

Special Issue on Parallel and Distributed Data Processing

Guest Editorial

As an exponentially increasing amount of data is generated everyday, parallel and distributed data processing has emerged as a key enabling technology and plays more and more important roles in modern computing and information technologies. In recent years, many researchers have become increasingly interested in the field of parallel and distributed data processing, and a large number of remarkable academic achievements have been published and applied to various domains.

In this special issue, we selected some excellent papers from the third International Symposium on Parallel Architectures, Algorithms and Programming (PAAP 2010), which was held in Dalian, China, December 18-20, 2010. In addition, we invited and selected some representative research papers in the broad area of parallel and distributed data processing.

The paper titled "A Translation Framework for Executing the Sequential Binary Code on CPU/GPU Based Architectures" proposes a novel translation framework for constructing the virtual execution environment aiming at accelerating the process of DBT on CPU/GPU based architectures. With parallelizable parts (hot spots) of binary code and their related information, the framework converts the sequential code into PTX form and executes them on GPUs.

The paper titled "A Dynamic-Static Combined Code Layout Reorganization Approach for Dynamic Binary Translation" presents a dynamic-static combined approach to reorganize the layout of software cache. Under this approach, we first employ an emulating execution to collect the profile information and the translated target code.

The paper titled "An Efficient Hybrid Clustering-PSO Algorithm for Anomaly Intrusion Detection" demonstrates IDCPSO algorithm, which combines an unsupervised clustering algorithm with the PSO technique to optimize the clustering results and obtain the optimal detection result with unlabeled data.

The paper titled "A Hybrid Method for XML Clustering by Structure and Content" presents an effective XML cluster method based both structural similarity calculated using method PFWLCS with position frequency weight, based on frequency path model, and content information contained in XML files.

The paper titled "Hybrid Distributed Shared Memory Space in Multi-core Processors" observes that it is unnecessary to perform V2P address translation for private data accesses and introduces hybrid DSM organization and run-time partitioning technique in order to improve the system performance by reducing V2P address translation overhead as much as possible.

The paper titled "Design and Evaluation of an Online Anomaly Detector for Distributed Storage Systems" exploits the stable relationship between workloads and system resource statistics to detect the performance anomaly and identify faulty sources which cause the performance anomaly in the system.

The paper titled "Multi-pattern matching with wildcards" presents efficient algorithms of multi-pattern matching with wildcards based on the fast Fourier transform and demonstrates an FFT implementation based on the modular arithmetic for machines with 64-bit word.

The paper titled "RPPA: A Remote Parallel Program Performance Analysis Tool" presents a remote parallel program performance analysis tool, RPPA (Remote Parallel Performance Analyzer), which is based on dynamic code instrumentation.

The paper titled "Topic Detection with Hypergraph Partition Algorithm" proposes SMHP (Similarity Matrix based Hypergraph Partition) algorithm, which aims at improving the efficiency of Topic Detection.

The paper titled "Query by Humming Systems Using Melody Matching Model Based on the Genetic Algorithm" discusses a system called Query by humming (QBH). QBH refers to music information retrieval systems where short audio clips of singing or humming act as queries. This paper proposes QBH using melody matching model based on the genetic algorithm and improving the ranking result by local sensitive hashing algorithm.

The paper titled "Segmenting Webpage with Gomory-Hu Tree Based Clustering" proposes a novel web page segmentation algorithm based on finding the Gomory-Hu tree in a planar graph. The algorithm firstly distills vision and structure information from a web page to construct a weighted undirected graph, whose vertices are the leaf nodes of the DOM tree and the edges represent the visible position relationship between vertices. Then it partitions the graph with Gomory-Hu tree based clustering algorithm.

The paper titled "A Two-Dimension XML Encoding Method based on Variable Length Binary Code" considers the issue of the insertion operation of intermediate nodes into XML documents using update-friendly XML labeling schemes, and proposes an efficient two-dimension XML encoding method based on the variable length binary code, DO-VLEI code.

The paper titled "A Hop-by-hop Cross-layer Congestion Control Scheme for Wireless Sensor Networks" addresses congestions in wireless sensor networks, and presents a hop-by-hop cross-layer congestion control scheme built on contention-based MAC protocol. Simulations have been conducted to compare the proposed scheme against closely-related existing schemes.

The paper titled "Role Assorted Community Discovery for Weighted Networks" considers the difficulties in

community discovery. The authors show that this problem can be solved according to the role assorted method which gives distinguish labels to vertices in the same community. The method leads to a number of possible algorithms for detecting community structures in both unweighted and weighted networks.

The paper titled “A Survey on Particle Swarm Optimization Algorithms for Multimodal Function Optimization” gives a comprehensive review of existing works done in the field of multimodal function optimization as well as a critical analysis of the existing methods.

It has been a great pleasure to run this special issue, which reveals important research results in the field of parallel and distributed data processing. We would like to thank Prof. Kassem Saleh, Editor-in-Chief of Journal of Software, and Dr. George J. Sun, Executive Editor of Academy Publisher, for giving us the opportunity to organize this special issue and for their great help in the organization of this issue. We thank all authors for their submissions and all reviewers for their diligent work in evaluating these submissions. We sincerely hope that you enjoy reading these distinguished papers.

Guest Editors

Xianchao Zhang

School of Software, Dalian University of Technology, China

Email: xczhang@dlut.edu.cn

Feng Xia

School of Software, Dalian University of Technology, China

Email: f.xia@ieee.org



Xianchao Zhang is a Full Professor at Dalian University of Technology, China. He received his B.S degree in Applied Mathematics and M.S. degree in Computational Mathematics from National university of Defense Technology in 1994 and 1998, respectively. He received his Ph.D. in Computer Theory and Software from University of Science and Technology of China in 2000. He joined Dalian University of Technology in 2003 after 2 years of industrial working experience at international companies. He worked as Visiting Scholar at The Australian National University and The City University of Hongkong in 2005 and 2009, respectively. His research interests include Algorithms, Machine Learning, Data Mining and Information Retrieval.



Feng Xia received the B.E. and Ph.D. degrees from Zhejiang University, China, in 2001 and 2006, respectively. He was a Research Fellow at Queensland University of Technology, Australia, from 2007 to 2008. Before joining Dalian University of Technology (DUT) in Mar 2009, with which he is currently an Associate Professor and PhD Supervisor, Dr. Xia was an Assistant Professor at Zhejiang University. He is on the Editorial Board of several international journals. He serves as General Chair, PC Chair, Workshop Chair, Publicity Chair, or PC Member of a number of conferences. He is also the Guest Editor of several journal special issues. Dr. Xia has (co-)authored one book and over 80 scientific papers. His research interests include cyber-physical systems (CPS), mobile computing, and social computing. He is a member of IEEE and ACM.