Measuring the performance of FCM versus PSO for fuzzy clustering problems

Hamid Reza Jafari[a*], Amir Reza Soltani[b] and Mohammad Reza Soltani[c]

[a]*PhD student in Automation, Department of Industrial Engineering, Science and Research Branch, Islamic Azad University, Ponak, Tehran, Iran*
[b]*Masters of Science student, South Tehran Branch, Islamic Azad University, Tehran, Iran*
[c]*Department of Industrial Engineering, Abhar Branch, Islamic Azad University, Abhar, Iran*

C H R O N I C L E	A B S T R A C T
	Clustering cellular manufacturing plays an important role in many industrial engineering problems. This paper investigates the performance of two methods of heuristic and metaheuristics fuzzy clustering. The proposed method investigates heuristic well-known FCM and particle swarm optimization (PSO) on some well-known benchmarks. We use two criteria of J(P) as well as Xie-Beni to compare the results. Three parameters of PSO method is tuned using design of experiment and then the results of PSO are compared versus FCM method in terms of two mentioned criteria. The proposed models are run for each instance 10 different times and, using ANOVA test, the means of two methods are compared. While the results of ANOVA do not indicate any meaningful difference between PSO and FCM in terms of J(P), we have found some meaningful differences between PSO and FCM in terms of Xie-Beni criterion. In other words, PSO performs better than FCM in terms of Xie-Beni.

1. Introduction

Clustering cellular manufacturing plays an important role in many industrial engineering problems (Singh, 1993; Venugopal, 1999; Yin & Yasuda, 2006; Ghosh et al., 2010; Feizollahi et al., 2012). According to Ballakur and Harold (1987), the initial stage in the facilities design of cellular manufacturing systems incorporates the identification of part families and machine groups and forming cells possessing specific manufacturing capabilities. Clustering has been widely used in different fields of engineering and science, for instance, Alizadeh et al. (2011) used clustering along with ANFIS modeling for stock price prediction. JabalAmeli and Mortezaei (2011) introduced capacitated facility location/network design problem with two separate objective functions in forms of multi-objective with limited capacity and solved their problem using a new hybrid algorithm where there were two stages. In the first stage, locations of facilities and design of fundamental network were determined and, in the second stage, demands were allocated to the facilities.

Corresponding author. Tel: +989192425395
E-mail: hr.ariyayee2009@gmail.com (H. R. Jafari)

Selim et al. (1998) performed an investigation on a fundamental issue in cellular manufacturing—cell formation, which influences the fundamental structure and the overall layout of a cellular manufacturing system. They first provided a mathematical model of the cell formation problem and then proposed a methodology-based classification of prior research, used in reviewing the most recent literature on the cell formation problem. Molleman et al. (2002) explained the evolution of a cellular manufacturing system in a medium-sized company over a 13-year period. They analyzed the arguments, which gave rise to the nearly continuous readjustment of the design of the cellular manufacturing system of this firm and the direction in which these adjustments happened. They argued that market developments, new manufacturing technology and modern production control systems would probably constrain the application area of cellular manufacturing.

Gindy et al. (1995) presented a method for improving the practical applicability of the fuzzy clustering technique for family formation in cellular manufacturing environments. The developed grouping methodology was experimentally illustrated using an industrial case study and several well- known component grouping instances from the published literature. Güngör and Arıkan (2000) used an application of fuzzy decision making in part-machine grouping and used fuzzy set theory (FST) to set out the cell layout. Josien and Liao (2000) presented an integrated implementation of fuzzy c-means and fuzzy KNN for GT part family and machine cell formation. Al-Ahmari (2002) presented a fuzzy analysis approach for part-machine grouping in cellular manufacturing systems.

2. The proposed study

Fuzzy clustering problem in an empirical study for fuzzy cell formation (FCF) was investigated by Li et al. (2007). Suppose, there are n parts and m machines and the purpose of cell formation is to cluster parts into C families of parts and assign machines to appropriate machines. The results of clustering can be specified in a matrix $U=[\mu_{ik}]$, $i=1, ..., c$ and $k=1, ..., n$ where μ_{ik} is the membership of part k in group i and the following constraints must hold,

$$0 \leq \mu_{ik} \leq 1 \tag{1}$$

$$\sum_{i=1}^{c} \mu_{ik} = 1 \tag{2}$$

$$0 \leq \sum_{k=1}^{n} \mu_{ik} \leq n \tag{3}$$

In addition, the objective is to minimize the following,

$$J(p) = \sum_{k=1}^{n} \sum_{i=1}^{c} \left[\mu_{ik} \right]^{m} \left\| x_k - V_i \right\|^{2}, \tag{4}$$

where $m>1$ is a real number and controls membership, V_i is the center of the set of family part i, x_k is vector of part k. The necessary condition to minimize $J(p)$ is as follows,

$$V_i = \frac{\sum_{k=1}^{n} [\mu_{ij}]^{m} x_k}{\sum_{k=1}^{n} [\mu_{ij}]^{m}}, \tag{5}$$

$$\mu_{ik}^{(t+1)} = \left[\sum_{j=1}^{c} \left(\frac{\left\| x_k - V_i^{(t)} \right\|^2}{\left\| x_k - V_j^{(t)} \right\|^2} \right)^{\frac{1}{m-1}} \right]^{-1}, \tag{6}$$

where $\left\| x_k - V_i \right\|^2$ is the Euclidian distance between x_k and v_i and $\mu_{ik(t+1)}$ is the membership of part k in group i. Next, we present of the most popular technique for FCM, where C, m and ε are the number of clusters, a real number and termination criteria, respectively (Bezdek, 1981).

Algorithm 1: Fuzzy FCM

Step 1: *Set t=0, choose P(o),*

Step 2: Calculate center of clusters $V_1(t)$, ..., $V_c(t)$ based on Eq. (5) and choose a real value for m,

Step 3: Update μ_i *(t+1)* based on Eq. (6) and update $p(t+1)$,

Step 4: Compare $p(t)$ with $p(t+1)$, if $\left| p^{(t+1)} - p^{(t)} \right| \le \varepsilon$ stop, otherwise increase t by one unit and go to step 2.

Xu and Wang (1989) first introduced this method and it was modified in other studies. Chu and Hayya (1991) considered part family based on membership function and setup FCM and then setup cell-machine problem based on the centers of FCM. Al-Ahmari (2002), in another assignment, considered a fuzzy analysis approach for part-machine grouping in cellular manufacturing systems. Yang et al. (2006) considered mixed-variable fuzzy clustering approach to part family and machine cell formation for GT applications. Chen and Ye (2006) proposed an adaptive hyper-fuzzy partition particle swarm optimization clustering algorithm. Li et al (2007) presented an improved fuzzy clustering method for cellular manufacturing.

3. Particle swarm optimization

Particle swarm optimization (PSO) is one of the most popular metaheuristics, which is designed based on social behaviors of birds. There are literally different versions of PSO and Andres and Lozano (2006) presented a PSO algorithm for part–machine grouping, which is described here. If i^{th} member of a vector D is denoted as $x_i=(x_{i1}, x_{i2}, ..., x_{iD})$ and P_g is the best member of group $p_i=(\rho_{i1}, \rho_{i2}, ..., \rho_{iD})$ and change in velocity for i^{th} member is $Vel_i =(Vel_{i1}, Vel_{i2}, ..., Vel_{iD})$. Therefore, we have,

$$Vel_{id}(t+1) = \chi\{wVel_{id}(t) + c_1\phi_1[\rho_{id}(t) - x_{id}(t)] + c_2\phi_2[(\rho_{gd}(t) - x_{id}(t)]\}, \tag{7}$$

$$x_{id}(t+1) = x_{id}(t) + Vel_{id}(t+1), \tag{8}$$

where $d=1,2,...,D$, $i=1,2,...,n$ and n is the size of swarm, w is inertia weight, c_1 and c_2 are positive numbers, ϕ_1 and ϕ_2 are random numbers between zero and one and finally χ is a control coefficient, which is equal to one for large-scale problems.

Algorithm 2: PSO method for clustering

Step 1. Set appropriate values for c_1, c_2, w, χ, $t=0$ and consider the number of clusters C and consider a termination criteria for ε and the number of particles n,

Step 2. Generate n initial values, randomly,
Step 3. Calculate $\mu_{ik}^{(t)}$ based on Eq. (5) and update $p^{(t+1)}$,
Step 4. Calculate the fitness value based on Eq. (4),
Step 5. Update local and best solutions,
Step 6. Update $Vel_l^{(t)}$ and $V_l^{(t)}$ based Eq. (7) and Eq. (8),
Step 7. Update $p^{(t+1)}$ using Step 3,
Step 8. If $\left| p^{(t+1)} - p^{(t)} \right| \le \varepsilon$ Stop, otherwise, increase t by one unit and go to step 4.

In this paper, we use different criteria for measuring the quality of clustering. The first criterion is as follows,

$$j_m(u,v) = \sum_i \sum_k (u_{ik})^m d^2(x_k, v_i),$$ (9)

where Eq. (9) measures Euclidian distance and, obviously, smaller values represent better solutions. The other criterion uses Xie-Beni method, which uses *j(p)* as follows,

$$Xie-Beni = \frac{\sum_{ii=1}^{c} \sum_{k=1}^{n} \left[(u_{ik})^m \|x_k - v_i\|^2 \right]}{n \left(\min_{i,j} \|v_i - v_j\|^2 \right)} = \frac{J_m(U,V:X)}{nSep(V)}.$$ (10)

4. The results

In this section, we present details of the implementation of our proposed method on 11 benchmark problems. We report the mean of results after 10 runs for every instance. Table 1 summarizes the results of J(P) as follows,

Table 1
The summary of testing the proposed method based on FCM and Xie-Beni for different instances

Row	Problem	Dimensions	Number of clusters	J(P) FCM	J(P) Xie-Beni
1	Glass	185×10	6	160.07	0.8382
2	Teaching	151×6	3	10615.83	0.3032
3	Iris	150×4	3	60.57	0.1371
4	Zoo	101×17	7	86.29	0.4536
5	Medium	99×2	5	655.25	1.7210
6	Ikaco	31×18	3	37.47	0.7843
7	Small 1	15×2	3	17.73	0.2145
8	Small 2	15×2	3	27.13	0.1423
9	Small 3	5×2	2	5.56	0.0783
10	Small 4	5×2	2	3.25	0.0641
11	Random	500×200	6	4158.40	9.98×10^{12}

We also examine the performance of the proposed PSO on the same instances but we first need to tune all parameters using design of experiment by examining 3^k plans. We consider three values of 1, 1.5 and 2 for c_1 and c_2, respectively. We also consider three values of 0.5, 07 and 0.9 for w. The parameters are tuned for 27 runs on an instance, IKACO, by minimizing J(P) as a criterion. Next, we have performed least square technique to fit a quadratic function and the results are as follows,

$$J(p) = -53.81c_1 - 52.22c_2 - 80.90w + 15.78c_1^2 + 15.71c_2^2 + 54.44w^2 + 3.57c_1c_2 + 217.209$$ (11)

In order to measure the quality of the final solution, we verify normal probability plot as well as variance of changes shown in Fig. 1 and Fig. 2.
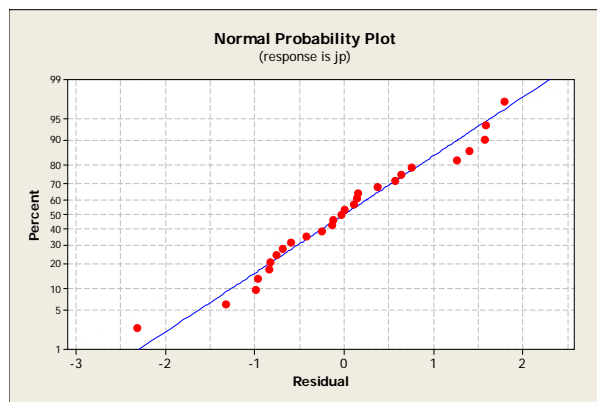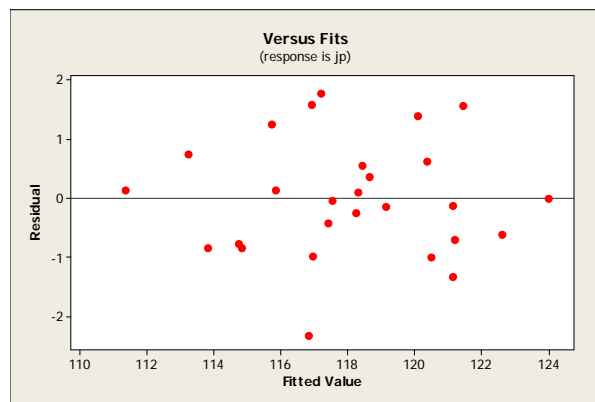

Fig. 1. Normal probability plot

Fig. 2. Versus Fits

As we can observe from the figures, it seems that the results are good enough for the implementation of PSO method. We have optimized Eq. (11) by considering some limitations on variables and the local optimal solution is $c_1 = 1.537221$, $c_1 = 1.487389$ and $w = 0.7429652$. Table 2 shows the performance of PSO for the same instances.

Table 2
The summary of testing the proposed method (PSO) for different instances

Row	Problem	Dimensions	# of clusters	J(P) FCM	J(P) Xie-Beni
1	Glass	185×10	6	161.05	0.0026
2	Teaching	151×6	3	10665.47	0.00052
3	Iris	150×4	3	60.57	0.00053
4	Zoo	101×17	7	87.25	0.0034
5	Medium	99×2	5	658.76	0.0041
6	Ikaco	31×18	3	37.54	0.0054
7	Small 1	15×2	3	17.73	0.1843
8	Small 2	15×2	3	27.13	0.0924
9	Small 3	5×2	2	5.56	0.0328
10	Small 4	5×2	2	3.25	0.0429
11	Random	500×200	6	4013.63	2870.35

Note that the results of Table 2 are the mean of 10 different runs. We have performed ANOVA to verify whether statistically there is any differences between *J(P)* results obtained from FCM and PSO. Table 3 shows ANOVA results,

Table 3
The results of ANOVA test

Source of change	Sum of Squares	df	Mean of Squares	F	P-Value
Regression	152	1	152	0.000	0.197
Residual	200027516	18	11112640		
Total	200027668	19	-		

Based on the results of Table 3, we can conclude that there is no meaningful difference between two methods. We have also compared the results of the proposed FCM and PSO in terms of Xie-Beni criterion and Table 4 shows details of our findings,

Table 4
The results of ANOVA test

Source of change	Sum of Squares	df	Mean of Squares	F	P-Value
Regression	0.954	1	0.954	6.98	0.017
Residual	2.461	18	0.137		
Total	3.415	19	-		

The results of Table 4 indicate that there is a meaningful difference between FCM and PSO in terms of Xie-Beni criterion. In other words, PSO method seems to perform better than FCM.

5. Conclusion

In this paper, we have presented an empirical study to compare the performance of FCM versus PSO. The proposed model of this paper implemented two criteria as an objective functions and using some benchmarks compared their performances. The study used design of experiments to tune the parameters and built some optimal design frameworks for the implementation of PSO. The results of our study indicated that while the results of ANOVA did not indicate any meaningful difference between PSO and FCM in terms of J(P), we have found some meaningful differences between PSO and FCM in terms of Xie-Beni criterion. In other words, PSO performed better than FCM in terms of Xie-Beni.

References

Al-Ahmari, A.M.A.(2002). A fuzzy analysis approach for part-machine grouping in cellular manufacturing systems. *Integrated Manufacturing System*, 13(7), 489-497.

Alizadeh, M., Gharakhani, M., Fotoohi, E., & Rada, R. (2011). Design and analysis of experiments in ANFIS modeling for stock price prediction. *International Journal of Industrial Engineering Computations*, *2*(2), 409-418.

Al-Ahmari, A. M. A. (2002). A fuzzy analysis approach for part-machine grouping in cellular manufacturing systems. *Integrated Manufacturing Systems*,*13*(7), 489-497.

Andres, A., & Lozano, S. (2006). A particle swarm optimization algorithm for part–machine grouping. *Robotics and Computer-Integrated Manufacturing*, 22, 468-474.

Ballakur, A., & Harold, J. S. (1987). A within-cell utilization based heuristic for designing cellular manufacturing systems. *International Journal of Production Research*, *25*(5), 639-665.

Bezdek, J.C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press: New York.

Chen, C.Y., & Ye, F. (2006). Adaptive hyper-fuzzy partition particle swarm optimization clustering algorithm. *Cybernetics and Systems*, 37, 463-479.

Chu, C.H., & Hayya, J.C. (1991). A fuzzy clustering approach to manufacturing cell formation. *International Journal of Production Research*, 29, 1475-1487.

Feizollahi, S., Shirmohammadi, A., & Kahreh, Z. S. (2012). Investigation the requirements of supply and distribution emergency logistics management and categorization its sub-criteria using AHP: a case study. *Management Science Letters*, 2, 2335–2340.

Gindy, N. N. Z., Ratchev, T. M., & Case, K. (1995). Component grouping for GT applications—a fuzzy clustering approach with validity measure. *International Journal of Production Research*, *33*(9), 2493-2509.

Ghosh, T., Sengupta, S., Chattopadhyay, M., & Dan, P. K. (2010). Meta-heuristics in cellular manufacturing: A state-of-the-art review. *International Journal of Industrial Engineering Computations*, 2, 87-122.

Güngör, Z., & Arıkan, F. (2000). Application of fuzzy decision making in part-machine grouping. *International Journal of Production Economics*, *63*(2), 181-193.

JabalAmeli, M. S., & Mortezaei, M. (2011). A hybrid model for multi-objective capacitated facility location network design problem. *International Journal of Industrial Engineering Computations,* 2, 509-524.

Josien, K., & Liao, T. W. (2000). Integrated use of fuzzy c-means and fuzzy KNN for GT part family and machine cell formation. *International Journal of Production Research*, *38*(15), 3513-3536.

Li, J., Chu, C.H., Wang, Y., & Yan, W. (2007). An improved fuzzy clustering method for cellular manufacturing. *International Journal of Production Research*, 45, 1049-1062.

Molleman, E., Slomp, J., & Rolefes, S. (2002). The evolution of a cellular manufacturing system–a longitudinal case study. *International Journal of Production Economics*, *75*(3), 305-322.

Selim, H. M., Askin, R. G., & Vakharia, A. J. (1998). Cell formation in group technology: review, evaluation and directions for future research. *Computers & Industrial Engineering*, *34*(1), 3-20.

Singh, N. (1993). Design of cellular manufacturing systems: an invited review. *European Journal of Operational Research*, *69*(3), 284-291.

Venugopal, V. (1999). Soft-computing-based approaches to the group technology problem: a state-of-the-art review. *International journal of production research*, *37*(14), 3335-3357.

Xu, H., & Wang, H.P. (1989). Part family formation for GT applications based on fuzzy mathematics. *International Journal of Production Research*, 27, 1637-1651.

Yang, M.S., Hung, W.L., & Cheng, F.C. (2006). Mixed-variable fuzzy clustering approach to part family and machine cell formation for GT applications. *International Journal of Production Economics*, 103, 185-198.

Yin, Y., & Yasuda, K. (2006). Similarity coefficient methods applied to the cell formation problem: A taxonomy and review. *International Journal of Production Economics*, *101*(2), 329-352.