

บทที่ 5

ทฤษฎีและบทวิเคราะห์วิธีนิชชิงเจเนติกอัลกอริทึม

เนื้อหาในบทนี้จะกล่าวถึงที่มาและความสำคัญของวิธีนิชชิงเจเนติกอัลกอริทึมที่ถูกพัฒนาขึ้นจากวิธีเจเนติกอัลกอริทึมเดิมเพื่อให้มีความสามารถในการรักษาพื้นที่คำตอบที่ดี รวมถึงทฤษฎีและสมการของวิธีนิชชิงเจเนติกอัลกอริทึมจากวิธีต่างๆ ที่ใช้อยู่ในปัจจุบัน

5.1 วิธีนิชชิงเจเนติกอัลกอริทึม

วิธีเจเนติกอัลกอริทึมพื้นฐานนั้น มีความเหมาะสมในการใช้ค้นหาคำตอบของฟังก์ชันที่มีจุดยอดเดียวในพื้นที่คำตอบทั้งหมด แต่ในปัญหาที่มีอยู่ในความเป็นจริงแล้ว ปัญหาส่วนใหญ่จะเป็นฟังก์ชันที่มีหลายจุดยอด ซึ่งวิธีเจเนติกอัลกอริทึมพื้นฐานไม่สามารถค้นหาคำตอบหลายค่าของปัญหาดังกล่าวได้ [1, 67] โดยต้องใช้กลไกการสร้างและการรักษากลุ่มประชากรย่อย ภายในพื้นที่คำตอบ เพื่อที่จะแยกแยะคำตอบ โกลบอลจากกลุ่มคำตอบ โลคอล และสามารถคำนวณหาคำตอบที่ดีที่สุดจากกลุ่มประชากรย่อยนั้นได้

วิธีนิชชิงเจเนติกอัลกอริทึมถูกพัฒนาขึ้นเพื่อลดผลกระทบของเจเนติกดริฟท์ที่เกิดจากขบวนการคัดเลือกในวิธีเจเนติกอัลกอริทึมพื้นฐาน วิธีนี้จะรักษาความหลากหลายทางพันธุศาสตร์ในประชากรและค้นหาจุดยอดอื่นๆที่มีอยู่ ในลักษณะการทำงานแบบขนาน ลดโอกาสที่คำตอบสุดท้ายที่ได้จะเป็นคำตอบโลคอล

วิธีนิชชิงเจเนติกอัลกอริทึมมีพื้นฐานมาจากการแบ่งทรัพยากรที่มีอยู่จำกัดภายในกลุ่มประชากรย่อย โดยที่ประชากรทั้งหมดจะถูกจัดแบ่งเป็นกลุ่มต่างๆ ตามลักษณะที่คล้ายคลึงกัน สมาชิกที่อยู่ในแต่ละกลุ่มจะแข่งขันกันเพื่อแย่งชิงทรัพยากรของกลุ่มนั้นๆ โดยที่ถ้ามีจำนวนกลุ่มประชากรย่อยมาก ก็จะทำให้ทรัพยากรถูกแบ่งไปสู่กลุ่มต่างๆ มากขึ้น และทรัพยากรในแต่ละส่วนก็จะมีจำนวนน้อยลง ซึ่งนำไปสู่สถานะคงตัวที่ซึ่งขนาดของประชากรย่อย จะเป็นสัดส่วนกับจำนวนของทรัพยากรรวมที่สามารถจัดหาได้

5.2 วิธีการใช้ฟังก์ชันแฮร์ริง

วิธีนิชชิงวิธีหนึ่งซึ่งนิยมใช้กันอย่างแพร่หลายคือ การใช้ฟังก์ชันแฮร์ริง วิธีนี้ที่มีการทดสอบแล้วว่ามีประสิทธิภาพในการหาคำตอบที่เหมาะสมสำหรับฟังก์ชันที่มีหลายจุดยอด ได้ [67] วิธีการใช้ฟังก์ชันแฮร์ริงจะทำให้แน่ใจได้ว่าประชากรจะกระจายกันอยู่ตลอดในพื้นที่คำตอบ และรักษาพื้นที่คำตอบที่ดีไว้ได้

วิธีการใช้ฟังก์ชันแฮร์ริงทำการแบ่งค่าความเหมาะสมจากชุดโครโมโซมต่างๆ ในกลุ่มชุดโครโมโซมที่คล้ายกัน โดยการทำให้ค่าความเหมาะสมนั้นลดลงตามสัดส่วนของความคล้ายคลึงกันกับชุดโครโมโซมอื่น ดังสมการต่อไปนี้

$$f'_i = \frac{f_i}{m_i} \quad (5.1)$$

$$f'_i = \frac{f_i}{\sum_{j=1}^n sh(d_{ij})} \quad (5.2)$$

เทอมของส่วนในสมการอาจเรียกได้ว่า จำนวนนิช m_i ซึ่งมีค่าเท่ากับผลรวมของฟังก์ชันการแฮร์ $sh()$ ระหว่างชุดโครโมโซมและทุกชุดโครโมโซมอื่นในประชากร ฟังก์ชันการแฮร์คือฟังก์ชันของระยะทาง d ระหว่างชุดโครโมโซม ซึ่งจะให้ค่าเป็น 1 ถ้าชุดโครโมโซมเหมือนกัน และจะมีค่า 0 ถ้ามีค่าน้อยกว่าค่าเรสโลต์ของความไม่เหมือนกันที่กำหนดไว้ σ_{share} ดังสมการต่อไปนี้

$$sh(d_{ij}) = \begin{cases} 1 - \left(\frac{d_{ij}}{\sigma_{share}} \right)^\alpha; & \text{if } d < \sigma_{share} \\ 0; & \text{otherwise} \end{cases} \quad (5.3)$$

โดยที่ d_{ij} คือความแตกต่างสมบูรณ์ (ในลักษณะค่าเจโนไทป์หรือฟีโนไทป์) ระหว่างชุดโครโมโซม i และ j ความเหมือนกันในลักษณะเจโนไทป์จะเกี่ยวข้องกับการแสดงค่าในบิตสตริงและใช้ระยะทางแฮมมิงเป็นตัววัด ส่วนความเหมือนกันในลักษณะฟีโนไทป์จะเกี่ยวข้องกับพารามิเตอร์จริงในพื้นที่คำตอบโดยตรงและใช้ระยะทางยูคลิเดียนเป็นตัววัด โดยที่การวัดระยะทางจากลักษณะฟีโนไทป์จะให้ผลลัพธ์ที่ดีกว่าเพียงเล็กน้อย α คือค่าคงที่ที่ใช้ในการปรับรูปร่างของฟังก์ชันแฮร์ริง ซึ่ง α จะนิยามกำหนดให้มีค่าเท่ากับ 1 เพื่อที่ผลจากฟังก์ชันที่ได้จะมีลักษณะคล้ายกับฟังก์ชันสามเหลี่ยม

วิธีการแฮร์ริงส่วนมากจะใช้ร่วมกับวิธีการคัดเลือกส่วนที่เหลืออยู่แบบสุ่ม และวิธีการคัดเลือกทั้งหมดแบบสุ่มโดยที่ถ้าใช้ร่วมกับวิธีการคัดเลือกแบบทัวนาเมนต์จะต้องคำนึงถึงปัญหาของเสถียรภาพ [54]

ข้อเสียของวิธีการใช้ฟังก์ชันแฮร์ริงคือเวลาที่ใช้ในการคำนวณทั้งนี้เนื่องจากการวัดค่าระยะทางดังกล่าวจะต้องเทียบกับชุดประชากรที่มีอยู่ทั้งหมดในประชากร ทำให้ถ้าขนาดของจำนวนประชากรมีค่ามากทำให้เวลาที่ใช้มีค่าสูงขึ้นอย่างมาก $\approx T(n^2)$ อีกทั้งวิธีนี้ไม่สามารถแยกแยะความแตกต่างระหว่างคำตอบโลกอลที่มีค่าใกล้เคียงกับคำตอบโลกอลอื่นมากกว่าค่ารัศมีนิช

5.3 วิธีคร่าว์ดิง

De Jong (ค.ศ. 1975) ได้เสนออัลกอริทึมที่เรียกว่า คราว์ดิงแฟกเตอร์ ที่มีลักษณะคล้ายกับวิธีเจเนติกอัลกอริทึมสถานะคงตัวจาก Syswerda (ค.ศ. 1989) เนื่องจากส่วนของประชากรจะสร้างขึ้นใหม่และล้มตายลงในแต่ละรุ่น โดยที่สมาชิกของประชากรใหม่ที่สร้างขึ้นจะนำไปแทนสมาชิกของประชากรรุ่นเก่าที่มีลักษณะคล้ายกันมากที่สุด

วิธีคร่าว์ดิงเริ่มจากคำนวณค่าเงินเนอเรนเบิร์ก (GG) โดยการคัดเลือกชุดโครโมโซมตามค่าความเหมาะสม และผ่านขบวนการแลกเปลี่ยนยีนและการผ่าเหล่ายีน โดยภายหลังจากขบวนการแลกเปลี่ยนยีนและการผ่าเหล่ายีน, ชุดโครโมโซมจำนวน $GG \times n$ จากประชากรรุ่นเดิมจะถูกเลือกขึ้นมาเพื่อทำลายทิ้งโดยการเพิ่มชุดโครโมโซมที่เกิดขึ้นใหม่แทนที่ ชุดโครโมโซมใหม่จะค้นหาชุดโครโมโซมที่จะทำให้ลายจากค่าตัวอย่างแบบสุ่มจากแฟกเตอร์คร่าว์ดิง (CF) และแทนที่ชุดโครโมโซมที่คล้ายคลึงกันที่สุดจากกลุ่มตัวอย่างนั้น ความคล้ายคลึงดังกล่าวจะกำหนดโดยใช้การเทียบบิตในพื้นที่แฮมมิง หรือการวัดระยะทางในลักษณะเจโนไทป์ และโดยทั่วไป CF จะมีค่าประมาณ 2 หรือ 3 และ GG มีค่าประมาณ 10%

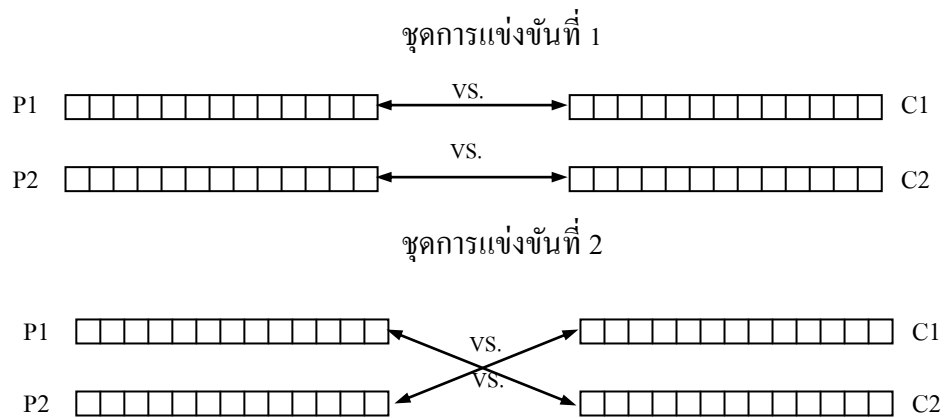
วิธีคร่าว์ดิงไม่ได้เป็นวิธีที่ซึ่งลักษณะประชากรจะมีเสถียรภาพในสปีชีชที่รวมกันอยู่ แต่จะช่วยรักษาความหลากหลายที่มีอยู่ก่อนนั้น เป้าหมายของ De Jong ในการออกแบบคร่าว์ดิงคือการรักษาความหลากหลายของแอลลีเพื่อที่จะป้องกันสภาวะการลู่เข้าที่เร็วเกินไป วิธีคร่าว์ดิงนี้ประสบผลสำเร็จมากในการรักษาความหลากหลายของแอลลี แต่จะมีความผิดพลาดแบบสุ่มเกิดขึ้นจากการกำหนดค่า CF ที่ต่ำ และจากแฟกเตอร์อื่นที่เป็นผลจากอัลกอริทึมที่เปลี่ยนแปลงอย่างช้าๆ ไปสู่การกำหนดตำแหน่งแอลลีอย่างถาวร

5.4 วิธีดิเทอร์มินิสติกคร่าว์ดิง

Mahfoud (ค.ศ. 1995) ได้ปรับปรุงการทำงานของวิธีคร่าว์ดิงพื้นฐานโดยนำเสนอแนวคิดการแข่งขันระหว่างพ่อแม่และลูกที่เกิดขึ้น ในกลุ่มสภาวะแวดล้อมเดียวกัน โดยที่เมื่อชุดโครโมโซมลูกผ่านขบวนการแลกเปลี่ยนยีนและการผ่าเหล่ายีน ชุดโครโมโซมลูกแต่ละชุดจะนำมาแทนที่ชุดโครโมโซมพ่อแม่ที่มีลักษณะใกล้เคียงกันมากที่สุด ถ้าชุดโครโมโซมลูกมีความเหมาะสมสูงกว่า วิธีดิเทอร์มินิสติกคร่าว์ดิงจะพิจารณา (ในลักษณะที่คล้ายกับวิธีทัวนาเมนต์) การเปรียบเทียบระหว่างชุดโครโมโซมดังกล่าว โดยวิธีการวัดความคล้ายคลึงกันระหว่างชุดโครโมโซมระยะทางตามค่าฟีโนไทป์

ในการวัดค่าระยะห่างจากค่าฟีโนไทป์ของชุดโครโมโซม ระหว่างเวกเตอร์ $x = (x_1, x_2, \dots, x_n)^T$ และ $x' = (x'_1, x'_2, \dots, x'_n)^T$ จะใช้ระยะห่างยูคลิเดียน ดังสมการต่อไปนี้

$$d(x, x') = [(x_1 - x'_1)^2 + (x_2 - x'_2)^2 + \dots + (x_n - x'_n)^2]^{1/2} \quad (5.4)$$



รูปที่ 5.1 ชุดการแข่งขันที่เกิดขึ้นจากวิธีดิวอร์มินนิสติกคราว์ดิง

(REPEAT for g generation)

DO $m/2$ times:

1. Select 2 parents, p_1 and p_2 , randomly, no replacement
2. Cross them, yielding c_1 and c_2
3. Apply mutation / other operators, yield c'_1 and c'_2
4. If $[d(p_1, c'_1) + d(p_2, c'_2)] \leq [d(p_1, c'_2) + d(p_2, c'_1)]$
 - ☐ If $f(c'_1) > f(p_1)$ then replace p_1 with c'_1
 - ☐ If $f(c'_2) > f(p_2)$ then replace p_2 with c'_2

Else

- ☐ If $f(c'_2) > f(p_1)$ then replace p_1 with c'_2
- ☐ If $f(c'_1) > f(p_2)$ then replace p_2 with c'_1

รูปที่ 5.2 การทำงานของวิธีดิวอร์มินนิสติกคราว์ดิง