความเข้าใจและวิเคราะห์ได้ง่าย งานวิจัยหลายชิ้นใน GA นิยมทดลองกับปัญหา OneMax จึงถือได้ว่าปัญหา OneMax สามารถนำมาใช้อ้างอิงเพื่อเปรียบเทียบประสิทธิภาพการทำงานของขั้นตอนวิธีเชิงพันธุกรรมแบบต่าง ๆ ได้

ปัญหา OneMax ที่ใช้ในการทดลองมีขนาดหนึ่งล้านบิต เปรียบเทียบประสิทธิภาพวิธีการไขว้เปลี่ยนแบบจุดเดียว วิธีการไขว้เปลี่ยนแบบสองจุด และวิธีการไขว้เปลี่ยนแบบยูนิฟอร์มตัดการกลายพันธุ์ออกและวัดจากค่าความเหมาะสมเฉลี่ยในการค้นหา

4.2 ปัญหา Royal Road

ปัญหา Royal Road [3] คือปัญหาที่โครโมโซมของเลขฐานสองถูกแบ่งออกเป็นบล็อกย่อย ๆ ที่มีขนาดเท่ากัน ค่าความเหมาะสมของโครโมโซมคือผลรวมของค่าความเหมาะสมของแต่ละบล็อก โดยค่าความเหมาะสมของแต่ละบล็อกจะเท่ากับ 0 ถ้าบิตใดบิตหนึ่งในบล็อกนั้นเป็น 0 แต่ถ้าทุกบิตในบล็อกนั้นเป็น 1 หมด บล็อกนั้นจะมีค่าความเหมาะสมเท่ากับจำนวนบิตในบล็อก

ยกตัวอย่างปัญหาขนาด 12 บิต กำหนดขนาดของบล็อกย่อยเท่ากับ 4 บิต { 0100 1011 0011}, { 0000 1111 0001}, {1111 0010 1111}, และ {1111 1111 1111} มีค่าความเหมาะสมคือ 0, 4, 8 และ 12 ตามลำดับ ปัญหา Royal Road เป็นปัญหาที่บิตในโครโมโซมมีความสัมพันธ์กัน ค่าความเหมาะสมจะเพิ่มขึ้นได้ก็ต่อเมื่อบิตที่อยู่ในบล็อกเดียวกันมีค่าเป็นหนึ่งทั้งหมด

ปัญหา Royal Road ที่ใช้ในการทดลองมีขนาดหนึ่งล้านบิต ทดลองกับบล็อกขนาดสิบบิตและหนึ่งร้อยบิต เปรียบเทียบประสิทธิภาพวิธีการไขว้เปลี่ยนแบบจุดเดียว วิธีการไขว้เปลี่ยนแบบสองจุด และวิธีการไขว้เปลี่ยนแบบยูนิฟอร์ม โดยตัดการกลายพันธุ์ออกและวัดจากค่าความเหมาะสมเฉลี่ยในการค้นหา

พารามิเตอร์ที่ใช้ในการทดลองอ้างอิงจาก [4] แต่ตัดส่วนของการกลายพันธุ์ออกแสดงในตารางที่ 1 และ 2

ตารางที่ 1 พารามิเตอร์ที่ใช้ในการแก้ปัญหา OneMax ขนาดหนึ่งล้านบิต

พารามิเตอร์	ค่า
จำนวนประชากร	100 ตัว
ความยาวโครโมโซมที่ถูกบีบอัด	40,000 ตำแหน่ง
ความยาวโครโมโซมที่คลายแล้ว	100,000 บิต
วิธีการคัดเลือก	แบบประลองโครโมโซมที่ประลองกันที่ 4 ตัว
อัตราวิธีการไขว้เปลี่ยน	แบบจุดเดียว แบบสองจุด แบบยูนิฟอร์ม
อัตราการกลายพันธุ์	0 %
เก็บโครโมโซมที่ดีที่สุดไว้	10 ตัว
จำนวนรุ่นจำกัดไว้ที่	100 รุ่น

ตารางที่ 2 พารามิเตอร์ที่ใช้ในการแก้ปัญหา Royal Road ขนาดหนึ่งล้านบิต

พารามิเตอร์	ค่า
จำนวนประชากร	100 ตัว
ความยาวโครโมโซมที่ถูกบีบอัด	40,000 ตำแหน่ง
ความยาวโครโมโซมที่คลายแล้ว	100,000 บิต
วิธีการคัดเลือก	แบบประลองโครโมโซมที่ประลองกันที่ 4 ตัว
อัตราวิธีการไขว้เปลี่ยน	แบบจุดเดียว แบบสองจุด แบบยูนิฟอร์ม
อัตราการกลายพันธุ์	0 %
ขนาดบล็อก	10,100 บิต
เก็บโครโมโซมที่ดีที่สุดไว้	10 ตัว
จำนวนรุ่นจำกัดไว้ที่	100 รุ่น

5. ผลการทดลอง

ผลการทดลองแสดงให้เห็นว่าวิธีการไขว้เปลี่ยนแบบยูนิฟอร์มให้ประสิทธิภาพที่ดีที่สุด หลังจากการวิวัฒนาการดำเนินไป 100 รุ่น (ภาพที่ 5, 6 และ 7) ตารางที่ 3 เปรียบเทียบค่าความเหมาะสมเฉลี่ย ระหว่างวิธีการไขว้เปลี่ยนในแต่ละปัญหา

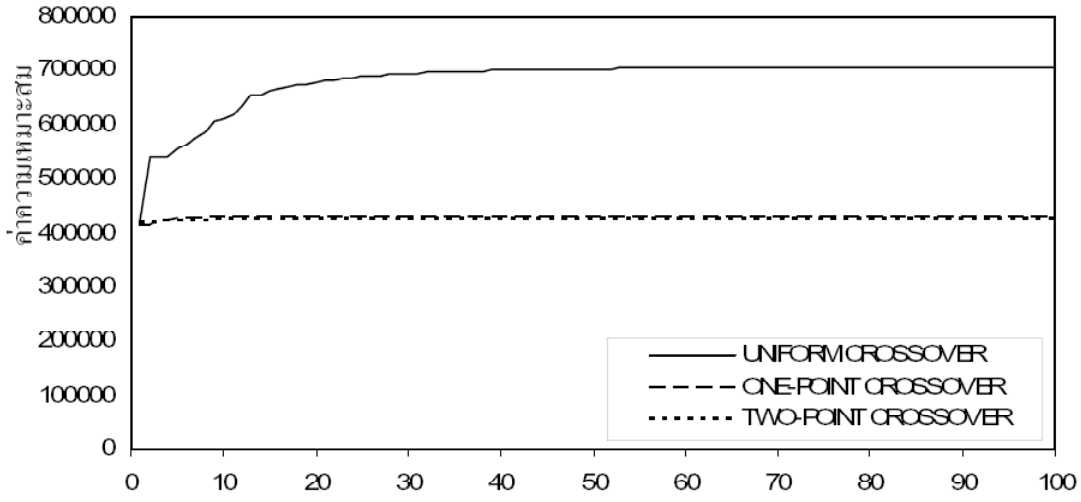
ตารางที่ 3 เปรียบเทียบค่าความเหมาะสมเฉลี่ย ระหว่างวิธีการไขว้เปลี่ยนในแต่ละปัญหา

วิธีการไขว้เปลี่ยน	ปัญหา OneMax	ปัญหา Royal Road บล็อก 10 บิต	ปัญหา Royal Road บล็อก 100 บิต
วิธีการไขว้เปลี่ยนแบบยูนิฟอร์ม	682,433.	672,061.	613,925.
วิธีการไขว้เปลี่ยนแบบจุดเดียว	428,808.	389,213.	253,750.
วิธีการไขว้เปลี่ยนแบบสองจุด	424,478.	352,464.	188,047.

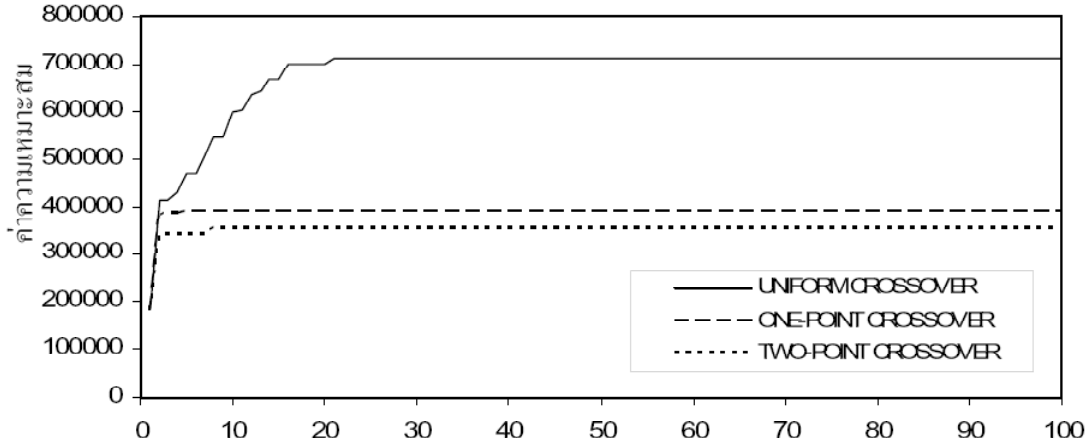
ผลการทดลองกับปัญหา OneMax พบว่าวิธีการไขว้เปลี่ยนแบบจุดเดียวและวิธีการไขว้เปลี่ยนแบบสองจุดให้ค่าความเหมาะสมเฉลี่ยใกล้เคียงกันผลจากการทดลองกับปัญหา Royal Road พบว่าเมื่อบล็อกมีขนาดมากขึ้นค่าความเหมาะสมเฉลี่ยจะน้อยลงรวมทั้งวิธีการไขว้เปลี่ยนแบบจุดเดียวจะให้ค่าความเหมาะสมเฉลี่ยมากขึ้นกว่าวิธีการไขว้เปลี่ยนแบบสองจุด

6. แนวทางการวิจัยในอนาคต

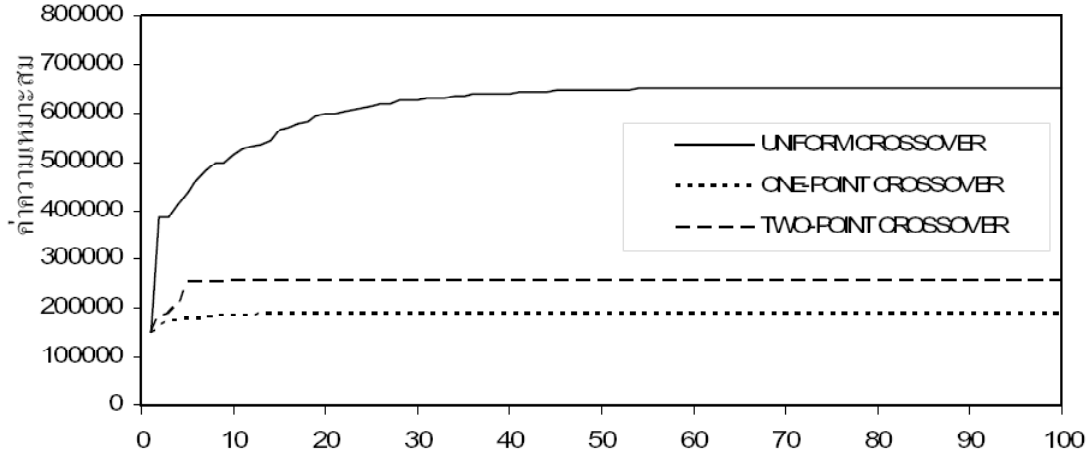
LZWGA เป็นอัลกอริทึมที่สามารถแก้ปัญหาขนาดใหญ่มากได้ ซึ่งวิธีการไขว้เปลี่ยนของ LZWGA เป็นส่วนหนึ่งของ



ภาพที่ 5 เปรียบเทียบวิธีการไขว้เปลี่ยนแบบจุดเดียว วิธีไขว้เปลี่ยนแบบสองจุด และวิธีไขว้เปลี่ยนแบบยูนิฟอร์มกับปัญหา OneMax ขนาดหนึ่งล้านบิต



ภาพที่ 6 เปรียบเทียบวิธีการไขว้เปลี่ยนแบบจุดเดียว วิธีไขว้เปลี่ยนแบบสองจุด และวิธีไขว้เปลี่ยนแบบยูนิฟอร์มกับปัญหา Royal Road ขนาดหนึ่งล้านบิต (บล็อกขนาดสิบบิต)



ภาพที่ 7 เปรียบเทียบวิธีการไขว้เปลี่ยนแบบจุดเดียว วิธีไขว้เปลี่ยนแบบสองจุด และวิธีไขว้เปลี่ยนแบบยูนิฟอร์ม กับปัญหา Royal Road ขนาดหนึ่งล้านบิต (บล็อกขนาดหนึ่งร้อยบิต)

การสร้างประชากรในรุ่นถัดไป จากการทดลองเรายังหาคำตอบไม่ได้ว่าทำไมการไขว้เปลี่ยนแบบยูนิฟอร์มจึงให้ประสิทธิภาพดีกว่าการไขว้เปลี่ยนแบบอื่น โดย [5] แสดงให้เห็นว่าตำแหน่งโครโมโซมที่ถูกเบียนอัดมีความสัมพันธุ์กับการคลายโครโมโซม จึงน่าสนใจว่าการที่การไขว้เปลี่ยนแบบยูนิฟอร์มให้ประสิทธิภาพที่ดีนั้นเกี่ยวข้องกับความสัมพันธ์ระหว่างตำแหน่งโครโมโซมที่ถูกบีบอัดหรือไม่ ซึ่งจะเป็นแนวทางในการวิจัยต่อไป

7. สรุป

งานวิจัยนี้เปรียบเทียบประสิทธิภาพวิธีการไขว้เปลี่ยนใน LZWGA โดยเปรียบเทียบวิธีการไขว้เปลี่ยนแบบจุดเดียว วิธีการไขว้เปลี่ยนแบบสองจุด และวิธีการไขว้เปลี่ยนแบบยูนิฟอร์ม ได้ทดลองกับปัญหา OneMax ขนาดหนึ่งล้านบิต และปัญหา Royal Road ขนาดหนึ่งล้านบิต ผลการทดลองแสดงให้เห็นว่าวิธีการไขว้เปลี่ยนแบบยูนิฟอร์มให้ประสิทธิภาพที่ดีที่สุด โดยหลังจากการวิวัฒนาการดำเนินไป 100 รุ่น วิธีการไขว้เปลี่ยนแบบยูนิฟอร์มให้ค่าความเหมาะสมเฉลี่ยดีกว่าวิธีการไขว้เปลี่ยนแบบจุดเดียว 1.59 เท่า และ 2.0 เท่า และดีกว่าวิธีการไขว้เปลี่ยนแบบสองจุด 1.6 เท่า และ 2.2 เท่า ในปัญหา OneMax และ Royal Road ตามลำดับ

8. อ้างอิง

[1] Goldberg, D.E., “The Design of Innovation, “ Kluwer Academic Publishers, USA, 2002.

[2] นริศ กุณาศล, วรเศรษฐ สุวรรณิก และ ประภาส จงสถิตย์

วัฒนา, “ การเข้ารหัสแบบ LZW ในขั้นตอนวิธีเชิงพันธุกรรม,” Proceedings of Electrical Engineering Conference (EECON-28), หน้า 861-864, October 20-21, 2005.

[3] M. Mitchell, S. Forrest ,and J.H. Holland, “The Royal Road for Genetic Algorithms: Fitness Landscapes and GA Performance,” Proceedings of the First European Conference on Artificial Life Cambridge, MA:MIT Press, 1992, pp. 245-254.

[4] Naris Kunasol, Worasait Suwannik, Prabhas Chongstitvatana, “Solving One-Million-Bit Problems Using LZWGA,” Proceedings of International Symposium on Communications and Information Technologies (ISCIT), October 18-20, 2006, pp. 32-36.

[5] สมรักษ์ นุ่มนาค และ วรเศรษฐ สุวรรณิก, “การปรับปรุงการกลายพันธุ์ ใน LZWGA,” Proceedings of International Joint Conference on Computer Science and Software Engineering (JCSSE) Thai Track, May 2-4, 2007.