Measuring Metaheuristic Performance over Timetabling Problem Instances Using Fitness Distance Correlation Method

Abu Bakar Md Sultan[†], Ramlan Mahmud[†], Muhammad Nasir Sulaiman[†],'Muhammad Rizam Abu Bakar[††]

[†]Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 Malaysia
[††] Faculty of Science, Universiti Putra Malaysia, 43400 Malaysia

Summary
The Timetabling Problem is a combinatorial optimization problem. Over the last decade variant of Metaheuristic approaches have been used to solve various type of Timetabling Problem with great success. Metaheuristic are approximate algorithm used when the size of the search and spaces becomes unmanageable. However the stochastic nature of metaheuristic make it difficult to measure the performance of the algorithms. In this paper we analyze the performance of Hybrid Evolutionary Metaheuristic for a set of benchmark timetabling instances through an analysis of the fitness-distance correlation (FDC). FDC is a statistical measure of a search difficulty in relation to Genetic Algorithm. We propose new ways of FDC analysis. Initial result indicates that FDC can also be used in different ways of analysis as well as different problem instances.

Key words:
Timetabling; Metaheuristic, Fitness Distance Correlation, Performance Prediction.

1. Introduction

The University Course Timetabling Problems (UCTP) deals with the scheduling of weekly timetable for a university. Lectures have to take place in a given number of time slots and rooms, so that a number of constraints are satisfied. Different versions of the problems arise at different institution. Comprehensive review on the timetabling problem and a number of research work can be found in [1] and [6]. The problem that we have studied in this paper is the UCTP [5]. We want to show that Fitness Distance Correlation (FDC) can be used to predict the performance of metaheuristic algorithm over the UCTP instances.

FDC is a measure of Genetic Algorithm performance introduced by [3]. Since the introduction, a few researchers have come forward to further analyze of the FDC. But as far as the authors is concern none of them has tried to discover the FDC in real world problem such as UCTP. In this papers we further shows that FDC can be also be employed to others algorithm rather than GA with several adjustments on the result obtained.

The motivation of the research presented in this paper came from the initial idea proposed by [4]. They present a study to better understand what make certain particular UCTP instances hard by employed linear statistical models. We are not going discuss about their model instead we are proposing alternative model to be consideration.

The paper is organized as follows. Section 2 presents the details of the FDC analysis. Section 3 briefly discusses the differences of our approach and the experiment being conducted. Section 4 present the result and we conclude in Section 5.

2. Fitness Distance Correlation (FDC)

A measure of search difficulty, Fitness Distance Correlation (FDC) is used to examine the performance of the genetic algorithm (GA) performance. The values can be used to predict the performance of GA on problems with known global optima. [3] ideas are to measure the extent to which fitness function values correlated with distance to global optimum. Given a set of F = $\{f_1, f_2,.....f_n\}$ of n individual 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

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

Manuscript revised May 30, 2006.
Manuscript revised June 25, 2006.

is the covariance of F and D, and S_F, S_D, $\bar{f}$ and $\bar{d}$ are the standard deviations and means of F and D respectively. For maximization problems, the assumption was that the fitness increases as distance decreases [4]. With and ideal fitness function, FDC will therefore be equal –1.0. The result of FDC indicated the performance of GA in the three different categorized as presented by [3]:

- *Misleading* (FDC ≥ 0.15) in which fitness increases with distance from global optimum.
- *Difficult* (-0.15<FDC<0.15*)* in which there is virtually no correlation between fitness distances.
- *Straightforward* (FDC ≤ -0.15) in which fitness tend to increase as the distance approach global optimum.

Jones proved that FDC is reliable although not infallible, indicator of GA performance on a wide range of the performance.

3. FDC FOR TIMETABLING INSTANCES.

In this section, we briefly discuss our approach of using FDC to measure the performance of our Hybrid Evolutionary metaheuristic algorithm to a set of benchmarking problem instances of timetabling problem. The problem instances were taken from metaheuristic research group (www.metaheuristic.net). It was a reduction reflect aspects of Napier University's real timetabling problem. The problem instance was generated by using a generator with different characteristic for different values of given parameters (Rossi-Doria, 2003). All instances produced have a perfect solution. For the purpose of this research we choose five small size instances.

The FDC was used to measure the algorithm performance over timetabling problem instances. The major differences with Jones approaches lies on information captured during the execution of the algorithm. Here we captured the value of every changes cost function and their respective time. The value of fitness now become the value of the cost function and the value of distance become the value of time to reach the respective cost function. The algorithm was executed over each problem instances for several times and FDC value was calculated.

The algorithm was a single point algorithm and we executed the algorithm for a few time with each regards to some predefined time to evaluate the performance. Table I below shows the differences of Jones approach and our approach in order to obtained performance result.

.

Table 1 : The differences of approach.

Items	Jones (1995)	Our approach
Algorithm	Multiple Point (GA)	Single Point Metaheuristic
Time	None	Predefined Time
Fitness Values	Each individual Fitness recorded	Last Fitness recorded
Distances	Distance regards to Fitness over nearest global optimum	Distance regards to time left to reach global optimum.
Problem Instances	Deceptive Function	Real World Problem (Timetabling)
Representation	binary	Decimal
Problem Difficulty	Known	Unpredictable
Value obtained	Every Iteration	Each trial

4. Result and Discussion

The result obtained from experimental process shows that FDC successfully determine the performance of the algorithm over each five instances. Table 2 below lists the result for instance01.

Table 2. Result for Test Instances01.

Number of Running	FDC
1	0.345
2	0.353
3	0.355
4	0.344
5	0.348

The calculated FDC value for combination number of trials gives us more accurate prediction about the performance of the algorithm. For instances the FDC value for five combination of trials for instan01 are 0.355. The same value is obtained for another five combination of trials on similar problem instances. This scenario also happens to the rest of problem instances. This indicated that the more

combination we used the more accurate prediction could be obtained. The result indicated that the FDC could be used to measure the performance of different types algorithm over different types of problem instances as well as different ways of values captured.

5. Conclusion

This paper evaluates the usage of FDC to the metaheuristic algorithm and timetabling problem instances. Initial result indicates that FDC are useful tools to measure problem difficulty and the performance of algorithm for better understanding their search behaviors. This is ongoing work. In order to further understanding these statistical measurement we try to analyze the search landscape and relate these to metaheuristic performance. Our future efforts are looking to these matters.

Acknowledgment. This research is sponsored by grants IRPA-04-02-04-0802-EA001 from Ministry of Science, Technology and Innovation Malaysia (MOSTI).

References:

[1] Carter, M.W., Laporte, G. (1997). "Recent Development in Practical Course Timetabling". In Edmund Burke and Carter W., Editors, Practice and Theory of Automated Timetabling, pages 3-19 LNCS 1408, Springer.

[2] Collard, P., Gaspar, A., Clergue, M., & Escazut. C. (1998) "Fitness Distance Correlation, as statistical measure of Genetic Algorithm difficulty, revisited", ECAI 1998, pg 650-654.

[3] Jones, T., (1995) "Evolutionary Algorithms, Fitness Landscape and Search". Ph.D dissertation, University of The new Mexico.

[4] Kostuch, P., Socha, K. (2004). "Hardness Prediction For University Course Timetabling Problem." Proce eding of The Evolutionary Computation in Combinatorial Optimization (EvoCOP 2004), Coimbra Portugal.

[5] Rossi-Doria, O., Sampels, M., Birattari, M., Chiarandini, M., Dorigo, M., Gambardella, L.M., Knowles, J., Manfrin, Max., Mastrolilli, M., Paechter, B., Paquete, L., & Stutzle, T. (2003) "A comparison of the Performance of Diffrent Metaheuristics on The Timetabling Problem". In E.Burke and P. De Causmaecker (Eds): PATAT 2002, LNCS 2740, pp 329-351, Springer.

[6] Schaerf, A.(1999) "A Survey of Automated Timetabling". Artificial Intelligent Review. 13, 87-127.

Abu Bakar Md Sultan is a PhD student at faculty of Computer Science and information Technology, Universiti Putra Malaysia (UPM). He holds a Master Degree in Software Engineering from UPM. His research interest includes artificial intelligence, scheduling and software agents.

Ramlan Mahmod hold a PhD from University of Bradford, United Kingdom. Currently, he is a Associate Professor at Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, His research area are artificial intelligence.

Md Nasir Sulaiman obtained his PhD from Loughborough University in 1994. Currently he is a Associate Professor in Department of Computer Science, Faculty of Computer Science and Information Technology, University Putra Malaysia. His research interest is artificial intelligence, neural networks, data mining, pattern recognition and parallel computing

Mohd Rizam Abu Bakar obtained her PhD from the University of Bradford. Currently, he is a lecturer at the Faculty of Science, Universiti Putra Malaysia.