Hybrid Differential Evolution Algorithm for Optimal Clustering

Abstract: The problem of optimal non-hierarchical clustering is addressed. A new algorithm combining differential evolution and $k$-means is proposed and tested on eight well-known real-world data sets. The classification of objects to be optimized is encoded by the cluster centres in differential evolution (DE) algorithm. A new efficient heuristic for this rearrangement was also proposed. The plain DE variants with and without the rearrangement were compared with corresponding hybrid $k$-means variants. The experimental results showed that hybrid variants with $k$-means algorithm are essentially more efficient than the non-hybrid ones. Compared to a standard $k$-means algorithm with restart, the new hybrid algorithm appeared more reliable and efficient, especially in more difficult tasks. The results for TRW and VCR criterion were compared. Both criteria provided the same optimal partitions and no significant differences were found in efficiency of the algorithms using these criteria.

Doc. Ing. Josef Tvrdík, CSc., KIP PřF OU
Prof. RNDr. Ing. Ivan Křivý, CSc., KIP PřF OU