

Guest Editorial Preface

Special Issue on Heterogeneous Big Data Analytics and Cloud Computing, Part 1

Ruomei Wang, Sun Yat-sen University, Guangzhou, China

Xiangjian He, University of Technology Sydney, Sydney, Australia

Songhua Xu, New Jersey Institute of Technology, Newark, USA

Propelled by the rapid advancement of massive storage devices and delivery technologies, both the amount and the diversity of consumer-produced data on the web are growing at an incredible rate. Heterogeneous data such as videos, images, graphics and text from multiple sources embodies different feature spaces, which leads to poor accessibility without using efficient and effective technologies and solutions. Grid and cloud computing have been one of most popular research topics in recent year. These technologies are very important since they provide a reliable, fault-tolerant and scalable environment to perform specific tasks and access geographically distributed and heterogeneous data on demand. The special issue is in two parts. This issue (part 1) includes six papers which can be grouped in three categories.

On the theoretical basis and applications of heterogeneous big data, the work *A Hybrid Imputation Method Based on Denoising Restricted Boltzmann Machine* by Jiang Xu, Siqian Liu, Zhikui Chen and Yonglin Leng proposes a hybrid method to fill incomplete data. Jinfei Yang and others, in the paper *A highly efficient big data mining algorithm based on stock market*, present a framework for predicting economic development level by mining big stock data.

In the field of computer vision and image processing, Ying Zhu and others introduce a new geometric method for 3D reconstruction based on images. A new surface is first developed named as rational W-spline surface via tensor product. By giving the new control matrix and the change of the parameters based on the images, 3D reconstruction effect can be obtained. The method has the advantage of low data storage requirement and high computation process efficiency. Xiaohong Shi and others present a regional color editing method based on multi-cues manipulation, including interactive segmentation, inpainting and gradient-preserving optimization. The proposed method achieves a visual satisfactory local color propagation results, but also preserve the texture details well.

How to manage services and heterogeneous big data efficiently is an important issue in cloud computing. To achieve workload consolidation, it is often necessary to detect the overloaded hosts by calculating upper threshold in cloud computing. In the paper, *An Approach for Detection of Overloaded Host to Consolidate Workload in Cloud Datacenter*, the authors use moving range (MR) method of variables control charts to obtain upper threshold, which shows better MR performance in terms of reduction in SLA violation and minimization in VM migration. Jun Li and Mengshu Hou, in a paper *Improving Data Availability for Deduplication in Cloud Storage*, introduce a deduplication

technology in the cloud storage to reduce the amount of data. It is based on the data chunk reference count and access frequency, and increases redundant information for the data chunks to ensure data availability and minimize storage overhead.

Ruomei Wang
Xiangjian He
Songhua Xu
Guest Editors
IJGHPC