HYBRID PSO-GA WITH CROSSOVER FOR DOCUMENT CLUSTERING

K. PREMALATHA AND A. M. NATARAJAN

Abstract

This paper presents a new hybrid PSO-GA with crossover approach for the document clustering problem. This algorithm is basically a population-based heuristic search technique, which can be used to solve combinatorial optimization problems, modeled on the hybrid of cultural and social rules derived from the analysis of the swarm intelligence (PSO) and the concepts of natural selection and evolution (GA). To obtain an optimal solution using Genetic Algorithm, operation such as selection, reproduction, and mutation procedures are used to generate for the next generations. In this case, it is possible to obtain local solution because chromosomes or individuals which have only a close similarity can convergent. In standard PSO the non-oscillatory route can quickly cause a particle to stagnate and also it may prematurely converge on suboptimal solutions that are not even guaranteed to local optimal solution. This work proposes a new hybrid model that enhances the search process by crossover operation between the stagnated particles and chromosomes for improving the diversity, and the convergence toward the preferred solution for the document clustering problem. The approach efficiency is verified and tested using a set of document corpus. Results indicate that the approach is highly forceful and can be considered as a feasible alternative to solve document clustering problems.