

Editorial

Recent advances in semantic computing and personalization



With the rapid growth of social communities and intelligent web services in recent years, there has been a huge volume of user-generated data in the internet every day [1]. To exploit such a large collection of web data, it is essential to identify the underlying high-level semantics of data in multiple modalities and sources. Such semantically-rich information facilitates the understanding of user intentions, needs and preferences. Furthermore, the paradigm of personalized information access, which has been increasingly employed in various mobile and web-based systems to avoid information overload and better satisfy end users' information needs, is supported by various semantic computing techniques such as ontology [2], user profiling [3], social annotations [4] and so on. Many research questions and challenges need to be addressed for the realization of powerful personalized models with semantic computing techniques.

This special issue aims to investigate (i) how high-level semantics are extracted and exploited from web and social data sources via state-of-the-art data mining techniques (e.g., support vector machines [5], Adaboost [6], deep neural networks [7], etc) and (ii) how personalized models are facilitated and supported by underlying semantics. In addition, the research issues in applying semantic computing and personalization in web-based social communities and interactive platforms are covered and discussed.

To balance the quality and coverage of user reviews, **'More Focus on What You Care About: Personalized Top Reviews Set'** presents a review recommendation model which identifies important aspects of the review and selects a top personalized reviews set according to user preferences. The effectiveness of the proposed model is further verified by conducting experiments on the datasets crawled from two e-commerce sites (i.e., Yelp and TripAdvisor).

In **'Real-time Personalized Twitter Search based on Semantic Expansion and Quality Model'**, authors propose a personalized search framework for real-time twitter data stream by implementing the semantic expansion based on user preferences and employing the quality model based on social features. The evaluation is based on a real twitter data stream consisted of 51,770,318 tweets. The experimental findings showed that the framework can improve ranking effectiveness and identify user preferences appropriately.

By analyzing the semantic features in multi-aspect vocal ratings and karaoke machine ratings, a personalized song recommender system based on a joint model is proposed in the article **'Karaoke Song Recommendation Using Multiple Kernel Learning Approximation'**. Specifically, the latent features of the vocal ratings are learnt by a multiple kernel learning method and then fed into the

training classifier for predicting the song ratings. The prediction model demonstrates its effectiveness in song recommendations.

The article **'Document Representation and Feature Combination for Deceptive Spam Review Detection'** tackles the research question of how to detect the deceptive spam reviews. Rather than constructing a set of hand-crafted features based on the linguistic patterns, the latent semantic features such as sentence representation are obtained through the convolution neural networks. In their cross-domain experiments, the learnt semantic features gain 86.1% improvements in F1 values compared to state-of-the-art baselines.

A supervised learning model based on a semantic joint model is proposed in **'A Multi-Relational Term Scheme for First Story Detection'** to identify the first story in the online event stream. Specifically, the semantic joint model extracts the local, global and topical information from the corpus, documents and topics respectively. The proposed schema outperforms extant term models in both retrospective and online experimental datasets.

In **'An Algorithm for Event Detection Based on Social Media Data'**, an event detection model for the foodborne disease in social media is presented. Based on the term features of twitters in a foodborne disease, a supervised learning classifier based on support vector machines is trained. Furthermore, some practical applications such as restaurant recommendations and disease location determination are facilitated by analyzing the latent semantic features of tweets.

To facilitate learning in collaborative environments, **'Discover Learning Path for Group Users: A Profile-based Approach'** develops a learning path discovering model for group learners based on the user profiles. Due to the knowledge diversity and preference variety of learners, it is very challenging to identify an optimal learning path for each group member when they attempt to complete a group task. Through the control and experimental group studies, the effectiveness of the proposed framework is validated. In other words, a framework, which consolidates the various semantic features (including the pre-knowledge, learning preferences, learning topology, and the temporal constraints), is capable to address this research issue.

To improve user perceptions in the virtual reality systems, the interactive semantics with modality technology are integrated in immersive glasses to learn the geographical information as shown in the article **'Virtual Reality Geographical Interactive Scene Semantics Research for Immersive Geography Learning'**. The proposed 3D geography learning model includes 3D topology analysis, profile analysis, cutting analysis, iso-surface and iso-body extraction. Through the case studies in some real scenarios, the extracted

semantics enable the 3D geography learning model to be a powerful tool for building immersive environments.

The article ‘**Semantics-Aware Content-Based Recommender Systems: Design and Architecture Guidelines**’ reviews the extant architecture of semantics-aware content-based recommender systems and elaborates the research issues such as ‘Magic barrier problem’ or ‘Incoherent items in a user profile’. The semantic representation of items, user profiles, filters and feedback of state-of-the-art systems are discussed and summarized. Furthermore, a set of guidelines is developed to address the shortages of these systems by analyzing possible scenarios substantially. The updated architecture with profile cleaner provides a feasible solution to the above problems.

Haoran Xie*

Department of Mathematics and Information Technology, The Education University of Hong Kong, New Territories, Hong Kong, China

Fu Lee Wang

Caritas Institute of Higher Education, New Territories, Hong Kong, China

Xudong Mao

Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong, China

Ke Li

College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, The United Kingdom

Qing Li

Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong, China

Handing Wang

Department of Computer Science, University of Surrey, Guildford, The United Kingdom

*Corresponding author.

E-mail address: hrxie2@gmail.com (H. Xie)

Received 11 February 2017

Accepted 19 February 2017

References

- [1] X. Wu, X. Zhu, G.-Q. Wu, W. Ding, Data mining with big data, *IEEE Trans. Knowl. Data Eng.* 26 (1) (2014) 97–107.
- [2] S. Calegari, G. Pasi, Personal ontologies: generation of user profiles based on the yago ontology, *Inf. Process. Manag.* 49 (3) (2013) 640–658.
- [3] H. Xie, Q. Li, X. Mao, X. Li, Y. Cai, Y. Rao, Community-aware user profile enrichment in folksonomy, *Neural Netw.* 58 (2014) 111–121.
- [4] A. Shraer, M. Gurevich, M. Fontoura, V. Josifovski, Top-k publish-subscribe for social annotation of news, *Proc. VLDB Endow.* 6 (6) (2013) 385–396.
- [5] B. Gu, V.S. Sheng, A robust regularization path algorithm for ν -support vector classification, *IEEE Trans. Neural Netw. Learn. Syst.* (2016), doi:10.1109/TNNLS.2016.2527796.
- [6] X. Wen, L. Shao, Y. Xue, W. Fang, A rapid learning algorithm for vehicle classification, *Inf. Sci.* 295 (2015) 395–406.
- [7] Y. Shen, X. He, J. Gao, L. Deng, G. Mesnil, Learning semantic representations using convolutional neural networks for web search, in: *Proceedings of the 23rd International Conference on World Wide Web*, ACM, 2014, pp. 373–374.



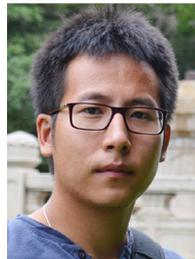
Haoran Xie is an Assistant Professor at The Education University of Hong Kong. He received his PhD in Computer Science from the City University of Hong Kong. His research interests include deep learning, big data, financial and educational data mining. He has published over 80 publications including AAAI, WWW, DASFAA, INTSYS, NEUNET, etc. He is served as guest editors in JMLC, NEUCOM, IJDET and WIJ. He has also served as a co-chair/committee member of WI, CPSCom, GCCCE, U-Media, WISE and ICWL.



Fu Lee Wang is the Vice-President (Research and Advancement) at the Caritas Institute of Higher Education. He received his PhD in Systems Engineering and Engineering Management from The Chinese University of Hong Kong. Prior to joining Caritas, he was a faculty member at the City University of Hong Kong. His research interests include e-business, e-learning, financial engineering, and information retrieval. Professor Wang has over 150 publications and has received 16 grants with a total of more than \$20 million Hong Kong dollars.



Xudong Mao is currently a Hong Kong PhD Fellowship Researcher at Department of Computer Sciences, City University of Hong Kong. He received the BSc degree in Information Security from Nankai University, Tianjin, China and the MPhil degree in City University of Hong Kong, Hong Kong. Before joining CityU HK, he was a research scientist at Alibaba (China), one of largest e-commerce companies. His research interests include deep generative models, object detection and deep learning.



Ke Li is an Assistant Professor at University of Exeter. He earned my PhD degree at the Department of Computer Sciences, City University of Hong Kong and received BSc and MSc degrees both in Computer Sciences and Technology from College of Information Engineering, in Xiangtan University. Before joining Exeter, he was a research fellow in the University of Birmingham. During 2014 to 2015, he was a postdoctoral research scholar in the Michigan State University. His research interests include computational intelligence, particularly, evolutionary computation and machine learning.



Qing Li is a Professor at the City University of Hong Kong. His research interests include object modeling, multimedia databases, social media and recommender systems. He is a Fellow of IET, a senior member of IEEE, a member of ACM SIGMOD and IEEE Technical Committee on Data Engineering. He is the chairperson of the Hong Kong Web Society, and is a steering committee member of DASFAA, ICWL, and WISE Society.



Handing Wang is a Research Fellow at Department of Computer Science, University of Surrey, UK. She is an active researcher in computational intelligence, nature-inspired computation, multi-objective optimization, and surrogate-assisted evolutionary algorithm. She has top-tier publications including *IEEE Transactions on Evolutionary Computation*, *IEEE Transactions on Cybernetics and Information Sciences*.