

DOCUMENT CLUSTERING BASED ON GENETIC ALGORITHM WITH DYNAMIC AND ADAPTIVE MUTATION

K. PREMALATHA AND A. M. NATARAJAN

Abstract

This paper presents Genetic algorithm with Dynamic mutation operator and Adaptive mutation rate at gene level (DAGA) for clustering documents. The mutation operation is essential to the success of genetic algorithms since it extends the search directions and avoids convergence to local optima. In each stage of the genetic process in a single problem, may require appropriately different mutation operators and mutation rate for best results. This paper proposes a genetic algorithm with dynamic & adaptive mutation to resolve these difficulties. In dynamic mutation the genetic algorithm simultaneously uses several mutation operators in producing the next generation. The mutation ratio of each operator changes according to evaluation results from the respective offspring it produces. In adaptation scheme that adapts the mutation rate separately for each gene location on the chromosome based on the feedback taken from the success and failure rates of the individuals in the earlier and current populations. Experiments results are examined with document corpus. It demonstrates that the proposed algorithm statistically outperforms the Simple GA and K-Means.

Keywords : Genetic Algorithm, DAGA, K-Means, Convergence, tf-idf