

บทที่ 4

ผลการทดลอง

จากการดำเนินการทดลองด้วยค่าพารามิเตอร์ที่ได้ ได้ผลการทดลองดังนี้

ตารางที่ 7

ตารางเปรียบเทียบระหว่าง MSA-PSO Clustering กับ PSO-Clustering และ EPSO-Clustering ของชุดข้อมูล Ruspini (ค่าที่ดีที่สุดในการทดลองซ้ำ 20 ครั้ง)

อัลกอริทึม	จำนวนของ กลุ่มข้อมูล	กลุ่มที่	จำนวน ข้อมูล	ระยะห่าง จากจุด กึ่งกลาง	ระยะห่างเฉลี่ย ของกลุ่มข้อมูล ภายใน
PSO-Clustering	4	1	20	12.5347	11.4961
		2	23	10.3747	
		3	17	14.0998	
		4	15	8.9750	
EPSO-Clustering	4	1	20	12.5293	11.4948
		2	23	10.3792	
		3	17	14.0955	
		4	15	8.9751	
MSA-PSO Clustering	4	1	20	12.5667	11.4905
		2	23	10.4902	
		3	17	13.9487	
		4	15	9.0372	

ตารางที่ 8

ตารางเปรียบเทียบระหว่าง MSA-PSO Clustering กับ PSO-Clustering และ EPSO-Clustering
ของชุดข้อมูล Ruspini ชนิดไม่คงที่ (ค่าที่ดีที่สุดในการทดลองซ้ำ 20 ครั้ง)

อัลกอริทึม	จำนวนของ กลุ่มข้อมูล	กลุ่มที่	จำนวน ข้อมูล	ระยะห่าง จากจุด กึ่งกลาง	ระยะห่างเฉลี่ย ของกลุ่มข้อมูล ภายใน
PSO-Clustering	4	1	20	12.5248	11.4959
		2	23	10.6075	
		3	17	13.8844	
		4	15	8.9668	
EPSO-Clustering	4	1	20	12.5376	11.4896
		2	23	10.3702	
		3	17	14.0745	
		4	15	8.9760	
MSA-PSO Clustering	4	1	20	12.5884	11.4657
		2	23	10.3483	
		3	17	14.0269	
		4	15	8.9293	

ตารางที่ 9

ตารางเปรียบเทียบค่าเฉลี่ยเลขคณิต และค่าเบี่ยงเบนมาตรฐานระหว่าง MSA-PSO Clustering กับ PSO-Clustering และ EPSO-Clustering ของชุดข้อมูล Ruspini ทั้งสองชนิดจากการทดลองซ้ำ 20 ครั้ง

อัลกอริทึม	ข้อมูลชนิดคงที่		ข้อมูลชนิดไม่คงที่	
	μ	σ	μ	σ
PSO-Clustering	11.8153	0.5334	12.1294	1.1673
EPSO-Clustering	11.5234	0.0182	11.5271	0.0224
MSA-PSO Clustering	11.5103	0.0153	11.5101	0.0269

ตารางที่ 10

ตารางเปรียบเทียบผลรวมระยะห่างเฉลี่ยของกลุ่มข้อมูล ระหว่าง MSA-PSO Clustering กับ PSO-Clustering และ EPSO-Clustering (ชุดข้อมูล Ruspini ทั้งสองชนิด ทดลองซ้ำ 20 ครั้ง สำหรับชุดข้อมูลอื่น ๆ ทดลองซ้ำ 10 ครั้ง)

ชุดข้อมูล	ค่าที่สนใจ	PSO	EPSO	MSA-PSO
Ruspini คงที่	ผลรวมระยะห่าง	889.34	867.34	864.65
	ค่าเบี่ยงเบนฯ	39.44	1.65	1.03
	ค่าที่ดีที่สุด	863.63	863.54	863.43
Ruspini ไม่คงที่	ผลรวมระยะห่าง	930.60	867.60	864.68
	ค่าเบี่ยงเบนฯ	163.26	1.99	1.94
	ค่าที่ดีที่สุด	863.78	863.17	861.58
Cancer	ผลรวมระยะห่าง	3,381.72	3,252.24	3,074.98
	ค่าเบี่ยงเบนฯ	275.00	244.57	9.66
	ค่าที่ดีที่สุด	3,125.24	3,082.45	3,066.88

ตารางที่ 10 (ต่อ)

ชุดข้อมูล	ค่าที่สนใจ	PSO	EPSO	MSA-PSO
Iris	ผลรวมระยะห่าง	102.19	100.27	97.39
	ค่าเบี่ยงเบนฯ	13.33	2.64	0.11
	ค่าที่ดีที่สุด	97.57	97.84	97.27
Vowel	ผลรวมระยะห่าง	163,662.16	163,883.14	151,766.48
	ค่าเบี่ยงเบนฯ	11,309.34	806.58	1,128.31
	ค่าที่ดีที่สุด	151,299.81	163,488.75	150,043.00
Wine	ผลรวมระยะห่าง	17,076.20	17,768.54	16,939.51
	ค่าเบี่ยงเบนฯ	211.96	119.93	9.4382
	ค่าที่ดีที่สุด	16,947.43	17,652.81	16,924.51

ตารางที่ 11

ตารางเปรียบเทียบเปอร์เซ็นต์ผิดพลาดของการจัดกลุ่มข้อมูล ระหว่าง MSA-PSO Clustering กับ PSO-Clustering และ EPSO-Clustering (ชุดข้อมูล Ruspini ทั้งสองชนิด ทดลองซ้ำ 20 ครั้ง สำหรับชุดข้อมูลอื่น ๆ ทดลองซ้ำ 10 ครั้ง)

ชุดข้อมูล	%	PSO	EPSO	MSA-PSO
Ruspini คงที่	ค่าเฉลี่ยเลขคณิต	0	0	0
	ค่าเบี่ยงเบนฯ	0	0	0
	ค่าผิดพลาดน้อยที่สุด	0	0	0
Ruspini ไม่คงที่	ค่าเฉลี่ยเลขคณิต	0	0	0
	ค่าเบี่ยงเบนฯ	0	0	0
	ค่าผิดพลาดน้อยที่สุด	0	0	0
Cancer	ค่าเฉลี่ยเลขคณิต	42.17	53.63	53.02
	ค่าเบี่ยงเบนฯ	0.90	1.39	0.05
	ค่าผิดพลาดน้อยที่สุด	50.95	51.98	53.00

ตารางที่ 11 (ต่อ)

ชุดข้อมูล	%	PSO	EPSO	MSA-PSO
Iris	ค่าเฉลี่ยเลขคณิต	10.87	10.33	10.87
	ค่าเบี่ยงเบนฯ	0.83	1.41	0.32
	ค่าผิดพลาดน้อยที่สุด	10.00	8.00	10.67
Vowel	ค่าเฉลี่ยเลขคณิต	47.00	41.03	56.25
	ค่าเบี่ยงเบนฯ	6.26	2.47	11.98
	ค่าผิดพลาดน้อยที่สุด	42.94	39.84	41.33
Wine	ค่าเฉลี่ยเลขคณิต	29.78	29.78	29.78
	ค่าเบี่ยงเบนฯ	0	0	0
	ค่าผิดพลาดน้อยที่สุด	29.78	29.78	29.78

จากผลการทดลองแสดงว่าอัลกอริทึม MSA-PSO Clustering ที่บทความนี้นำเสนอ สามารถนำมาประยุกต์กับการจัดกลุ่มข้อมูลได้ ได้ค่าเฉลี่ยของกลุ่มข้อมูลภายในน้อยที่สุดจาก ตารางที่ 7 และตารางที่ 8 และยังมีค่าเบี่ยงเบนมาตรฐานน้อยมากเมื่อเทียบกับ PSO-Clustering และ EPSO-Clustering แสดงว่า MSA-PSO Clustering มีประสิทธิภาพในการจัดกลุ่มข้อมูล Ruspini ดีที่สุดในอัลกอริทึมที่นำมาเปรียบเทียบ เมื่อนำไปทดสอบกับข้อมูลชนิดอื่น ทั้งสามอัลกอริทึมไม่สามารถจัดกลุ่มข้อมูลได้ถูกต้อง แต่ถ้าจะพิจารณา เฉพาะความสามารถในการจัดกลุ่มด้วยการวัดระยะห่างระหว่างข้อมูลกับจุดกึ่งกลางที่ได้จะเห็นว่า MSA-PSO Clustering ก็ได้ผลรวมระยะห่างเฉลี่ยที่ดีที่สุดกว่าทั้งสองอัลกอริทึม และมีค่าเบี่ยงเบนมาตรฐานที่มีนัยสำคัญดีมาก ดังแสดงในตารางที่ 10 การทำซ้ำหลาย ๆ รอบเพื่อเปรียบเทียบผลทำให้ทราบได้ว่า MSA-PSO Clustering มีความเสถียรมากกว่า PSO-Clustering และ EPSO-Clustering จากข้อมูลใน ตารางที่ 7,8 และตารางที่ 9 ถึงแม้ผลที่ได้จะได้ค่าที่แตกต่างกันไม่มาก แต่ถ้าพิจารณาถึงการนำไปใช้จริงเมื่อข้อมูลจริงมีอยู่เป็นจำนวนมากและต้องการนำมาจัดกลุ่มเพื่อทำเหมืองข้อมูล ค่าของจุดกึ่งกลางของข้อมูลที่เปลี่ยนแปลงยิ่งได้ระยะห่างเฉลี่ยของกลุ่มข้อมูลภายในน้อยขึ้นเท่าใด ก็เท่ากับว่าได้ตัวแทนของกลุ่มที่มีความถูกต้องมากยิ่งขึ้น รวมทั้งการเปลี่ยนแปลงค่าของข้อมูล ถึงแม้จะเป็นเพียงเล็กน้อย ก็อาจส่งผลทำให้จุดกึ่งกลางที่ได้มีการเปลี่ยนแปลง แต่สำหรับ PSO นั้น การเปลี่ยนแปลงเพียงเล็กน้อยอัลกอริทึมไม่สนใจ อาจทำให้ค่าของจุดกึ่งกลางที่ได้ยังคงเป็นค่าเดิม

อยู่ ด้วยโครงสร้างที่ไม่คงที่ของกลุ่มอนุภาคของ MSA-PSO จะทำให้การเปลี่ยนแปลงค่าของข้อมูลใด ๆ ไม่ถูกละเลยเนื่องจากทำให้กลุ่มเคลื่อนที่มีพื้นที่ในการค้นหาของตนมากขึ้น และรวมทั้งการเพิ่มกระบวนการเร่งด้วยการสุ่ม กำหนดจุดกึ่งกลางของกลุ่มข้อมูลที่ถูกจัดแบ่งแล้วด้วยกลุ่มเคลื่อนที่ย่อย กลุ่มอนุภาค ทั้งสองวิธีการในอัลกอริทึมมีผลกับการเปลี่ยนแปลงของผลลัพธ์ที่ได้จากอัลกอริทึมทั้งสิ้น แต่เมื่ออัลกอริทึมนี้ซึ่งอาศัยการประเมินค่า fitness ด้วยระยะทาง Euclidian เพื่อหาระยะห่างของข้อมูลกับจุดศูนย์กลางแต่เพียงอย่างเดียว นั้นไม่เหมาะสมเมื่อนำมาใช้กับชุดข้อมูลจริง กล่าวคือสามารถจัดแบ่งกลุ่มข้อมูลได้ แต่ไม่สามารถกำหนดสมาชิกในแต่ละกลุ่มได้ถูกต้องเลย จึงเป็นงานที่น่าสนใจในอนาคตที่สามารถพัฒนาต่อยอดให้กับอัลกอริทึมนี้ได้อีก