

Shuffled Frog-Leaping Algorithm

Y su aplicación al Problema del Agrupamiento con Restricciones

Alejandro Alonso Membrilla

Introducción

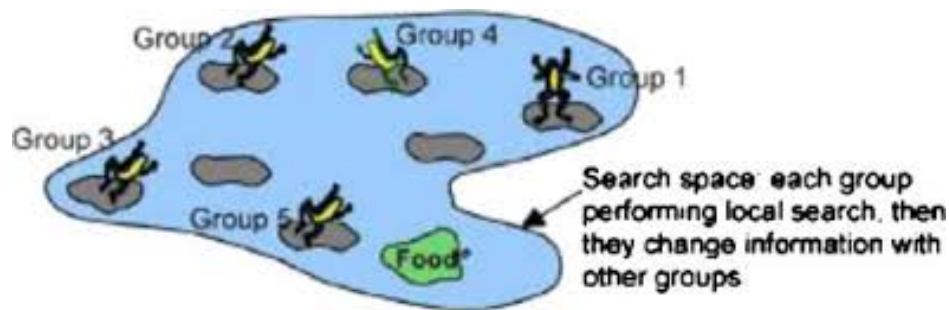
El SFLA (Shuffled Frog-Leaping Algorithm) se trata de un algoritmo memético que agrupa una población de distintas soluciones, denominadas **ranas**.



Este algoritmo combina elementos de Shuffled Complex Evolutionary Algorithm a nivel global con búsquedas locales basadas en el PSO.

Inspiración

El SFLA se inspira en el comportamiento de las ranas saltando entre las piedras de una charca en busca de comida.



Las ranas se asocian en grupos, de forma que los integrantes de cada grupo colaboran entre sí compartiendo información con sus compañeras cercanas.

La información compartida por las ranas se conoce con el nombre de **memes**, y el grupo de ranas que comparten información se denomina **memplex**.

Algoritmo

Como ya se ha mencionado, el SFLA combina una fase de distribución de las ranas a nivel global con una búsqueda local PSO, que se repite un cierto número de veces.



Los memeplexes se forman ordenando todas las ranas según lo buena que sea su posición, y se distribuyen alternadamente entre los memeplexes:

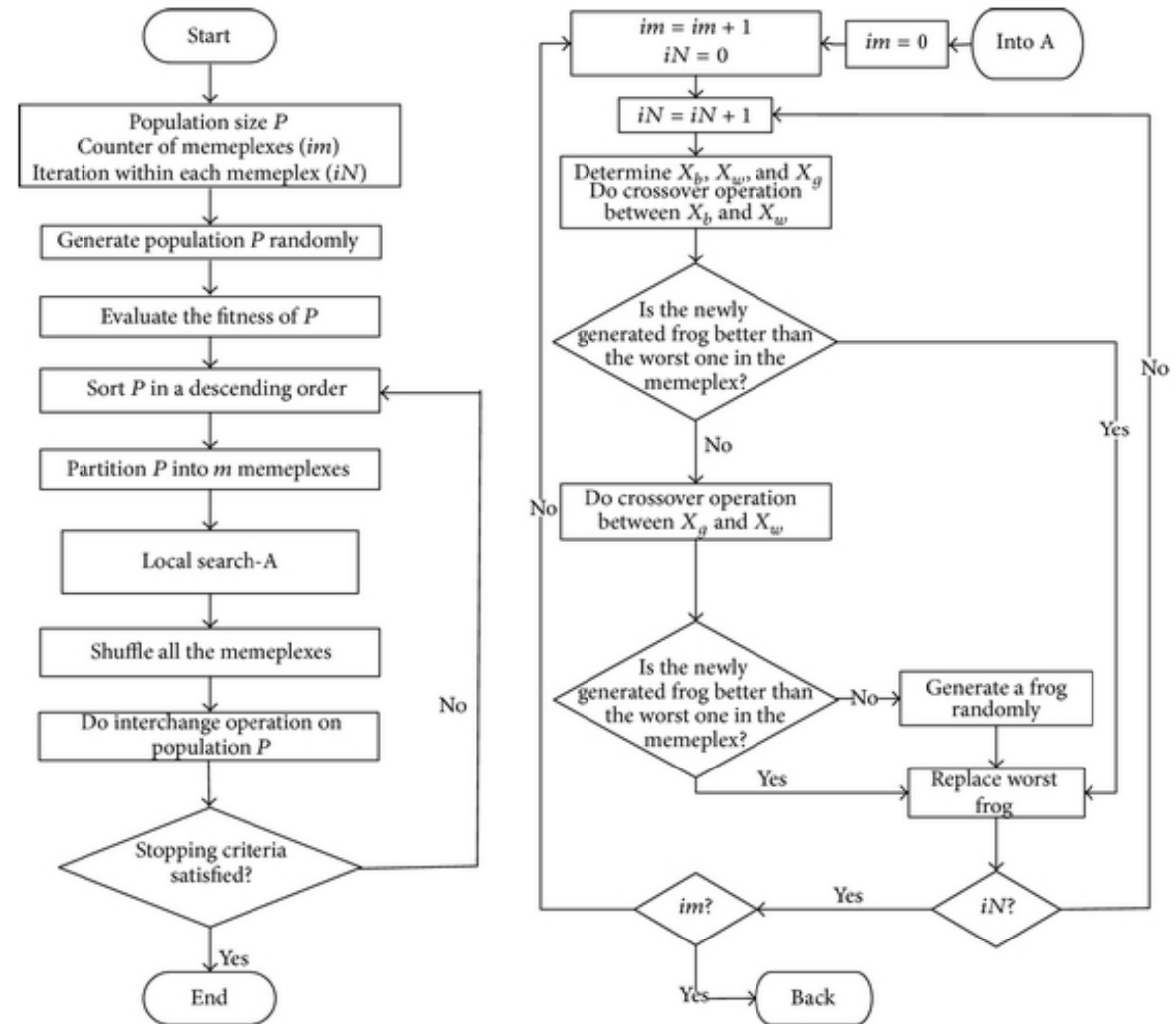
La mejor rana va al primer memeplex, la segunda al segundo... Cuando se acaben los memeplexes se vuelve a empezar.

Algoritmo

En cada memplex, tomamos un subconjunto de las ranas (un submemplex), y la rana del memplex en peor posición saltará en dirección a la rana en mejor posición.

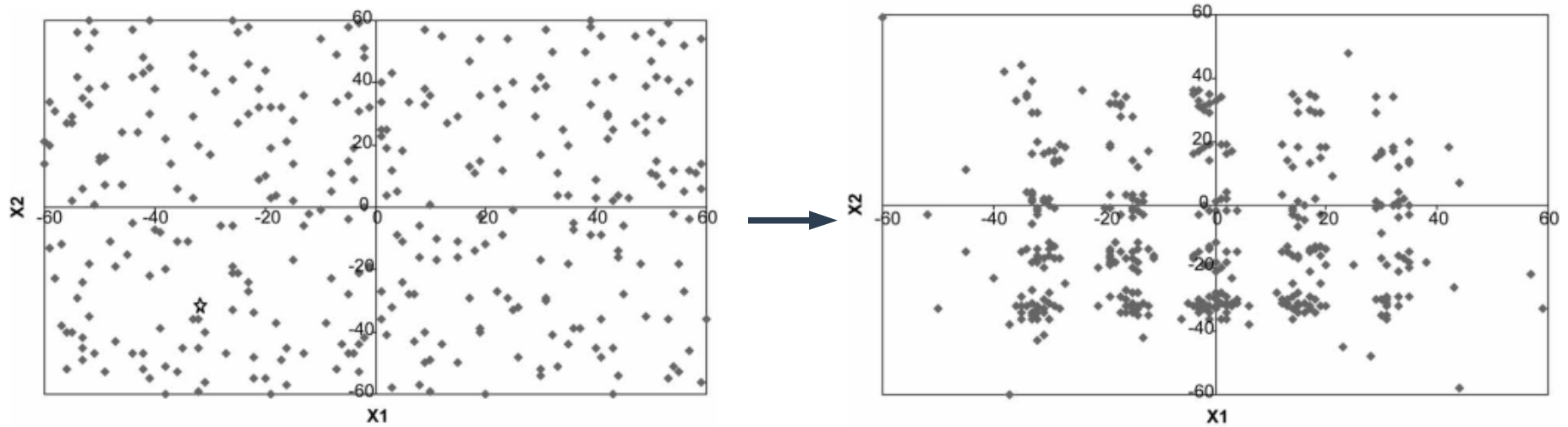
Este proceso se repite cierto número de veces, tras lo cual todas las ranas se redistribuyen en la charca. Si se desea continuar el algoritmo, las ranas vuelven a agruparse en memplexes y se repite el proceso.

La figura de la derecha representa el SFLA paso por paso.

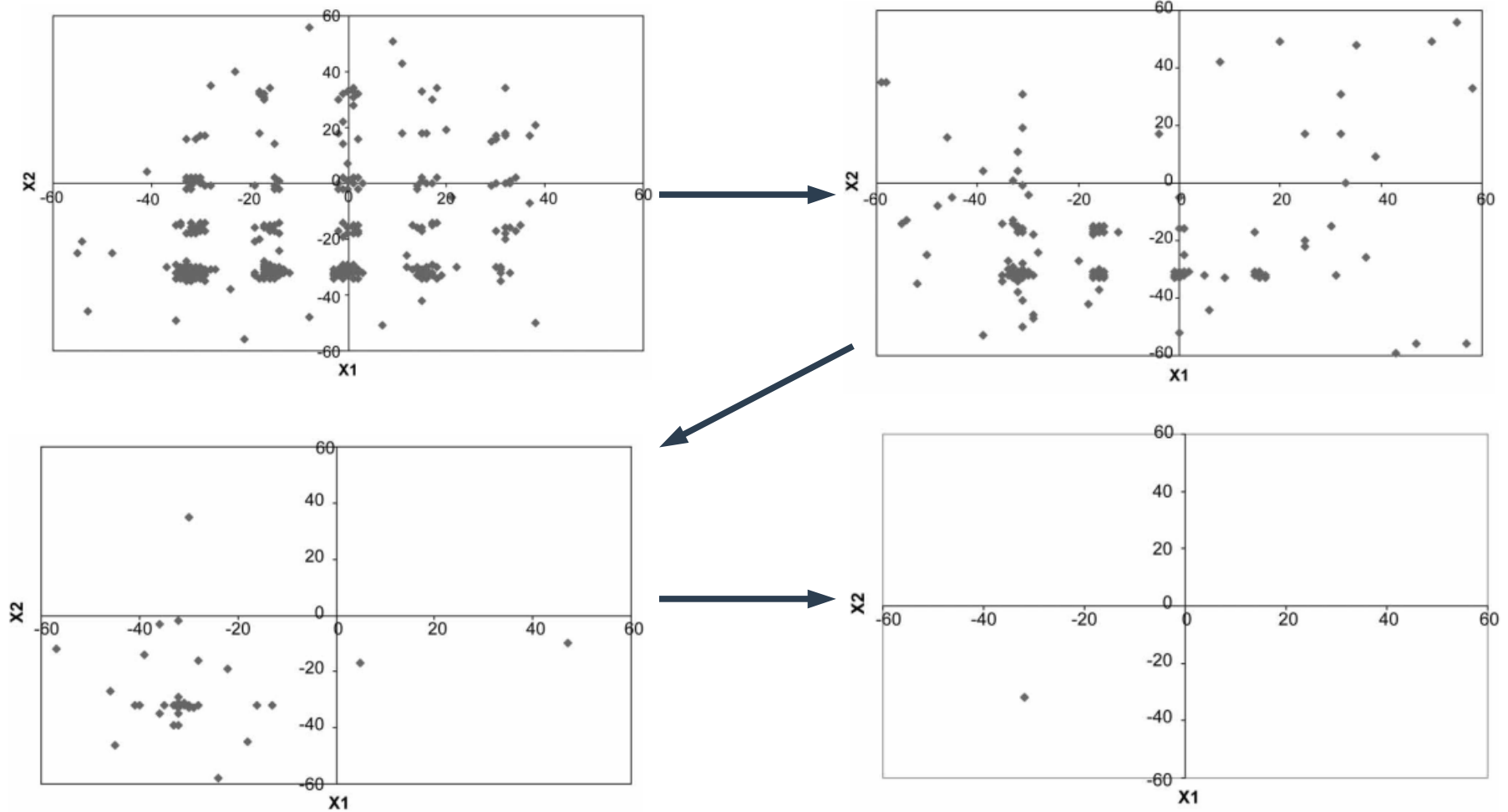


Ejemplo

A continuación mostramos un ejemplo de la evolución de la población de ranas en una aplicación del SFLA sacada del artículo *Shuffled frog-leaping algorithm: a memetic meta-heuristic for discrete optimization* (M Eusuff, K Lansey, F Pasha; 2006).



Ejemplo



Mejoras

1. Esquema de reinicialización: cada cierto número de ejecuciones de la función objetivo, reiniciamos la mayor parte de la población de la charca. Conservamos una cierta cantidad de las mejores ranas de la charca actual.

2. Generación de pseudocentroides aleatorios: cambiamos el esquema de inicialización aleatoria por otro en el que generamos una serie de puntos de forma aleatoria en el espacio factible, y asociamos cada elemento del dataset a su punto más cercano. Favorece generar soluciones con elementos cercanos dentro de cada cluster.

Para esta práctica hemos implementado el SFLA con el esquema de reinicialización y lo hemos probado para el método clásico de generación de soluciones (*SFLAv2-1*) y el método de pseudocentroides aleatorios (*SFLAv2-2*).

Mejoras

Además, hemos implementado una versión concurrente del *SFLAv2-2* (*SFLAv2-2parallel*) que aprovecha la evolución independiente de los distintos memplexes para dividir el trabajo entre los núcleos de la computadora, reduciendo los tiempos de ejecución.

Resultados

Media de los resultados obtenidos por cada algoritmo en el PAR con 10% de restricciones

	Iris				Ecoli			
	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr.</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr.</i>	<i>T</i>
COPKM	0.6640	12.8000	0.7480	0.3360	36.9980	222.2000	42.6600	13.6620
BL	0.7087	48.0000	1.0525	0.0468	31.9513	1059.6000	60.3805	4.0406
BMB	0.6693	0.0000	0.6693	0.4610	30.9789	1091.2000	60.2559	4.1538
ES	0.6693	0.0000	0.6693	0.1796	21.7699	78.6000	23.8788	2.6034
ILS_BL	0.6693	0.0000	0.6693	0.2120	36.8628	1128.0000	67.1272	3.8272
ILS_ES	0.6693	0.0000	0.6693	0.2088	39.2603	1187.6000	71.1238	1.0324
ILS_ESv2	0.7518	69.8000	1.2518	0.1198	27.1459	380.2000	37.3467	3.2306
AGE-CU	1.8289	433.2000	4.9645	395.4912	45.1246	1728.6000	91.8093	1266.7649
AGE-CSF	0.8961	179.4000	2.1947	395.7278	43.1012	1558.0000	85.1784	1254.1372
AGG-CU	1.3802	319.0000	3.6892	356.2720	42.9482	1454.6000	82.2329	1125.3195
AGG-CSF	0.9792	187.8000	2.3386	360.6781	42.1086	1429.0000	80.7019	1144.2363
AM-(10,1.0)	1.4801	322.4000	3.8137	376.1226	44.6770	1653.0000	89.3200	1189.8078
AM-(10,0.1)	0.9950	189.6000	2.3674	378.3358	43.3957	1490.8000	83.6581	1165.9391
AM-(10,0.1MEJ)	1.1810	229.4000	2.8415	380.5959	43.1913	1496.4000	83.6049	1177.8385
SFLAv1	0.7144	53.2000	1.0954	2.0030	25.5042	706.0000	44.4463	6.3310
SFLAv2-1	0.7072	56.0000	1.1083	2.3432	26.2064	762.6000	46.6670	6.6238
SFLAv2-2	0.6694	0.0000	0.6780	2.3332	23.6877	606.2000	39.9521	6.7384
SFLAv2-2parallel	0.6706	6.4000	0.7165	1.7822	24.2322	527.8000	38.3931	3.9402

Resultados

COPKM	Newthyroid				Rand			
	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr.</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr.</i>	<i>T</i>
	x	x	x	x	0.7500	0.0000	0.7500	0.3240
BL	12.5763	43.2000	14.1728	0.4404	0.8471	47.4000	1.2290	0.0526
BMB	11.9972	134.4000	16.3226	1.4482	0.7156	0.0000	0.7156	0.2892
ES	12.6504	42.8000	14.2321	0.3530	0.7156	0.0000	0.7156	0.1504
ILS_BL	12.6504	42.8000	14.2321	0.8454	0.7156	0.0000	0.7156	0.1324
ILS_ES	13.8421	86.4000	17.0351	0.4906	0.7156	0.0000	0.7156	0.1792
ILS_ESv2	12.4152	90.4000	15.7561	0.6350	0.7978	29.0000	1.0314	0.1184
AGE-CU	13.4154	1038.8000	51.8573	633.7665	2.6531	426.6000	6.1073	390.3711
AGE-CSF	15.4570	560.4000	36.1952	646.3932	1.3098	165.2000	2.6474	397.6981
AGG-CU	14.2847	683.0000	39.5599	576.6832	1.9655	281.0000	4.2407	356.2747
AGG-CSF	14.6835	544.8000	34.8444	586.1826	1.4330	192.8000	2.9941	362.8217
AM-(10,1.0)	13.9925	803.8000	43.7379	600.9357	2.0743	310.0000	4.5843	373.4074
AM-(10,0.1)	14.0753	704.2000	40.1350	607.7462	1.5731	207.4000	3.2524	378.5930
AM-(10,0.1MEJ)	14.9616	583.0000	36.5361	626.2960	1.5597	205.2000	3.2212	391.4706
SFLAv1	13.5723	177.6000	20.1358	3.4536	0.8809	64.8000	1.4030	1.8668
SFLAv2-1	14.2574	66.4000	16.7113	3.7604	0.8809	50.4000	1.2869	2.2476
SFLAv2-2	13.1012	64.4000	15.4811	3.8030	0.7156	0.0000	0.7156	2.2486
SFLAv2-2parallel	13.6424	70.2000	16.2367	2.4590	0.7156	0.0000	0.7156	1.7666

Resultados

Media de los resultados obtenidos por cada algoritmo en el PAR con 20% de restricciones

	Iris				Ecoli			
	<i>Tasa C</i>	<i>Tasa inf</i>	<i>Agr.</i>	<i>T</i>	<i>Tasa C</i>	<i>Tasa inf</i>	<i>Agr.</i>	<i>T</i>
COPKM	0.6620	2.0000	0.6680	0.3080	36.1200	181.4000	38.4980	6.5308
BL	0.7048	100.8000	1.0656	0.0928	31.9257	2041.0000	59.3058	8.4040
BMB	0.6693	0.0000	0.6693	0.7616	30.7502	2183.0000	60.0353	8.9296
ES	0.6693	0.0000	0.6693	0.3212	21.7674	168.6000	24.0292	5.7378
ILS_BL	0.6693	0.0000	0.6693	0.3770	36.5980	2256.4000	66.8678	7.9742
ILS_ES	0.6693	0.0000	0.6693	0.3422	37.1773	2622.2000	72.3543	1.8396
ILS_ESv2	0.7345	0.0000	1.0810	0.2480	27.2157	677.8000	36.3084	6.6278
AGE-CU	1.7605	892.4000	4.9888	519.6885	45.1190	3499.0000	92.3681	1884.1881
AGE-CSF	0.9441	376.4000	2.3058	524.5151	42.2472	3169.0000	85.0401	1882.9124
AGG-CU	1.1444	542.6000	3.1072	438.9558	41.7953	3066.0000	83.1973	1583.8081
AGG-CSF	0.9326	368.4000	2.2653	431.5933	40.2405	2872.2000	79.0256	1545.9742
AM-(10,1.0)	1.4224	658.0000	3.8027	521.6618	44.6353	3382.0000	90.3046	1833.8258
AM-(10,0.1)	1.0489	426.0000	2.5900	493.9535	42.1419	3074.8000	83.6628	1762.7859
AM-(10,0.1MEJ)	1.0107	407.8000	2.4859	468.9191	41.4123	3090.8000	83.1493	1642.9583
SFLAv1	0.7044	122.2000	1.1418	3.0698	25.0300	1602.0000	46.5210	10.4258
SFLAv2-1	0.7059	105.2000	1.0824	3.3634	26.3709	1616.0000	48.0496	10.7348
SFLAv2-2	0.6684	3.4000	0.6806	3.4590	24.8063	1077.0000	39.2543	11.0674
SFLAv2-2parallel	0.6693	0.0000	0.6693	2.3984	24.2903	1123.8000	39.3662	6.3348

Resultados

COPKM	Newthyroid				Rand		
	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr.</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr.</i>
	x	x	x	x	0.7500	0.0000	0.7500
BL	12.8428	91.8000	14.5387	0.8972	0.8414	98.8000	1.2392
BMB	12.0509	249.2000	16.6547	3.3726	0.7156	0.0000	0.7156
ES	13.5906	52.8000	14.5660	0.8560	0.7156	0.0000	0.7156
ILS_BL	13.5906	52.8000	14.5660	1.9664	0.7156	0.0000	0.7156
ILS_ES	13.5048	169.2000	16.6306	1.0406	0.7156	0.0000	0.7156
ILS_ESv2	14.1890	56.6000	15.2346	1.3328	0.7741	52.6000	0.9859
AGE-CU	13.9083	1981.0000	50.5548	899.9752	2.6088	890.4000	6.2119
AGE-CSF	15.4180	1090.6000	35.5930	906.5044	1.2264	316.4000	2.5068
AGG-CU	15.3512	1140.4000	36.4475	759.7317	1.8450	570.4000	4.1532
AGG-CSF	15.3210	892.2000	31.8259	740.3376	1.3435	387.0000	2.9095
AM-(10,1.0)	15.0514	1379.0000	40.5615	830.9825	1.8289	569.4000	4.1330
AM-(10,0.1)	15.5643	918.8000	32.5612	860.2795	1.3035	362.0000	2.7684
AM-(10,0.1MEJ)	15.1769	1047.2000	34.5491	820.9502	1.2961	338.2000	2.6646
SFLAv1	13.5087	348.2000	19.9414	5.5304	0.8624	120.0000	1.3456
SFLAv2-1	13.9083	160.0000	16.8642	5.8110	0.8723	123.2000	1.3684
SFLAv2-2	13.9373	110.2000	15.9732	5.8654	0.7156	0.0000	0.7156
SFLAv2-2parallel	13.8190	117.8000	15.9953	3.7320	0.7156	0.0000	0.7156

Conclusiones

- Hemos obtenido resultados relativamente buenos en comparación con los algoritmos de prácticas anteriores, especialmente para el dataset Ecoli, y mucho mejores que los de los algoritmos genéticos estudiados hasta ahora.
- La versión mejorada del SFLA ha demostrado obtener mejores soluciones que la original.
- La implementación paralela del SFLA ha obtenido ganancias satisfactorias, además de tiempos más bajos que los de la BL y muchos de sus derivados.
- Su abundancia de parámetros y su estructura paralela lo hacen una metaheurística prometedora para afrontar todo tipo de problemas de optimización combinatoria de forma eficiente.