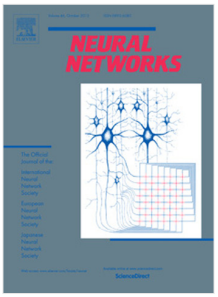
Accepted Manuscript



Novel density-based and hierarchical density-based clustering algorithms for uncertain data

Xianchao Zhang, Han Liu, Xiaotong Zhang

PII: S0893-6080(17)30140-5
DOI: http://dx.doi.org/10.1016/j.neunet.2017.06.004
Reference: NN 3769

To appear in: *Neural Networks*

Received date : 10 March 2017
Revised date : 22 May 2017
Accepted date : 8 June 2017

Please cite this article as: Zhang, X., Liu, H., Zhang, X., Novel density-based and hierarchical density-based clustering algorithms for uncertain data. *Neural Networks* (2017), http://dx.doi.org/10.1016/j.neunet.2017.06.004

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

*Manuscript
Click here to view linked References

ACCEPTED MANUSCRIPT

Novel Density-based and Hierarchical Density-based Clustering Algorithms for Uncertain Data

Xianchao Zhang*[a,b], Han Liu[a,b], Xiaotong Zhang[a,b]

[a] *Dalian University of Technology, Dalian 116024, China*
[b] *Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, Dalian 116024, China*

Abstract

Uncertain data has posed a great challenge to traditional clustering algorithms. Recently, several algorithms have been proposed for clustering uncertain data, and among them density-based techniques seem promising for handling data uncertainty. However, some issues like losing uncertain information, high time complexity and nonadaptive threshold have not been addressed well in the previous density-based algorithm FDBSCAN and hierarchical density-based algorithm FOPTICS. In this paper, we firstly propose a novel density-based algorithm PDBSCAN, which improves the previous FDBSCAN from the following aspects: (1) it employs a more accurate method to compute the probability that the distance between two uncertain objects is less than or equal to a boundary value, instead of the sampling-based method in FDBSCAN; (2) it introduces new definitions of probability neighborhood, support degree, core object probability, direct reachability probability, thus reducing the complexity and solving the issue of nonadaptive threshold (for core object judgement) in FDBSCAN. Then, we modify the algorithm PDBSCAN to an improved version (PDBSCANi), by using a better cluster assignment strategy to ensure that every object will be assigned to the most appropriate cluster, thus solving the issue of nonadaptive threshold (for direct density reachability judgement) in FDBSCAN. Furthermore, as PDBSCAN and PDBSCANi have difficulties for clustering uncertain data with non-uniform cluster density, we propose a novel hierarchical density-based algorithm POPTICS by extending the definitions of PDBSCAN, adding new definitions of fuzzy core distance and fuzzy reachability distance, and employing a new clustering framework. POPTICS can reveal the cluster structures of the datasets with different local densities in different regions better than PDBSCAN and PDBSCANi, and it addresses the issues in FOPTICS. Experimental results demonstrate the superiority of our proposed algorithms over the existing algorithms in accuracy and efficiency.

Keywords: Clustering, uncertain data, density-based algorithm, hierarchical density-based algorithm

☆A preliminary version of this paper has been presented in proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI 2014) [1].

*Corresponding author.
Email addresses: xczhang@dlut.edu.cn (Xianchao Zhang*), liu.han.dut@gmail.com (Han Liu), zxt.dut@hotmail.com (Xiaotong Zhang)

دریافت فوری ← متن کامل مقاله

ISIArticles
مرجع مقالات تخصصی ایران

✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
✓ امکان دانلود نسخه ترجمه شده مقالات
✓ پذیرش سفارش ترجمه تخصصی
✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
✓ دانلود فوری مقاله پس از پرداخت آنلاین
✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات