

Divide-and-Conquer Signal Processing, Feature Extraction, and Machine Learning for Big Data

The demand for distributed analysis gradually increases as the size of data has become increasingly large. Take social network services for example, the number of users on Facebook[®] and Twitter respectively reaches 1.3 billion and 0.6 billion by the end of 2013 according to the survey at www.statisticbrain.com. In the Very-Large-Scale Integration (VLSI) area, a modern design in 2013 contains approximately 5.0 billion transistors in a single chip. Such a scale of data sizes is way beyond the capability of a single computer or even a workstation. In view of this change, many famous and off-the-shelf tools, such as Apache[™] Hadoop[®] and Apache[™] Spark[®], are developed to handle large-scale analysis by using computer clusters. These tools usually employ divide-and-conquer architectures in their implementations. With such architectures, when a large-scale dataset is segmented into subsets, the original problem can accordingly be divided into subproblems, separately processed in each computer. The concept of distributed processing corresponds with the essence of divide-and-conquer designs. This concept is also called “Map and Reduce,” especially when described in Hadoop[®]. Although these tools provide developers with a convenient and distributed communication protocol between computers, they do not automatically generate divide-and-conquer algorithms. Distribution of computing subtasks and collection of the results still involve programming by developers. Moreover, not every algorithm can be converted into a divide-and-conquer version and gives optimal solutions. This somehow decreases the development of big data processing. Therefore, how to design algorithms for the divide-and-conquer architecture is of priority concerns.

In response to the aforementioned problem, this special issue particularly focuses on divide-and-conquer signal processing, feature extraction, and machine learning approaches for large-scale data. Meanwhile, divide-and-conquer methods for generating approximate solutions, or local optimal results in contrast with global solutions, are also highlighted in this special issue.

Via this issue, we call upon specialists in the science and engineering domains, which will advance the state-of-the-art technologies in big data, to contribute their creativity to this domain. Research areas relevant to the special issue include, but are not limited to, the following topics.

- Distributed audiovisual signal processing and feature extraction based on divide-and-conquer architectures
- Distributed machine learning approaches based on divide-and-conquer architectures
- Data mining algorithms based on divide-and-conquer architectures
- Hadoop[®] applications based on big data

Submissions must not have been previously published, with the exception that substantial extensions of conference papers can be considered. The authors will be required to follow the Author's Guide for manuscript submission to the Neurocomputing at: <http://www.journals.elsevier.com/neurocomputing/>. Full manuscripts should be submitted electronically through the manuscript track system. Authors need to choose “**DC-ML for Big Data**” when selecting the article type.

Proposed Schedule:

Submission deadline:	December 1, 2014
Notification of acceptance:	February 1, 2015
Final manuscript due:	March 15, 2015
Tentative publication date:	April 15, 2015

Lead Guest Editor:

Daniel Bo-Wei Chen, Department of Electrical Engineering, Princeton University, USA (corresponding author) (dennisbwc@gmail.com; boweic@princeton.edu)

Guest Editors:

1. Wen Ji, Institute of Computing Technology, Chinese Academic of Sciences, China (jiwen@ict.ac.cn)
2. Seungmin Rho, Department of Multimedia, Sungkyul University, Korea (korea.smrho@gmail.com)