

عنوان مقاله:

Clustering a Big Mobility Dataset Using an Automatic Swarm Intelligence-Based Clustering Method

محل انتشار:

مجله نوآوری های مهندسی برق و کامپیوتر، دوره 6، شماره 2 (سال: 1397)

تعداد صفحات اصل مقاله: 20

نویسندگان:

Iman Behravan - *Department of Electrical Engineering, PhD student, University of Birjand, i.behravan@birjand.ac.ir*

Seyed Hamid Zahiri - *Department of Electrical Engineering, Faculty of Engineering, University of Birjand, hzahiri@birjand.ac.ir*

,Seyed Mohammad Razavi - *Department of Electrical Engineering, Faculty of Engineering, University of Birjand*

Roberto Trasarti - *KDD lab, ISTI-CNR, Pisa, Italy, roberto.trasarti@isti.cnr.it*

خلاصه مقاله:

Big data referred to huge datasets with high number of objects and high number of dimensions. Mining and extracting big datasets is beyond the capability of conventional data mining algorithms including clustering algorithms, classification algorithms, feature selection methods and etc. Clustering, which is the process of dividing the data points of a dataset into different groups (clusters) based on their similarities and dissimilarities, is an unsupervised learning method discovers useful information and hidden patterns from raw data. K-means yet is an efficient clustering algorithm but it suffers from some drawbacks. It has a tendency to converge to a local optimum point, its output result depends on its initial value of cluster centers and it is unable in finding the number of clusters. In this research a new clustering method for big datasets is introduced based on Particle Swarm Optimization (PSO) algorithm. PSO is a heuristic algorithm with high ability in searching the solution space and finding the global optimum point. The proposed method is a two-stage algorithm which first searches the solution space for proper number of clusters and then searches to find the position of the centroids. Its performance is evaluated on 13 synthetics and a biological microarray dataset. Finally, 2 real big mobility datasets, are investigated and analyzed using the proposed clustering method.

کلمات کلیدی:

Big data clustering, Bobility dataset, K-means, Swarm intelligence, Particle swarm optimization

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/929881>

