



Guest Editorial: WWWJ Special Issue of the 22th International Conference on Web Information Systems Engineering (WISE 2021)

Hua Wang¹ · Wenjie Zhang² · Lei Zou³ · Zakaria Maamar⁴

Published online: 13 December 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

1 WISE2021

The 22th International Conference on Web Information Systems Engineering (WISE 2021) was held in Melbourne, Australia, October 26–29, 2021. It was processed through on-line and off-line for participants as some countries have closed board due to the global pandemic situation. However, WISE2021 was very successful. Building on the success of its predecessors, WISE 2021 continues to be a major international forum for researchers, professionals, and industrial practitioners to share their knowledge in the rapidly growing area of Web technologies, methodologies, and applications. The first WISE event took place in Hong Kong, China (2000). Then the trip continued to Kyoto, Japan (2001); Singapore (2002); Rome, Italy (2003); Brisbane, Australia (2004); New York, USA (2005); Wuhan, China (2006); Nancy, France (2007); Auckland, New Zealand (2008); Poznan, Poland (2009); Hong Kong, China (2010); Sydney, Australia (2011); Paphos, Cyprus (2012); Nanjing, China (2013); Thessaloniki, Greece (2014); Miami, USA (2015); Shanghai, China (2016); Puschino, Russia (2017); Dubai, UAE (2018); Hong Kong, China (2019) and Amsterdam, Netherlands (2020). This year, WISE was held in, supported by University of New South Wales (Australia), Victoria University (Australia), Peking University (China) and Zayed University (United Arab Emirates).

A total of 229 research papers were submitted by 390 authors from various countries to the conference for consideration, and each paper was double-blind reviewed by at least three reviewers. Finally, 55 submissions were selected as full papers (with an acceptance rate of 24% approximately), plus 29 as short papers. The research papers cover the areas of social network data analysis, recommender systems [1], topic modeling [2], data diversity, data similarity, context-aware recommendation, prediction [3, 4], big data processing [5],

✉ Hua Wang
Hua.Wang@vu.edu.au

¹ Victoria University, Footscray, Australia

² University of New South Wales, Sydney, Australia

³ Peking University, Beijing, China

⁴ Zayed University, Dubai, United Arab Emirates

cloud computing, event detection [6], data mining [7], sentiment analysis, ranking in social networks, microblog data analysis, query processing [8], spatial and temporal data, graph theory and non-traditional environments [9, 10]. We are honored to have several of the world's leading experts in the field join us as distinguished keynote speakers and invited speakers.

2 The special issue

Six top ranked papers out of 55 full papers at WISE 2021 have been selected for the special issue of World Wide Web Journal (WWWJ). The selected papers underwent a rigorous extra refereeing and revision process. In particular, the six special issues papers have been extended with at least 60% new and unpublished material. Note that, adding more related work or extending the introduction was not considered in the 60%; rather the new content often includes more technical and implementation details, improved algorithms, more experiment results, etc. The revised papers again went through the review and revision processes in accordance with WWWJ guidelines.

The paper titled “Critical Nodes Identification in Large Networks: The Inclined and Detached Models” by Chen Chen, et al. aims to propose and investigate the inclined anchored k-core and minimum detached k-core problems. Authors prove that both problems are NP-hard, which implies that it is non-trivial to solve them within polynomial time. Thus, following the routine in previous studies, they plan to design heuristic strategies that can return competitive results efficiently. In addition, due to the different properties of the objective functions, lots of pruning rules in previous studies no longer hold for the problems. Secondly, in real-world social networks, the search space is usually quite large, which is time-consuming to conduct the exploration. To enhance the processing, different search approaches and pruning techniques are considered. Comprehensive experiments are conducted on nine networks to demonstrate the effectiveness and efficiency of the proposed techniques.

The paper “CupMar: A Deep Learning Model for Personalized News Recommendation based on Contextual User-Profile and Multi-Aspect Article Representation” by Dai Hoang Tran, et al. proposes a neural deep news recommendation model called CupMar. CupMar is not only able to learn the user-profile representation in different contexts, but also is able to leverage the multi-aspects properties of a news article to provide accurate, personalized news recommendations to users. Making recommendation for news articles poses a great challenge due to vast amount of online information. However, providing personalized recommendations from news articles, which are the sources of condense textual information is not a trivial task. A recommendation system needs to understand both the textual information of a news article, and the user contexts in terms of long-term and temporary preferences via the user's historic records. Unfortunately, many existing methods do not possess the capability to meet conditions needed. The main components of the CupMar approach include the News Encoder and the User-Profile Encoder. Specifically, the News Encoder uses multiple properties such as news category, knowledge entity, title and body content with advanced neural network layers to derive informative news representation, while the User-Profile Encoder looks through a user's browsed news, infers both of her long-term and recent preference contexts to encode a user representation, and finds the most relevant

candidate news for her. Authors evaluate the CupMar model with extensive experiments on the popular Microsoft News Dataset (MIND), and demonstrate the strong performance of the approach.

The paper titled “Optimizing Subgraph Matching over Distributed Knowledge Graphs Using Partial Evaluation” by Xin Wang et al. presents an optimal partial evaluation algorithm and a filter method to reduce partial matches by exploring the computing characteristics of partial evaluation and assembly framework. The partial evaluation and assembly framework has recently been applied for processing subgraph matching queries over large-scale knowledge graphs in the distributed environment. The framework is implemented on the master–slave architecture, endowed with outstanding scalability. However, there are two drawbacks of partial evaluation: if the volume of intermediate results is large, a large number of repeated partial matches will be generated; and the assembly computation handled by the master would be a bottleneck. The developed method includes three parts: (1) An index structure named inner boundary node index (IBN-Index) is constructed to prune for graph exploration to improve the searching efficiency of the partial evaluation phase. (2) The boundary characteristics of local partial matches are utilized to construct a boundary node index (BN-Index) to reduce the number of local partial matches, and (3) The experimental results over benchmark datasets show that the approach outperforms the state-of-the-art methods.

The paper “Text Style Transfer between Classical and Modern Chinese through Prompt-based Reinforcement Learning” by Min Peng et al. studies an unsupervised prompt-based reinforcement learning (PBRL) framework to transfer text between classical and modern Chinese styles via an entangled approach. Text style transfer aims at converting the stylistic features of a sentence to another style while preserving its content. Despite the remarkable progress achieved in English style transfer, Chinese style transfer still relies heavily on manual processing. Taking classical and modern Chinese style transfer as an example, most of the existing method cannot carry out this task due to the lack of sufficient parallel corpus for supervised learning and the special language phenomenon in Chinese. The PBRL framework mainly consists of two stages, i.e., a prompt-based fine-tuning stage and a bi-directional reinforcement learning stage. In the first stage, authors leverage a priori knowledge-based synonym dictionary to build a pseudo-parallel corpus for prompt learning to provide the system a warm start. Then the style-transfer accuracy reward and content-preservation reward are specially designed for bidirectional-reinforcement optimization. Experimental evaluations show that the developed model outperforms state-of-art networks by a large margin.

The paper “Improving Medical Experts’ Efficiency of Misinformation Detection: An Exploratory Study” by Mikolaj Morzy et al. aims to optimize the utilization of medical experts’ time. Fighting medical disinformation in the era of the pandemic is an increasingly important problem. Today, automatic systems for assessing the credibility of medical information do not offer sufficient precision, so human supervision and the involvement of medical expert annotators are required.

Authors equip them with tools for semi-automatic initial verification of the credibility of the annotated content. They introduce a general framework for filtering medical statements that do not require manual evaluation by medical experts, thus focusing annotation efforts on non-credible medical statements. A developed framework is based on the construction of filtering classifiers adapted to narrow thematic categories. This allows medical experts to fact-check and identify over two times more noncredible medical statements in a given time interval without applying any changes to the annotation flow. Authors verify the results across a broad spectrum of medical topic areas and perform quantitative, as well

as exploratory analysis on our output data. They also point out how those filtering classifiers can be modified to provide experts with different types of feedback without any loss of performance.

The paper titled “A Text and GNN Based Controversy Detection Method On Social Media” by Benslimane Samy, M. D. et al. aims to combine both user interactions present in the graph structure of a discussion and the discussion text features to detect controversy. Expressed opinions on social media frequently cause a controversy. Controversial content refers to content that attracts different opinions and interrogations, implying interaction between communities. Its automatic identification remains a challenging task. Most of the existing approaches rely on the graph structure of discussion and/or the content of messages but didn’t deeply explore the recent advances on Graph Neural Network (gnn) to predict if a discussion is controversial or not. The proposed approach can work with any social network and relies on gnn techniques to encode the whole graph representation (including its texts) in an embedding vector, before performing a graph classification task. Two controversy detection strategies are proposed in the paper. The first one is based on a hierarchical graph representation learning to take advantage of hierarchical relationships that could exist between users. The second one is based on the attention mechanism, which allows each user node to give more or less importance to its neighbors when computing node embeddings. Authors present different experiments conducted with data sources collected from both Reddit and Twitter to show the applicability of the approach to different social networks. Conducted experiments show the positive impact of combining textual features and structural information in terms of performance and accuracy.

Acknowledgements We are very indebted to the reviewers who reviewed the papers very carefully. We would also like to thank all the authors who submitted their papers to the special issue. Special thanks to the journal editors for their great help and support in organizing the issue.

References

1. Li, J.Y., et al.: Distributed differential evolution with adaptive resource allocation. *IEEE Trans. Cybernet* 1–14 (2022). <https://doi.org/10.1109/TCYB.2022.3153964>
2. Zhang, F., et al.: Decision-based evasion attacks on tree ensemble classifiers. *World Wide Web* **23**(5), 2957–2977 (2020)
3. Yin, J., et al.: Vulnerability exploitation time prediction: an integrated framework for dynamic imbalanced learning. *World Wide Web* **25**(1), 401–423 (2022)
4. Sarki, R., et al.: Automated detection of mild and multi-class diabetic eye diseases using deep learning. *Health Inform. Sci. Syst.* **8**(1), 1–9 (2020)
5. Pandey, D., et al.: Automatic breast lesion segmentation in phase preserved DCE-MRIs. *Health Inform. Sci. Syst.* **10**(1), 1–19 (2022)
6. Yin, J., et al.: Knowledge-Driven Cybersecurity intelligence: Software Vulnerability Co-exploitation Behaviour Discovery. *IEEE Trans. Ind. Inform.* 1–9 (2022). <https://doi.org/10.1109/TII.2022.3192027>
7. Supriya, S., et al.: Automated epilepsy detection techniques from electroencephalogram signals: a review study. *Health Inform. Sci. Syst.* **8**(1), 1–15 (2020)
8. You, M., et al.: A knowledge graph empowered online learning framework for access control decision-making. *World Wide Web* 1–22 (2022). <https://link.springer.com/article/10.1007/s11280-022-01076-5>
9. Ge, Y.F., et al.: MDDE: multitasking distributed differential evolution for privacy-preserving database fragmentation. *VLDB J* **31**, 957–975 (2022). <https://doi.org/10.1007/s00778-021-00718-w>
10. He, J., et al.: A framework for cardiac arrhythmia detection from IoT-based ECGs. *World Wide Web* **23**(5), 2835–2850 (2020)

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