Development and Tuning of Bacteria Foraging Optimization Algorithm on Cell Formation in Cellular Manufacturing System

H. Nouri
Islamic Azad University, Sanandaj Branch, Sanandaj, Iran
Email: ro_eagle@yahoo.com

Tang S. H, M. K. A. Ariffin, B.T. Hang Tuah, and R. Samin
Universiti Putra Malaysia, Malaysia
Emails: {saihong, tuah, khairol, zali}@eng.upm.edu.my

***Abstract*—Ever since Kevin M. Passino invented the bacteria foraging optimization algorithm, one of the main challenges has been employment of the algorithm to problem areas other than those for which the algorithm was proposed. This research work inquires the applications of designed experiments aided by multiple regression analysis for tuning of this emerging novel optimization algorithm parameters to the cell formation (CF) problem considering operation sequence. In this paper, an attempt is made to tuning chemo tactic and swarming steps parameters meanwhile taking into consideration bacteria foraging optimization algorithm convergence speed and performance. The factorial designed experiment is suggested to create treatments of experiment. The adequacy of the proposed model is analyzed based on some commonly statistical criteria. The results lie in favor of adequacy of the proposed model.**

***Index Terms*—Cell formation, Bacteria foraging optimization algorithm, Designed experiments, Cellular manufacturing system.**

I. Introduction

In this paper, an attempt has been made to tackle the problem with operation sequence of the parts using bacteria foraging optimization algorithm**.** Systematic choice of bacteria foraging parameters speeds up the convergence of the algorithm without falling prey of premature convergence. By making the bacteria foraging coefficient adaptive, the convergence speed is improved .To choose the parameters of swarming, designed experiment methods like multiple regression methodology is carried out. The model was tested using a wide variety of statistical criteria like assessing utility of model, making inferences about the model's parameters, the coefficient of correlation and assessing model adequacy. The model found to be consistent in producing good results. The major purpose of this work is to develop a simple, yet efficient, methodology capable of producing quick solutions for estimation and prediction.

Manuscript received July 2, 2012; revised December 17, 2012.

The Bacterial Foraging Optimization (BFO) is invented by Passino [3] is swarmming evolutionary computational approach. It is inspired by the foraging behavior of Escherichia *coli* bacteria in human intestines. According to this approach foraging is considered as an optimization process whereby the bacterium strives to maximize the energy gained per unit foraging time.

II. Proposed Bacteria Foraging Optimization: A Brief Overview

Actually BFO is invented firstly for continuous domain and movement of bacteria on domain is continuously, but in contrary, in our problem we should move bacteria position representation by MCIM (Machine Cell Incidence Matrix) matrix for find better solution in term of machine-cell membership. Thus, we have matrix which row member represent machine-cell membership and column mean cell number, 1' one means machine is belonged to cell and 0' zero means otherwise. For example in MCIM initial machines assignment represented by Fig. 1 is [1122332], that means machines 1,2 to cell 1, machines 3,4,7 to cell 2 and machines 5,6 to cell 3 are assigned respectively,. Therefore for tumbling and swimming of this bacteria position representation, we use global and local mutation, that means in global mutation we change position of column with other and in local mutation we change position of row with other based on probability conditions. Consequence matrix after some tumbling and swimming is shown in Fig. 2, [3312132], that means, just machines 3, 5 to cell 1, machines 4, 7 to cell 2 and machines 1, 2, 6 to cell 3 are assigned respectively. Therefore for determine of probability of global and local mutation in tumbling and swimming we define $P_{Gm-Tumbling}$ for probability of global mutation in tumbling (1-$P_{Gm-Tumbling}$, probability of local mutation in tumbling) and $P_{Gm-Swimming}$ for probability of global mutation in swimming (1- $P_{Gm-Swimming}$, probability of local mutation in swimming).

Here, we will also have cell-to-cell signaling via an attractant and will represent that with $J_{cc}{}^{i}$ *(θ,* θ^{i}*(j, k, l)), i*

doi: 10.12720/joace.1.1.36-39

=1, 2, *S*, for the *i*-th bacterium [3]. Let $d_{attract}$ be the depth of the attractant released by the the cell and $w_{attract}$ be a measure of the width of the attractant signal. The cell also repels a nearby cell in the sense that it consumes nearby nutrients and it is not physically possible to have two cells at the same location. To model this, we let $h_{repellant} = d_{attract}$ be the height of the repellant effect and $w_{repellant}$ be a measure of the width of the repellant. Let

$$J_{cc}(\theta, P(j,k,l)) = \sum_{i=1}^{S} J_{cc}^{i}\left(\theta, \theta^{i}(j,k,l)\right) = \sum_{i=1}^{S}[-d_{attract}\exp(-w_{attract}\sum_{m=1}^{p}(\theta_m - \theta_m^i)^2)] + \sum_{i=1}^{S}[h_{repellant}\exp\left(-w_{repellant}\sum_{m=1}^{p}(\theta_m - \theta_m^i)^2\right) \quad (1)$$

	C1	C2	C3	
M1	1			$\sum M = 1$
M2	1			$\sum M = 1$
M3		1		$\sum M = 1$
M4		1		$\sum M = 1$
M5			1	$\sum M = 1$
M6			1	$\sum M = 1$
M7		1		$\sum M = 1$
	$\sum M \geq 2$	$\sum M \geq 2$	$\sum M \geq 2$	

Figure 1. Initial machines assignment [1122332]

	C1	C2	C3	
M1			1	$\sum M = 1$
M2			1	$\sum M = 1$
M3	1			$\sum M = 1$
M4		1		$\sum M = 1$
M5	1			$\sum M = 1$
M6			1	$\sum M = 1$
M7		1		$\sum M = 1$
	$\sum M \geq 2$	$\sum M \geq 2$	$\sum M \geq 2$	

Figure 2. Consequent Machines Assignment [3312132]

To apply the BFA algorithm for the cell formation problem, we chose cell membership matrix $X = [x_{ik}]$ instead of the position of bacteria, θ(*i*,*j*,*k*,*l*), and accordingly performance measure like group technology efficiency instead of nutrition function *J* (*i*, *j*,*k*,*l*). Group technology efficiency (GTE) given by Harhalakis et al. [1] is formulated in following equation (4). Additionally, we adopted Norm (X_m - X_{im}) instead of (θ_m - θ_{im}) for calculation of Jcc (θ, P (j, k, and l)).

$$I_p = \sum_{j=1}^{N}(n-1) \quad (2)$$

$$I_r = \sum_{j=1}^{N}\sum_{w=1}^{n-1} t_{njw} \quad (3)$$

$$GTE = \frac{I_p - I_r}{I_p} \quad (4)$$

where

t_{njw} = 0 if the operations w; w + 1 are performed in the same cell, = 1 otherwise

I_p = Maximum number of inter-cell travels possible in the system

I_r = Number of inter-cell travels actually required by the system

n = Number of operations (*w* = 1, 2, 3, *n*)

N = Number of parts

GTE = Group technology efficiency

III. The Proposed Designed Experiments Model

The proposed ov-erall model in this research work is:

$$E(y) = \sum_{m=0}^{npr}\sum_{n=0}^{npr}\sum_{p=0}^{npr}\sum_{i=1}^{npr-2}\sum_{j=i+1}^{npr-1}\sum_{k=j+1}^{npr}\beta_{mnp}x_i^{\,m}x_j^{\,n}x_k^{\,p} \quad (5)$$

where npr =number of predictors

For instance, if two predictors are considered the model is obtained:

$$E(y) = \beta_{00} + \beta_{01}x2 + \beta_{02}x2^2 + \beta_{10}x1 + \beta_{11}x1x2 + \beta_{12}x1x2^2 + \beta_{20}x1^2 + \beta_{21}x1^2x2 + \beta_{22}x1^2x2^2$$

Regression analysis because of the prediction equation found by it contributes other information not provided by traditional analysis of variance in this research work is proposed.

Test of an individual parameter coefficient in the multiple regression models is [2]:

$$|t| > t_{\alpha/2} \quad (6)$$

$$t = \frac{\widehat{\beta_i}}{s_{\widehat{\beta_i}}} \quad (7)$$

and $t_{\alpha/2}$ is based on [*n*-(*k*+1)] DF

where:

n=number of observations

k=number of independent variables in model

Conducting t tests on each β parameter in a model with a large number of terms is not a good way to determine whether a model is contributing information for the prediction of *y*. If we were to conduct a series of t tests to determine whether the independent variables are contributing to the predictive relationship, it is very likely that we would make one or more errors in deciding which terms to retain in the model and which to exclude. We begin with the easier problem-finding a measure of how well a multiple regression models fit a set of data. For this we use the multiple coefficient of determination [2]:

$$R_a^2 = 1 - \frac{(n-1)}{n-(k+1)}\left(\frac{SSE}{SS_{yy}}\right) \quad (8)$$

where $SSE = (y_i - \widehat{y_i})^2$, $SS_{yy} = (y_i - \overline{y})^2$ and $\widehat{y_i}$ is the predicted value of y_i for the multiple regression model.

The fact that R_a^2 is a sample statistic implies that it can be used to make inferences about the statistical utility of the model for predicting y values for specific settings of the independent variables. In particular, for the work measurement data, the test [2]:

$$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$$

H_a: At least one of the parameters β_k is nonzero

Would formally test the utility of the overall model. The test statistic used to test this null hypothesis is:

$$F = \frac{SS(model)/k}{SSE/[n-(k+1)]} \quad (9)$$

Rejection rejoins: F>F_α where $v_1 = k\ and\ v_2 = [n-(k+1)]$
and F test for comparing nested models is [2]:

$$F = \frac{(SSE_R - SSE_C)/\#of\ \beta's\ tested\ in\ H_0}{MSE_C} \quad (10)$$

where
SSE_R = Sum of squared errors for the reduced model
SSE_C = Sum of squared errors for the complete model
MSE_C = Mean square error (s^2) for the complete model
Rejection rejoins: F> F_α where $v_1 = \#of\ \beta' stested\ in\ H_0\ and\ v_2 = [n-(k+1)]$

IV. Results and Discussion

In this study, an efficient algorithm based on bacteria foraging algorithm is proposed for cell formation problem. It takes into consideration the operational sequence of the parts. The algorithm was coded in PC MATLAB and tested on the Pentium IV machine.

The model developed using BFA has been tested with many benchmark problems of various sizes ranging from 20×20 to 40×25 drawn from the open literature. Problems have been tested by varying the number of cells from 4 to 8. The factorial designed experiments are proposed for defining of treatments. Experiment 1 is prepared for determining of probability of global mutation in tumbling and swimming. Experiment 2 is prepared for determining cell to cell attraction parameters .Switch key is defined for contributing on/off of cell to cell attraction function on BFA algorithm execution, therefore in Experiment 1 this switch key is turned off for omitting of cell to cell attraction function contribution on effects of global or local mutation in tumbling and swimming.

For any treatment 100 replication of observation for any of data set is considered. Data is entered into a computer and a MINITAB statistical software package is used to estimate the unknown parameters in the deterministic component of hypothesized models. For Experiment-1 after many stepwise regressions based on nested models and comparing reduced model with completed model by applying of equation (14) following model is obtained:

$$y = \overline{GTE_1} = 64.7 - 6.29x_1 + 2.32x_2 - 4.43x_1^2 \quad (11)$$

This function is maximized in x_1=-0.711 and x_2=+1. The encoded amounts of these results are:

$$P_{Gm-Tumbling}=0.216$$
$$P_{Gm-swimming}=.9$$

According Table I, the value R_a^2 for this model is $R_a^2 = 97.4$ .This value of R_a^2 implies that 97.4% of sample variation in GTE is attributable to, or explained by,one or more of independent variables x_1 and x_2 .Therefore R^2 and R_a^2 are samples statistic that represent adequacy of the overall model is minimum 97.4%, and so we could arrive at the same decision by checking the observed significance level (p-value) of F test, given as 0.000.This value indicates that we will reject H_0 according equation(9) for any α greater than 0.000.The MINITAB printout shown in Table I also gives the two-tiled observed significance level(i.e., p-value) for each t test. These values that we would reject H_0 for any of $\beta_k = 0$ in favor of H_a for any of $\beta_k \neq 0$ at any α larger than 0.000,0.003 and 0.002 for β's of x_1 , x_2 and ${x_1}^2$ respectively.This small p-values leave little doubt that x_1, x_2 contributes information for the prediction of GTE. The value of $s = \sqrt{MSE} = 1.02$ implies that the most of observed GTE values will fall within approximately $2s$ =2.04 of their respective predicted values.

For Experiment-2 after many stepwise regressions based on nested models and comparing reduced model with completed model by applying of equation (10), finally following model is obtained:

$$y = GTE_2 = 65.9 - 2.96x_1 + 3.54x_2 - 2.53x_2^2 + 2.62x_2x_3 - 2.58x_2x_3^2 + 1.12x_1x_3^3 + 2.61x_1x_2x_3 \quad (12)$$

This function is maximized in x_1=-1, x_2=0.7 and x_3=0.0.The encoded amount of this results are:
$d_{attract} = .01$
$w_{attract} = 8.5$
$w_{repellant} = 5$
$h_{repellant} = d_{attract} = .01$

TABLE I. Factorial Experiments Results Print Out of Experiment 1

Regression Analysis: y versus x1, x2, x1^2

The regression equation is
y = 64.7 - 6.29 x1 + 2.32 x2 - 4.43 x1^2

Predictor	Coef	SE Coef	T	P
Constant	64.6633	0.5877	110.04	0.000
x1	-6.2950	0.4155	-15.15	0.000
x2	2.3233	0.4155	5.59	0.003
x1^2	-4.4250	0.7197	-6.15	0.002

S = 1.01785 R-Sq = 98.4% R-Sq(adj) = 97.4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	309.31	103.10	99.52	0.000
Residual Error	5	5.18	1.04		
Total	8	314.4			

According Table II, the value R_a^2 for the model is $R_a^2 = 79.9$ .This value of R_a^2 implies that 79.9% of sample variation in GTE is attributable to,or explained by,one or more of independent variables x1,x2 and x3.Thus , R^2 and R_a^2 are samples statistic that represent adequacy of the overall model is minimum 79.9%.we could arrive at the same decision by checking the observed significance level (p-value) of F test, given as 0.000.This value indicates that we will reject H_0 according equation(9) for any α greater than 0.000.the MINITAB printout shown in Table 2 also gives the two-tiled observed significance level(i.e., p-value) for each t test. These values that we would reject H_0 for any of $\beta_k = 0$ in favor of H_a for any

of $\beta_k \neq 0$ at any α larger than 0.02 for all β's except β's of $x_1 \times x_3^3$ and α larger than 0.067 for β's of $x_1 \times x_3^3$.this small p-values leave little doubt that x_1 , x_2 and x_3 contributes information for the prediction of GTE .The value of $s = \sqrt{MSE} = 1.986$ implies that the most of observed GTE values will fall within approximately $2s$ =3.972 of their respective predicted values.

TABLE II. FACTORIAL EXPERIMENT RESULTS PRINT OUT OF EXPERIMENT 2

Regression Analysis: y versus x1, x2, ...

The regression equation is
y = 65.9 - 2.96 x1 + 3.54 x2 - 2.53 x2^2 + 2.62 x2*x3 - 2.58 x2*x3^2 + 1.12 x1*x3^3 + 2.61 x1*x2*x3

Predictor	Coef	SE Coef	T	P
Constant	65.9011	0.6621	99.54	0.000
x1	-2.9611	0.4682	-6.32	0.000
x2	3.5433	0.8109	4.37	0.000
x2^2	-2.5272	0.8109	-3.12	0.006
x2*x3	2.6208	0.5734	4.57	0.000
x2*x3^2	-2.5825	0.9931	-2.60	0.018
x1*x3^3	1.1158	0.5734	1.95	0.067
x1*x2*x3	2.6100	0.7022	3.72	0.001

S = 1.98624 R-Sq = 85.3% R-Sq(adj) = 79.9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	7	434.421	62.060	15.73	0.000
Residual Error	19	74.958	3.945		

No evidence of lack of fit (P >= 0.1).

V. CONCLUSION

In this work a BFA based algorithm was proposed to solve the cell formation problem using non binary, real valued work load data as input matrices contain operation sequence of parts data. Moreover, the designed experiment aided by multiple regression methodology is carried out for tuning of BFA parameters in tumbling, swimming and swarming steps. The adequacy of model by wide variety of statistical criteria assessing is proved. The Once appropriately modified, this methodology may be employed for tuning of other BFA parameters in favor of better performance of this algorithm such a number of bacteria in the population, number of chemo tactic steps per bacteria lifetime, Limits the length of a swim, the number of bacteria reproductions (splits) per generation and The probability that each bacteria will be eliminated/dispersed.

REFERENCES

[1]. G. Harhalakis, R. Nagi, and J. M. Proth, "An efficient heuristic in manufacturing cell formation for group technology applications," *International Journal of Production Research*, vol. 28, no. 1, pp. 185–198, 1990.

[2]. W. Mendenhal and T. Sincich, *Statistics for Engineering and Sciences*, 4th Edition. Printice-Hall, Upper Saddle River, 1995.

[3]. K. M. Passino, "Biomimicry of bacterial foraging for distributed optimization and control," *IEEE Control Systems Magazine*, vol. 22, no. 3, pp. 52–67, 2002.

Hossein Nouri received his Bachelor and Master degrees from Iranian Universities, while obtaining his PhD from Universiti Putra Malaysia. He has years of academic experience from Islamic Azad University, Qazvin Branch. Currently, he is an academic of Islamic Azad University, Sanandaj Branch. His research focus includes manufacturing optimization.

Tang Sai Hong received his PhD and BEng from Dublin City University and Universiti Pertanian Malaysia, respectively. He is an Associate Professor and attaches with the Department of Mechanical & Manufacturing Engineering, Universiti Putra Malaysia since 1997. Currently, he works in the fields of robotics, operations research and artificial intelligence.

Mohd Khairol Anuar Mohd Ariffin is an Associate Professor in the Department of Mechanical & Manufacturing Engineering, Universiti Putra Malaysia. He received his BEng and PhD from United Kingdom Universities, and obtained his MSc from Malaysia. His research areas are manufacturing engineering and optimization.

B.T. Hang Tuah Baharudin is a Senior Lecturer in the Department of Mechanical & Manufacturing Engineering, Universiti Putra Malaysia. He received his BEng and PhD from United Kingdom Universities. His research interests include high speed machining, micro machining, rapid prototyping, reverse engineering, product development and mechanical design.

Razali Samin is a Professor Engineer and also a Senior Lecturer in the Department of Mechanical & Manufacturing Engineering, Universiti Putra Malaysia. He obtained his BEng and MSc from Universiti Putra Malaysia. His research interests include automation, robotics, dynamics and control.