



Big Data Analytics for Sustainable Computing

H . Anandakumar¹ · R. Arulmurugan² · Chow Chee Onn³

Published online: 7 October 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

1 Editorial

Big Data analytics is the process of collecting, organizing and analyzing large sets of data. To analyze such a large volume of data, Big Data analytics is typically performed using specialized software tools and applications for predictive analytics, data mining, text mining, forecasting and data optimization. Big Data analytics can help organizations to better understand the information contained within the data and will also help identify the data that is most important to the business and future business decisions. Collectively these processes are separate but highly integrated functions of high-performance analytics. Using Big Data tools and software enables an organization to process extremely large volumes of data that a business has collected to determine which data is relevant and can be analyzed to drive better business decisions in the future. This issue focused high quality research papers that address significant and new big data application and related system development issues in the emerging sustainable application domains. We anticipate this issue will open new entrance for further research and technology improvements in this important area.

This issue features fourteen selected papers with high quality. The first article, “An Attempt to Design Improved and Fool Proof Safe Distribution of Personal Healthcare Records for Cloud Computing”, presents the hitches of storing

healthcare related data on cloud and safeguarding in an astounding way. Diverse technology applications of cloud can be reinforced with wrapping the information of the user’s unsigned verification based on the element based encoding (EBE) is employed and fine grained data access control based on the advanced encryption scheme are tailored. The author researched that the key generation based on various set of elements plays a vital role in hiding the information of the users while gaining access to the data present over the network. Moreover for reducing the setbacks in safeguarding the key of the data containers the personal health records are classified into several associated fields. Thereby, proposed research work focus on effectiveness in terms of safety and secrecy of data over cloud.

The second article, “Prediction of Individual’s Character in Social Media Using Contextual Semantic Sentiment Analysis”, proposed Sentiment Analysis Prediction for text mining as well as natural language processing. Paper discussed about people’s opinion towards product, service, tourism, movies, political issues, education systems opinion through social media like Twitter, Facebook etc., The key focus of this paper is to introduce Opinion COW (Opinion Co-Occurrence Word) method for Opinion Circle Method using contextual semantic sentiment analysis and Hybrid method to categorize the twitter users based on Maslow theory and identify the sentiment of each tweet.

The third article, “Ensembled Population Rescaled Differential Evolution with Weighted Boosting for Early Breast Cancer Detection”, proposed early detection of breast cancer assists in increase of survival rate. Here an ensemble classifier method called Population Rescaled Differential Evolution with Weighted Boosting (PRDE-WB) is presented for early detection. At first, the given input breast images are subjected to pre-processing using Logarithmic Cube-root Shift technique. Where the ROI is extracted according to the cube root of the given breast input images. Then Population Rescaled Differential Evolution Optimization (PRDE) is implemented on the extracted ROI for obtaining breast cancer regions. PRDE process undergoes initialization, mutation, crossover and selection, optimally sketching the contour of the detected tumor using Population Rescaling Factor.

✉ H . Anandakumar
anandakumar.psgtech@gmail.com

R. Arulmurugan
arulmr@gmail.com

Chow Chee Onn
cochow@um.edu.my

- ¹ Computer Science and Engineering, Sri Eshwar College of Engineering, Coimbatore, Tamilnadu, India
- ² Computer Science and Engineering, Presidency University, Bangalore, Karnataka, India
- ³ Electrical Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur, Malaysia

The fourth article, “Filter Bank Multicarrier based Generalized Discrete Fourier Transform for Peak to Average Power Ratio Reduction”, The GDFT has been recommended to develop the performance of PAPR in communication frameworks applying multi-carrier methodologies to developed algorithm centered on the GDFT that considers more PAPR reductions eliminating any further computational complexities. Therefore, the recommended algorithm has been formulated to utilized Side Information (SI) represents a single bit of every data segment. This developed algorithm is also applicable in more enhanced DFT-spreading frameworks aimed at comparing data. Based on the collected simulation findings, it is evident that varied number of sub-carriers GDFT-spreading, which considers the availability or the presence of the developed algorithm, indicates the recommendable framework signifying that the enhanced algorithms received in most instances is an additional PAPR amount reductions based on other DFT-spreading frameworks.

The fifth article, “Modeling Displacement and Direction Aware Ad Hoc On-Demand Distance Vector Routing Standard for Mobile Ad Hoc Networks”, proposed a new routing protocol DDC – AODV to handle the displacement and direction factors in ad-hoc networks. For minimizing the connection breakdowns and to acquire a constant path a fresh reactive routing standard is designed based on hierarchy based displacement conscious. The designed displacement and magnitude conscious ad hoc on demand distance vector routing standard intends to address the displacements and magnitude features over the ad hoc networks. The DDC – AODV directs the path identification and path acknowledgment based on the displacement of the contributing nodes and their magnitudes. The NS2 offers two load assessment employing assessments to explore the consequences and benefits of DDC – AODV over the AODV standards.

The sixth article, “A Statistical Lower UWB Channel Model for In-Body Communications”, proposes a frequency dependent on-body to in-body statistical pathless model to determine the pathless in the human abdominal region. Due to the growing need for developing implantable wireless transceivers in Body Area Networks, it becomes necessary to characterize the in-body wireless channel in Ultra Wide Band frequency range. In case of BAN with implantable sensors, in-vivo measurement is not practically feasible. Hence, developments of radio wave propagation prediction models that predict the in-body channel pathless becoming necessary nowadays. This paper also introduces honey as an easily available and less expensive base material for preparation of the body mimicking liquid phantoms, based on the results of its dielectric property analysis. Though this model is developed for human abdominal region, it could be extended to other parts of the human body using the same procedure.

The seventh work titled, “Analyzing Customer Outage Cost in a Microgrid”, presents an ideally modeled Microgrid consisting of PV panels, Wind turbines, Fuel cell, Battery storage and Microturbines capable of operating both in standalone and

grid-connected modes. This work focuses on reduction of the expenses involved in supplying power from Microgrid and intermittency of power generation. Hence a heuristic Multi-objective Particle Swarm optimization technique is used to obtain optimal solution for Customer outage cost and operating cost of Microgrid under islanded and grid-connected mode of operations for different scenarios.

The eighth article “Optimal Charging Scheduling of Electric Vehicles in Micro Grids Using Priority Algorithms and Particle Swarm Optimization” has proposed a systematic charging scheme to reduce the charging station electricity consumption cost from the grid. The proposed systematic charging scheme is based on “Particle Swarm Optimization (PSO)”. It is compared with well-established algorithms such as “arrival time-based priority (ATP) algorithm” and “SOC-based priority (SBP) algorithm”. In addition, a microgrid scenario is further considered for reducing the charging cost by properly shifting the EV load. Based on the study carried out for a sample test cases considered, it was found that the proposed scheme has better performance compared to the existing schemes.

The ninth article “An improved low complex offset min-sum based decoding algorithm for LDPC codes” introduces an efficient and robust offset min-sum decoding scheme for optimal decoding of LDPC codes. This improved approach introduces a new offset correction factor to suppress the error propagation during the approximation of high precision soft values within the given range boundary of signal strength-to-background noise ratio (SNR). The proposed offset correction factor for LDPC decoding is obtained using the concept of random process and estimation theorem. Through theoretical derivation, it is shown that the proposed offset correction factor can be obtained using a simple method which requires only a few, less complex arithmetic operations. Experimental performances showed that proposed algorithm has the advantage of good coding gain improvement and comparable convergence speed when validated with several popular LDPC decoding algorithms. Therefore, the proposed algorithm can be an appropriate candidate for practical implementation of LDPC decoding architectures compatible for future generation high-speed communication systems.

The tenth work titled “Service Level Agreement Based Catalogue Management and Resource Provisioning in Cloud for Optimal Resource Utilization” points out the Service Level Agreement based knowledge base made as a catalogue for resource management in cloud. This catalogue is frequently monitored and resource availability and utilization is calculated and responses are made properly. Good quality of literature surveys has been made which points out the ranking based resource allocation which concerns on performance, future prediction model which fails on future dynamic demands and the other similar surveys. But this work is unique in identifying the dynamic changes in utilization of virtual machines. The work captures the dynamic changes in the virtual machines and this history is saved as a catalog for further reference. The observation in

the utilization of system, it is clear that for the equal distribution of machines and virtual machines, when the cloudlet is run with minimal utilization, that is whenever the machine is less accessed by users, the power consumption of the machine is slightly reduced. This happens for average user access also. When the same is taken for real servers the power consumption of the servers can be considerably reduced. Also there is future direction, where a cost based system can be modeled where whenever the system is running in underutilized mode a minimal virtual machine can be provided for service which in turn will be less cost effective and whenever there is a heavy workload, the machines can be scaled at the mean time price can be scaled up.

The eleventh article titled ‘Performing Data Assessment in terms of Sensor Node Positioning over Three Dimensional Wireless Sensor Networks’ presents the modelling of sensing technique for the three dimensional wireless sensor networks based on the standard termed as three dimensional scattered grouping (3D – SG). It supports extraction of precise information among the clusters to model a three dimensional information analysis (3D – DA) scheme. Moreover the node deployment policies depict the significant feature in improving the data accuracy over the three dimensional wireless sensor networks. Therefore the deployment of nodes in their accurate position over the three dimensional sensor networks is quite difficult. The goal is to model a three dimensional node positioning scheme for positioning the possible nodes and their deployment policies for improving the information accuracy within the network.

The Twelfth paper titled ‘Efficient Machine Learning Model for Movie Recommender Systems Using Multi-Cloud Environment’, presents the recommender system or a recommendation system is a subclass of information filtering system that seeks to predict the rating a user would give to an item. Recommender systems are utilized in a variety of areas including movies, music, news, books, research articles, search queries, social tags, and products in general. A multi-cloud strategy is the use of two or more cloud computing services. While a multi-cloud deployment can refer to any implementation of multiple software as a service (SaaS) or platform as a service (PaaS) cloud offerings, today, it generally refers to a mix of public infrastructure as a service (IaaS) environments, such as Amazon Web Services and Microsoft Azure. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

The next paper titled ‘Extreme gradient boost classification based interesting user patterns discovery for web service composition’, presents a methodology to extract interesting actionable patterns of web users for web service composition through Best First Decision Tree Based Extreme Gradient Boost Classification (BFDT-XGBC) technique. Web mining

usually deals with the application of various data mining techniques in the web servers to discover the interesting patterns by accessing the behavior of web users. The interesting patterns are identified by ensemble classification. At first, the user accessed patterns are extracted from the web server log. Then Best First Decision Trees (BFDT) are used as base learners to classify similar web patterns accessed by the same user based on correlation measure. Extreme gradient boost is applied to calculate the loss functions of all best first decision trees for providing strong classification results. Therefore BFDT-XGBC technique effectively discovers the user interesting patterns for web service composition.

The last paper titled ‘SDNFV Based Threat Monitoring and Security Framework for Mobile Edge Computing Infrastructure’, presents DoS botnet attacks such as Advanced Persistent & Ransom DoS assaults, Botnets and Application DDoS flood attacks. The main objective of research is to demonstrate the feasibility and advantages of finer networking abstractions that the SDN/NFV brings to the MEC context and to solve security issues at the edge networks, without incurring major overhead. The performance of DTARS framework within three practical SDN enabled Mobile LTE MEC network, SDN enabled IoT MEC network and Software-Defined Datacenter Edge network. The proposed solution enables security against DDoS attacks on different levels.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Anandakumar Haldorai Professor (Associate) and Research Head in Department of Computer Science and Engineering, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India. He has received his Master’s in Software Engineering from PSG College of Technology, Coimbatore and PhD in Information and Communication Engineering from PSG College of Technology under, Anna University, Chennai. His research areas include Big Data, Cognitive Radio Networks, Mobile Communications and Networking

Protocols. He has authored more than 82 research papers in reputed International Journals and IEEE conferences. He has authored 7 books and many book chapters with reputed publishers such as Springer and IGI. He is editor of Inderscience IJISC and served as a reviewer for IEEE, IET, Springer, Inderscience and Elsevier journals. He is also the guest editor of many journals with Elsevier, Springer, Inderscience, etc. He has been the General Chair, Session Chair, and Panelist in several conferences. He is senior member of IEEE, IET, ACM and Fellow member of EAI research group.



Arulmurugan Ramu is a Professor, Presidency University, Bangalore, India. His research focuses on the automatic interpretation of images and related problems in machine learning and optimization. His main research interest is in vision, particularly high-level visual recognition. He has authored more than 35 papers in major computer vision and machine learning conferences and journals. He is the recipient of Ph.D. degrees in Information and Communication Engineering from

the Anna University at Chennai, M.Tech in Information Technology Anna University of Technology and B.Tech degree in Information Technology. He is guided many Ph.D. research scholar under the area of Image Processing using Machine Learning. He is an Associate Editor of Inderscience IJISC journal. He is awarded as Best Young Faculty Award 2018 and nominated for Best Young Researcher Award (Male) by International Academic and Research Excellence Awards (IARE-2019).



Chee-Onn Chow received his Bachelor of Engineering (Hons) and Master of Engineering Science degrees from University of Malaya, Malaysia in 1999 and 2001, respectively. He received his Doctorate of Engineering from the Tokai University, Japan in 2008. He joined the Department of Electrical Engineering as tutor in 1999, and subsequently been offered a lecturer position in 2001. He is currently an Associate Professor in the same department

since 2015. His research interests include various issues related to wireless communications. He is a Chartered Engineer (IET, UK), a Professor Engineer (BEM, Malaysia) and a Senior Members of IEEE.