

บทที่ 4

ทฤษฎีและบทวิเคราะห์วิธีเจเนติกอัลกอริทึม

ในบทนี้กล่าวถึงที่มาและแนวคิดพื้นฐานของวิธีเจเนติกอัลกอริทึม ที่สามารถนำมาใช้ในการแก้ปัญหาต่างๆ ได้เป็นอย่างดี รวมถึงบทวิเคราะห์การพฤติกรรมทำงานและการปรับปรุงตัวปฏิบัติการทางพันธุศาสตร์ที่จะช่วยทำให้ประสิทธิภาพการทำงานของวิธีเจเนติกอัลกอริทึมโดยรวมสูงขึ้นเป็นอย่างมาก

วิธีเจเนติกอัลกอริทึมเป็นวิธีการค้นหาตอบที่มีพื้นฐานมาจากกลไกของขบวนการคัดเลือกและการถ่ายทอดพันธุกรรมตามธรรมชาติ ซึ่งจากหลักการที่ว่า สิ่งมีชีวิตที่แข็งแรงมากกว่าจะมีโอกาสอยู่รอดและเป็นผู้ผลิตทายาทในรุ่นต่อไปได้มากกว่า ซึ่งในปัจจุบันวิธีเจเนติกอัลกอริทึมได้นำมาใช้กันอย่างกว้างขวาง ในลักษณะการคำนวณค่าที่เหมาะสมและปัญหาการเรียนรู้ วิธีเจเนติกอัลกอริทึมประสบความสำเร็จอย่างสูงสำหรับการแก้ปัญหาที่มีรูปแบบที่ซับซ้อน ปัญหาที่มีหลายคำตอบโลคอล (ที่ใกล้เคียงกับคำตอบโกลบอล) หรือปัญหาที่ไม่สามารถหาอนุพันธ์ได้ เมื่อเทียบกับเทคนิคการคำนวณที่มีพื้นฐานจากการหาอนุพันธ์และวิธีการไล่เขา (การหาคำตอบที่ดีขึ้นจากตำแหน่งคำตอบเดิม)

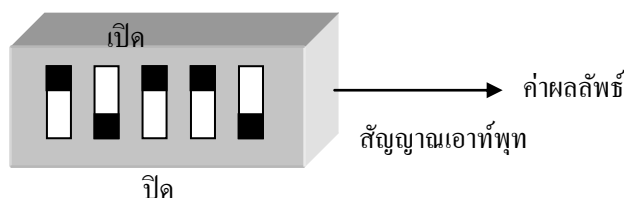
จุดเด่นของวิธีเจเนติกอัลกอริทึมที่เหนือกว่าวิธีการค้นหาแบบเดิม เช่นวิธีที่มีพื้นฐานบนการคำนวณแคลคูลัสทั้งวิธีตรงและวิธีอ้อม วิธีการค้นหาโดยการแจกแจงคำตอบ และ วิธีการค้นหาโดยการสุ่ม คือ

1. วิธีเจเนติกอัลกอริทึม ทำงานโดยอาศัยการเข้ารหัสของค่าพารามิเตอร์ของปัญหา ซึ่งไม่ใช้การค้นหาคำตอบจากค่าพารามิเตอร์เหล่านั้นโดยตรง
2. วิธีเจเนติกอัลกอริทึม ค้นหาผลลัพธ์ในการคำนวณจากกลุ่มของคำตอบ ซึ่งไม่ใช้การค้นหาจากคำตอบเดียว
3. วิธีเจเนติกอัลกอริทึม ใช้เฉพาะค่าความเหมาะสมจากฟังก์ชันเป้าหมาย มาใช้เป็นข้อมูลสำหรับการค้นหาคำตอบ ซึ่งไม่จำเป็นต้องใช้ความรู้ในการหาค่าอนุพันธ์ หรือความรู้ในทางคณิตศาสตร์อื่นๆ
4. วิธีเจเนติกอัลกอริทึม ค้นหาผลลัพธ์จากวิธีการจากความน่าจะเป็น ซึ่งไม่ใช้กรรมวิธีการคำนวณโดยตรง

4.1 แนวคิดพื้นฐานของวิธีเจเนติกอัลกอริทึม

ถ้าพิจารณาปัญหาการหาคำตอบที่เหมาะสม ในลักษณะกล่องดำที่มีดิฟฟิวทซ์ในการปรับค่า ดังรูปที่ 4.1 โดยถ้ามีจำนวนสัญญาณอินพุตเท่ากับ s ซึ่งการปรับค่าสวิทช์จะให้สัญญาณเอาต์พุตที่แตกต่างกันออกไป ถ้ากำหนดให้ฟังก์ชัน $f(s)$ เป็นฟังก์ชันการทำงานของกล่องดำ และ s เป็นลักษณะการปรับตั้งสวิทช์ทั้งหมด เป้าหมายของปัญหานี้คือการปรับตั้งสวิทช์เพื่อให้มีสัญญาณเอาต์พุตที่สูงที่สุด ซึ่ง

ถ้าใช้วิธีการแก้ปัญหาโดยวิธีอื่นจำเป็นจะต้องทราบถึงฟังก์ชันการทำงานที่แท้จริงของกล่องดำเพื่อใช้คำนวณหาผลลัพธ์จากการปรับตั้งสวิทช์ วิธีเจเนติกอัลกอริทึมจะพิจารณาเพียงลักษณะการปรับตั้งสวิทช์ (อินพุต) และค่าผลลัพธ์ที่ได้ (เอาต์พุต) สำหรับการค้นหาคำตอบที่ดีที่สุดในการปรับตั้งสวิทช์



รูปที่ 4.1 แนวคิดของการเข้ารหัสและการวัดค่าเอาต์พุต สำหรับปัญหาการหาค่าที่เหมาะสม

สำหรับวิธีเจเนติกอัลกอริทึมนั้น ขบวนการค้นหาคำตอบจะเริ่มจากการเข้ารหัสสถานะของสวิทช์ทั้งหมดเป็นไบนารีบิตสตริง โดยจากปัญหาดังกล่าวสามารถใช้บิตสตริงขนาด 1 บิต แทนสถานะของอินพุตในแต่ละค่า ซึ่งเมื่อนำมาเรียงต่อกันทั้งหมดจะได้เป็นชุดรหัส 10110 โดยที่รหัส 0 แสดงสถานะ “ปิด” ของสวิทช์ และรหัส 1 แสดงสถานะ “เปิด” ของสวิทช์ จากนั้นสร้างกลุ่มของรหัสทั้งหมดขึ้นมาโดยการสุ่ม ซึ่งถ้ากำหนดขนาดของจำนวนรหัสทั้งหมดเท่ากับ 4 จะได้เป็นดังนี้

01101

11000

01000

10011

วิธีเจเนติกอัลกอริทึม จะค้นหาคำตอบจากที่ดีที่สุดจากกลุ่มของชุดรหัสเหล่านี้โดยคำนวณหาผลลัพธ์จากชุดรหัสทั้งหมด เพื่อเป็นข้อมูลเริ่มต้นสำหรับกรรมวิธีการค้นหาคำตอบ

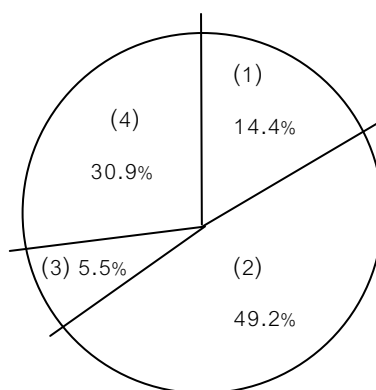
4.1.1 การสร้างชุดรหัสขึ้นมาใหม่

การสร้างชุดรหัสขึ้นมาใหม่เป็นขบวนการคัดลอกชุดรหัสโดยสอดคล้องกับค่าความเหมาะสม (คำนวณจากฟังก์ชันเป้าหมายของปัญหา) จากชุดรหัสที่มีอยู่เดิม ซึ่งอาจจะพิจารณาได้ชุดรหัสใดที่มีค่าความเหมาะสมที่สูง ก็มีโอกาที่จะถูกเลือกเป็นชุดรหัสต้นแบบในการคัดลอกรหัส สำหรับกลุ่มของชุดรหัสชุดใหม่

ขบวนการคัดเลือกชุดรหัสเพื่อเป็นต้นแบบนั้นมีหลากหลายวิธี ซึ่งวิธีหนึ่งที่ใช้ในการคัดเลือกชุดรหัสอย่างง่ายคือ การพิจารณาค่าเปอร์เซ็นต์ของค่าความเหมาะสมจากชุดรหัสเหล่านั้น หรือเรียกอีกอย่างหนึ่งว่าวิธีการคัดเลือกโดยใช้กงล้อรูเล็ต โดยถ้าชุดสตริงใดมีเปอร์เซ็นต์ของค่าความเหมาะสมสูงก็จะมีส่วนพื้นที่ในกงล้อรูเล็ตมาก ซึ่งจากปัญหาดังกล่าวมาสามารถเขียนแสดงได้ดังนี้

ตารางที่ 4.1 ตัวอย่างค่าความเหมาะสมของปัญหาการปรับตั้งค่าสวิตช์

ลำดับที่	สตริง	ค่าความเหมาะสม	% ผลรวม
1	01101	169	14.4
2	11000	576	49.2
3	01000	64	5.5
4	10011	361	30.9
ผลรวม		1170	100.0



รูปที่ 4.2 กงล้อรูเลตที่ใช้ในขั้นตอนการคัดเลือกชุดรหัสอย่างง่าย

โดยเมื่อหมุนกงล้อรูเลต และทำการเลือกชุดรหัสต้นแบบจะพบว่าชุดรหัสที่ 1 ซึ่งมีค่าเปอร์เซ็นต์ของค่าความเหมาะสมเท่ากับ 14.4 จะมีค่าความน่าจะเป็นที่จะถูกได้รับการคัดเลือกเท่ากับ 0.144 และเมื่อชุดรหัสใดได้รับการคัดเลือกก็ถูกคัดลอกรหัสทั้งหมดและสร้างเป็นชุดรหัสต้นแบบ

4.1.2 การแลกเปลี่ยนรหัส

หลังจากที่ได้ชุดรหัสต้นแบบคู่หนึ่ง ขั้นตอนการแลกเปลี่ยนรหัสอย่างง่าย หรือการแลกเปลี่ยนรหัสชนิดหนึ่งจุด จะเริ่มขึ้นจากการสุ่มตำแหน่งของรหัสที่ต้องการแลกเปลี่ยน ที่อยู่ในช่วง 1 ถึง $l-1$ โดยที่ l คือขนาดของชุดรหัส จากนั้นทำการแลกเปลี่ยนรหัสจากตำแหน่งดังกล่าว และสร้างเป็นชุดรหัสขึ้นใหม่ขึ้นมามีดังรูป

ตำแหน่งของการแลกเปลี่ยน = 4

ชุดรหัสต้นแบบที่ 1 = 0 1 1 0 | 1

ชุดรหัสต้นแบบที่ 2 = 1 1 0 0 | 0

ชุดรหัสใหม่ที่ 1 = 0 1 1 0 | 0

ชุดรหัสใหม่ที่ 2 = 1 1 0 0 | 1

รูปที่ 4.3 ตัวอย่างการแลกเปลี่ยนรหัสอย่างง่าย

4.1.3 การผ่าเหล่ารหัส

การผ่าเหล่ารหัสเป็นวิธีการเปลี่ยนค่ารหัสในแต่ละบิตโดยตรง เนื่องจากการแลกเปลี่ยนรหัสอาจจะสูญเสียรหัสบางตัวที่มีความสำคัญ อีกทั้งการแลกเปลี่ยนดังกล่าวไม่ได้ทำให้เกิดส่วนของรหัสขึ้นมาใหม่ ดังนั้นเพื่อชดเชยผลกระทบดังกล่าว ชุดรหัสใหม่ที่เกิดขึ้นจะถูกเปลี่ยนค่ารหัส จากตำแหน่งที่ได้จากการสุ่ม โดยมีค่าอยู่ในช่วง 1 ถึง I ซึ่งโดยทั่วไปนั้น จะการผ่าเหล่ารหัสนั้นจะมีโอกาสเกิดขึ้นต่ำมาก จากการกำหนดค่าความน่าจะเป็นในการผ่าเหล่า โดยชุดรหัสใหม่ที่ผ่านขบวนการผ่าเหล่าจะมีลักษณะดังรูป

ตำแหน่งของการผ่าเหล่า = 3

ชุดรหัสเดิม = 0 1 1 0 1

ชุดรหัสใหม่ = 0 1 0 0 1

รูปที่ 4.4 ตัวอย่างการผ่าเหล่ารหัสอย่างง่าย

เจเนติกอัลกอริทึมอย่างง่าย()

```
{
    สร้างประชากรเริ่มต้นโดยการสุ่ม;
    กำหนดค่าความเหมาะสมของชุดโครโมโซม;
    ถ้าการคำนวณยังไม่ลู่เข้า
    {
        ปรับแต่งค่าความเหมาะสม;
        คัดเลือกชุดโครโมโซมต้นแบบ;
        ดำเนินการแลกเปลี่ยนยีนและการผ่าเหล่ายีนโดยการสุ่ม;
        กำหนดค่าความเหมาะสมของชุดโครโมโซมใหม่;
    }
}
```

รูปที่ 4.5 โครงสร้างของวิธีเจเนติกอัลกอริทึมอย่างง่าย

เพื่อความเข้าใจที่ดีขึ้น ถ้าสมมุติให้การทำงานของกล่องดำนี้คือ $f(x) = x^2$ โดยที่ค่า x อยู่ในช่วง 0 ถึง 31 และทำการเข้ารหัสค่า x เป็นชุดสตริงไบนารีขนาด 5 บิต จากตัวอย่างข้างต้น ถ้าชุดรหัสมีค่าเป็น 10011 จะสามารถคำนวณค่าฟังก์ชันเป้าหมายได้ดังนี้

$$1 \cdot 2^4 + 0 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 16 + 0 + 0 + 2 + 1 = 19$$

โดยที่ชุดรหัสจะสามารถแสดงค่าฟังก์ชันเป้าหมายในช่วงระหว่าง 0 (00000) ถึง 31 (11111)

ตัวอย่างการคำนวณการหาค่าสูงสุดของฟังก์ชัน $f(x) = x^2$ ด้วยวิธีเจเนติกอัลกอริทึมที่สามารถแสดงได้ดังตารางต่อไปนี้

ตารางที่ 4.2 ตัวอย่างการคำนวณในช่วงขบวนการคัดเลือก

ชุดรหัสที่	ชุดรหัสเริ่มต้น (จากการสุ่มค่า)	ค่าตัวแปร x	$f(x)$	ค่าความน่าจะเป็นในการ คัดเลือก $f_i / \sum_i f_i$	ค่าการคัดเลือก ที่คาดหวัง $f_i / \bar{f}_i$	ค่าการคัดเลือก ที่เกิดขึ้นจากวงล้อ รูเล็ต
1	0 1 1 0 1	13	169	0.14	0.58	1
2	1 1 0 0 0	24	576	0.49	1.97	2
3	0 1 0 0 0	8	64	0.06	0.22	0
4	1 0 0 1 1	19	361	0.31	1.23	1
ผลรวม			1170	1.00	4.00	4.0
ค่าเฉลี่ย			293	0.25	1.00	1.0
ค่าที่มากที่สุด			576	0.49	1.97	2.0

ตารางที่ 4.3 ตัวอย่างการคำนวณในช่วงการดำเนินการทางพันธุศาสตร์

ชุดรหัสที่	ชุดรหัสจาก การสร้างชุด รหัสขึ้นใหม่	คู่ของชุด รหัส	ตำแหน่งการ แลกเปลี่ยนรหัส	ชุดรหัสใหม่	ค่าตัวแปร x	$f(x)$
1	0 1 1 0 1	2	4	0 1 1 0 0	12	144
2	1 1 0 0 0	1	4	1 1 0 0 1	25	625
3	1 1 0 0 0	4	2	1 1 0 1 1	27	729
4	1 0 0 1 1	3	2	1 0 0 0 0	16	256
ผลรวม						1754
ค่าเฉลี่ย						439
ค่าที่มากที่สุด						729

4.2 การวิเคราะห์การทำงานของวิธีเจเนติกอัลกอริทึม

4.2.1 ขบวนการคัดเลือก

ถ้าพิจารณาการทำงานของขบวนการคัดเลือกทั้งหมด สามารถแยกออกเป็นส่วนของการคัดเลือกชุดโครโมโซมและส่วนของการสุ่มตัวอย่าง ในส่วนของการคัดเลือกนั้นจะเป็นการคัดเลือกโครโมโซมต้นแบบจากประชากรที่มีอยู่ โดยจะกำหนดค่าอัตราการสุ่มเป้าหมาย $tsr(x, t)$ จากชุดโครโมโซม x เป็นตัวชี้บอกค่าคาดการณ์จำนวนชุดโครโมโซมใหม่ที่เกิดขึ้นจากชุดโครโมโซม x นี้ ในช่วงรุ่นประชากร t และในส่วนของการสุ่มตัวอย่างนั้น จะสร้างประชากรรุ่นใหม่โดยการคัดลอกชุดของโครโมโซม จากสัดส่วนของค่าอัตราการสุ่ม ซึ่งจะต้องพิจารณาวิธีการสุ่มตัวอย่างด้วยความ

ระมัดระวัง เพื่อไม่ให้เกิดข้อผิดพลาดจากการสุ่มค่า เช่น ผลการแปลงค่าจำนวนจริง tsr ไปเป็นค่าจำนวนเต็มจากจำนวนการคัดลอกชุดโครโมโซมที่เกิดขึ้นจริง โดยค่าอัตราการสุ่มเป้าหมาย $tsr(x, t)$ แสดงได้ดังสมการ

$$tsr(H, t) =_{def} \sum_{i=1}^{m(H, t)} \frac{tsr(x_i, t)}{m(H, t)} \quad (4.1)$$

โดยที่ $x_i \in H$ และ $m(H, t)$ แทนจำนวนของตัวแทนของไฮเปอร์เพลน H ในประชากร $P(t)$ ซึ่งโดยปกติ ค่าอัตราการสุ่มเป้าหมาย อาจพิจารณาได้เป็นค่าอัตราการเติบโตของไฮเปอร์เพลน H ในช่วงเวลา t และไม่คำนึงถึงผลของความสูญเสียจากตัวปฏิบัติการรวมโครโมโซม

ขบวนการคัดเลือกตามสัดส่วนของค่าความเหมาะสมนั้น เป็นวิธีการที่ใช้โดยทั่วไป ซึ่งสามารถคำนวณได้อย่างง่ายแต่ก็มีเหตุผลพอเพียงสำหรับใช้ในการคัดเลือกชุดโครโมโซม ดังสมการ

$$tsr(x, t) = \frac{u(x)}{\bar{u}(t)} \quad (4.2)$$

โดยที่ $u()$ คือค่าฟังก์ชันความเหมาะสม และ $\bar{u}()$ คือค่าความเหมาะสมเฉลี่ยในประชากร $P(t)$ ซึ่งถ้าพิจารณาขบวนการคัดเลือกจากสัดส่วนค่าความเหมาะสม สิ่งที่น่าสนใจคือวิธีนี้จะสร้างอัตราการสุ่มเป้าหมายสำหรับไฮเปอร์เพลนในกลุ่มประชากรที่พิจารณา

$$\begin{aligned} tsr(H, t) &= \sum_{i=1}^{m(H, t)} \frac{tsr(x_i, t)}{m(H, t)} \\ &= \sum_{i=1}^{m(H, t)} \frac{u(x_i)}{\bar{u}(t)m(H, t)} \\ &= \frac{u(H, t)}{\bar{u}(t)} \end{aligned} \quad (4.3)$$

โดยที่ $u(H, t)$ คือค่าความเหมาะสมเฉลี่ยของตัวแทนของไฮเปอร์เพลน H ในประชากร $P(t)$ ซึ่งผลที่ได้นี้คือหัวใจของทฤษฎีสกิม่า (Holland ปี 1975) และถูกเรียกว่าเป็นทฤษฎีพื้นฐานของวิธีเจเนติกอัลกอริทึม (Goldberg, 1989)

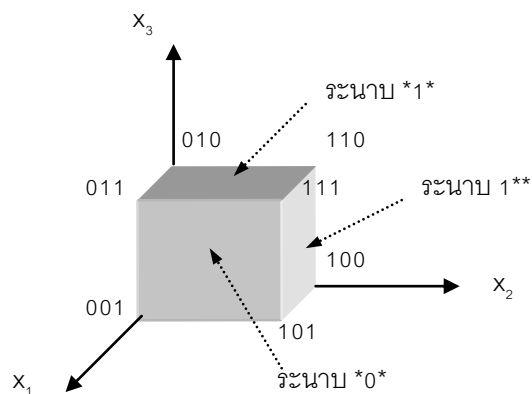
4.2.2 ทฤษฎีสกิม่า

4.2.2.1 สกิม่า

สกีมาคือกลุ่มของยีนต้นแบบในบิตสตริงที่มีความคล้ายคลึงกันหลายตำแหน่งบิต ถ้ากำหนดให้สัญลักษณ์ * แสดงถึงค่าของบิตที่อาจจะมีความเป็น 0 หรือ 1 (ในไบนารีบิต) จะสามารถสร้างสกีมาตา (สตริงย่อยหลายชุด) จากค่าอักขระ {0, 1, *} ดังตัวอย่างดังต่อไปนี้

ถ้ากำหนดให้สตริงและสกีมาตามความยาวเท่ากับ 5 สกีมาตา 00000 จะสัมพันธ์กับสตริง 2 ชุดคือ {10000, 00000} ในทำนองเดียวกันสกีมา *111* จะสัมพันธ์กับสตริง 4 ชุดคือ {01110, 01111, 11110, 11111} ซึ่งจะพบว่า ท่ามกลางสตริงที่มีความยาวนั้นและอักขระที่เป็นไปได้จำนวนมากภายในกลุ่มประชากร จะมีสกีมาตาที่คล้ายคลึงกันปรากฏอยู่อย่างมากมาย

ถ้ากำหนดให้สกีมาที่มีความยาวเท่ากับ 3 จะสามารถพิจารณาคล้ายกับไฮเปอร์เพลนดังรูป



รูปที่ 4.6 ไฮเปอร์เพลนของสกีมาขนาดความยาวเท่ากับ 3

4.2.2.2 บิวติงบล็อก

บิวติงบล็อกจะนิยามได้จากสกีมาที่มีขนาดสั้นและมีค่าความเหมาะสมสูง ซึ่งจะสามารถอยู่รอดได้จากผลของขบวนการคัดเลือกและขบวนการสร้างชุดโครโมโซมไปเป็นประชากรรุ่นใหม่ในระดับการเติบโตในอัตราเอ็กโปเนนเชียล ซึ่งเป็นลักษณะเฉพาะตัวของวิวัฒนาการอัลกอริทึมในการค้นหาแบบขนาน

4.2.2.3 การวิเคราะห์สกีมาด้วยสมการทางคณิตศาสตร์

ถ้ากำหนดให้ $\delta(H)$ คือความยาวของสกีมา H ที่แสดงระยะระหว่างตำแหน่งของบิตแรกและบิตสุดท้ายที่ระบุค่าไว้อย่างชัดเจน เช่นสกีมา 011*1** จะมีค่าความยาวที่กำหนดเท่ากับ 4 จากการหาระยะระหว่างลำดับของบิตแรก $O_f(H) = 1$ และลำดับของบิตสุดท้าย $O_l(H) = 5$ โดยถ้ามีจำนวนตัวอย่าง m ของสกีมา H ที่กระจายอยู่ในประชากร $A(t)$ จะสามารถเขียนได้ดังสมการ $m = m(H, t)$

ในช่วงของการสร้างชุดโครโมโซมใหม่นั้น สตริงต่างๆ จะมีโอกาสที่จะถูกคัดเลือกจากอัตรา $p_i = f_i / \sum f_j$ ซึ่งค่าคาดหวังตัวอย่าง m ของสกีมา H ที่กระจายอยู่ในประชากรในรุ่นต่อไป $m(H, t+1)$ จะมีค่าดังสมการ

$$m(H, t+1) = m(H, t) * n * \frac{f(H)}{\sum f_j} \quad (4.4)$$

ถ้ากำหนดให้ค่าความเหมาะสมเฉลี่ย $\bar{f} = \sum f_j / n$ จะสามารถแสดงสมการการเติบโตของสกีมาได้ดังนี้

$$m(H, t+1) = m(H, t) * \frac{f(H)}{\bar{f}} \quad (4.5)$$

จากสมการดังกล่าวพบว่าสกีมาจะเติบโตในอัตราของค่าความเหมาะสมเฉลี่ยของสกีมาต่อค่าความเหมาะสมเฉลี่ยของประชากร และสกีมาค่าที่มีค่าความเหมาะสมสูงกว่าค่าอัตราความเหมาะสมเฉลี่ยของประชากรจะมีจำนวนมากขึ้นในกลุ่มประชากรรุ่นต่อไป ในขณะที่สกีมาค่าที่มีค่าความเหมาะสมต่ำกว่าค่าอัตราความเหมาะสมเฉลี่ยของประชากรจะมีจำนวนต่ำลงในกลุ่มประชากรรุ่นต่อไป โดยที่ค่าที่คาดหวังนี้จะเกิดขึ้นทุกๆ สกีมาค่า H ในลักษณะขนาน

ถ้าสกีมา H มีค่าสูงกว่าค่าเฉลี่ย $c\bar{f}$ เมื่อ c คือค่าคงที่ สมการข้างต้นสามารถจัดรูปใหม่เป็น

$$m(H, t+1) = m(H, t) * \frac{(\bar{f} + c\bar{f})}{\bar{f}} = (1+c) * m(H, t) \quad (4.6)$$

ถ้าประชากรรุ่นแรก $t = 0$ และค่าอัตราดังกล่าวมีค่าคงที่ตลอด จะพบว่าในประชากรรุ่นใดๆ นั้น

$$m(H, t) = m(H, 0) * (1+c)^t \quad (4.7)$$

4.2.2.4 ผลกระทบของการแลกเปลี่ยนยีน

และเนื่องจากสกีมาจะอยู่รอดเมื่อตำแหน่งการแลกเปลี่ยนยีนอยู่นอกช่วงความยาว ดังนั้นความน่าจะเป็นภายใต้วิธีการแลกเปลี่ยนยีนหนึ่งจุดคือ $p_s = 1 - \delta(H)/(l-1)$ และถ้าการแลกเปลี่ยนยีนจะเกิดขึ้นจากความน่าจะเป็น p_c ดังนั้นความน่าจะเป็นในการอยู่รอด แสดงได้ใหม่ดังนี้

$$p_s \geq 1 - p_c \frac{\delta(H)}{l-1} \quad (4.8)$$

และจำนวนของสกีมา H ที่คาดว่าจะเกิดขึ้นในประชากรรุ่นต่อไปคือ

$$m(H, t+1) \geq m(H, t) * \frac{f(H)}{\bar{f}} * \left[1 - p_c * \frac{\delta(H)}{l-1} \right] \quad (4.9)$$

จากสมการข้างต้น พบว่าสกิม H จะเติบโตหรือล้มตายลงจากแฟคเตอร์การคูณที่เกิดขึ้น โดยทั้งแฟคเตอร์ของการแลกเปลี่ยนและการสร้างชุดโครโมโซมใหม่ดังกล่าวจะขึ้นอยู่กับลักษณะสกิมที่มีค่าความเหมาะสมสูง/ต่ำกว่าค่าเฉลี่ยในประชากร และลักษณะความยาวของสกิมที่กำหนดไว้

4.2.2.5 ผลกระทบของการผ่าเหล่ายีน

ทฤษฎีสกิมสามารถปรับปรุงให้ดีขึ้นได้โดยการเพิ่มเทอมที่แสดงผลด้านลบของการผ่าเหล่ายีน และตัวปฏิบัติการทางพันธุศาสตร์อื่นๆ แต่หัวใจสำคัญของทฤษฎีนี้แสดงให้เห็นว่าการค้นหาที่เกิดขึ้นจะเป็นการทำงานแบบขนาน โดยค้นหาจากไฮเปอร์เพลนที่มีอยู่มากมายในกลุ่มประชากร และจากส่วนไฮเปอร์เพลนขนาดเล็ก ซึ่งสอดคล้องกับสมการอัตราการสุ่มข้างต้น เมื่อความสูญเสียจากตัวปฏิบัติการรวมโครโมโซมมีค่าน้อย

ถ้าพิจารณาถึงผลกระทบของการผ่าเหล่ายีนจากอัตราความน่าจะเป็น p_m ดังนั้นถ้าสกิม H อยู่รอดได้ ถ้าการผ่าเหล่าที่เกิดขึ้น ไม่อยู่ในตำแหน่งของสกิมที่ระบุไว้ทั้งหมด ดังนั้นการหาค่าตำแหน่งดังกล่าวหรือแอลลีจะอยู่รอดในอัตรา $1 - p_m$ และสกิมจะอยู่รอดเมื่อลำดับ $O(H)$ ทั้งหมดที่ระบุในสกีมนั้นคงอยู่ โดยถ้าคุณความน่าจะเป็นในการอยู่รอด $(1 - p_m)$ ด้วยลำดับ $O(H)$ ของตัวมัน ค่าความน่าจะเป็นของการอยู่รอดจากการผ่าเหล่าคือ $(1 - p_m)^{O(H)}$ ซึ่งสำหรับความน่าจะเป็นของการผ่าเหล่าที่มีค่าน้อยมาก $p_m \ll 1$ ความน่าจะเป็นที่สกีมจะอยู่รอดจะสามารถประมาณเป็น $(1 - O(H)p_m)$ ดังนั้นจะมีจำนวนคาดหวังของการคัดลอกสกีม H ในประชากรรุ่นต่อไปภายใต้ขบวนการสร้างชุดโครโมโซมใหม่ การแลกเปลี่ยน และการผ่าเหล่ายีน ดังสมการต่อไปนี้

$$m(H, t+1) \geq m(H, t) * \frac{f(H)}{\bar{f}} * \left[1 - p_c \frac{\delta(H)}{l-1} - O(H)p_m \right] \quad (4.10)$$

จากผลของการคัดเลือกที่เกิดขึ้นนี้ จะพบว่าจำนวนการค้นหาที่ถูกแบ่งออกไปตามไฮเปอร์เพลนที่มีขนาดสั้น และมีค่าความเหมาะสมที่สูงกว่าค่าเฉลี่ย จะมีจำนวนมากขึ้นอย่างต่อเนื่องในอัตราเอ็กโปเนนเชียล โดยที่การค้นหาที่มีค่าความเหมาะสมที่ต่ำกว่าค่าเฉลี่ยจะมีจำนวนลดลงอย่างต่อเนื่องในทำนองกลับกัน

ลักษณะการคัดเลือกตามสัดส่วนของค่าความเหมาะสมอีกรูปแบบหนึ่งคือ การคัดเลือกโดยการเรียงเป็นเชิงเส้น

$$tsr(x, t) = \min + (\max - \min) \frac{rank(x, t) - 1}{N - 1} \quad (4.11)$$

โดยที่ $rank(x, t)$ คือตัวชี้ไปยัง x เมื่อเรียงลำดับชุดโครโมโซมในประชากรตามค่า $u(x)$ ที่เพิ่มขึ้น ชุดโครโมโซมที่แย่ที่สุดจะกำหนดให้ tsr มีค่าน้อยที่สุด (ค่า $\min$ ตามสมการ) และชุดโครโมโซมที่ดีที่สุด

จะกำหนดให้ tsr มีค่าสูงที่สุด (ค่า $\max$ ตามสมการ) ซึ่งค่า tsr สำหรับชุดโครโมโซมอื่นจะกำหนดตามการเรียงลำดับที่เกิดขึ้น โดยที่มีข้อกำหนดสำหรับเงื่อนไข $0 \leq tsr(x,t)$ และ $\sum tsr(x,t) = N$ ว่า $1 \leq \max \leq 2$ และ $\min = 2 - \max$

ประโยชน์ที่สำคัญของการเรียงลำดับตลอดการคัดเลือกนี้ ส่งผลต่อการเข้าสู่คำตอบ โดยที่จะลดผลของแนวโน้มการเข้าสู่คำตอบที่เร็วมากเกินไป

4.2.2.6 ผลกระทบของฟังก์ชันความเหมาะสม

โดยปกติแล้ว ค่าฟังก์ชันค่าความเหมาะสม u จะประกอบด้วย 2 ฟังก์ชันซ้อนกัน ดังสมการ

$$u(x) = g(f(x)) \quad (4.12)$$

โดยที่ f คือค่าฟังก์ชันเป้าหมาย และ g คือฟังก์ชันความเหมาะสม ซึ่งจะแปลงค่าของฟังก์ชันเป้าหมายจากปัญหาที่พิจารณาไปสู่ค่าจำนวนจริงค่าบวก โดยการแปลงค่าที่เกิดขึ้น จะขึ้นอยู่กับปัญหาที่ต้องการหาค่าเหมาะสม โดยฟังก์ชันเป้าหมายที่มีค่าต่ำ จะต้องแปลงไปเป็นค่าความเหมาะสมที่สูง หรือเมื่อฟังก์ชันเป้าหมายอาจจะให้ค่าเป็นลบ ซึ่งฟังก์ชันค่าความเหมาะสม u โดยทั่วไปจะทำการแปลงฟังก์ชันเป้าหมายในลักษณะเชิงเส้น ดังสมการ

$$u(x) = af(x) + b \quad (4.13)$$

โดยที่ a จะมีค่าเป็นบวกถ้าต้องการหาค่าที่เหมาะสมที่สุดจากฟังก์ชัน f และมีค่าเป็นลบเมื่อต้องการหาค่าที่เหมาะสมที่ต่ำสุดจากฟังก์ชัน f และ b เป็นค่าที่ชดเชยเพื่อให้ผลลัพธ์จากฟังก์ชันค่าความเหมาะสมไม่เป็นค่าลบ

ผลกระทบจากการแปลงค่าดังกล่าวส่งผลต่อทฤษฎีสุกิมา ตัวอย่างเช่น ถ้าพิจารณาวิธีเจเนติกอัลกอริทึม 2 วิธี ที่ใช้การคัดเลือกโดยสัดส่วนค่าความเหมาะสม แต่แตกต่างกันตรงที่ วิธีหนึ่งใช้ฟังก์ชันค่าความเหมาะสมเป็น

$$u_1(x) = af(x) + b \quad (4.14)$$

และอีกวิธีหนึ่งใช้ฟังก์ชันค่าความเหมาะสมเป็น

$$u_2(x) = af(x) + b + c, \quad c \neq 0 \quad (4.15)$$

สำหรับไฮเปอร์เพลน H ที่มีขนาดสั้นนั้น จะมีอัตราการเติบโตในประชากรรุ่นแรกเป็น

$$tsr_1(H,t) = \frac{u_1(H,t)}{\bar{u}_1(t)} \quad (4.16)$$

ในขณะที่อัตราการเติบโตในประชากรรุ่นที่ 2 คือ

$$\begin{aligned} tsr_2(H,t) &= \frac{u_2(H,t)}{\bar{u}_2(t)} \\ &= \frac{u_1(H,t) + c}{\bar{u}_1(t) + c} \end{aligned} \quad (4.17)$$

ถึงแม้ว่าวิธีเจเนติก อัลกอริทึมทั้ง 2 วิธี จะมีพฤติกรรมการค้นหาตามทฤษฎีสกิมมา แต่วิธีทั้งสองจะแบ่งการค้นหาไปสู่ไฮเปอร์เพลนในอัตราที่ต่างกัน ซึ่งการตีความของทฤษฎีสกิมมานั้นจะเปลี่ยนแปลงไปถ้ามีการปรับระดับฟังก์ชันค่าความเหมาะสมในช่วงระหว่างการค้นหา โดยสาเหตุในการปรับระดับฟังก์ชันค่าความเหมาะสม มาจากแนวคิดที่วิธีเจเนติกอัลกอริทึมที่จะเปลี่ยนแปลงลักษณะประชากรอย่างรวดเร็ว โดยค่าของ $f(x)$ จะมีแนวโน้มที่จะเข้าสู่ช่วงแคบๆ ซึ่งถ้าปรับค่าคงที่ในการแปลง (ในสมการที่ 4.8) ผลของแรงกดดันในการคัดเลือก จะลดลงเป็นอย่างมาก ในลักษณะที่ชุดโครโมโซมในประชากรจะเข้าสู่ระดับค่าความเหมาะสมที่เท่าๆกัน ตัวอย่างของการแปลงค่าความเหมาะสมคือ ฟังก์ชันค่าความเหมาะสมเชิงเส้นพลวัต [1]

$$u(x) = af(x) + b(t) \quad (4.18)$$

โดยที่ $b(t)$ จะกำหนดจาก ผลกระทบของการปรับเปลี่ยน ตัวอย่างเช่นถ้ากำหนดให้ปัญหาที่กำลังพิจารณาเป็นการหาค่าเหมาะสมที่สุดของ $f(x)$

$$b(t) = -\min\{f(x) \mid x \in P(t)\} \quad (4.19)$$

โดยทั่วไปนั้น ควรที่จะปรับเปลี่ยน $b(t)$ อย่างเหมาะสม เพื่อที่จะทำให้เกิดฟังก์ชันค่าความเหมาะสมที่มีเสถียรภาพ

4.2.3 การทำงานที่ขนานกันภายในวิธีเจเนติกอัลกอริทึม

ลักษณะที่แตกต่างของวิธีเจเนติก อัลกอริทึม คือความสามารถในการดำเนินการสกิมมาแบบขนาน โดยที่สกิมมาที่มีอยู่ในประชากรจะถูกนำมาเป็นต้นแบบการคัดเลือกใหม่ที่เกิดขึ้น และสกิมมาที่มีค่าสูงกว่าค่าเฉลี่ยจะเติบโตในอัตราอีกโปเนนเชียล

สิ่งสำคัญอย่างยิ่งที่ควรพิจารณาในระหว่างการแลกเปลี่ยนยีนคือผลของการรักษาความหลากหลายทางพันธุศาสตร์ในกลุ่มประชากรและการค้นหาพื้นที่คำตอบใหม่จากกลุ่มประชากรที่มีอยู่

ซึ่งสำหรับการประยุกต์ใช้งานในทางปฏิบัติ เช่นการแก้ปัญหาการหาค่าที่เหมาะสม และการเรียนรู้ การสร้างจำนวนประชากรที่มาก อาจจะไม่ก่อให้เกิดประโยชน์ในการค้นหามากนัก ซึ่งถ้าใช้จำนวนประชากรที่น้อย ตัวปฏิบัติการทางพันธุศาสตร์ อาจจะทำให้เกิดแนวโน้มการลู่เข้าที่เร็วมากเกินไป ซึ่งเกิดความล้มเหลวในการค้นหาสมาชิกมาใหม่ ๆ

4.2.4 เจเนติกดริฟต์

ถ้าพิจารณาปัญหาที่มีจุดยอดหลายจุดแล้วนั้น ผลลัพธ์ที่ดีที่สุดจากวิธีเจเนติกอัลกอริทึมคือการลู่เข้าสู่คำตอบที่ดีที่สุดของประชากร เจเนติกดริฟต์จะนิยามได้จากสถานะการลู่เข้าของประชากรจำนวนอนันต์เมื่อมีผลกระทบจากการแลกเปลี่ยนยีน และขบวนการทางพันธุศาสตร์อื่นๆ ในการลดผลกระทบจากเจเนติกดริฟต์นั้น จะต้องพัฒนาวิธีเจเนติกอัลกอริทึมให้สามารถสร้างและรักษากลุ่มประชากรย่อยที่มีเสถียรภาพได้ วิธีเจเนติกอัลกอริทึมที่ใช้กลไกนี้ซึ่งเจเนติกอัลกอริทึมจะมีความสามารถในการหาคำตอบที่เป็นไปได้หลายคำตอบได้เป็นอย่างดี ซึ่งจะกล่าวในรายละเอียดต่อไป

4.3 การปรับปรุงประสิทธิภาพการทำงานของวิธีเจเนติกอัลกอริทึม

4.3.1 การเข้ารหัสโครโมโซมที่แข็งแกร่ง

การเข้ารหัสโครโมโซมคือการแปลงค่าพารามิเตอร์ควบคุมไปเป็นสตริงของอักขระหรือบิตต่างๆ ตามประเภทของการเข้ารหัส ซึ่งมีแนวคิดมากจากการเลียนแบบลักษณะทางพันธุศาสตร์ วิธีการเข้ารหัสที่ใช้กันโดยทั่วไปคือแบบไบนารี แบบค่าจำนวนจริง และแบบการเปลี่ยนลำดับ

การวิเคราะห์การเข้ารหัสนั้นสามารถพิจารณาในเรื่องของเจโนไทป์ที่ไม่แสดงพื้นที่คำตอบได้ทั้งหมด การเทียบค่าในการจำลองระหว่างเจโนไทป์และฟีโนไทป์ ตัวปฏิบัติการทางพันธุศาสตร์ บิตดิงบลิค ไอโซมอร์ฟิซึมในการเปลี่ยนรูปร่าง และการเข้ารหัสแบบผสม

วิธีเจเนติกอัลกอริทึมนั้นเป็นแบบจำลองในการคำนวณที่ได้แนวคิดจากทางชีววิทยา และขบวนการทำงานส่วนใหญ่ นั้น ได้จากลักษณะที่เกิดขึ้นในทางพันธุศาสตร์ ซึ่งในทางพันธุศาสตร์นั้น กลุ่มยีนจะประกอบขึ้นเป็นโมเลกุลที่มีความยาว ซึ่งเรียกว่า Deoxyribonucleic acid หรือ DNA ในแต่ละเซลล์ จะเกิดการจัดรูปร่างดีเอ็นเอร่วมกับโปรตีน เป็นโครงสร้างที่เรียกว่าเจโนไทป์ ซึ่งสำหรับการวิธีเจเนติกอัลกอริทึมนั้น จะพิจารณาเพียงเพียงให้สตริงของยีนแสดงถึงเจโนไทป์ โดยที่ยีนเหล่านั้นจะอยู่ในตำแหน่ง (โลคัส) และมีค่า (แอลลี) ที่แตกต่างกันออกไปในเจโนไทป์ โครงสร้างที่ประกอบไปด้วยยีนเหล่านี้ จะเรียกว่าโครโมโซม โดยในธรรมชาตินั้น การเทียบเคียงระหว่างเจโนไทป์และคุณลักษณะของสิ่งมีชีวิต (ฟีโนไทป์) จะมีความซับซ้อนสูง ยีนในแต่ตัวจะมีผลกับคุณภาพของฟีโนไทป์ ซึ่งเรียกว่า Pleiotropy การเทียบเคียงอื่นๆ จะพิจารณาจากอักขระในฟีโนไทป์ที่ถูกควบคุมหรือผลจากยีนหลายตัว โดยในวิธีเจเนติกอัลกอริทึมโดยทั่วไปจะเป็นการแปลงค่า (แบบหนึ่งต่อหนึ่ง) ไปสู่คุณลักษณะทางฟีโนไทป์ ซึ่งสะท้อนถึงวิวัฒนาการของสิ่งมีชีวิตทางธรรมชาติอย่างง่าย

การเข้ารหัสจะพิจารณาตามคุณสมบัติดังต่อไปนี้

1. การเข้ารหัสจะรวมเอาบิตดิงบล็อคลำดับหลักจำนวนหลายส่วนเข้ารวมกัน ซึ่งเป็นสิ่งที่มีความสำคัญสำหรับประเภทของปัญหา
2. การเข้ารหัสจะสอดคล้องกับกลุ่มของตัวปฏิบัติการทางพันธุศาสตร์ซึ่งสามารถส่งผ่านบิตดิงบล็อคจากเจโนไทป์พ่อแม่ไปสู่เจโนไทป์ของลูกได้เป็นอย่างดี
3. การเข้ารหัสจะทำให้ฮิพทาซีสมีค่าน้อยที่สุด โดยที่ผลกระทบจากยีนหนึ่งจะลดผลการกระทำของยีนอื่นๆ
4. การเข้ารหัสจะสามารถจัดการการเทียบค่าไปสู่ฟีโนไทป์นั้นได้อย่างง่าย และสามารถแปลงค่าความเหมาะสมโดยใช้การคำนวณไม่มากนัก
5. การเข้ารหัสจะใช้ขบวนการเทียบค่าเจโนไทป์และฟีโนไทป์ที่เหมาะสมให้เกิดประโยชน์ ถ้าการเทียบค่าอย่างง่ายไปสู่เจโนไทป์นั้นไม่สามารถกระทำได้โดยตรง
6. การเข้ารหัสจะต้องรวมเอาคำตอบที่เป็นไปได้เข้าไว้ด้วยกัน (ถ้าเป็นไปได้) หรืออาจจะใช้วิธีการปรับโทษหรือเทคนิคอื่นเข้าช่วย ถ้ามีคำตอบที่เป็นไปไม่ได้ (ไม่อยู่ในข้อจำกัด) เกิดขึ้น
7. การเข้ารหัสจะลดผลรูปร่างไอโซมอร์ฟิก
8. การเข้ารหัสจะใช้ค่ายีนที่ได้จากอักขระของคาไดนอลิตีที่เป็นไปได้ที่น้อยที่สุด โดยการเข้ารหัสไบนารีจะถูกพิจารณาว่าเป็นวิธีที่ดีที่สุดถ้าเหมาะสมกับปัญหาที่พิจารณา
9. การเข้ารหัสจะก่อให้เกิดปัญหาจากช่วงตำแหน่งในพื้นที่คำตอบ และกลุ่มคำตอบที่เป็นไปได้

ปัญหาวิธีเจเนติกอัลกอริทึมที่ต้องการการเข้ารหัสที่มีความเฉพาะตัว ซึ่งก็ไม่มีวิธีพื้นฐานที่จะเหมาะสมกับทุกปัญหา ทั้งนี้จะต้องพิจารณาลักษณะเฉพาะของแต่ละปัญหาเป็นสำคัญ

4.3.1.1 การเข้ารหัสไบนารี

การเข้ารหัสไบนารีเป็นวิธีที่ดีสำหรับปัญหาที่ซึ่งตัวแปรของปัญหานั้นสามารถเทียบค่าไปสู่สตริงของค่าศูนย์หรือหนึ่ง คล้ายกับลักษณะทางธรรมชาติ ไบนารีสตริงส่วนใหญ่จะใช้ในการแสดงการเข้ารหัส เช่น $E = [10001010]$ ประกอบด้วย 8 ยีน โดยที่โลคัส (ตำแหน่ง) ของยีนลำดับ i แสดงถึงบิตลำดับ i ในสตริงไบนารีโดยตรง และแอสลีสของบิตต่างๆ แสดงได้ด้วย $E[i]$

สำหรับปัญหาอื่นนั้น การเข้ารหัสไบนารีอาจจะไม่เหมาะสมเนื่องจากว่า

1. ฮิพทาซีส โดยค่าของบิตอาจจะมีความสัมพันธ์กับการแสดงความเหมาะสมจากบิตอื่น ซึ่งสามารถที่จะส่งผลไปสู่ความเหมาะสมของเจโนไทป์โดยไม่มีอิทธิพลจากค่าของบิตอื่น และอาจจะเป็นผลทางลบในการแลกเปลี่ยนสกีมาตาของวิธีเจเนติกอัลกอริทึม
2. ลักษณะทางธรรมชาติสำหรับปัญหาที่กำลังพิจารณานั้น จะต้องการบิตที่มีลำดับความหมายสูงกว่าการใช้ไบนารี และบิตนั้นสามารถจัดเรียงเข้าด้วยกันเพื่อสร้างบิตดิงบล็อคที่เทียบไปสู่คุณลักษณะเฉพาะของปัญหาที่สามารถแยกออกเป็นส่วนตัว ซึ่งในกรณีนี้อักขระที่มีลำดับสูงจะเป็นวิธีที่มีประสิทธิภาพมาก

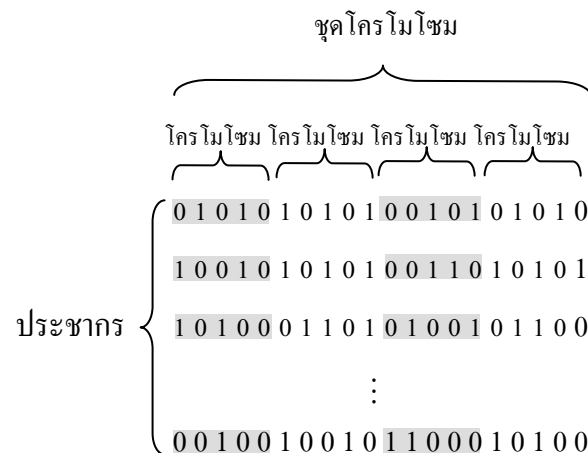
3. คำตอบที่เป็นไปไม่ได้ ตัวปฏิบัติการทางพันธุศาสตร์นั้นสามารถสร้างคำตอบที่เป็นไปไม่ได้ เกิดขึ้นจากการเข้ารหัสไบนารี

ถ้าพิจารณาปัญหาการหาค่าที่เหมาะสมคือ การหาค่าที่มากที่สุดของ $f(x_1, x_2, x_3, \dots, x_m)$ เมื่อตัวแปรควบคุม x_i เป็นค่าจำนวนจริง ตามสมการข้อจำกัด $a_i \leq x_i \leq b_i$ วิธีการเข้ารหัสไบนารีโดยทั่วไปจะใช้สตริงไบนารีขนาด n บิต แสดงค่าจำนวนจริงในช่วง

$$a_i + k \frac{(b_i - a_i)}{2^n} \leq x_i \leq a_i + (k+1) \frac{b_i - a_i}{2^n} \quad (4.20)$$

ซึ่งสอดคล้องกับรหัสไบนารีพื้นฐานสำหรับค่าจำนวนเต็ม k โดยที่ $0 \leq k \leq 2^n$

ถ้าไม่พิจารณาถึงช่วงค่าจำนวนจริงที่ได้จากรหัสไบนารี จะสามารถเทียบเคียงกับรหัสไบนารีสำหรับค่าจำนวนเต็ม k ได้จาก $k \frac{(b_i - a_i)}{2^n}$



รูปที่ 4.7 การจัดเรียงตัวของประชากร

4.3.1.2 การเข้ารหัสเกรย์

การเข้ารหัสเกรย์เป็นวิธีในการเข้ารหัสไบนารีที่มีลักษณะพิเศษ ถ้ามีการเพิ่มค่าตัวแปรควบคุมจากช่วงค่าของการเข้ารหัส จะเกิดการเปลี่ยนค่าบิตเพียงแคในตำแหน่งเดียว สมการการแปลงจากค่าไบนารีไปเป็นค่าเกรย์แสดงได้ดังนี้

$$\gamma_k = \begin{cases} \beta_1; & \text{if } k=1 \\ \text{comp}(\beta_k); & \text{if } k > 1 \end{cases} \quad (4.21)$$

โดยที่ r_k คือรหัสเกรย์ในลำดับ k และ β_k คือรหัสไบนารีในลำดับ k และบิตที่ 1 จะเริ่มจากตำแหน่งซ้าย และ $comp()$ คือฟังก์ชันการหาค่าคอมพลิเมนต์

การเข้ารหัสเกรย์นั้นนอกจากที่จะลดผลของแฮมมิงคลิฟฟ์ (เกิดขึ้นเมื่อค่าจำนวนเลขที่อยู่ติดกันเป็นค่าคอมพลิเมนต์ของไบนารี) แล้ว จะยังคงมีความเป็นไปได้ในการลดจำนวนคำตอบโบลอลในพื้นที่คำตอบใหม่ การเข้ารหัสเกรย์จะสร้างฟังก์ชันการเทียบค่าที่แตกต่างออกไปจากเดิม ทำให้เกิดความสัมพันธ์ไฮเปอร์เพลนในลักษณะใหม่ ซึ่งเปลี่ยนแปลงการแข่งขันสกีมาตาในการค้นหาทางพันธุศาสตร์

4.3.1.3 การเข้ารหัสพารามิเตอร์

ในปัญหาการหาค่าที่เหมาะสมของฟังก์ชันขนาด n มิติ $f(x,y,z)$ นั้น เป้าหมายของการหาค่าที่เหมาะสมโดยปกติคือการหาค่าที่มาก/น้อยที่สุด ของฟังก์ชันในโดเมนที่พิจารณา ซึ่งสำหรับวิธีเจเนติก อัลกอริทึมเดิม ตัวแปรควบคุมของปัญหาจะถูกนำมาเข้ารหัสเป็นบิตสตริง และถ้าพิจารณาฟังก์ชันขนาด n มิติ ค่า x,y,z จะนำมาเข้ารหัสเป็นสตริงไบนารีย่อย 3 ส่วน โดยที่สตริงย่อยเหล่านี้จะถูกนำมาเรียงต่อกันเป็นเจโนไทป์บิตสตริงชุดเดียว โดยพารามิเตอร์อาจจะถูกเข้ารหัสโดยสตริงไบนารีหรือจำนวนค่าจำนวนจริงเดิมจากการกำหนดบิตเครื่องหมาย บิตค่าเลขฐาน และบิตค่ากำลัง

4.3.1.4 การเข้ารหัสโดยการเปลี่ยนลำดับ

ปัญหาการเปลี่ยนลำดับนั้นจะต้องการการจัดเรียงชุดของบิตในโครงสร้างการเชื่อมโยงให้มีความเหมาะสม เช่นปัญหาการเดินทางของพนักงานขายของ การเข้ารหัสโดยการเปลี่ยนลำดับอย่างเดียวยังไม่สามารถรวบรวมบิตดิงบล็อคที่มีอยู่ได้ดัดนัก ในปัญหาการเปลี่ยนลำดับนั้น บิตดิงบล็อคที่เกิดขึ้นจะถูกกำหนดจากลักษณะทางธรรมชาติของฟังก์ชันความเหมาะสมของปัญหา ตัวปฏิบัติการทางพันธุศาสตร์และการเข้ารหัส

4.3.2 การปรับค่าความเหมาะสม

4.3.2.1 การปรับค่าความเหมาะสมโดยการตัดค่าซิกมา

วิธีการการปรับค่าความเหมาะสมเป็นส่วนหนึ่งที่จะช่วยลดผลกระทบจากชุดโครโมโซมที่มีค่าความเหมาะสมที่สูงมาก เมื่อเทียบกับชุดโครโมโซมอื่น โดยการปรับค่าความเหมาะสมของชุดโครโมโซมทั้งหมดก่อนที่จะวนการคัดเลือกจะเริ่มขึ้น เทคนิคการปรับค่าความเหมาะสมโดยการตัดส่วนซิกมาจะพิจารณาถึงความผันแปรของค่าความเหมาะสมในประชากรมาใช้ในการปรับค่าความเหมาะสมเดิม ดังสมการต่อไปนี้

$$f' = f - (\bar{f} - c\sigma) \quad (4.22)$$

ค่าคงที่ c จะกำหนดให้อยู่ในช่วง 1 ถึง 3 สำหรับการปรับค่าเบี่ยงเบนมาตรฐาน และถ้าผลลัพธ์จากสมการ f' ที่ได้มีค่าเป็นลบ ค่า f' จะถูกกำหนดให้มีค่าเท่ากับ 0

4.3.3 วิธีการคัดเลือกที่ดี

4.3.3.1 วิธีการคัดเลือกแบบทัวร์นาเมนต์

ถ้าเทียบวิธีการคัดเลือกที่ใช้อยู่กันโดยทั่วไปแล้ว วิธีการคัดเลือกแบบทัวร์นาเมนต์เป็นวิธีที่มีจุดเด่นเหนือกว่าวิธีอื่นๆ เนื่องจากวิธีนี้ส่วนใหญ่จะเกี่ยวข้องกับการคำนวณระหว่างชุดโครโมโซม ซึ่งเป็นการจำลองวิถีทางธรรมชาติในการคัดเลือกจากการเรียงลำดับค่าความเหมาะสม อีกทั้งวิธีนี้จะมีผลกระทบจากการรบกวนของขบวนการคัดเลือก น้อยที่สุด และไม่ขึ้นกับจำนวนขนาดของประชากรมากนัก

วิธีการคัดเลือกแบบทัวร์นาเมนต์นิยมใช้มากในวิวิเจเนติกอัลกอริทึมประเภทสถานะคงตัว โดยจะคัดเลือกทัวร์นาเมนต์ของชุดโครโมโซมขึ้นมาอย่างสุ่ม และชุดโครโมโซมที่ดีที่สุดในทัวร์นาเมนต์นั้นจะถูกกำหนดให้เป็นชุดโครโมโซมต้นแบบ และกระทำซ้ำจนกระทั่งได้ชุดโครโมโซมต้นแบบที่ต้องการครบตามจำนวน โดยที่พารามิเตอร์ขนาดทัวร์นาเมนต์จะมีค่าอยู่ในช่วง 2 ถึง จำนวนชุดโครโมโซมทั้งหมดในประชากร

Tournament Selection

Input: The population $\vec{P} = (P_1, \dots, P_N)$,

The fitness values of the population $\vec{\phi} = (\phi_1, \dots, \phi_N)$.

The tournament size $t = \{1, 2, \dots, N\}$

Output: The population after selection $\vec{P}' = (P'_1, \dots, P'_N)$

Tournament($t, \vec{J}, \vec{\phi}$):

For $i \leftarrow 1$ to N do

$P'_i \leftarrow$ The best fit individual out of t uniform randomly

Pick individuals from $\vec{P}$

End

Return $\vec{P}'$

รูปที่ 4.8 การทำงานของการคัดเลือกแบบทัวร์นาเมนต์

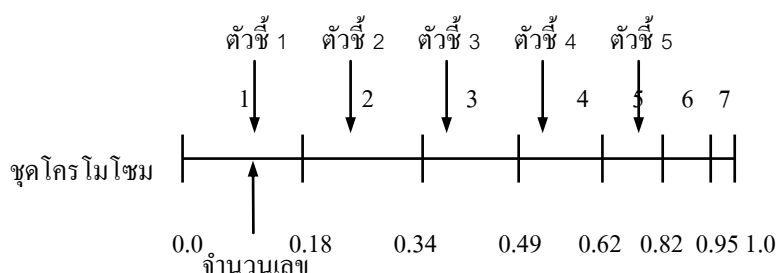
4.3.3.2 วิธีการคัดเลือกทั้งหมดแบบสุ่ม

วิธีการคัดเลือกทั้งหมดแบบสุ่มถูกพัฒนาขึ้นมาเพื่อลดความโน้มเอียงในขบวนการคัดเลือกจาก

วิธีกึ่งสุ่ม และให้ชุดโครโมโซมที่กระจายไปทั่วประชากร โดยที่ค่าของความเหมาะสมจากแต่ละชุดโครโมโซมทั้งหมดจะนำมาเรียงต่อกันคล้ายกับวิธีกึ่งสุ่ม จากนั้นสุ่มค่าขึ้นมาค่าหนึ่งและใช้ค่าสุ่มนั้นในการแบ่งระยะค่าความเหมาะสมรวมทั้งหมดออกเป็นส่วนเท่าๆกัน ดังรูป โดยตำแหน่งจุดแบ่งในแต่ละส่วนจะแทนถึงชุดโครโมโซมที่ถูกคัดเลือกขึ้นมา

ถ้ากำหนดให้ N_{ptr} คือจำนวนชุดโครโมโซมที่ต้องการคัดเลือก ระยะทางระหว่างตัวชี้คือ $\frac{1}{N_{ptr}}$

โดยที่ตำแหน่งของตัวชี้แรก จะได้จากค่าจำนวนสุ่มในช่วง $\left[0, \frac{1}{N_{ptr}}\right]$



รูปที่ 4.9 วิธีการคัดเลือกทั้งหมดแบบสุ่ม

4.3.4 การแลกเปลี่ยนยีน

4.3.4 การแลกเปลี่ยนยีนอย่างสม่ำเสมอ

การแลกเปลี่ยนยีนหนึ่งจุด และสองจุด เป็นวิธีที่ใช้ในการดำเนินการทางพันธุศาสตร์ของวิธีเจเนติกอัลกอริทึมโดยทั่วไป ซึ่งแสดงถึงจำนวนจุดที่เกิดการแลกเปลี่ยนกลุ่มยีนระหว่างคู่โครโมโซมต้นแบบ

ถ้ากำหนดตำแหน่งที่เกิดการแลกเปลี่ยนยีนโดยตรง สำหรับขบวนการแลกเปลี่ยนยีน จะสามารถแสดงได้เป็นสตริงไบนารีที่มีความยาวเท่ากับขนาดของชุดโครโมโซม ตำแหน่งที่ต้องการแลกเปลี่ยนยีนจะกำหนดให้มีค่าเท่ากับ 1 ดังรูป

ชุดโครโมโซมต้นแบบที่ 1: 0 0 1 1 1 1

ชุดโครโมโซมต้นแบบที่ 2: 1 1 1 1 0 0

ตำแหน่งการแลกเปลี่ยนยีน: 0 1 0 1 0 1

ชุดโครโมโซมใหม่ที่ 1: 0 1 1 1 1 0

ชุดโครโมโซมใหม่ที่ 2: 1 0 1 1 0 1

รูปที่ 4.10 การแลกเปลี่ยนยีนด้วยวิธีการแลกเปลี่ยนยีนอย่างสม่ำเสมอ

หรือถ้าพิจารณาตำแหน่งที่เกิดการแลกเปลี่ยนยีนโดยตรง เทียบกับวิธีการแลกเปลี่ยนยีนหนึ่งจุด และสองจุด จะมีลักษณะดังรูป

ตำแหน่งการแลกเปลี่ยนยีนด้วยวิธี

การแลกเปลี่ยนยีนหนึ่งจุด: 1 1 1 1 1 0 0 0 0 0

การแลกเปลี่ยนยีนสองจุด: 0 0 1 1 1 1 1 0 0 0

การแลกเปลี่ยนยีนอย่างสม่ำเสมอ: 0 1 0 0 1 0 1 1 0 1

รูปที่ 4.11 ตำแหน่งที่เกิดการแลกเปลี่ยนยีนโดยตรงเทียบกับวิธีการแลกเปลี่ยนยีนชนิดต่างๆ

สำหรับการแลกเปลี่ยนยีนอย่างสม่ำเสมอ นั้น จะสร้างสตริงตำแหน่งการแลกเปลี่ยนยีน โดยการสุ่มค่าบิตตลอดความยาวของชุดโครโมโซมต้นแบบ ด้วยความน่าจะเป็นเท่ากับ 0.5 และจากทฤษฎีสกิมมาในรูปแบบทั่วไป

$$m(S, t+1) = m(S, t) \frac{f(S)}{\bar{f}} (1 - \varepsilon) \quad (4.16)$$

โดยที่ ε คือค่าความสูญเสียที่สะท้อนถึงความเสียหายจากตัวปฏิบัติการการแลกเปลี่ยนและการผ่าเหล่า ยีน และเทอม $1 - \varepsilon$ หมายถึงอัตราการอยู่รอด

4.3.4.2 การอยู่รอดของสกีมาภายหลังการแลกเปลี่ยนยีน

สกีมา A สามารถที่จะอยู่รอดภายหลังการแลกเปลี่ยนยีน ถ้าสกีมานั้นปรากฏอยู่ในชุดโครโมโซมลูกที่เกิดขึ้น (ในชุดโครโมโซมชุดใดชุดหนึ่ง) ซึ่งถ้ากำหนด $P_1(n_A)$ และ $P_2(n_A)$ แสดงโอกาสที่ สกีมา A สามารถอยู่รอด เมื่อ จำนวนบิต n_A จะถูกกำหนดโดยตำแหน่งที่เกิดการแลกเปลี่ยนยีน และ $P_S(n_A)$ แสดงโอกาสการอยู่รอดทั้งหมด เมื่อจำนวนบิต n_A ของสกีมา A ถูกกำหนดขึ้น ถ้าพิจารณาโอกาส P_{ok} ของบิตที่ถูกกำหนดไว้จะสามารถอยู่รอดได้ดังสมการ

$$P_1(n_A) = P_{ok}^{(n_A)} \quad (4.17)$$

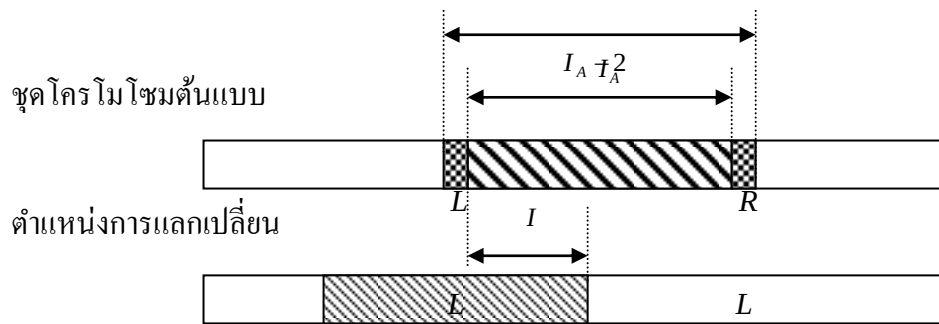
$$P_2(n_A) = P_{ok}^{(b_A - n_A)}$$

และ

$$P_S(n_A) = P_1(n_A) + [1 - P_1(n_A)]P_2(n_A) \quad (4.18)$$

โดยที่ P_{ok} คือโอกาสของบิตในชุดโครโมโซมต้นแบบจะตรงกับบิตของสกีมา

ถ้ากำหนดให้ $M(m)$ เป็นจำนวนของวิธีที่จะกำหนดการแลกเปลี่ยนจำนวน m บิต บนชุดโครโมโซมใดๆ จะพบว่าสำหรับการแลกเปลี่ยนหนึ่งจุด จะมีค่าเท่ากับ $l-1$ ตำแหน่งการแลกเปลี่ยน สำหรับการแลกเปลี่ยนสองจุดจะมีค่าเท่ากับ l วิธี ถ้ามีตำแหน่งการแลกเปลี่ยน 1 บิต และจะมีค่าเท่ากับ $l-n-1$ วิธี ถ้ามีตำแหน่งการแลกเปลี่ยน n บิต และสำหรับการแลกเปลี่ยนสมมาตรจะมีค่าเท่ากับ $\binom{l}{i}$ สำหรับจำนวน i บิต



รูปที่ 4.12 ความสัมพันธ์ระหว่างชุดโครโมโซม, สกิม่า A , และตำแหน่งการแลกเปลี่ยนจากวิธีการแลกเปลี่ยนสองจุด

จากการทดสอบ [51] พบว่าถ้าพิจารณาอัตราการอยู่รอดเฉลี่ยสำหรับสกิมตา ด้วยค่าความยาวที่กำหนดไว้ในโครโมโซมขนาด 30 บิต พบว่า อัตราการอยู่รอดเฉลี่ยจากวิธีการแลกเปลี่ยนสมมาตรจะมีค่าน้อยกว่าอัตราการอยู่รอดเฉลี่ยจากวิธีการแลกเปลี่ยนหนึ่งจุดและสองจุด และอัตราทั้งสามจะมีค่าลดลงเมื่อจำนวนความยาวที่กำหนดมีค่าสูงขึ้น โดยถ้าพิจารณาสกิมตาที่มีขนาด 5 บิต ในโครโมโซมขนาด 30 บิต พบว่าอัตราการอยู่รอดเฉลี่ยจากวิธีการแลกเปลี่ยนสมมาตรจะมีค่าค่อนข้างคงที่ แต่อัตราการอยู่รอดเฉลี่ยจากวิธีการแลกเปลี่ยนหนึ่งจุดจะมีค่าลดลงเรื่อยๆ เมื่อจำนวนความยาวที่กำหนดมีค่ามากขึ้น อัตราการอยู่รอดเฉลี่ยจากวิธีการแลกเปลี่ยนสองจุดจะลดลงถึงจุดหนึ่งแล้วจะคงตัว

และถ้าพิจารณาอัตราการรวมตัวของสกิมตา ด้วยค่าความยาวที่กำหนดไว้ในโครโมโซมขนาด 30 บิต พบว่า อัตราการอยู่รอดเฉลี่ยจากวิธีการแลกเปลี่ยนสมมาตรจะมีค่าสูงกว่าอัตราการอยู่รอดเฉลี่ยจากวิธีการแลกเปลี่ยนหนึ่งจุดและสองจุดเพียงเล็กน้อย และอัตราทั้งสามจะมีค่าลดลงเมื่อจำนวนความยาวที่กำหนดมีค่าสูงขึ้น โดยถ้าพิจารณาสกิมตาที่มีขนาด 5 บิต ในโครโมโซมขนาด 30 บิต พบว่าอัตราการรวมตัว จากวิธีการแลกเปลี่ยนสมมาตรจะมีค่าค่อนข้างคงที่ แต่อัตราการอยู่รอดเฉลี่ยจากวิธีการแลกเปลี่ยนหนึ่งจุดและสองจุดจะมีค่าลดลงเรื่อยๆ เมื่อจำนวนความยาวที่กำหนดมีค่ามากขึ้น แล้วจะคงตัวสำหรับค่าความยาวที่กำหนดค่าหนึ่ง (ประมาณ 23)

ซึ่งจากผลการทดสอบและการเฝ้าสังเกต [51] พบว่า ในกรณีส่วนใหญ่ นั้น วิธีการแลกเปลี่ยนสมมาตรจะประสิทธิภาพในการรวมสกิมตาเหนือกว่าการแลกเปลี่ยนหนึ่งจุดและสองจุด วิธีการแลกเปลี่ยนสมมาตรถูกแสดงให้เห็นแล้วว่ามีประสิทธิภาพสูงกว่าสำหรับการใช้ในปัญหาการหาค่าที่เหมาะสม

4.3.5 เทคนิคการปรับเปลี่ยนค่าความน่าจะเป็นในการแลกเปลี่ยนยีนและการผ่าเหล่ายีน

วิธีเจเนติกอัลกอริทึมทำงานโดยอาศัยความน่าจะเป็นสำหรับการดำเนินงานทางพันธุศาสตร์สำหรับชุดโครโมโซมในประชากรรุ่นต่างๆ โดยที่ชุดโครโมโซมจะเปลี่ยนแปลงไปในแต่ละรุ่น ในช่วงการคัดเลือก ชุดโครโมโซมจะถูกคัดเลือกตามโอกาสจากค่าความเหมาะสม ซึ่งชุดโครโมโซมที่ดีจะถูกเลือกเป็นชุดโครโมโซมต้นแบบ ในขณะที่ชุดโครโมโซมที่ไม่ดี จะไม่ได้รับการคัดเลือก ตามกฎของการคัดเลือกทางธรรมชาติ ถ้าพิจารณาจากเหตุผลข้างต้น สิ่งที่ใช้วัดชุดโครโมโซมว่า “ดี” หรือ “ไม่ดี” คือค่าความเหมาะสมของชุดโครโมโซมนั้น

การแลกเปลี่ยนยีนของชุดโครโมโซมต้นแบบ มีความสำคัญอย่างยิ่งสำหรับวิธีเจเนติกอัลกอริทึม เนื่องจากจะเป็นการแลกเปลี่ยนส่วนประกอบของโครโมโซมแบบสุ่ม ซึ่งเป็นไปได้ว่าถ้ามีชุดโครโมโซมต้นแบบที่ดีจะทำให้ชุดโครโมโซมใหม่ที่ดีกว่า ซึ่งโดยปกติแล้วการแลกเปลี่ยนยีนจะเกิดขึ้นด้วยความน่าจะเป็น p_c โดยถ้า p_c มีค่าสูง ก็จะทำให้มีโอกาสที่จะทำให้เกิดชุดโครโมโซมใหม่ได้เร็วขึ้น แต่ชุดโครโมโซมที่มีอยู่เดิมก็จะถูกทำลายเร็วขึ้นกว่าขบวนการคัดเลือกจะพิจารณาชุดโครโมโซมนั้น โดยทั่วไปนั้นจะกำหนดค่า p_c ในช่วง 0.5-1.0

การผ่าเหล่ายีน เป็นอีกส่วนหนึ่งของการดำเนินงานทางพันธุศาสตร์ โดยจะช่วยนำส่วนประกอบของโครโมโซมที่สูญเสียไป กลับคืนมา อีกทั้งยังช่วยเพิ่มส่วนประกอบของโครโมโซมที่ไม่มีอยู่เดิม ให้เกิดขึ้นในประชากรใหม่อีกด้วย โดยที่การผ่าเหล่ายีนเป็นการทำงานที่สำคัญในการป้องกันการเข้าสู่คำตอบที่เร็วมากเกินไป และป้องกันไม่ให้วิธีเจเนติกอัลกอริทึมเข้าสู่คำตอบโลคอล ซึ่งโดยปกติแล้วการผ่าเหล่ายีนจะเกิดขึ้นด้วยความน่าจะเป็น p_m โดยถ้า p_m มีค่าสูงก็จะทำให้การค้นหาของวิธีเจเนติกอัลกอริทึมมีลักษณะคล้ายกับการค้นหาแบบสุ่ม โดยทั่วไปนั้นจะกำหนดค่า p_m ในช่วง 0.001-0.05

การกำหนดค่า p_c และ p_m ให้เหมาะสมนั้น ควรที่จะมีการปรับเปลี่ยนค่าก่อนการดำเนินการทางพันธุศาสตร์ทุกครั้ง แทนการใช้เป็นค่าคงที่จากวิธีโดยทั่วไป ซึ่ง Srinivas และพวก [53] ได้นำเสนอสมการการปรับค่าไว้ดังนี้

$$p_c = k_1 (f_{\max} - f_{lrg}) / (f_{\max} - f_{ave}), \quad f_{lrg} \geq f_{ave}, \quad (4.19)$$

$$p_c = k_3, \quad f_{lrg} \leq f_{ave}, \quad (4.20)$$

$$p_m = k_2 (f_{\max} - f) / (f_{\max} - f_{ave}), \quad f_{lrg} \geq f_{ave}, \quad (4.21)$$

$$p_m = k_4, \quad f < f_{ave}, \quad (4.22)$$

โดยที่ $k_1, k_2, k_3, k_4 \leq 1.0$ และ f_{lrg} คือค่าความเหมาะสมที่มากที่สุดจากชุดโครโมโซมต้นแบบ

จากสมการข้างต้น จะพบว่า กลไกการปรับค่า p_c และ p_m จะทำ 2 หน้าที่หลัก คือ เมื่อการคำนวณเข้าสู่จุดต่ำสุดโลคอล ($f_{\max}$ มีค่าใกล้เคียงกับ f_{ave}) ค่า p_c และ p_m จะถูกปรับให้มีค่าเพิ่มขึ้นเพื่อสร้างส่วนประกอบชุดโครโมโซมใหม่ๆ สำหรับการค้นหาคำตอบอื่น อีกส่วนหนึ่งคือถ้าชุดโครโมโซมต้นแบบ “ดี” ($f_{\max}$ มีค่าใกล้เคียงกับ f_{lrg}) ค่า p_c จะถูกปรับให้มีค่าลดลง เพื่อที่จะรักษาชุดโครโมโซมกลุ่มนั้นเอาไว้ และในทำนองเดียวกัน ค่า p_m จะถูกปรับให้มีค่าลดลง ตามสัดส่วนของ $f_{\max} - f$

จากการทดสอบ [53] การใช้วิธีการปรับค่า p_c และ p_m นี้จะช่วยลดอัตราการเข้าสู่จุดต่ำสุดแบบโลคอล เมื่อเทียบกับการใช้ค่า p_c และ p_m แบบคงที่ และค่าแฟกเตอร์ k_1, k_2, k_3, k_4 ที่แนะนำ คือ 1.0, 0.5, 1.0, 0.5 ตามลำดับ เพื่อที่จะให้เกิดการแลกเปลี่ยนยีนเสมอเมื่อชุดโครโมโซมต้นแบบมีค่าความเหมาะสมน้อยกว่าค่าความเหมาะสมเฉลี่ย

4.3.6 การป้องกันการลู่เข้าที่เร็วเกินไป

สถานะการลู่เข้าที่เร็วเกินไป อาจพิจารณาได้จากการสูญเสียความหลากหลายทางพันธุศาสตร์ในประชากร ก่อนที่จะได้คำตอบที่เหมาะสม โดยวิธีการลดผลกระทบที่เกิดขึ้น มีดังนี้

1. การป้องกันการแลกเปลี่ยนยีนระหว่างชุดโครโมโซมที่มาจากชุดโครโมโซมพ่อ-แม่เดียวกัน
2. การแลกเปลี่ยนยีนสม่ำเสมอ
3. การกำจัดชุดโครโมโซมที่มีลักษณะเหมือนกัน

กลยุทธ์การรักษาความหลากหลายทางพันธุศาสตร์ในประชากรสามารถแยกออกพิจารณาได้ดังนี้

4.3.6.1 วิธีการคัดเลือกกลุ่มชุดโครโมโซมสำหรับการปฏิบัติการทางพันธุศาสตร์

สำหรับกลยุทธ์การคัดเลือกกลุ่มชุดโครโมโซมโดยทั่วไปแล้วจะมีเป้าหมายในการป้องกันชุดโครโมโซมที่ไม่เหมือนกัน อย่างสิ้นเชิง ทฤษฎีฟังก์ชันการแบ่งส่วนค่าความเหมาะสม (Goldberg [1]) จะลดค่าความเหมาะสมของชุดโครโมโซม จากฟังก์ชันที่เทียบค่าความเหมือนกันของชุดโครโมโซมกับชุดโครโมโซมอื่นในประชากร ซึ่งวิธีดังกล่าวจะลดผลการคัดเลือกชุดโครโมโซมในสปีชีส์เดียวกันจากสปีชีส์ที่มีอิทธิพลต่อขบวนการคัดเลือกสูง

วิธีการป้องกันการแลกเปลี่ยนยีนระหว่างชุดโครโมโซมที่มาจากชุดโครโมโซมพ่อ-แม่เดียวกันของ Eshelman เป็นอีกวิธีการหนึ่งโดยการจับคู่ชุดโครโมโซมแบบสุ่ม แต่จะพิจารณาระยะทางแฮมมิงระหว่างคู่โครโมโซมนั้น โดยจะต้องมีค่ามากกว่าค่าเทรสโฮลด์ที่กำหนด ซึ่งในช่วงเริ่มต้น จะปรับค่าเทรสโฮลด์ให้เท่ากับค่าระยะทางแฮมมิงเฉลี่ยคาดหวังของประชากรรุ่นแรก และจะมีค่าลดลงเมื่อประชากรเข้าสู่กลุ่มคำตอบ กลยุทธ์ในการคัดเลือกกลุ่มโครโมโซมพ่อ-แม่ที่เลือกชุดโครโมโซมที่แตกต่างกัน จะมีผลข้างเคียงเกิดขึ้นจากการทำลายสปีชีส์จากการแลกเปลี่ยนยีน เนื่องจากมีเพียงสปีชีส์เพียงส่วนน้อยที่แลกเปลี่ยนกัน

4.3.6.2 วิธีการสร้างชุดโครโมโซมใหม่

กลยุทธ์นี้จะใช้การแลกเปลี่ยนยีนช่วยการส่งเสริมความหลากหลายในทางพันธุศาสตร์ในประชากร โดยที่วิธีโดยตรงคือการเพิ่มอัตราของการแลกเปลี่ยนยีน และวิธีอื่นคือการทำให้การแลกเปลี่ยนยีนที่เกิดขึ้นทำงานบนค่าบิตที่แตกต่างกัน เช่นถ้าคู่โครโมโซมพ่อแม่แตกต่างกัน 2 บิต ชุดโครโมโซมลูกก็จะแตกต่างไปจากชุดโครโมโซมพ่อแม่ โดยที่วิธีหนึ่งที่จะช่วยได้มากกว่าคือการแลกเปลี่ยนยีนสม่ำเสมอ ซึ่งไม่เหมือนกับการแลกเปลี่ยนยีนชนิดเดิมที่จะสร้างชุดโครโมโซมลูก (ทั้ง 2 ชุดโครโมโซม) ที่เหมือนกัน โดยมีแนวโน้มที่จะแลกเปลี่ยนยีนที่แตกต่างกันในจำนวนประมาณจำนวนครึ่งหนึ่งของความยาวชุดโครโมโซม ดังนั้นจะรักษาสกิมตาได้น้อยกว่าการแลกเปลี่ยนยีนหนึ่งจุดหรือสองจุด ซึ่งถ้าพิจารณาโดยรวมแล้ว วิธีนี้จะสร้างชุดโครโมโซมลูกที่แตกต่างกับชุดโครโมโซมพ่อแม่ที่ใกล้เคียงที่สุด ในจำนวนประมาณเศษหนึ่งส่วนสี่ของบิตที่แตกต่างกัน

4.3.6.3 วิธีการนำชุดโครโมโซมใหม่แทนที่ชุดโครโมโซมเดิม

De Jong ได้นำเสนอวิธีการวาดัง สำหรับการแทนที่ชุดโครโมโซมเดิม โดยที่ชุดโครโมโซมใหม่ที่เกิดขึ้นจะนำไปแทนชุดโครโมโซมที่ใกล้เคียงกันในประชากรชุดเดิม ซึ่งวิธีของ Syswerda และ Whitley จะแทนที่ชุดโครโมโซมใหม่ลงไปในประชากรชุดเดิม ถ้าไม่มีชุดโครโมโซมที่เหมือนกับชุดโครโมโซมใหม่อยู่ในประชากร อย่างไรก็ตามวิธีนี้ยังคงมีผลต่อสกิมตาที่ถูกทำลายจากการแลกเปลี่ยนยีน ถึงแม้ว่าผลกระทบที่เกิดขึ้นจะไม่รุนแรงเท่ากันก็ตาม