

Northumbria Research Link

Citation: Yang, Longzhi, Hu, Jia and Hung, Che-Lun (2021) Special issue on recent advances in data science and systems. Expert Systems, 38 (5). e12735. ISSN 0266-4720

Published by: Wiley-Blackwell

URL: <https://doi.org/10.1111/exsy.12735> <<https://doi.org/10.1111/exsy.12735>>

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/id/eprint/48044/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)



**Northumbria
University**
NEWCASTLE



UniversityLibrary

Special Issue on Recent advances in Data Science and Systems

As an interdisciplinary area, Data Science draws scientific inquiry from a broad range of subject areas such as statistics, mathematics, computer science, machine learning, optimisation, signal processing, information retrieval, databases, cloud computing, computer vision, natural language processing, etc. Data Science is on the essence of deriving valuable insights from data. It is emerging to meet the challenges of processing very large datasets, i.e. Big Data, with the explosion of new data continuously generated from various channels, such as smart devices, web, mobile and social media.

Data Systems are posing many challenges in exploiting parallelism of current and upcoming computer architectures. Data volumes of applications in the fields of sciences and engineering, finance, media, online information resources, etc. are expected to double every two years over the next decade and beyond. The importance of data intensive systems has been raising and will continue to be the foremost field of research. This raise brings up many research issues, in forms of capturing and accessing data effectively and fast, processing it whilst still achieving high performance and throughput, and storing it efficiently for future use.

This special issue focuses on recent advances in Data Science (e.g. Knowledge Discovery, Data Mining, Machine Learning, Big Data Analytics, Deep Learning, etc) and data systems, and innovative real-world applications of such technologies to deliver effective and efficient solutions for current and future challenges. This special issue has attracted more than 20 submissions and 6 manuscripts were selected based on review reports. Each paper was reviewed by at least two reviewers and went through at least two rounds of reviews. The contributions of these papers are briefed below.

The first contribution by Li et al. reports a novel weighted probabilistic frequent itemset mining algorithm in uncertain databases (i.e. w-PFI), which is implemented by an efficient candidate generation and validation paradigm similar to the working principle of Apriori. This work additionally presents a new probability model to support w-PFI candidate, and three pruning techniques to effectively remove the unpromising candidates immediately to improve system efficiency. The experimental results show that the proposed algorithm w-PFI yields the best performance amongst the referenced competitors in terms of running time and scalability.

The second paper by Sadhukham and Palit presents a novel neighbourhood-based multi-label classifier based on the principles of reverse k-nearest neighbourhood. That is, the neighbourhood was estimated using the reverse k-nearest neighbourhood. This adaptive neighbourhood estimation with the support of implicit handling of the local imbalance works particularly well for multiple-label datasets with imbalanced labels. The proposed approach improves the efficacy of the compared methods based on the experimentation as evidenced by its competitive performance.

The third publication by Tsinaslanidis and Guijarro considers chart pattern recognition for trading purposes. In particular, this work proposes the design of a trading system using generic pattern recognition technique which takes proven generic profitable patterns based on historical data as system inputs rather than restricting the search to specific technical patterns. The effectiveness of the proposed system was validated and evaluated by applying the approach to 560 NYSE stocks with generally promising results demonstrated.

The article produced by Hu et al. documents an adaptive network with a stacked hourglass network and SSD for video pose estimation especially for videos with joint occlusion. The proposed network is supported by the optimisation of time series motion data using an outlier detection and a Kalman filter. The work was evaluated by applying the proposed adaptive network on two well-known benchmark human pose estimation datasets. The results show higher accuracy and good practicality.

The next article by Naik et al. proposes a cognizant honeypot for active fingerprinting attack detection using dynamic fuzzy rule interpolation. This project firstly actively collected data using simulated attacks on honeypots and extracted the most influential attributes from the collected data as the signatures of active fingerprinting attacks. Then, the selected attributes were utilised to devise the dynamic fuzzy rule interpolation system and subsequently to implement

the cognizant honeypot. The proposed system is featured by its dynamic rule base for more accurate and efficient detection.

The final contribution by Gao et al. reports a hand gesture recognition approach using multimodal data fusion and a multiscale parallel convolutional neural network for human robot interaction. Ten hand gestures were considered in this project and the multiscale parallel convolutional neural network was trained using a dataset generated by this project. The proposed method was implemented on a seven-degree-of-freedom bionic manipulator and promising results were demonstrated based on the experiments using this manipulator.

ACKNOWLEDGEMENTS

We would like to express our sincere thanks to Dr. Jon G. Hall (Editor-in-Chief of the Wiley-Blackwell Journal Expert Systems: The Journal of Knowledge Engineering) for providing the opportunity to edit this special issue. Additional thanks to the editorial staff for their excellent support. Finally, guest editors would also like to thank all the referees for their thorough and constructive comments.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

Longzhi Yang¹

Jia Hu²

Allen Hung³

¹Department of Computer and Information Sciences, Northumbria University,
Newcastle upon Tyne, NE1 8ST, UK

²University of Exeter, UK

³Providence University, Taiwan

Correspondence

Longzhi Yang, ¹Department of Computer and Information Sciences, Northumbria University,
Newcastle upon Tyne, NE1 8ST, UK