Jurnal Teknologi, 49(D) Dis. 2008: 87–92

FITNESS DISTANCE CORRELATION(FDC) AS A HARDNESS PREDICTION FOR UNIVERSITY COURSE TIMETABLING PROBLEM

ABU BAKAR MD SULTAN[1], RAMLAN MAHMOD[2], MD NASIR SULAIMAN[3], MOHD RIZAM ABU BAKAR[4] & MOHD TAUFIK ABDULLAH[5]

Abstract. The timetabling problem is a combinatorial optimization problem. Frequently, metaheuristic techniques have been employed to solve this problem. Metaheuristic is a stochastic types algorithm and its performance usually difficult to predict. This paper described the usage of Fitness Distance Correlation (FDC) method to predict the performance of metaheuristic algorithm through statistical techniques. FDC is a statistical measure of search hardness in relation to Genetic Algorithm. Initial result from the experiment by hybrid algorithm over standard timetabling instances is very promising. We propose new ways of FDC analysis. The result indicates that FDC could be expanding in different ways of analysis as well as different instances.

Keywords: Timetabling; metaheuristic; genetic algorithm; fitness distance correlation; hardness prediction

Abstrak: Masalah penjadulan adalah masalah pengoptimuman kombinatorial. Kerap kali teknik-teknik metaheuristik digunakan untuk menyelesaikan masalah sebegini. Metaheuristik adalah algoritma bercirikan stokastik dan biasanya prestasi sukar diramal. Kertas ini menerangkan kaedah kolerasi Jarak Kecerdasan (FDC) untuk meramal prestasi algoritma metaheuristik melalui teknik statistik. FDC adalah pengukur statistik kepada kesukaran carian bagi algoritma genetik. Keputusan awal yang dihasilkan dari pengujian oleh algoritma hibrid ke atas contoh masalah penjadualan piawai amat menggalakkan. Keputusan menunjukkan FDC boleh diperluaskan kepada kaedah analisis dan masalah lain.

Kata kunci: Penjadualan; metaheuristik; algoritma genetik; korelasi jarak kecergasan; ramalan kesukaran

1.0 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. The constraints in

[1,2,3&5]Faculty of Computer Science and Information Technology, University Putra Malaysia, 43400 UPM Serdang Selangor, Malaysia
Email: abakar@fskstm.upm.edu.my, ramlan@fskstm.upm.edu.my, nasir@fskstm.upm.edu.my
[4] Department of Mathematic, Faculty of Science, University Putra Malaysia, 43400 UPM Serdang Selangor, Malaysia.
Email: rizam@fsas.upm.edu.my

timetabling are classified in two: hard and soft constraints. Hard constraints must be satisfied to yielded feasible timetable and the number soft constraint solve determined the quality of timetabling produced. Different versions of the problems arise at different institution. Comprehensive review on the timetabling problem and a number of research works can be found in [1 – 2]. The size and complexity of modern university timetabling problems encouraged research in metaheuristic techniques.

Metaheuristic techniques begin with one or more initial solution and iteratively employ search strategies to avoid local optima. Metaheuristics are stochastic methods used when the size of the search spaces become unmanageable for exact methods and no effective algorithm capable of finding optimal solution [3]. Some of the techniques for timetabling have been reported in [1 – 2].

FDC is a measurement of Genetic Algorithm performance introduced by [3]. Since the introduction, a few researchers have come forward to further analyze of the FDC. Up to knowledge, there is no FDC work to measure performance of GA in real world problem.

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 discusses the metaheuristic algorithm used for the experiment. Section 4 briefly discusses the FDC approach the experiment being conducted. Section 5 present the result and we conclude the finding in Section 6.

2.0 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 with known global optima. Ideas to measure the extent to which fitness function values correlated with distance to global optimum is given by [5]. Given a set of F = $\{f_1, f_2, ..., f_n\}$ of n individual in the population and the corresponding set D = $\{d_1, d_2, ... d_n\}$ of Hamming Distances to the nearest global optimum, [5] computes the correlation coefficient FDC as:

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

where

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

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 [5]:

- *Misleading* (FDC $\geq$ 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 $\leq$ –0.15) in which fitness tend to increase as the distance approach global optimum.

Jones [5] proved that FDC is reliable although not infallible indicator of GA performance on a wide range of the performance.

3.0 HYBRID EVOLUTIONARY ALGORITHM (HEA)

The Hybridization of HEA consists of components from different metaheuristics. The original idea of HEA came from the concepts of Memetic Algorithm (MA) combined with the acceptance criteria borrowed from Simulated Annealing.

The evolutionary step began with two initial random solutions, and then it underwent the improvement process under local search, mutation and crossover. Mutation as a key element in intensification stage occurred in every iteration, whereas crossover, a key element in diversification, occurs only on certain prescribed conditions. The reason was to avoid premature convergence and the solution was always at a higher rate of diversity. The algorithm is described Figure 1. The local search [5] was used for stochastic process improvement in two phases. The first phase improved infeasible timetable so that it became feasible by reducing the number of timeslots used. The second phase was to increase the quality of a feasible timetable by reducing the number of soft constraints violations.

4.0 FDC FOR TIMETABLING INSTANCES

This section briefly discuss the approach of using FDC to measure the performance of our Hybrid Evolutionary metaheuristic algorithm (HEA) 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 of a typical university course timetabling problem. It has been introduced to 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 [6]. All instances produced have a perfect solution.

```
Tp_HEA main

Generate 2 Solution (S1, S2)
Improve → local Search (S1, S2)
Best_Solution → S1

  while (not(termination condition)){

        if (S1 > S2){
                Best_Solution → S2
          else (S1 → S2)}

                Best_Solution → Mutate( )
                S2 → Best_Solution

        if (rnd_num < 0.01){
          Best_Solution → Crossover (Best_Solution, S1)}

                S1 → local Search( )
                Best_Solution → local Search( )

        if (S1 > Best_Solution) && Diff_Rate > 0.3)
                S1 → Best_Solution
                }
        }
```

Figure 1 Hybrid evolutionary algorithm

The HEA was run 15 times on each instances and the value of time together with the value of the cost function for respective changes was captured for calculating FDC. Table 1 listed the differences of Jones approach and proposed approach in this paper. This paper only discusses the result obtained from the implementation of HEA on selected instances mentioned in previous paragraph.

Table 1 The differences of two approaches

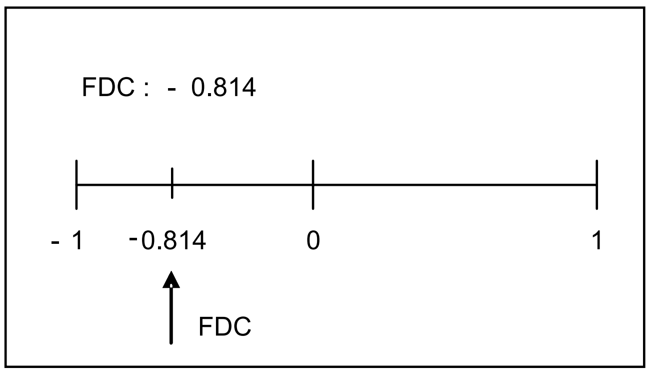
Items	Jones (1995)	Our approach
Algorithm	Multiple Point (GA)	Single Point Metaheuristic
Time	None	Predefined Time
Fitness values	Each individual Fitness recorded	Last Fitness recorded
Distances	Distance regards to Fitness over nearest global optimum	Distance regards to time left to reach global optimum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

5.0 RESULT AND DISCUSSION

To conform the result, the experiment to class of easy instances that considered being straightforward and easy for the algorithm to solve. Initial prediction, the FDC result should fall under the straightforward category. The algorithm was executed for 15 times, each trial was given 90 second as a predefined time to complete and result for every trial was recorded. Figure 2 below illustrated the process of calculating FDC for each trial.

The result shows that the FDC was –0.814 and proves our prediction. In fact the FDC value –0.184 close to –1.0 thus we consider the fitness function to be ideal. 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. In addition the FDC gave indicator the hardness of specific metaheuristic in solving certain particular instances.

Cost
15
13
12
10
9
8
7
6
4



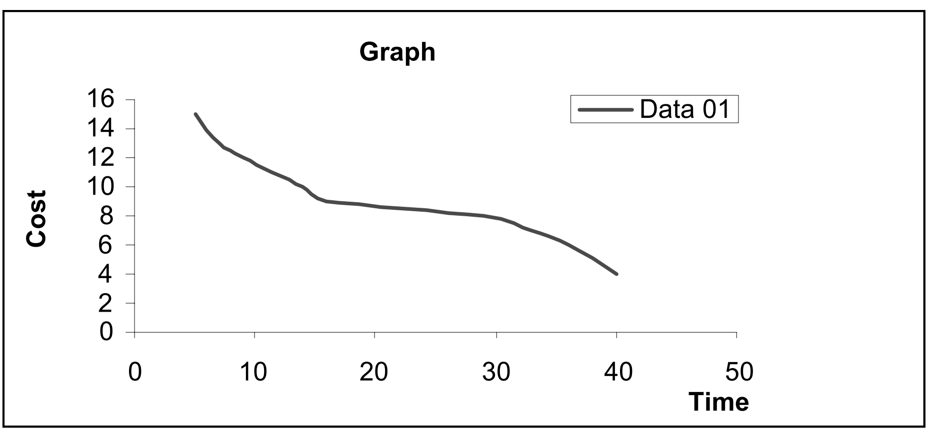


Figure 2 The calculation of FDC

6.0 CONCLUSION

This paper evaluates the usage of FDC to the metaheuristic algorithm and timetabling problem instances. Initial result indicates that FDC is a useful tools to measure problem difficulty and the performance of algorithm in order better understanding their search behaviors. This is ongoing work. Our future efforts are to further understand this statistical measurement by analyzing the search landscape and relate these to FDC.

ACKNOWLEDGEMENT

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. and G. Laporte. 1997. Recent Development in Practical Course Timetabling. In E. Burke and W. Carter W. Eds. *Practice and Theory of Automated Timetabling*. Berlin, German: Springer Verlag.

[2] Schaerf, A.1999. A Survey of Automated Timetabling. *Artificial Intelligent Review*. 13: 87-127.

[3] Rossi-Doria, O. and B. Paechter. 2003. A Memetic Algorithm for the University Course Timetabling. Proceeding of Combinatorial Optimization (CO) 2004.

[4] Kostuch, P. and K. Socha. 2004. Hardness Prediction for University Course Timetabling Problem. Proceeding of the Evolutionary Computation in Combinatorial Optimization (EvoCOP 2004), Coimbra Portugal.

[5] Jones, T. 1995. Evolutionary Algorithms, Fitness Landscape and Search. Ph.D dissertation. University of The New Mexico.

[6] Rossi-Doria, O., M. Sampels, M. Birattari, M. Chiarandini, M. Dorigo, L. M. Gambardella, J. Knowles, M. Manfrin, M. Mastrolilli, B. Paechter, L. Paquete and T. Stutzle. 2003. *A comparison of the Performance of Diffrent Metaheuristics on The Timetabling Problem*. In E.Burke and P. De Causmaecker (Eds). Springer Verlag. Berlin: 329-351.