Chapter 9
Guided Search-Based Multi-Objective Evolutionary Algorithm for Grid Workflow Scheduling

Ritu Garg
National Institute of Technology Kurukshetra, India

ABSTRACT

The computational grid provides the global computing infrastructure for users to access the services over a network. However, grid service providers charge users for the services based on their usage and QoS level specified. Therefore, in order to optimize the grid workflow execution, a robust multi-objective scheduling algorithm is needed considering economic cost along with execution performance. Generally, in multi-objective problems, simulations rely on running large number of evaluations to obtain the accurate results. However, algorithms that consider the preferences of decision maker, convergence to optimal tradeoff solutions is faster. Thus, in this chapter, the author proposed the preference-based guided search mechanism into MOEAs. To obtain solutions near the pre-specified regions of interest, the author has considered two MOEAs, namely R-NSGA-II and R-ε-MOEA. Further, to improve the diversity of solutions, a modified form called M-R-NSGA-II is used. Finally, the experimental settings and performance metrics are presented for the evaluation of the algorithms.

INTRODUCTION

Real world optimization problems very often involve multiple objectives that have to considered simultaneously. The applications scheduling problem in computational grid environment often requires multiple objectives to be considered like makespan, economic cost, reliability etc. (Topcuoglu, Hariri, & Wu, 2002; Kumar, Dutta, & Mookerjee, 2009). As in many real world multi objective problems (MOP), it is obvious that these objectives are conflicting in nature and it is very difficult to handle number of conflicting objectives at the same time. Since enhancement in one objective may cause the deterioration

DOI: 10.4018/978-1-5225-5832-3.ch009

in another and no solution exist that is best with respect to all the objectives; rather every solution is a tradeoff among the identified objectives. There are mainly two methods to find out the tradeoff: weighted sum method and Pareto optimal method. The MOPs are generally solved by priori approaches which basically transform the multiple objectives into single objecting or weighted sum function. In weighted sum method, each objective function is linked with a weighting coefficient and then minimization of weighted sum of all the objectives is performed in order to obtain the single preferred solution. However, this is usually not applicable if the decision maker (DM) does not explicitly know how to weight the different objectives before the optimal alternatives are known. At the same time, it is difficult to estimate the weights, as small change in weights may change the solution drastically.

On the other hand, in posteriori approaches, the aim is to find the set of Pareto optimal solutions. Subsequently, the set of Pareto optimal solutions are passed to user\DM that makes a single choice according to his/her preferences, which is also called as posteriori articulation of preferences. A good Pareto optimal front (POF) must provide convergence (solutions close to optimal) as well as diversity (uniformly cover all possible ranges of optimal solutions) Usually, the Pareto optimal solutions are preferred over single solution in the real life applications. However, it becomes a critical issue if the optimization process involves computationally expensive function evaluations corresponding to large scale complex problems.

In order to reduce the total number of evaluations for faster convergence towards the efficient POF, a guided search that takes into account the user/DM preferences or reference points (RP) during optimization seems to be promising. In reference point approach, DM requires to specify reference points\preferences in terms of aspiration and reservation levels for all objective functions. Moreover, the user may have some specific preferences for the solution or they may be having the rough idea about what he/she would prefer. Thus, in this case, instead of providing Pareto optimal front in the entire solution space, it is preferable to provide Pareto solutions close to user\DM preferences, which is also called prior articulation of preferences. Further, RP approach makes it possible to find Pareto optimal solutions corresponding to multiple preferences simultaneously. Additionally, RP based approaches are applicable to MOPs with large number of objectives because preference relation gives a finer order of vectors of objective space in comparison to Pareto dominance relation. With this motivation, several algorithms with the focus to search towards the reference areas in objective space have been proposed. Some applications using RP based evolutionary multi objective optimization particularly R-NSGA-II are logistic network design (Cheshmehgaz, Islam, & Desa, 2014) reversible logic circuit design (Wang, 2014) and workflow scheduling in cloud (Verma & Buyya, 2005) etc.

Application Workflows in Grid

Grid computing is a novel paradigm that provides the global computing infrastructure for users to access the services over a network. The term Grid is analogous to an electric power grid that provides consistent, pervasive, dependable and transparent access to electricity irrespective of its source. Grid computing is a variant of parallel and distributed computing that involves the integrated and collaborative use of wide range of heterogeneous and distributed resources for executing large-scale computing applications. The resources may include expensive computational systems, high-speed networks, storage devices, databases, scientific instruments, softwares etc. owned and managed by different organizations. However, in order to meet the computational requirements of large and diverse groups of users, the grid computing systems need to address various challenging issues that are inherent to the grid environment.