UNIVERSITI PUTRA MALAYSIA

HYBRID METAHEURISTIC ALGORITHM AND METAHEURISTIC PERFORMANCE MEASUREMENT FOR SOLVING UNIVERSITY COURSE TIMETABLING PROBLEM

ABU BAKAR MD SULTAN

FSKTM 2007 16

HYBRID METAHEURISTIC ALGORITHM AND METAHEURISTIC PERFORMANCE MEASUREMENT FOR SOLVING UNIVERSITY COURSE TIMETABLING PROBLEM

By

ABU BAKAR MD SULTAN

Thesis Submitted to School of Graduate Studies, Universiti Putra Malaysia in Fulfilment of the Requirement for the Degree of Doctor of Philosophy

August 2007

UPM

DEDICATION

This thesis is dedicated to my parents, my wife and to anyone who thought me more about forbearance, fortitude and forgiveness

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

HYBRID METAHEURISTIC ALGORITHM AND METAHEURISTIC PERFORMANCE MEASUREMENT FOR SOLVING UNIVERSITY COURSE TIMETABLING PROBLEM.

By

ABU BAKAR MD SULTAN

August 2007

Chairman: Associate Professor Ramlan Mahmod, PhD

Faculty: Computer Science and Information Technology

Metaheuristics have received considerable interest in the fields of applied artificial intelligence and combinatorial optimization such as university course timetabling problem (UCTP). Metaheuristics begin with one or more initial solutions and iteratively employ search strategies to avoid local optima.

Recently, it was observed that the combination of concepts of different metaheuristics, called hybrid metaheuristics, can provide a more efficient behavior and higher flexibility in dealing with real-world and large-scale problems. Frequently, hybridizing the metaheuristic components lie on how we can effectively structure metaheuristic components to efficiently explore

and exploite search space. Acquiring the proper balance between intensification and diversification strategies is the crucial factor in obtaining an effective metaheuristic. This research focused on the implementation of an hybrid evolutionary metaheuristic namely Two_point Hybrid Evolutionary Algorithm (Tp_HEA) on university course timetabling problem instances (UCTP). Tp_HEA is based on two solutions that represent intensification at one point and diversification on the other point. Systematic exchange of information between these two points is to ensure the proper management of the balance between intensification and diversification.

The proposed Tp_HEA was tested on twelve standard UCTP instances according to the specified experimental procedure. The result obtained from the average point analysis and percentage of invalid solution was very promising. Out of twelve datasets, eight produced better performance when comparison was made against five other metaheuristics. The performance was measured in terms of constraints solved. Experimental results revealed that the arrangement of the Tp_HEA component would affect the search landscape of most UCTP problem instances.

The stochastic nature of metaheuristic including the Tp_HEA, results in inconsistent performance and the difficulty in obtaining accurate prediction from average point analyses. Thus, the second contribution of this research

is the introduction of Metaheuristic Performance Measurement (MPM). MPM is the attempt of measuring metaheuristic performance statistically, thus accurate indices can be obtained.

The validity of MPM as a new measuring technique was tested using selected results obtained from proposed Tp_HEA together with the result produced by genetic algorithm (GA). The analysis showed that MPM values obtained from both algorithms almost in line with the result obtained from average point analysis. The specific indices of performance produced by MPM were the major elements that differentiate MPM from average point analysis. The indices gave values for the performance, and thus the performance was more easily estimated. The reliability of MPM could be further observed when the analysis of variance showed that MPM values obtained from different independent runs were not significantly varied. Therefore, MPM was able to obtain a good estimation as compared to other commonly used metaheuristic measuring techniques.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

ALGORITMA METAHEURISTIK HIBRID DAN PENGUKUR PRESTASI METAHEURISTIK UNTUK PENYELESAIAN MASALAH PENJADUALAN KURSUS UNIVERSITI

Oleh

ABU BAKAR MD SULTAN

Ogos 2007

Pengerusi: **Profesor Madya Ramlan Mahmod, PhD**

Fakulti: **Sains Komputer dan Teknologi Maklumat**

Metaheuristik telah menarik banyak perhatian dalam bidang aplikasi kecerdasan buatan dan pengoptimuman kombinatorial seperti masalah penjadualan kursus universiti. Kaedah metaheuristik bermula dengan satu atau lebih penyelesaian awal dan secara lelaran menggunakan strategi-strategi carian untuk menghindar berlakunya optima awal.

Terbaru, didapati pengabungan konsep dari metaheuristik yang berbeza dipanggil penghibridan metaheuristik boleh menghasilkan gelagat yang effisien dan fleksibal dalam menyelesaikan masalah sebenar dan besar. Selalunya penghibridan metaheuristik terletak kepada bagaimana kita boleh menstrukturkan secara efektif komponen-komponen metaheuristik supaya

penjelajahan dan penyusupan ruang carian berlaku dengan efisien. Mendapatkan imbangan yang bersesuaian antara intensifikasi dan diversifikasi adalah isu penting menghasilkan metaheuristik yang efektif. Penyelidikan ini memfokus kepada perlaksanaan metaheuristik lelaran hybrid yang dinamakan Algoritma Hibrid Dua Titik Metaheuristik (Tp_HEA) ke atas masalah penjadualan kursus universiti (UCTP). Tp_HEA berasaskan kepada dua penyelesaian yang mewakili intensifikasi di satu titik dan diversifikasi di satu titik lain. Pertukaran maklumat bersistematik antara dua titik ini untuk memastikan imbangan antara intensifikasi dan diversifikasi diurus sebaiknya.

Tp_HEA yang diperkenalkan diuji ke atas duabelas UCTP piawai menuruti prosedur pengujian dinyatakan. Keputusan yang diperolehi dari analisa purata dan peratusan penyelesaian yang tidak sah amat menggalakkan. Dari duabelas dataset, lapan menghasilkan prestasi lebih baik bila perbandingan dibuat dengan lima metaheuristik yang lain. Proses-proses pengujian yang mendalam mendapati pengstrukturan komponen-komponen yang berbeza-beza memberi kesan ke atas landskap carian masalah UCTP dibawah kajian. komponen yang berbeza memberikan kesan ke atas landskap carian kepada semua masalah UCTP yang dikaji.

Ciri stokastik metaheuristik menjadikan ianya sukar diukur dari segi prestasi. Dengan itu sumbangan kedua kajian ini adalah pengenalan kepada

Pengukuran Prestasi Metaheuristik (MPM). MPM merupakan pendekatan mengukur prestasi metaheuristik melalui kaedah statistik bagi membolehkan indek tepat diperolehi.

Kebolehan MPM sebagai teknik pengukur baru diuji mengunakan keputusan terpilih yang diperolehi dari Tp_HEA bersama keputusan yang dihasilkan oleh algoritma genetic (GA). Analisa menunjukkan nilai-nilai MPM yang diperolehi dari kedua-dua algoritma adalah selari dengan yang didapati dari analisa purata, hanya MPM memberikan index sebagai penunjuk kepada prestasi. Kebolehpercayaan MPM selanjutnya boleh dilihat dari analisa varian yang menunjukkan nilai-nilai MPM yang diperolehi dari larian berasingan tidak begitu berbeza. Oleh itu MPM berkebolehan mendapatkan anggaran yang baik jika dibandingkan dengan lain-lain teknik pengukuran metaheuristik yang biasa digunakan.

ACKNOWLEDGEMENTS

In the name of Allah, the beneficent, the Merciful,

I would like to express my gratitude to the supervisory committee led by Assoc. Prof. Dr. Ramlan Mahmod and committee members, Assoc. Prof. Dr. Hj. Md Nasir Sulaiman and Assoc. Prof. Dr. Mohd Rizam Abu Bakar for their guidance, motivating and patient supervision throughout my research.

I deeply appreciate the help of Assoc. Prof. Hassan Selamat and Assoc. Prof. Dr. Abdul Azim Abdul Ghani who had made it possible for me to continue my study to PhD.

My deepest appreciation is to my parents, wife and children for their continued support, love and prayers over the past years. Finally, my thanks are also extended to my friends and colleagues, sharing experiences throughout the years.

I certify that an Examination Committee has met on 13th August 2007 to conduct the final examination of Abu Bakar Md Sultan on his Doctor of Philosophy thesis entitled "Two Point Hybrid Metaheuristic and Metaheuristic Performance Measurement for University Course Timetabling Problem" in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the candidate be awarded the relevant degree. Members of the Examination Committee are as follows:

FATIMAH AHMAD, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

ABDUL AZIM ABD GHANI, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

HAMIDAH IBRAHIM, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

SAFAAI DERIS, PhD
Professor
Faculty of Computer Science and Information System
Universiti Teknologi Malaysia
(External Examiner)

_____
HASANAH MOHD GHAZALI, PhD
Professor/Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy.
The members of the Supervisory Committee were as follows:

Ramlan Mahmod, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Md Nasir Sulaiman, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Mohd Rizam Abu Bakar, PhD
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Member)

AINI IDERIS, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 15 November 2007

DECLARATION

I hereby declare that the thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UPM or other institutions.

_____

ABU BAKAR BIN MD SULTAN

Date:

TABLE OF CONTENTS

	Page
DEDICATION	ii
ABSTRACT	iii
ABSTRAK	vi
ACKNOWLEDGEMENTS	ix
APPROVAL	x
DECLARATION	xii
LIST OF TABLES	xvii
LIST OF FIGURES	xix
LISTS OF ABBREVIATIONS	xxii

CHAPTER

1 INTRODUCTION

1.1	Background to the Research	1
1.2	Problem Statement	4
1.3	Research Objectives	6
1.4	Scope of The Research	7
1.5	Contribution of the Research	8
1.6	Thesis Outline	8

2 METAHEURISTICS

2.1	Introduction	11
2.2	Combinatorial Optimization Problem (COP)	11
2.3	Metaheuristic Approaches	13
2.3.1	Fundamental Properties of Metaheuristic	17
2.3.2	Classification of Metaheuristics	17
2.4	Constructive and Local Search	20
2.4.1	Constructive Heuristics	20
2.4.2	Local Search	21
2.5	Evolutionary Algorithms (EA)	22
2.6	Simulated Annealing (SA)	25
2.7	Tabu Search (TS)	27
2.8	Ant Colony Optimization (ACO)	28
2.9	Iterated Local Search (ILS)	30
2.10	Memetic Algorithm	31
2.11	Other Metaheuristics	32
2.12	Hybridizing Metaheuristics	34
2.13	Intensification and Diversification (I&D)	35
2.14	Hypherheuristics	36
2.15	Summary	38

3 TECHNIQUES FOR TIMETABLING

3.1 Introduction 40
3.2 Timetabling Problems 41
3.2.1 Introduction 41
3.2.2 Course Timetabling Problems 42
3.2.3 Methods for Timetabling Problem 43
3.2.4 Issues in Timetabling Research 45
3.3 Application of Metaheuristic to Timetabling Problem 48
3.3.1 EvolutionaryAlgorithms (EA) 48
3.3.2 Simulated Annealing (SA) 51
3.3.3 Tabu Search (TS) 52
3.3.4 Ant Colony Optimization (ACO) 53
3.3.5 Iterated Local Search (ILS) 54
3.3.6 Hybrid Metaheuristics 55
3.4 Local Search 56
3.5 Metaheuristics Performance Evaluation 60
3.5.1 Background 60
3.5.2 Searching Performance Evaluation 62
3.5.3 Search Landscape 63
3.5.4 Fitness Distance Correlation (FDC) 65
3.6 Summary 65

4 RESEARCH METHODOLOGY

4.1 Introduction 68
4.2 Research Overview 68
4.2.1 Identification of The Problem 70
4.2.2 Design and Implementation 70
4.2.3 Measuring Proposed Metaheuristic 71
4.2.4 Testing New Measuring Technique 71
4.3 Data Sets 72
4.3.1 Datasets Category 72
4.3.2 Description of Instances 74
4.4 System Design 75
4.4.1 Computing Environment 77
4.4.2 System Development 78
4.4.3 Overall System Architecture 79
4.4.4 Direct Solution Representation 80
4.4.5 Common Search Landscape 81
4.4.6 Initial Random Seeding 82
4.5 System Implementation 82
4.5.1 Running of The System 82
4.5.2 Design of Experiment 83
4.5.3 Comparison of Relative Performance (Tp_HEA) 86
4.5.4 Validation (MPM) 87
4.6 Analysis 88
4.6.1 Average-point Evaluation 89

4.6.2 Percentage of Invalid Solution 89
4.6.3 Validation of Proposed Metaheuristic Measuring Performance 90
4.6.4 Measuring Instances Effect on Metaheuristic Performance 90
4.6.5 Testing the Consistency of MPM Values 91
4.7 Summary 92

5 A MODEL OF HYBRID METAHEURISTIC
5.1 Introduction 93
5.2 Two-Point Hybrid Evolutionary Metaheuristic 93
5.2.1 Background 94
5.2.2 Overview of Tp_HEA 95
5.2.3 Intensification Stage 100
5.2.4 Diversification Stage 100
5.3 Metaheuristic Performance Measurement (MPM) 101
5.3.1 Motivation 102
5.3.2 Statistical Model for MPM 102
5.3.3 Innovation 104
5.3.4 MPM and Search Landscape Analysis 105
5.4 Summary 106

6 RESULTS AND DISCUSSION
6.1 Introduction 108
6.2 Average Point Evaluation 110
6.2.1 Small UCTP Instances 112
6.2.2 Medium UCTP Instances 115
6.2.3 Large UCTP Instances 118
6.2.4 Discussion of Experiment Results 119
6.3 Percentage of Invalid Solution 120
6.3.1 Small UCTP Instances 121
6.3.2 Medium UCTP Instances 121
6.3.3 Large UCTP Instances 125
6.3.4 Discussion of Experiment Results 126
6.4 Confirmation of Metaheuristic Performance Measurement 127
6.4.1 Measuring Selected Running Test 129
6.4.2 Comparison of MPM Values (Tp_HEA) 132
6.4.3 Comparison of Best Solution against MPM Values 139
6.4.4 Discussion of Experiment Results 140
6.5 Analysis of Variance (ANOVA) on MPM Values 141
6.5.1 Instances Effect on Algorithm Performance 143
6.5.2 Consistency Checking on Different Independent Run 145
6.5.3 Discussion of Experiment Results 147
6.6 Summary 149

7 CONCLUSION AND FUTURE WORK
7.1 Introduction 151
7.2 Conclusion 152
7.3 Future Work 155

BIBLIOGRAPHY 157

APPENDICES 168

BIODATA OF THE AUTHOR 176

LIST OF PUBLICATIONS 177

LIST OF TABLES

Table		Page
2.1	Fundamental Properties of Metaheuristic	17
2.2	I&D Components to the Basic Metaheuristics	36
4.1	Instances Parameters	73
4.2	Computing Requirement	77
4.3	Experimental Information	84
5.1	The differences of approach MPM vs FDC.	105
6.1	Percentage of Invalid Solution for Small Instances	121
6.2	Percentage of Invalid Solution for Medium Instances	122
6.3	Percentage of Invalid Solution for Large Instances	125
6.4	Summary of MPM Value for Tp_HEA	130
6.5	Summary of MPM Value for GA	131
6.6	MPM of Tp_HEA vs GA (test1)	133
6.7	MPM of Tp_HEA vs GA (test2)	134
6.8	MPM of Tp_HEA vs GA (test3)	135
6.9	MPM of Tp_HEA vs GA (test4)	136
6.10	MPM of Tp_HEA vs GA (test5)	137
6.11	MPM of Tp_HEA vs GA (Average)	138
6.12	Comparison for small01.tim (Tp_HEA)	139
6.13	Comparison for small01.tim (GA)	139
6.14	Analysis of Variance for Small Instance Data	144

	(*Tp_HEA*)	
6.15	Analysis of Variance for Medium Instance Data (*Tp_HEA*)	144
6.16	Analysis of Variance for Small Instance Data (*GA*)	144
6.17	Analysis of Variance for Medium Instance Data (*GA*)	144
6.18	Analysis of Variance for Small Instance Data (*Tp_HEA*)	146
6.19	Analysis of Variance for Medium Instance Data (*Tp_HEA*)	146
6.20	Analysis of Variance for Small Instance Data (*GA*)	146
6.21	Analysis of Variance for Medium Instance Data (*GA*)	146
6.22	Summarized of ANOVA Results	148
6.23	Summarized of ANOVA Results	149

LIST OF FIGURES

Figure		Page
2.1	Evolutionary Algorithm (EA)	23
2.2	Simulated Annealing Algorithm	26
2.3	Simple Tabu Search Algorithm	27
2.4	Ant Colony Optimization	29
2.5	Iterated Local Search Metaheuristic	31
2.6	Memetic Algorithm	32
2.7	General Hypherheuristic Framework	38
3.1	Local Search Algorithm	59
4.1	Research Phases	69
4.2	System Architecture	80
4.4	System Command Line	83
4.5	Overview of Experimental Processes	85
5.1	Tp_HEA Algorithm	97
5.2	Flowchart	99
5.3	Correlation for Cost and Time for Maximize and Minimize Problem.	103
6.1	Average Point for Small01.tim	112
6.2	Average Point for Small02.tim	113
6.3	Average Point for Small03.tim	113
6.4	Average Point for Small04.tim	114
6.5	Average Point for Small05.tim	114

6.6	Average Point for Medium01.tim	115
6.7	Average Point for Medium02.tim	116
6.8	Average Point for Medium03.tim	116
6.9	Average Point for Medium04.tim	117
6.10	Average Point for Medium05.tim	117
6.11	Average Point for Large01.tim	118
6.12	Average Point for Large02.tim	119
6.13	Percentage of Invalid Solution For Medium01.tim	122
6.14	Percentage of Invalid Solution For Medium02.tim	123
6.15	Percentage of Invalid Solution For Medium03.tim	123
6.16	Percentage of Invalid Solution For Medium04.tim	124
6.17	Percentage of Invalid Solution For Medium05.tim	124
6.18	Percentage of Invalid Solution For Large01.tim	125
6.19	Percentage of Invalid Solution For Large02.tim	126
6.20	MPM value for One Independent Run	128
6.21	Performance Pattern (Tp_HEA)	130
6.22	Performance Pattern (GA)	131
6.23	MPM of Tp_HEA against GA (test1)	133
6.24	MPM of Tp_HEA against GA (test2)	134
6.25	MPM of Tp_HEA against GA (test3)	135
6.26	MPM of Tp_HEA against GA (test4)	136
6.27	MPM of Tp_HEA against GA (test5)	137
6.28	MPM of Tp_HEA against GA (Average)	138

6.29	The Reference Distribution ($F_{4,20}$) for the Test Statistic F_0	145
6.30	The Reference Distribution ($F_{4,20}$) for the Test Statistic F_0	147

LIST OF ABBREVIATIONS

ACO	Ants Colony Optimization
AI	Artificial Intelligence
ANOVA	Analysis of Variance
COP	Combinatorial Optimization Problem
CRD	Completely Randomized Design
EA	Evolutionary Algorithms
FDC	Fitness Distance Correlation
GA	Genetic Algorithm
GLS	Guided Local Serach
GRASP	Greedy Randomized Adaptive Search Procedure
ILS	Iterated Local Search
MA	Memetic Algorithm
MN	Metaheuristic Network
MPM	Metaheuristic Performance Measurement
NP	Nondeterministic Polynomial
PATAT	Practice and Theory of Automated Timetabling
SA	Simulated Annealing
TS	Tabu Search
TSP	Traveling Salesman Problem
Tp_HEA	Two_point Hybrid Evolutionary Algorithm
UCTP	University Course Timetabling Problem
QAP	Quadratic Application Problem

CHAPTER 1

INTRODUCTION

1.1 Background to the Research

Timetabling problems have attracted the continuous interest of researchers mainly because the problems provide the opportunity of testing combinatorial solution methods in formulations that represent difficult practical problems (Dimopoulou and Miliotois, 2001). Reis and Oliviera (2001) and He *et al.* (2005) reported that scientific community has given considerable amount of attention to automated timetabling during the last four decades according to the variants of timetabling literature published since then.

Timetable scheduling is an activity of assigning subjects to time and space such that all constraints are satisfied (Deris *et al.*, 2000). It can be categorized into several types and the most common type is academic university, college and school timetabling. The problem of constructing course timetables for higher learning institutions consist of allocating the set of courses offered by a university to a time period and classrooms in such a way that no teacher, student or room is used more than once at a specific time.