Computing (2020) 102:741–743
https://doi.org/10.1007/s00607-020-00793-x

EDITORIAL

Editorial for the special issue on big time series data

Lizhe Wang[1] · Dan Chen[2] · Albert Zomaya[3] · Laurence T. Yang[4] · Dimitrios Georgakopoulos[5]

Published online: 13 February 2020

Most phenomena in the world exhibit certain degree of complexity as a contrast to the simplicity of the basic laws of physics. From earth to human brain, the complex systems behind such phenomenon evolve in a nonlinear and non-stationary manner. Big time series data record the evolvement of a complex system(s) in large temporal and spatial scales with great details of the interactions amongst different parts of the system. Those generally appear in the form of a list of multiple variables through time, e.g., time sequences of scalp potentials from many electrodes with numerous subjects in a city.

However, exploration of the overall mechanism has been hindered by the notoriously high dimensionality and scale of big time series data as well as the enormously complicated interdependencies amongst data elements. The issue becomes even more challenging under the circumstance of insufficient a priori knowledge. The "Big Mechanism" program initiated by DARPA is a milestone

✉ Dan Chen
dan.chen@whu.edu.cn

Lizhe Wang
lizhe.wang@computer.org

Albert Zomaya
albert.zomaya@sydney.edu.au

Laurence T. Yang
ltyang@stfx.ca

Dimitrios Georgakopoulos
dgeorgakopoulos@swin.edu.au

[1] China University of Geosciences, 388 Lumo Road, Wuhan, Beijing, People's Republic of China

[2] School of Computer Science, Wuhan University, 299 Bayi Road, Wuhan, People's Republic of China

[3] Centre for Distributed and High Performance Computing (Director), School of Information Technologies, The University of Sydney, Sydney, Australia

[4] Department of Computer Science, St. Francis Xavier University, Antigonish, Canada

[5] School of Software and Electrical Engineering, Swinburne University of Technology, Hawthorn, VIC, Australia

centering on this problem, which aims at exploration of small and dispersed mechanisms to be integrated into a knowledge base to "understand the mechanisms of big and complex system."

The recent leap-forward of data centric approaches rooted from both academia and industry provides an unprecedented opportunity in finding the low-dimensional and "small" representations thus to predict the system's behavior via computational intelligence. Since the computing issues are highly interdisciplinary and cover various topics, a special issue of *Computing* provides the ideal forum for presenting and discussing the latest research. This special issue aims to present outstanding research results in regard to algorithms, methods and high-performance computing frameworks for underlying mechanism discovery of big time series data. We have received 17 manuscript submissions in total; of these, 5 papers have been accepted after several rounds of very constructive and deep reviews. These papers are recapped as the follows:

1. Modern early warning systems for natural hazards and urban vulnerabilities (e.g., Landslides) play a significant role in mitigation and minimizing loss of life and damage to infrastructure based on varieties of time series data sources. However, the automatic discovery of potential sources has become a challenge due to the complexity and high variety of data sources. To solve the research problem, Phengsuwan et al. propose a landslip ontology for automatic verification and prediction of landslide hazards. The ontology integrates time series data from multiple heterogeneous sources and establishes relationship between them. The method achieves an accurate discovery mechanism of potential sources and supports early warning of landslide.
2. The boom of big data in education has provided an unrivalled opportunity for educators to evaluate the learners' cognitive state. Previous research works on cognitive state analysis focus on attention, ignoring the roles of emotion in human learning. Therefore, Chen et al. develop an emotion-sensitive learning framework for cognitive state analysis. The framework adopts a multi-task learning mechanism for face detection, landmark location and head pose estimation, and ranking CNN for facial expression recognition and intensity evaluation. High correctness rates have been achieved in analysis of attention and emotion.
3. Effective indoor localization largely relies on the fingerprint database of received signal strength in connection with radio frequency sources. It is challenging to construct fingerprint database for accurate localization as received signal strength must be captured via a full space scan from one point to another every few meters in a certain period of time. In order to tackle this problem, Ai et al. introduce a fast fingerprints construction method. The method requires sparse sampling via Gauss process regression (GPR) and reduces outliers to improve reliability of fingerprint database. The method enables fast fingerprint database construction and accurate localization. It offers potential for the construction of large-scale indoor/outdoor fingerprint database in smart city.
4. With rapid development of sensor and satellite technologies, large amounts of remote sensing image data are available and image registration is the key process

for the generation of these image mosaics. However, mainstream techniques need to consume vast amounts of computing resources to process such images, and these result in inefficiency and low precision. Hence, Zeng et al. have proposed the mosaic framework for forestry image registration. In the framework, image features are adaptively obtained through deep convolutional neural networks, then similarity between features will be computed by correlation filter. High effectiveness and robustness are achieved in the cases of various weather and seasonal conditions.

5. Switched audio codec has been proved to be efficient for compressing a large range of audio signals at low bit rates. However, coding quality strongly relies on an exact classification of the input signals and existing works are suffered from either high computation complexity or unsatisfactory coding quality. Tu et al. has introduced RNN-based signal classification method for hybrid audio data compression, where temporal information of sequences is preserved via RNN and discriminative features can then be selected by coding parameters. The method ensures satisfactory performance in mode selection accuracy, coding quality and computational complexity.

For this special issue of *Computing*, we have selected the interesting papers described above to represent some important advances in predictive analysis of big time series data. In conclusion, it should be noted that research along this direction has been growing fast and many new formulations spanning multiple disciplines are being formed. Thus, innovative interdisciplinary techniques, together with high performance algorithms, are highly desirable in studying big time series data from a computing perspective. Critical applications in biosciences and medicines will be a hotspot.