8th International Congress on Civil Engineering,
May 11-13, 2009, Shiraz University, Shiraz, Iran

Utilizing Self Organizing Maps to Enhance Pareto Front Discovery in Large Data Sets

K. Nowruzi[1], G.R. Rakhshandehroo[2], P. Monadjemi[3]

1 - Ph.D. Student, Shiraz University, Shiraz, Iran
2 - Associate Prof. Shiraz University, Shiraz, Iran
3 - Assistant Prof. Shiraz University, Shiraz, Iran

Corresponding Author: norozi@shirazu.ac.ir

Abstract

Finding the Pareto front has become the goal of multiobjective optimization methods. Nondominated Sorting Algorithm (NSA) is mostly used by researchers to find the Pareto front. Although NSA has been used increasingly in recent years, one of its main drawbacks has been its computational efficiency. The Algorithm's computational cost has a cubic dependency to the number of individuals being searched for the Pareto front. For small data sets, the algorithm performs Pareto front discovery within a reasonable time. For larger data sets, however, the number of comparisons grow enormously, and hence, the time and effort that goes to Pareto front discovery increases very rapidly. In this paper, the topological ordering of Self Organizing Maps (SOM) was utilized to enhance nondominated sorting efficiency for Pareto Front discovery in large data sets. The method utilizes extracted clusters from application of SOM as a close approximation of the topology of the input data. Then, nondominated sorting is applied to the clusters which have a smaller population size. In other words, the proposed hybrid method eliminates the unnecessary comparison of individuals most likely to be a Pareto front member with individuals that are far from the front. As a case study, the method was applied to an inventory of about 7500 solutions of a benchmark water distribution optimization problem. Different population sizes were tried to demonstrate efficiency of the proposed method. It is verified that the method increased optimization efficiency considerably. Even in the data sets that were biased toward the Pareto front, the proposed method increased efficiency of nondominated sorting and reduced the computational effort by four folds.

Keywords: Pareto Front, Nondominated Sorting, Self Organizing Map, Water Distribution Networks, Optimization.

1. INTRODUCTION

Finding the Pareto front has become the goal of multi objective optimization problems. In realistic problems of multiobjective optimization, it is usually impossible to obtain solutions that are superior for all objectives. In fact, one solution may be best for one objective function and worst for many other objective functions. Therefore, multi objective optimization aims at finding a set of solutions that is the best with respect to predefined set of criteria's or tries to find solutions that are relatively better than other solutions according to a preference structure [1]. One such preference structure is the nondominated sorting. Nondominated sorting which aims at the discovery of Pareto front was first developed and utilized mostly by deb. At el [2]. Though the concept has been introduced long ago, by the Italian economist, Pareto, it was through the mentioned work that is now widely adapted and used by many researchers in optimization. Nondominated sorting is used widely for Water Distribution Network (WDN) optimization [3,4,5] and other engineering optimizations [1].

Design of an efficient WDN is a large-scale combinatorial, non-linear optimization problem, involving many complex implicit constraint sets (such as nodal mass balance and energy conservation) which are commonly satisfied through the use of hydraulic network solver[6]. Traditionally, WDN optimization has been mostly focused on the cost minimization [7]. However, it is well known now that there is other reliability and quality based objectives that are to be considered as well. In fact, optimization of WDNs has recently shifted from simple cost minimization toward multi objective optimization with common objectives such as, cost, reliability, quality, and hydraulic constraints [3,7].

The nonlinearity of multi objective optimization problems often requires very large-scale stochastic searching algorithms. A search algorithm mostly used by researchers to find the Pareto front is Nondominated Sorting Algorithm (NSA) within the population. Although NSA has been increasingly utilized in recent years, one of its main drawbacks is its computational efficiency. The algorithm's computational cost has a cubic dependency to the number of individuals being searched for the Pareto front[8]. For small data sets, the