International Journal of Mobile Computing and Multimedia Communications
Volume 12 • Issue 1 • January-March 2021

An Improved Gravitational Clustering Based on Local Density

Lei Chen, School of Information and Electrical Engineering, Hunan University of Science and Technology, China
Qinghua Guo, School of Information and Electrical Engineering, Hunan University of Science and Technology, China
Zhaohua Liu, State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, China
Long Chen, School of Information and Electrical Engineering, Hunan University of Science and Technology, China
HuiQin Ning, School of Information and Electrical Engineering, Hunan University of Science and Technology, China
Youwei Zhang, School of Information and Electrical Engineering, Hunan University of Science and Technology, China
Yu Jin, School of Information and Electrical Engineering, Hunan University of Science and Technology, China

ABSTRACT

Gravitational clustering algorithm (Gravc) is a novel and excellent dynamic clustering algorithm that can accurately cluster complex dataset with arbitrary shape and distribution. However, high time complexity is a key challenge to the gravitational clustering algorithm. To solve this problem, an improved gravitational clustering algorithm based on the local density is proposed in this paper, called FastGravc. The main contributions of this paper are as follows. First of all, a local density-based data compression strategy is designed to reduce the number of data objects and the number of neighbors of each object participating in the gravitational clustering algorithm. Secondly, the traditional gravity model is optimized to adapt to the quality differences of different objects caused by data compression strategy. And then, the improved gravitational clustering algorithm FastGravc is proposed by integrating the above optimization strategies. Finally, extensive experimental results on synthetic and real-world datasets verify the effectiveness and efficiency of FastGravc algorithm.

KEYWORDS

Data Compression, Dynamic Clustering, Gravitational Clustering, Local Density

INTRODUCTION

Clustering is an indispensable and very important method for mining complex real-world data. It uses an unsupervised way to reveal the hidden rules and patterns of human society. In the past 20 years, a large number of excellent clustering algorithms have been proposed, applied, improved and further optimized. Overall, these algorithms can be simply divided into the following categories: partitioned clustering, hierarchical clustering, density clustering, and dynamic clustering (Saxena A et al., 2017). (1) The partitioned clustering and hierarchical clustering is the most commonly and most widely used algorithms, K-Means and BIRCH are the typical cases. This algorithm can obtain excellent clustering accuracy on regular, noise-free dataset like a circle or ellipse, and the time efficiency is very high(Stevan N et al., 2014). But, on irregular and non-uniform datasets like non-circular, the clustering accuracy of this algorithm is not satisfactory. (2) The density clustering and graph clustering are also commonly used excellent clustering algorithm, DbScan and SpectralClustering are typical cases respectively. These two kinds of algorithms are suitable for various datasets, and can achieve

DOI: 10.4018/IJMCMC.2021010101

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

excellent clustering performance on irregular and uneven datasets. However, these algorithms require longer clustering time and cannot accurately identify the noise in the dataset. (3) Dynamic clustering algorithm is a novel and outstanding clustering algorithm (Bae J et al., 2020), Gravc is a typical case. The basic idea of this algorithm is to extract a dynamic process from natural phenomena such as gravitation, synchronization, and evolution, and use it to cluster complex dataset. This algorithm can achieve good clustering accuracy on irregular and uneven dataset, and can accurately identify noise and abnormal data in the dataset (Chen L et al., 2017). But, due to the dynamic clustering process, the time complexity of this algorithm is very high.

With the coming of the era of big data, more and more complex data have emerged, such as the wireless and mobile multimedia network(Ajay K et al., 2019), transportation, weather, wireless sensor network(Surender S et al., 2011), etc. These complex new data bring some new challenges to the traditional clustering algorithm from the clustering accuracy and time efficiency (Fahad A et al., 2014). For the clustering accuracy aspect, complex data in the era of big data present some new features such as irregularity, unevenness and high noise. These features will cause the clustering accuracy of traditional clustering algorithms to seriously deteriorate (Shirkhorshidi A S et al., 2014). For the time efficiency, the increasing scale of complex data puts forward higher demand on the time efficiency of traditional clustering algorithm (Mohebi A et al., 2016).

Based on the above challenges, it is not difficult to find that, the gravitational dynamic clustering algorithm is very suitable for the clustering needs of complex data in the era of big data. It can accurately cluster complex dataset of any shape, density and distribution, and can flexibly and correctly find noise and abnormal points in the dataset. However, the time efficiency of the gravitational clustering algorithm needs to be further optimized to adapt to the clustering needs of complex data in the era of big data. To solve this problem, an improved gravity clustering algorithm based on local density is proposed in this paper, called FastGravc. The basic idea of FastGravc algorithm is as follows. ***Firstly***, a local density-based data compression is designed to merge multiple dense data points into a "big" point with greater quality. Based on this strategy, the number of data objects participating in the gravitational clustering algorithm and the number of neighbors of each object will be greatly reduced, thereby improving the time efficiency of the algorithm. ***Secondly***, different data objects will have different qualities after data compression. For this reason, the gravity model is enhanced to deal with differences in the quality of data objects. ***Finally***, based on the above optimization strategy and improved model, a dynamic interaction process is started to get the final clustering result. **Figure 1** shows the base idea of local density-based data compression strategy. In the figure, the data distribution of a synthetic dataset containing two clusters is presented as shown in **Figure 1(a)**, where data points of the same color belong to the same cluster. It is not difficult to find that, this dataset has multiple dense areas bounded by blue solid in **Figure 1(a)**. Each dense area contains a large number of data points, and these data points must belong to the same cluster. That is to say, all data points in a dense area can be compressed into a "big" data point with higher-quality, as shown by P1, P2, P3 in **Figure 1(b)**. When the dense area is compressed, the number of data points in the data set will be greatly reduced, and the number of objects participating in the gravitational clustering algorithm will also be greatly reduced, thereby effectively improving the time efficiency of the algorithm.

In summary, the main contributions of this paper are as follows:

- A local density-based data compression strategy is designed to compress each dense area into a "big" data point with higher-quality, thereby reducing the number of objects participating in the gravitational clustering algorithm and improving the time efficiency of the algorithm.
- The traditional gravity model is enhanced to solve the difference in the quality of data points caused by data compression strategy.
- Integrating the above two optimizations, an improved gravitational clustering algorithm called FastGravc is proposed to use a dynamic interactive process to get the final clustering of complex dataset.

20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/an-improved-gravitational-clustering-based-on-local-density/271385

Related Content

Supporting the Mobile In-situ Authoring of Locative Media in Rural Places: Design and Expert Evaluation of the SMAT app
Keith Cheverst, Trien V. Doand Dan Fitton (2015). *International Journal of Handheld Computing Research (pp. 1-19).*
www.irma-international.org/article/supporting-the-mobile-in-situ-authoring-of-locative-media-in-rural-places/138112

Jammer Location-Oriented Noise Node Elimination Method for MHWN
Jianhua Fan, Qiping Wang, Xianglin Weiand Tongxiang Wang (2014). *International Journal of Mobile Computing and Multimedia Communications (pp. 1-19).*
www.irma-international.org/article/jammer-location-oriented-noise-node-elimination-method-for-mhwn/144442

DSOA: A Service Oriented Architecture for Ubiquitous Applications
Fabricio Nogueira Buzeto, Carlos Botelho de Paula Filho, Carla Denise Castanhoand Ricardo Pezzuol Jacobi (2011). *International Journal of Handheld Computing Research (pp. 47-64).*
www.irma-international.org/article/dsoa-service-oriented-architecture-ubiquitous/53856

Application of Fuzzy Logic for Slice QoS in 5G Networks: A Comparison Study of Two Fuzzy-Based Schemes for Admission Control
Phudit Ampririt, Ermioni Qafzezi, Kevin Bylykbashi, Makoto Ikeda, Keita Matsuoand Leonard Barolli (2021). *International Journal of Mobile Computing and Multimedia Communications (pp. 18-35).*
www.irma-international.org/article/application-of-fuzzy-logic-for-slice-qos-in-5g-networks/277230

Multimodality in Mobile Applications and Services
M. Caschera (2007). *Encyclopedia of Mobile Computing and Commerce (pp. 675-681).*
www.irma-international.org/chapter/multimodality-mobile-applications-services/17155