Special Issue Editorial: Intelligent Data Analysis for Sustainable Computing

Yulei Wu, *Senior Member, IEEE*, Yi Pan, Nektarios Georgalas, and Geyong Min

1 Introduction

Recent years have witnessed a deluge of big data streams that contain a wealth of information relevant to sustainable development goals. The analysis of such data streams poses tremendous challenges in the current computing systems, due to the strong correlations between data streams and the emerging needs of real-time decision support in real-world problems.

To obtain this valuable information, there is an urgent demand for high-level computational intelligence based on emerging analytical techniques. This results in modern data analysis techniques having the potential to yield accurate, inexpensive, and high scalable models for providing intelligent and real-time decision support in creating effective computing systems. This will also result in addressing sustainability problems in computing and information processing environments at different levels of computational intelligence paradigms. Computational intelligent data analysis is playing an ever-increasingly important and critical role in achieving sustainable information and communication technology (ICT) in new computing paradigms of the current data-driven era.

This special issue is devoted to the most recent developments and research outcomes addressing the related theoretical and practical aspects of computational intelligence solutions in sustainable computing and aims at presenting latest innovative ideas targeted at the corresponding key challenges, either from a methodological or from an application perspective. On the basis of significance, originality, novelty and presentation, 10 articles were selected to be included in this Special Issue. We will now introduce the 10 accepted articles.

2 Cloud/Edge Computing and Cloud Data Centers

Cloud datacentres are turning out to be massive energy consumers and environment polluters, which necessitate the need for promoting sustainable computing approaches for achieving environment-friendly datacentre execution. Predicting the future workload demands and their respective behaviours at the datacentres are being the focus of recent research in the context of sustainable datacentres. Y. Lu, L. Liu, J. Panneerselvam, X. Zhai, X. Sun, and N. Antonopoulos in "Latency-Based Analytic Approach to Forecast Cloud Workload Trend for Sustainable Datacentres" propose a novel forecasting model named K-RVLBPNN combining intelligent data analysis methods and artificial intelligence techniques for predicting the future workload arrival trend, by exploiting the latency sensitivity characteristics of Cloud workloads, based on an improved K-means clustering algorithm and a backpropagation neural network algorithm. Experiments conducted on real-world Cloud datasets exhibit that the proposed model shows better prediction accuracy, outperforming traditional Hidden Markov Model, Naive Bayes Classifier and Rand Variable Learning Rate Backpropagation Neural Network model, respectively.

Virtual machines (VMs) scheduling is a critical issue for tasks offloading and computation in Mobile Edge Computing. Hao *et al*. pioneers the use of Formal Concept Analysis methodology for identifying the mapping from tasks to VMs in "Virtual Machines Scheduling in Mobile Edge Computing: A Formal Concept Analysis Approach." Specifically, the VMs profile and tasks descriptions are initially characterized as the formal contexts, respectively. The corresponding formal concepts which refer to the rules set are then generated. To better infuse the rules set of VMs and tasks, this paper addresses the matching problem from a given task to a VM according to the principle of maximum similarity degree between formal concepts of VM and task. Extensive simulations are conducted with a real dataset for the validation of feasibility and effectiveness of the proposed approach. Specifically, the proposed approach can significantly reduce the energy consumption around 28% comparing to the approach without consideration of energy consumption.

The number and scale of data centers continue to increase and expand, giving users convenient and

- *Y. Wu and G. Min are with the College of Engineering, Mathematics and Physical Sciences, University of Exeter, EX4 4QF, Exeter, UK. E-mail: y.l.wu@exeter.ac.uk.*
- *Y. Pan is with the Department of Computer Science, Georgia State University, Atlanta, 30302 GA.*
- *N. Georgalas is with the BT Research and Innovation, British Telecom, IP5 3RE, Ipswich, UK.*

(Corresponding author: Yulei Wu.)

Digital Object Identifier no. 10.1109/TSUSC.2019.2962575

efficient computing and storage, while also bring high operating and energy costs. Accurate monitoring and estimation of cloud servers power consumption is the first step to realize sustainable computing. In "An Artificial Neural Network Approach to Power Consumption Model Construction for Servers in Cloud Data Centers," the authors take a fine-grained and in-depth analysis of the power consumption of the server and the performance of its subcomponents, under different types of workloads. After having a better understanding of the relationship between the two, they apply three different artificial neural network (ANN) algorithms to build server power consumption models based on hardware counters. Experiments show that the proposed approach has better performance, compared with the traditional linear model and machine learning-based models, when the servers run with different types of workload.

3 Security

In the era of Industry 4.0, the intelligent embedded-enabled smart manufacturing is an important infrastructure. As a novel type of threat, an advanced persistent threat (APT) has the novel features of strong concealment, latency and long-term entanglement, which can penetrate the core systems of smart manufacturing, especially for intelligent embedded systems, and cause great destruction from the cyber side to physical side. To address this challenge, the authors of "Sustainable Secure Management Against APT Attacks for Intelligent Embedded-Enabled Smart Manufacturing" propose a sustainable secure management mechanism for smart manufacturing against APTs. The proposed mechanism includes two parts: sustainable threat intelligence analysis and sustainable secure resource management. The security defense capabilities against traditional attacks and APT are evaluated, which verifies the feasibilities and efficiency of the proposed mechanism. The proposed sustainable secure resource management mechanism provides deep and continuous protection for intelligent embedded systems in smart manufacturing.

Cloud security has become a vital issue within thousands of inter-connected servers in clouds, as malicious attacks or discovered vulnerabilities may spread more rapidly than ever. In "A Trust Verification Architecture with Hardware Root for Secure Clouds," Yu *et al.* design a novel cloud architecture with a special physical server named as the trust verification server (TVS) to provide trust services according to the trust platform module (TPM) specification, then the servers in the cloud can use TVS remotely as a high-performance TPM chip. The experiments show that the TVS can work efficiently with huge performance improvements at more than 100 times compared with the use of TPM in the cloud. This can be used to solve the complex cloud security problems such as VM sprawl and VM escape.

Many Internet giants such as Google, Amazon, and Microsoft now provide Machine Learning as a Service (MLaaS) to meet the increasing demand for machine learning services. However, the prediction results of training data and testing data with the same machine learning model in MLaaS have remarkable differences, and thus the attackers can leverage machine learning techniques to launch the so-called membership inference attacks, i.e., to infer whether a record is in the training data or not. In "MIASec: Enabling Data Indistinguishability Against Membership Inference Attacks in MLaaS," the authors propose MIASec that can guarantee the data indistinguishability of the training data and thereby has the ability to defend against membership inference attacks in MLaaS. Experiment results show that MIASec can defend the membership inference attacks effectively. In particular, MIASec can reduce the precision and recall of attacks respectively by 11.7 percent and 15.4 percent in average, and by 18.6 percent and 21.8 percent at best.

4 Water Supply Systems

Constructing Sustainable Smart Water Supply systems are facing serious challenges with the fast expansion of modern cities. Traditional urban water quality control based only on indicator tests may cause delay in warning, which posts great risks to the drinking water. There have been accidents of massive infections in some big cities. In "Quality Risk Analysis for Sustainable Smart Water Supply Using Data Perception," Wu *et al.* analyze these problems, and provide a feasible solution by building a data-driven risk analysis framework. An Adaptive Frequency Analysis (Adp-FA) method has been proposed using indicators' frequency domain information to resolve their inner relationships and individual predictions. They also investigate the scalability of this method from indicator, geography and time domains. For applications, the authors select industrial quality datasets collected from a Norwegian project in 4 different urban water supply systems, namely Oslo, Bergen, Strømmen, and Ålesund. The results show the proposed method performs better in most of the aspects. It facilitates explainable early warnings of water quality risks and supports further decisions in the quality control.

5 Online Education and e-Learning

Online education and e-learning have vigorously sprung up and produced massive educational data in a streaming way. It is significant to study how to acquire the appropriate learning resources and the suitable learning partners from the streaming-updated educational big data. Liu *et al.* in "Multi-Dimensional Correlative Recommendation and Adaptive Clustering via Incremental Tensor Decomposition for Sustainable Smart Education" aim to provide these sustainable smart educational services by correlatively analyzing the global educational data from multiple dimensions via incremental tensor decomposition. They present an incremental tensor-based correlative analysis and personalized recommendation (ITCA-PR) algorithm to recommend appropriate resources, as well as an incremental tensor-based adaptive clustering and community recommendation (ITAC-CR) algorithm to recommend suitable learning partners under various contexts and accordingly construct adaptive learning communities. Extensive experimental results demonstrate

that ITCA-PR and ITAC-CR algorithms outperform some state-of-the-art recommendation and clustering algorithms. The proposed ITCA-PR and ITAC-CR approaches are conducive to providing timely and precise educational services.

6 Renewable Energy

In "Species and Memory Enhanced Differential Evolution for Optimal Power Flow Under Double-Sided Uncertainties," the authors focus on dynamic optimal power flow (DOPF) problems with uncertainties on both the supply side and demand side. Most of existing works on DOPF require certain knowledge of these uncertain parameters in advance. Moreover, only a little work considered both uncertainties simultaneously. It might be because the combination of variable uncertainties could lead to a huge problem size for existing methods. Therefore, the idea of tracking the moving optimum, inspired by the ideas in the evolutionary dynamic optimization (EDO), is introduced. Based on this idea, a species and memory enhanced differential evolutionary algorithm is proposed, where a memory strategy and a multi-population strategy are specially designed to deal with the uncertainties. The experimental results on the modified IEEE 57-bus and 118-bus systems show that the proposed algorithms perform better than the comparison algorithms for most cases.

7 Healthcare

Multiclass classification of pathological brain resonance (MR) images is a more challenging task than binary classification. Nayak *et al.* in "Automated Diagnosis of Pathological Brain Using Fast Curvelet Entropy Features" propose a new feature extraction framework based on fast curvelet transform and entropy features. The entropy-based features have been computed from the various components (subbands) of unequally spaced fast Fourier transform based fast curvelet transform (FCT-USFFT) and wrapping based fast curvelet transform (FCT-WR) separately. The features are finally subjected to kernel extreme learning machine (KELM) for multi-class classification. The article examines the discriminatory potential of these two features in comparison to different wavelet-based entropy features. One-versus-all and multi-class classification results on Harvard Medical School University are shown which proves superiority of the proposed algorithm.

8 Conclusion

In conclusion, we have introduced the 10 accepted papers, categorized into cloud/edge computing and cloud data centers, security, water supply systems, online education and e-learning, renewable energy, and healthcare. Sustainable computing plays an important role in our society with pervasive intelligent data analysis. Therefore, there remains a need to keep a watchful brief of emerging challenges, as well as those that currently exist in the intelligent data analysis for sustainable computing.

Acknowledgments

The guest editors would like to express their deep thanks to the editor-in-chief, professor Albert Zomaya, for providing them with the opportunity to host this special issue in the *IEEE Transactions on Sustainable Computing*. They also thank all the authors who submitted their papers. Last but not least, they thank the thoughtful work of the many reviewers who have provided invaluable evaluations and recommendations.

Yulei Wu received the BSc degree (1st Class Hons.) in computer science and the PhD degree in computing and mathematics from the University of Bradford, UK, in 2006 and 2010, respectively. He is currently a senior lecturer in the Department of Computer Science, University of Exeter, UK. His expertise is on networking and his main research interests include autonomous networks, intelligent networking technologies, network slicing and softwarization, SDN/NFV, green networking, wireless networks, network security and privacy, and analytical modelling and performance optimization. He is an editor of the *IEEE Transactions on Network and Service Management*, *Elsevier Computer Networks*, and *IEEE Access*. He has contributed to major conferences on networking in various roles including the steering committee chair, the general chair, the program chair, and the technical program committee member. His research has been supported by the Engineering and Physical Sciences Research Council (EPSRC) of United Kingdom, National Natural Science Foundation of China, the university's innovation platform, and industry. He is a fellow of the Higher Education Academy (HEA), and a senior member of the IEEE.

Yi Pan received the BE and ME degrees in computer engineering from Tsinghua University, China, in 1982 and 1984, respectively, and the PhD degree in computer science from the University of Pittsburgh, USA, in 1991. He is currently a Regents' professor and the chair of computer science with Georgia State University, USA. He has served as an associate dean and chair of the Biology Department during 2013-2017 and chair of computer science during 2006-2013. His profile has been featured as a distinguished alumnus in both the *Tsinghua Alumni Newsletter* and *University of Pittsburgh CS Alumni Newsletter*. His research interests include parallel and cloud computing, big data, and bioinformatics. He has published more than 250 journal papers with more than 100 papers published in various IEEE/ACM Transactions/journals. In addition, he has published more than 150 papers in refereed conferences. He has also coauthored/coedited 43 books. His work has been cited more than 11,600 times in Google Scholar and his current H-index is 56. He has served as an editor-in-chief or editorial board member for 20 journals including seven IEEE Transactions. He is the recipient of many awards including the IEEE Transactions Best Paper Award, several other conference and journal best paper awards, four IBM Faculty Awards, two JSPS Senior Invitation Fellowships, IEEE BIBE Outstanding Achievement Award, NSF Research Opportunity Award, and AFOSR Summer Faculty Research Fellowship. He has organized many international conferences and delivered keynote speeches at more than 60 international conferences around the world.

Nektarios Georgalas is currently a principal researcher in the Applied Research Department of British Telecom. Currently, he leads two collaborative research programmes with key BT partners delivering innovations in the areas of cloud, data centres, network virtualisation, smart cities, IoT, and mobility. During his career with BT, since 1998, he has managed numerous collaborative and internal research projects in areas such as network management, market-driven data management systems, policy-based management, distributed information systems, SOA/Web Services, Model Driven Design and Development of telecoms OSS, Cloud, and NfV. He has led numerous international collaborations on the application of advanced techniques for design, development and operation of telecoms Networks and Operation Support Systems environments. In the past, he was very active leading and contributing to key programmes within the TeleManagement Forum, where he established international standards teams, led Catalysts, and influenced the Forum's strategy towards a model-driven and software-defined ecosystem of digital services in dynamic marketplaces. Recently, he was involved in TMF's Smart City Forum and related Catalysts helping establish the key capabilities and APIs for Smart Cities, as part of the "City as a Platform" programme. His work has been recognized several times by numerous international awards including the TMForum's Excellence Award for Innovation in 2010, Most Innovative Catalyst Award in 2014, Best New Catalyst Award in 2015 and Most Significant Contribution to Frameworx Award in 2015, Most Innovative Catalyst—Smart X Commercial in 2016, Outstanding Performance in the Catalyst Programme in 2017 and Smart City Innovator of the Year Excellence Award in 2017. Other recognition accolades include Global Telecoms Business's Business Service Innovation Award in 2010, 2012, and 2013. He has been Finalist in UK IT Industry Award for Best IT Innovation in 2013 and Highly Commended for the IET Innovation Award for Telecommunication in 2009. He has also achieved Best innovation for Large Enterprise and Best Customer Experience Innovation Finalists in BT Innovation Awards in 2010. He has been recognized in BT's TSO Brilliant People in 2015. He is an inventor and coinventor of 11 patents. He has also authored more than 60 papers in international journals and conferences. He has served as a general co-chair, programme co-chair, programme committee and keynote speaker or invited panellist in top international IEEE academic and TMForum conferences.

Geyong Min received the BSc degree in computer science from the Huazhong University of Science and Technology, China, in 1995, and the PhD degree in computing science from the University of Glasgow, UK, in 2003. He is currently a professor of high performance computing and networking with the Department of Computer Science within the College of Engineering, Mathematics and Physical Sciences at the University of Exeter, UK. His research interests include future Internet, computer networks, wireless communications, multimedia systems, information security, high performance computing, ubiquitous computing, modelling, and performance engineering.