

**A NOVEL GENETIC MULTI OBJECTIVE TASK
SCHEDULING ALGORITHM IN CLOUD
COMPUTING**

K.M.S.U.Bandaranayake

229308U

Masters of Science in Computer Science

Department Of Computer Science and Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

June 2024

A NOVEL GENETIC MULTI OBJECTIVE TASK SCHEDULING ALGORITHM IN CLOUD COMPUTING

K.M.S.U.Bandaranayake

229308U

Dissertation submitted in partial fulfillment of the requirements for the
degree
Masters of Science in Computer Science

Department Of Computer Science and Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

June 2024

DECLARATION

I declare that this is my own work and this Dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature:

Date: 09-07-2024

The supervisor should certify the Dissertation with the following declaration.

The above candidate has carried out research for the Masters of Science in Computer Science Dissertation under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of Supervisor: Dr. Sunimal Rathnayake

Signature of the Supervisor:

Date: 09/07/2024

ACKNOWLEDGEMENT

I wish to convey my deep appreciation to the Department of Computer Science and Engineering for their support and encouragement. My sincere thanks go to my research supervisor, Dr. Sunimal Rathnayake, for his unwavering dedication and insightful guidance. Additionally, I am profoundly grateful to my parents and family for their continuous support throughout my academic journey.

ABSTRACT

Cloud computing is a trending cutting-edge model that allows users to utilize services through the internet. It offers on-demand scalable resources and compensates customers based on their usage. The cloud platform provides extensive computing power with a variety of configurable options. Effective resource management is crucial for optimizing the underlying system and ensuring that both cloud providers and users meet the Quality-of-Service (QoS) requirements specified in Service Level Agreements (SLAs). Numerous strategies are employed for cloud resource management, with scheduling being a prominent approach for optimizing cloud systems. Task and workflow scheduling are key unresolved challenges within cloud computing scheduling. Assigning tasks to suitable cloud resources is known to be an NP-complete problem. There are various factors affecting the execution of different workloads. Hence there is no single optimal solution under a given polynomial time. Many factors and objectives need to be considered when trying to find an optimal schedule. The total execution time of virtual machines has not been considered by previous studies as a scheduling parameter in the quest for an optimal schedule using meta-heuristic methods such as genetic algorithms. This study introduces G-TRETA, a new multi-objective genetic algorithm that aims to find an optimal schedule considering both total resource execution time and cost. The proposed algorithm was compared with FCFS, MaxMin, MinMin and MCT algorithms for the degree of imbalance, makespan, cost, and system throughput. CloudSim 5.0 simulator and DEAP framework with synthetic data traces and Nasa Ames/iPSC 860 real-world traces were used to evaluate the algorithm. From the evaluation results the proposed algorithm gives better results. We can observe 36.37% and 50.46% for makespan, 57.16%, 101.88% for system throughput, 1.47 and 0.94 differences of index improvement for degree of imbalance for synthetic and real-world traces respectively while having efficient cost values.

Keywords: Multi-Objective Task Scheduling, Cloud Computing, Total Resource Execution Time Aware Scheduling, Genetic Algorithms

TABLE OF CONTENTS

Declaration of the Candidate & Supervisor	i
Acknowledgement	ii
Abstract	iii
Table of Contents	iv
List of Figures	vii
List of Tables	viii
List of Appendices	ix
1 Introduction	1
1.1 Background and Motivation	1
1.2 Problem Statement	1
1.3 Research Questions	2
1.4 Research Objectives	2
1.5 Summary	2
2 Literature Review	3
2.1 Overview	3
2.2 Cloud Computing	3
2.2.1 Cloud Service Models	3
2.2.2 Cloud Deployment Models	4
2.3 Resource Management in Cloud	5
2.3.1 Task Scheduling	6
2.3.2 Workflow Scheduling	7
2.4 Genetic Algorithms	7
2.5 Advantages of Genetic Algorithms	9
2.6 Related Work	10
2.6.1 Static Task Scheduling	11
2.6.2 Dynamic Task Scheduling	11
2.6.3 Multi-Objective Task Scheduling	12

2.6.4	Multi-Objective Hybrid Task Scheduling	13
2.6.5	Multi-Objective Workflow Scheduling	14
2.6.6	Multi-Objective Hybrid Workflow Scheduling	14
2.7	Summary	15
3	Proposed Research Method	16
3.1	Problem Formulation an Constraints	16
3.1.1	Constraints	17
3.2	Objectives and Evaluation Metrics	17
3.3	Proposed Algorithm	18
3.3.1	Individual Representation and Initial Population	19
3.3.2	Selection Operator	19
3.3.3	Crossover Operator	19
3.3.4	Mutation Operator	20
3.3.5	Fitness Function	21
3.3.6	Termination	21
3.3.7	Elistism	21
3.4	Summary	22
4	Evaluation and Results	23
4.1	Implementation Method and Tools	23
4.2	Workload Traces	23
4.3	Experimentation Setup and Specifications	24
4.4	Experimentation with Synthetic Workload Traces	25
4.4.1	Experimental Results Analysis	26
4.5	Experimentation with Real Workload Traces	34
4.5.1	Experimental Results Analysis	35
4.5.2	Evaluation and Analysis	35
4.6	Discussion	43
5	Conclusion and future work	44
5.1	Limitations and Challenges	44
5.1.1	Limitations	44
5.1.2	Challenges	44

5.2 Contributions of the study	45
5.3 Conclusion	45
5.4 Future work and Recommendations	46
References	47
Appendix A	50

LIST OF FIGURES

Figure	Description	Page
Figure 2.1	Shared Responsibility Model of Cloud Service Models.	4
Figure 2.2	Workflow Example.	7
Figure 2.3	Genetic Algorithm Workflow.	8
Figure 3.1	Task scheduling solution representation in genetic algorithm.	19
Figure 3.2	Tournament selection example.	20
Figure 3.3	Two-point crossover example.	20
Figure 4.1	Implementation Method.	24
Figure 4.2	Sample of Minimum/Average fitness over generations for synthetic workload trace.	27
Figure 4.3	Sample for diversity of individuals in the population for the synthetic workload trace.	28
Figure 4.4	Makespan results synthetic workload.	29
Figure 4.5	Throughput results for synthetic workload trace.	30
Figure 4.6	Results of Degree of Imbalance for synthetic workload trace.	31
Figure 4.7	Cost results for synthetic workload trace.	32
Figure 4.8	Cost results for synthetic workload trace after re-run with different weights for w_2 .	33
Figure 4.9	Minimum/Average fitness over generations for real workload traces.	36
Figure 4.10	Makespan results for real world workload trace.	37
Figure 4.11	System throughput results for real world workload trace.	38
Figure 4.12	Degree of Imbalance results for real world workload trace.	39
Figure 4.13	Cost results for real workload trace.	41
Figure 4.14	Cost results for real workload trace after cost improvement for real workload trace.	42

LIST OF TABLES

Table	Description	Page
Table 2.1	Cloud Resource Categories.	6
Table 3.1	Symbol Table.	16
Table 4.1	VM instance types and characteristics.	24
Table 4.2	CloudSim Simulator 5.0 Specifications for Synthetic Workload Trace.	25
Table 4.3	Configurations and Specifications for G-TRETA.	25
Table 4.4	Makespan improvement percentages of G-TRETA for synthetic workload trace	29
Table 4.5	System throughput improvement percentages of G-TRETA for synthetic workload trace.	30
Table 4.6	DI improvement percentages of G-TRETA for synthetic workload trace.	31
Table 4.7	Cost increase percentages of G-TRETA.	33
Table 4.8	CloudSim Simulator 5.0 Configurations Specifications for real world trace simulation.	34
Table 4.9	G-TRETA Specifications for real world traces.	34
Table 4.10	Makespan improvement percentages of G-TRETA for real world workload trace.	35
Table 4.11	System throughput improvement percentages of G-TRETA for real world workload trace.	36
Table 4.12	DI improvement percentages of G-TRETA for Real World Workload Trace.	37
Table 4.13	Cost improvement increase of G-TRETA.	40

LIST OF APPENDICES

Appendix	Description	Page
Appendix -A		50