BINARY PSO FOR DOCUMENT CLUSTERING WITH LOCAL SEARCH

K. PREMALATHA & A. M. NATARAJAN

Abstract

In recent years, there is a tremendous growth in the volume of text documents that are available on the internet, digital libraries, news sources, and company-wide intranets. This has led to an increased interest in developing methods that can help users to effectively navigate, summarize, and organize this information with the ultimate goal of helping them to find what they are looking for. Fast and high-quality document clustering algorithms play an important role towards the goal by organizing large amounts of information into a small number of meaningful clusters. Traditional clustering algorithms will search only a small sub-set of all possible clustering and consequently, there is no guarantee that the solution found will be optimal. This paper presents Binary Particle Swarm Optimization (BPSO) algorithm for document clustering with local search. Contrary to the localized searching for the traditional clustering algorithms, the proposed clustering algorithm performs a globalized search in the entire solution space and also it makes a local search to improve the quality of the solution or goodness of fitting. Experiment results on classical benchmarks are given. A marked improvement in the algorithm is observed with the traditional K-Means clustering algorithm.