

**Complex Methods Applied to Data Analysis, Processing, and
Visualisation**

Author

Garcia-Rodriguez, J, Angelopoulou, A, Tomás, D, Lewis, A

Published

2019

Journal Title

Complexity

Version

Version of Record (VoR)

DOI

[10.1155/2019/9316123](https://doi.org/10.1155/2019/9316123)

Rights statement

© 2019 Jose Garcia-Rodriguez et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from

<http://hdl.handle.net/10072/387009>

Griffith Research Online

<https://research-repository.griffith.edu.au>

Editorial

Complex Methods Applied to Data Analysis, Processing, and Visualisation

Jose Garcia-Rodriguez ¹, Anastasia Angelopoulou,² David Tomás ¹ and Andrew Lewis³

¹University of Alicante, Alicante, Spain

²University of Westminster, London, UK

³Griffith University, Brisbane, Australia

Correspondence should be addressed to Jose Garcia-Rodriguez; jgarcia@dtic.ua.es

Received 30 April 2019; Accepted 30 April 2019; Published 11 June 2019

Copyright © 2019 Jose Garcia-Rodriguez et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The amount of data available every day is not only enormous but growing at an exponential rate. Over the last ten years there has been an increasing interest in using complex methods to analyse and visualise massive datasets, gathered from very different sources and including many different features: social networks, surveillance systems, smart cities, medical diagnosis systems, business information, cyberphysical systems, and digital media data. Nowadays, there are a large number of researchers working in complex methods to process, analyse, and visualise all this information, which can be applied to a wide variety of open problems in different domains. This special issue presents a collection of research papers addressing theoretical, methodological, and practical aspects of data processing, focusing on algorithms that use complex methods (e.g., chaos, genetic algorithms, cellular automata, neural networks, and evolutionary game theory) in a variety of domains (e.g., software engineering, digital media data, bioinformatics, health care, imaging and video, social networks, and natural language processing). A total of 27 papers were received from different research fields, but sharing a common feature: they presented complex systems that process, analyse, and visualise large amounts of data. After the review process, 8 papers were accepted for publication (around 30% of acceptance ratio).

These papers can be organised in different groups. The focus of the first group of articles is time series. The paper titled “LMC and SDL Complexity Measures: A Tool to Explore Time Series” by J. Piqueira and S. Mattos presented a generalisation of LMC (López-Ruiz, Mancini and Calbet) and SDL (Shiner, Davison and Landsberg) complexity measures,

considering that the state of a system or process is represented by a continuous temporal series of a dynamical variable. As the two complexity measures are based on the calculation of informational entropy, an equivalent information source was defined by using partitions of the dynamical variable range. During the time intervals, the information associated with the measured dynamical variable was the seed to calculate instantaneous LMC and SDL measures. To show how the methodology worked generating indicators, two examples concerning meteorological data and economic data were presented and discussed. Another accepted work dealing with time series is “Improved Permutation Entropy for Measuring Complexity of Time Series under Noisy Condition”, presented by Z. Chen et al. This paper proposes an improved permutation entropy method (IPE) as a tool to measure and analyse complexity of time series combining some advantages of previous modifications of PE. Its effectiveness was validated through both synthetic and experimental analysis, overcoming PE limitations such as its low performance under noisy conditions.

The second group of publications includes works dealing with sensing data and image recognition. The paper by J. Guo et al., entitled “Activity Feature Solving Based on TF-IDF for Activity Recognition in Smart Homes”, presents an activity feature solving strategy based on TF-IDF. In smart homes based on the internet of things, daily activity recognition aims to know resident’s daily activity in a noninvasive manner. The performance of daily activity recognition heavily depends on solving strategy of activity feature. However, the current common employed solving strategy based on statistical

information of individual activity does not support well the activity recognition. The proposal by Guo et al. exploits statistical information related to both individual activity and the whole of activities. Two distinct datasets were commissioned to mitigate the effects of coupling between datasets and sensor configuration. A number of traditional machine learning and deep learning techniques were evaluated to assess the performance of the method proposed for residents activity recognition. The second paper in this group is “MI-based Robust Waveform Design in Radar and Jammer Games”, written by B. Wang et al. Due to the uncertainties of the radar target prior information in the actual scene, the waveform designed based on the radar target prior information cannot meet the needs of parameter estimation. To improve the performance of parameter estimation, Wang et al. presents a novel transmitted waveform design method under the hierarchical game model of radar and jammer. This approach maximises the mutual information between the radar target echo and the random target spectrum response. Another work in this group is “A Novel Semi-Supervised Learning Method Based on Fast Search and Density Peaks”. This paper by F. Gao et al. address the problem of radar image recognition. Recognition algorithms achieve good classification results under the condition of sufficiently labelled samples, but labelled samples are scarce and costly to obtain. The main issue faced in this paper is how to use unlabelled samples to improve the performance of a recognition algorithm when the number of available labelled samples is limited. Unlike previous semisupervised learning methods, this work does not use unlabelled samples directly, but looks for safe and reliable samples before using them. The authors proposed two new semisupervised learning methods: one based on fast search and density peaks (S2DP) and the other on iterative S2DP. Finally, F. Zhao et al. propose in “Two-Phase Incremental Kernel PCA for Learning Massive or Online Datasets” a specific kernel PCA (KPCA) that can incorporate data into KPCA in an incremental way. This fact overcame typical drawbacks of KPCA when handling massive or online datasets. They tested their proposal in a synthesised dataset and in the classical MNIST database of handwritten digits images.

The last group of papers includes research in social impact domains such as economics and education. A. Herrero et al. present in “Hybrid Unsupervised Exploratory Plots: a Case Study of Analysing Foreign Direct Investment” a new visualisation technique, called HUEP. This proposal for descriptive data analysis combines the outputs of exploratory projection pursuit and clustering methods in a novel and informative way. As a case study, HUEP was validated in a real-world context for analysing the internationalisation strategy of companies by taking into account bilateral distance between home and host countries. As a multifaceted concept, distance encompasses multiple dimensions. Together with data from both the countries and the companies, various psychic distances were analyzed by means of HUEP, gaining deep knowledge about the internationalization strategy of large Spanish companies. Informative visualizations were obtained from the analyzed dataset, leading to useful business implications and decision making. The last paper in this issue, written

by A. Hernández-Blanco et al., is focused on the educational domain. “A Systematic Review of Deep Learning Approaches to Educational Data Mining” surveys the research carried out in deep learning techniques applied to educational data mining (EDM) from its origins to the present day. The main goals of this study are to identify the EDM tasks that have benefited from deep learning techniques and those that are pending to be explored, to describe the main datasets used in this research area, to provide an overview of the key concepts, main architectures, and configurations of deep learning and its applications to EDM, and to discuss current state-of-the-art and future directions on this research field.

Conflicts of Interest

The authors declare that they have no conflicts of interest

Acknowledgments

This work has been funded by the Spanish Government TIN2016-76515-R grant for the COMBAHO project, supported with FEDER funds. We would like to thank Central University of Ecuador and in particular Jaime Salvador-Meneses and Zoila Ruiz for their participation and support in managing this special issue.

*Jose Garcia-Rodriguez
Anastasia Angelopoulou
David Tomás
Andrew Lewis*