International Journal of Computer Sciences and Engineering Systems
Vol. 11 No. 1 (June, 2017)

Evaluating Premature Convergence for Metaheuristic

Abu Bakar Md SULTAN[1], Ramlan MAHMOD[2], Mohd Taufik ABDULLAH[3] & Azizol ABDULLAH[4]

[1]Department of Information System, Universiti Putra Malaysia 43000 UPM Serdang, Selangor, Malaysia
E-mail: abakar@fsktm.upm.edu.my
[2]Department of Multimedia, Universiti Putra Malaysia 43000 UPM Serdang, Selangor, Malaysia
E-mail: ramlan@fsktm.upm.edu.my
[3]Department of Computer Science, Universiti Putra Malaysia 43000 UPM Serdang, Selangor, Malaysia
E-mail:mtaufik@fsktm.upm.edu.my
[4]Department of Networkinq, Universiti Putra Malaysia 43000 UPM Serdang, Selangor, Malaysia
E-mail: azizol@fsktm.upm.edu.my

Abstract: *Premature convergence is a common problem to population-based metaheurustic. The evaluation of premature convergence rate is difficult to obtain because the stochastic nature of metaheuristic. This paper presents a statistical effort to evaluate and predict the premature rate and performance of metaheuristic. The Fitness Distance Correlation technique was used to determine the premature rate and the memetic algorithm is tested on five selected timetabling datasets. The results shows that using relatively less effort, we can gain meaningful values of premature problems.*

Keywords: *Premature Convergence, Metaheuristic, Fitness Distance Correlation*

1. INTRODUCTION

Premature convergence is a common problem to population-based metaheuristic. The problem is caused by several algorithmic features, particularly selection pressure and too high gene flow between population member [1]. Diversity is undoubtedly a key issue in the performance of metaheuristic [2]. Many researchers agree that population diversity is largely depend of the interplay between the intensification and diversification strategies [3,4]. The problem of premature convergence could occur at any time during the optimization process. Many variants of evolutionary algorithms have been proposed to improve their performance as well as variant of measuring techniques to evaluate them. Thus, the motivation of this paper is the observation that the stochastic nature of most metaheuristics causes difficulty in determining the premature convergence rate. This paper proposed a measuring technique for evaluating and predicting premature convergence when it happens. The technique is applied to timetabling dataset using memetic algorithm. This measurement is important to evaluate the performance of existing algorithm and for future improvement.

The rest of the paper is organized as follows. In section 2 a description of the particular measuring technique is given and the evaluation method used for the experiment is presented in section 3. Finally, in section 4 we outline the results and the conclusion for our study is presented in section 5.

2. FITNESS DISTANCE CORRELATION

The work carried out by Jones [5] in his attempts to investigate the search difficulty for GA has produced the idea of FDC. The study determined how closely the relationship was between the fitness function and the heuristic function in order to obtain the search difficulty. A measure of search difficulty, Fitness Distance Correlation (FDC), was used to examine the performance of the Genetic Algorithm (GA) with or without incorporating knowledge of the GA operation. The values can be used to predict the performance of GA on problems with known global optima.

Jones ideas were to measure the extent to which fitness function values correlated with the distance to a global optimum. Given a set of $F = \{f_1, f_2, \ldots\ldots, f_n\}$ of n individuals in the population and the corresponding set $D = \{d_1, d_2, \ldots, d_n\}$ of Hamming Distances to the nearest global optimum, he computes the correlation coefficient FDC as:

$$FDC = \frac{C_{FD}}{S_F S_D}$$

where

$$C_{FD} = \frac{1}{n}\sum_{i=1}^{n}(f_i - \bar{f})\,(d_i - \bar{d})$$

C_{FD} is the covariance of F and D, S_F and S_D are the standard

deviations for F and D, and $\bar{f}$ and $\bar{d}$ means of F and D respectively.

3. METHOD OF EVALUATION

In this section, the method of using FDC to evaluate premature convergence is presented. The MA was run on the five different timetabling datasets and the value of the cost function with the respective time taken from one state to the next was captured. Then FDC was calculated for each independent run. The objective of the cost function is to reduce constraints for timetabling problem. The cost function is represented by F together with the corresponding time represent by D. Time determine the distance to arrive on the next solution.

Five-selected independent run was chosen for each datasets to present the premature convergence rate occurrences during the execution of MA on each datasets.

4. RESULT AND DISCUSSION

Table 1 shows the FDC values resulted from five independent run on five different datasets. The averages present the performance of the MA for each datasets. Premature values are the range toward maximum optimum (1-Average). Premature values determine the rate of premature convergence occurs during the execution of the datasets on datasets.

The result gives indicator performance of the algorithm and the premature convergence rate for future improvement.

5. CONCLUSION

The performance of the metaheuristic on optimization problem is largely depending on the diversity of the solution. However the stochastic nature of the metaheuristic make it difficult to estimate the convergence rate of metaheuristic. Thus this paper proposed the usage of FDC to evaluate the convergence rate. Initial results presented in this paper give promises indicator on premature rate. The rate is useful to detect the performance of the algorithm and for future improvement. The current research uses memetic algorithm and testing on five timetabling problem datasets. Future work will study different metaheuristic and also different optimization problem.

Table 1
Summary of FDC Values

Datasets	*Independent Run*					*Average*	*Premature*
	test 1	*test 2*	*test 3*	*test 4*	*test 5*		
small01	-0.814	-0.81	-0.819	-0.813	-0.804	-0.812	0.188
small02	-0.725	-0.733	-0.723	-0.719	-0.734	-0.727	0.273
small03	-0.906	-0.913	-0.877	-0.95	-0.95	-0.919	0.081
small04	-0.735	-0.726	-0.718	-0.728	-0.724	-0.726	0.274
small05	-0.823	-0.818	-0.822	-0.816	-0.823	-0.82	0.18

REFERENCES

[1] Ursem U. K. Diversity-guided Evolutionary Algorithms. LNCS (2002); 2439 462-471. Springer Verlag.

[2] Popovic D. Retaining Diversity of Search Point Distribution through a Breeder Genetic Algorithm For Neural Network Learning. *IEEE International Confrence on Neural Network* (1997); 1: 495-498.

[3] Ferland J. A., Marti R., Lavoie A., Gagne E. Scheduling using Tabu Search Solutions for the Linear Ordering Problems. *Computer and Operation Research* (2001); **28**: 1075-1092.

[4] Laguna M., Marti R, Canpos V. Intensification and Diversification with Elite Tabu Search Solution for the Linear Ordering Problem. Computers and Operation Research (1999); **26**: 1217-1230.

[5] Jones, T. (1995). Evolutionary Algorithms, Fitness Landscapes and Search, Ph.D Thesis, University of New Mexico.