

Editorial

Scientific Programming Techniques and Algorithms for Data-Intensive Engineering Environments

Giner Alor-Hernández ¹, Jezreel Mejía-Miranda,² and José María Álvarez-Rodríguez ³

¹Division of Research and Postgraduate Studies, Instituto Tecnológico de Orizaba, Orizaba, VER, Mexico

²Research Center in Mathematics (CIMAT, A. C.), Unit Zacatecas, Zacatecas, ZAC, Mexico

³Department of Engineering and Computer Science, Universidad Carlos III de Madrid, Madrid, Spain

Correspondence should be addressed to Giner Alor-Hernández; galor@itorizaba.edu.mx

Received 18 February 2018; Accepted 19 February 2018; Published 5 November 2018

Copyright © 2018 Giner Alor-Hernández 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.

1. Introduction

In recent years, the development, advancement, and use of Information and Communications Technology (ICT) have had a major impact on the operation, structure, and strategy of organizations around the world. Today, it is unthinkable to conceive an organization without the use of ICT, because it allows for the reduction of communication costs and operation while increasing flexibility, interactivity, performance, and productivity. As a result, digital technology fueled by ICT and ICT is currently embedded in any task, activity, and process that is done in any organization or even our daily life activities. This new digital age also implies that science, engineering, and business environments need to reshape their strategies and underlying technology to become a key player of the Industrial Revolution 4.0.

Engineering methods such as requirements engineering, systems modeling, complex network analysis, or simulation are currently applied to support the development of critical systems and decision-making processes in operational environments. As an example, cyberphysical systems featured by mechanical, electrical, and software components are a major challenge for the industry in which new and integrated science and engineering techniques are required to develop and operate these systems in a collaborative data-intensive environment.

Both the development processes and the operational environments of complex systems need the application of scientific and engineering methods to fulfill the management

of new multidisciplinary, data-intensive, and software-centric environments. Programming paradigms such as functional, symbolic, logic, linear, or reactive programming in conjunction with development platforms are considered a cornerstone for the proper development of intelligent and federated programming platforms to support continuous and collaborative engineering.

More specifically, the availability of huge amounts of data that are continuously generated by persons, tools, sensors, and any other smart connected device requires new architectures to address the challenge of solving complex problems such as pattern identification, process optimization, discovery of interactions, knowledge inference, execution of large simulations, or machine cooperation. This situation implies the rethinking and application of innovative scientific programming techniques for numerical, scientific, and engineering computation on top of well-defined hardware and software architectures to support the proper development and operation of complex systems.

In this context, the evolution and extension of engineering methods through scientific programming techniques in data-intensive environments are expected to take advantage of innovative algorithms implemented using different programming paradigms and execution platforms. The conjunction of scientific programming techniques and engineering techniques will support and enhance existing development and production environments to provide high quality, economical, reliable, and efficient data-centric software products and services. This advance in the field of scientific

programming methods will become a key enabler for the next wave of software systems and engineering.

Therefore, the main objective of this special issue was to collect and consolidate innovative and high quality research contributions regarding scientific programming techniques and algorithms applied to the enhancement and improvement of engineering methods to develop real and sustainable data-intensive science and engineering environments.

2. Papers

This special issue aims to provide insights into the recent advances in the aforementioned topics by soliciting original scientific contributions in the form of theoretical foundations, models, experimental research, surveys, and case studies for scientific programming techniques and algorithms in data-intensive environments. This special issue just contains one type of contribution: regular research papers. These works have been edited according to the norms and guidelines of the journal. Several call for papers were distributed among the main mailing lists of the field for researchers to submit their works to this issue. We received a total of 20 submissions which were subject to a rigorous review process to ensure their clarity, authenticity, and relevance to this special issue. At least three reviewers were assigned to every work to proceed with the peer review process. Seven papers were finally accepted for their publication after corrections requested by reviewers and editors were addressed. The seven regular research papers introduce new and interesting results in the form of theoretical and experimental research and case studies under new perspectives on scientific programming techniques and algorithms for data-intensive engineering environments.

One of the special issue's research papers is entitled "Design and Solution of a Surrogate Model for Portfolio Optimization Based on Project Ranking," where E. Fernandez et al. propose a knowledge-based decision support system for the Portfolio Selection Problem on a Set of Ordered Projects. The results show that the reduction of the dimensionality supports the decision-maker in choosing the best portfolio.

In another contribution, entitled "Sentiment Analysis in Spanish for Improvement of Products and Services: A Deep Learning Approach," M. A. Paredes-Valverde et al. propose a deep-learning-based approach that allows companies and organizations to detect opportunities for improving the quality of their products or services through sentiment analysis. The approach is based on Convolutional Neural Network (CNN) and word2vec. To determine the effectiveness of the approach for classifying tweets, some experiments were conducted with different sizes of a Twitter corpus composed of 100000 tweets. Results showed a precision of 88.7%, a recall of 88.7%, and an F -measure of 88.7% considering the complete dataset.

In a further paper, entitled *Scalable* "Parallel Distributed Coprocessor System for Graph Searching Problems with Massive Data," Y. Sun et al. propose a scalable and novel field programmable gate array-based heterogeneous multi-core system for scientific programming. Authors designed considerable parallelism and relatively low clock frequencies

to achieve high performance and customized memory architecture to deal with irregular memory access pattern.

In an additional contribution, entitled "Analysis of Medical Opinions about the Nonrealization of Autopsies in a Mexican Hospital Using Association Rules and Bayesian Networks," E. R. Delgado identifies the factors influencing the reduction of autopsies in a hospital in Veracruz. The study is based on the application of data mining techniques such as association rules and Bayesian networks in datasets created from the opinions of physicians. For the exploration and extraction of the knowledge, some algorithms like Apriori, FPGrowth, PredictiveApriori, Tertius, J48, Naive Bayes, MultilayerPerceptron, and BayesNet were analyzed. To generate mining models and present the new knowledge in natural language, a web-based application was developed. The results have been validated by specialists in the field of pathology.

The fifth paper, entitled "Applying Softcomputing for Copper Recovery in Leaching Process," C. Leiva et al. present a predictive modelling contrasting; a linear model, a quadratic model, a cubic model, and a model based on the use of an artificial neural network (ANN) for copper recovery in leaching process. The ANN was constructed with 9 input variables, 6 hidden layers, and a neuron in the output layer corresponding to copper leaching prediction. The validation of the models was performed with real information and these results were used by a mining company in northern Chile to improve copper mining processes.

In a further contribution, entitled "Semantic Annotation of Unstructured Documents Using Concepts Similarity," F. Pech et al. propose a semantic annotation strategy for unstructured documents as part of a semantic search engine. Ontologies are used to determine the context of the entities specified in the query. The strategy to extracting the context is focused on concepts similarity. Each relevant term of the document is associated with an instance in the ontology. The similarity between each of the explicit relationships is measured through the combination of two types of associations: the association between each pair of concepts and the calculation of the weight of the relationships.

Finally, a paper entitled "A Heterogeneous System Based on Latent Semantic Analysis Using GPU and Multi-CPU" is by G. A. Leon-Paredes et al. They introduce a heterogeneous Latent Semantic Analysis (hLSA) system, which has been developed using general-purpose computing on graphics processing units (GPGPU), which can solve large problems faster through the thousands of concurrent threads on the multiple-core multiprocessors of GPUs and multi-CPU architectures which offer a shared memory programming model in a multithreaded environment. The results of the experiments show that the acceleration reached by an hLSA system for large matrices with one hundred and fifty thousand million values is around eight times faster than the standard LSA version with an accuracy of 88% and a recall of 100%. The performance gain is achieved by using heterogeneous system architectures for matrix computation and text processing.

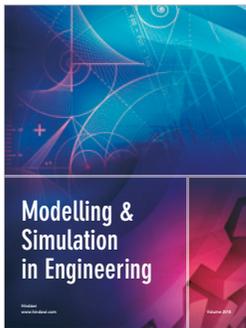
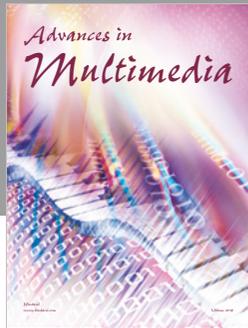
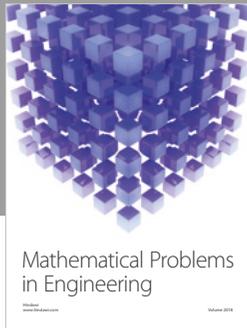
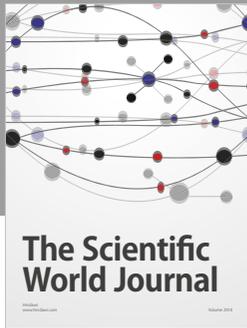
3. Conclusions

As can be seen, all accepted papers are aligned with the scope of the special issue, and all of them provide quite interesting research techniques, models, and studies directly applied to the area of scientific programming techniques and algorithms for data-intensive engineering Environments.

Acknowledgments

The preparation of this special collection has been partially supported by the Tecnológico Nacional de México (TECNM), National Council of Science and Technology (CONACYT), and the Public Education Secretary (SEP) through PRODEP. It has also been supported by the Research Agreement between the RTVE (the Spanish Radio and Television Corporation) and the UC3M to boost research in the fields of big data, linked data, complex network analysis, and natural language. Last but not least, we would also like to express our gratitude to the reviewers who kindly accepted to contribute in the evaluation of papers at all stages of the editing process. We equally and especially wish to thank the editorial board for the opportunity to edit this special issue and for providing valuable comments to improve the selection of research works.

Giner Alor-Hernández
Jezreel Mejía-Miranda
José María Álvarez-Rodríguez




Hindawi

Submit your manuscripts at
www.hindawi.com

