World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:16, No:08, 2022

Data Clustering Algorithm Based on Multi-Objective Periodic Bacterial Foraging Optimization with Two Learning Archives

Authors : Chen Guo, Heng Tang, Ben Niu

Abstract : Clustering splits objects into different groups based on similarity, making the objects have higher similarity in the same group and lower similarity in different groups. Thus, clustering can be treated as an optimization problem to maximize the intra-cluster similarity or inter-cluster dissimilarity. In real-world applications, the datasets often have some complex characteristics: sparse, overlap, high dimensionality, etc. When facing these datasets, simultaneously optimizing two or more objectives can obtain better clustering results than optimizing one objective. However, except for the objectives weighting methods, traditional clustering approaches have difficulty in solving multi-objective data clustering problems. Due to this, evolutionary multi-objective optimization algorithms are investigated by researchers to optimize multiple clustering objectives. In this paper, the Data Clustering algorithm based on Multi-objective Periodic Bacterial Foraging Optimization with two Learning Archives (DC-MPBFOLA) is proposed. Specifically, first, to reduce the high computing complexity of the original BFO, periodic BFO is employed as the basic algorithmic framework. Then transfer the periodic BFO into a multi-objective type. Second, two learning strategies are proposed based on the two learning archives to guide the bacterial swarm to move in a better direction. On the one hand, the global best is selected from the global learning archive according to the convergence index and diversity index. On the other hand, the personal best is selected from the personal learning archive according to the sum of weighted objectives. According to the aforementioned learning strategies, a chemotaxis operation is designed. Third, an elite learning strategy is designed to provide fresh power to the objects in two learning archives. When the objects in these two archives do not change for two consecutive times, randomly initializing one dimension of objects can prevent the proposed algorithm from falling into local optima. Fourth, to validate the performance of the proposed algorithm, DC-MPBFOLA is compared with four state-of-art evolutionary multi-objective optimization algorithms and one classical clustering algorithm on evaluation indexes of datasets. To further verify the effectiveness and feasibility of designed strategies in DC-MPBFOLA, variants of DC-MPBFOLA are also proposed. Experimental results demonstrate that DC-MPBFOLA outperforms its competitors regarding all evaluation indexes and clustering partitions. These results also indicate that the designed strategies positively influence the performance improvement of the original BFO.

Keywords : data clustering, multi-objective optimization, bacterial foraging optimization, learning archives
Conference Title : ICECDM 2022 : International Conference on Evolutionary Computation and Data Mining
Conference Location : Sydney, Australia
Conference Dates : August 30-31, 2022