

Editorial

Shuigeng Zhou^{ID}, Yi-Ping Phoebe Chen^{ID}, and Hiroshi Mamitsuka^{ID}

THIS special section consists of eight papers selected from the accepted papers of the 27th International Conference on Genome Informatics (GIW2016), which was held in Shanghai, China, October 3-5, 2016. Each paper was first reviewed by at least three reviewers of GIW 2016, and after revision, additional reviewing was conducted on the revised manuscript before it was formally accepted. These papers cover diverse topics, including gene clustering, protein-protein interaction network inference, essential proteins identification, glycan structure identification, lncRNA function prediction, lncRNA-disease association prediction, signal transduction network construction, and parallel algorithms.

In "Subspace Weighting Co-Clustering of Gene Expression Data," X. Chen, J.Z. Huang, Q. Wu, and M. Yang proposed a novel algorithm Subspace Weighting Co-Clustering (SWCC in short) for clustering high dimensional gene expression data. They introduced a gene subspace weight matrix to identify the contribution of gene objects in distinguishing different sample clusters, and designed a new co-clustering objective function to recover the co-clusters in the gene expression data. They also developed an iterative algorithm to solve the objective function. Their empirical study validates the proposed method.

In "Reinforce: An Ensemble Approach for Inferring PPI Network from AP-MS Data," B. Tian, Q. Duan, C. Zhao, B. Teng, and Z. He devised an ensemble method Reinforce to infer protein-protein interaction network from AP-MS data. The Reinforce method uses rank aggregation and false discovery rate control. By assuming the null hypothesis that the interaction scores from different scoring methods are randomly generated, it integrates multiple ranking results from different algorithms or different data sets in three steps. Their experimental results show that Reinforce can get more stable and accurate inference results than existing algorithms.

In "Essential Protein Detection by Random Walk on Weighted Protein-Protein Interaction Networks," B. Xu, J. Guan, Y. Wang, and Z. Wang introduced a novel method

Essentiality Ranking (EssRank in short) for improving the accuracy of essential protein detection. EssRank features in 1) evaluating the confidence scores of interactions by integrating various biological information to overcome the low quality problem of PPI data, 2) proposing weighted edge clustering coefficient (WECC) by considering both interaction confidence scores and network topology to calculate edge weights in PPI networks, and evaluating the weight of each node using the sum of WECC values of its linking edges, and 3) employing a random walk method that makes use of both direct and indirect protein interactions to compute protein essentiality iteratively. Their experimental results on the yeast PPI network show that EssRank outperforms most existing methods, including the most commonly-used centrality measures (SC, DC, BC, CC, IC, EC), topology based methods (DMNC and NC) and the data integrating method IEW.

In "An Improved Approach for N-Linked Glycan Structure Identification from HCD MS/MS Spectra," W. Sun, Y. Liu, G.A. Lajoie, B. Ma, and K. Zhang proposed a method to match the input spectra with glycan structures acquired from a glycan structure database by incorporating a de novo sequencing assisted ranking scheme. The proposed method can be used for automated glycan structure identification from HCD MS/MS of glycopeptides. Their experimental results showed that the proposed method is able to identify glycans effectively and performs better than two existing methods.

In "Integrating Multiple Heterogeneous Networks for Novel lncRNA-Disease Association Inference," J. Zhang, Z. Zhang, Z. Chen, and L. Deng developed a new global network-based framework LncRDNetFlow to prioritize disease-related lncRNAs. They used a flow propagation algorithm to integrate multiple networks reflecting various biological information including lncRNA similarity, protein-protein interactions, disease similarity, and the associations between them to infer lncRNA-disease associations. Their experimental results show that the proposed method outperforms significantly the existing approaches.

In "KATZLGO: Large-Scale Prediction of lncRNA Functions by Using the KATZ Measure Based on Multiple Networks," Z. Zhang, J. Zhang, C. Fan, Y. Tang, and L. Deng introduced a global network-based method to predict the functions of large scale human lncRNAs. They first constructed the global network by fusing three heterogeneous networks: lncRNA-lncRNA similarity network, lncRNA-protein association network and protein-protein interaction network, then employed the KATZ measure to calculate similarities between lncRNAs and proteins in the global network. Finally, they annotated lncRNAs with Gene

- S. Zhou is with the Shanghai Key Lab of Intelligent Information Processing, School of Computer Science, Fudan University, Shanghai 200433, China. E-mail: sgzhou@fudan.edu.cn.
- Y.-P. P. Chen is with the Department of Computer Science and Information Technology, La Trobe University, Melbourne, Victoria 3086, Australia. E-mail: phoebe.chen@latrobe.edu.au.
- H. Mamitsuka is with the Bioinformatics Center, Institute for Chemical Research, Kyoto University, Kyoto, 606-8501, Japan. E-mail: mami@kuicr.kyoto-u.ac.jp.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TCBB.2018.2827138

Ontology (GO) terms of their neighboring protein-coding genes according to the KATZ scores. They validated the proposed method by using a manually annotated lncRNA benchmark and a protein-coding gene benchmark with known function annotations, and experimental results show the advantage of the proposed method.

In "MGT-SM: A Method for Constructing Cellular Signal Transduction Networks," M. Li, R. Zheng, Y. Li, F.-X. Wu, and J. Wang developed a new multivariate Granger test-based framework to construct cellular signal transduction networks. They used SVD to evaluate the coefficient matrix from gene expression data and employed Monte Carlo simulation to estimate the significance of directed edges in the derived networks. They applied the proposed method to Yeast Synthetic Network and MDA-MB-468, and achieved better results than using three existing methods.

In "A CPU/MIC Collaborated Parallel Framework for GROMACS on Tianhe-2 Supercomputer," S. Peng, Y. Cui, S. Yang, W. Su, X. Zhang, T. Zhang, W. Liu, and X. Zhao proposed a CPU and Intel® Xeon Phi Many Integrated Core (MIC) collaborating parallel framework to speed up GROMACS by employing the offload mode on a MIC coprocessor. They also optimized GROMACS to run on both CPU and MIC simultaneously.

The publication of this special section is the result of many people's effort. We would like to thank the authors for contributing their works to this special section, and we also appreciate the reviewers' time in reviewing these manuscripts and giving constructive comments. Without their contributions, cooperation and support, this special section cannot come out.

Shuigeng Zhou
Yi-Ping Phoebe Chen
Hiroshi Mamitsuka

Guest Editors



Shuigeng Zhou (M'07) received the bachelor's degree from the Huazhong University of Science and Technology (HUST), in 1988, the master's degree from the University of Electronic Science and Technology of China (UESTC), in 1991, and the PhD degree in computer science from Fudan University, in 2000. He is now a professor with the School of Computer Science, Fudan University, Shanghai, China. He served in the Shanghai Academy of Spaceflight Technology from 1991 to 1997, as an engineer and a senior engineer, (since August 1995) respectively. He was a post-doctoral researcher in the State Key Lab of Software Engineering, Wuhan University, from 2000 to 2002. His research interests include data management, data mining, machine learning, and bioinformatics. He has extensively published in domestic and international journals (including the *IEEE Transactions on Knowledge and Data Engineering*, the *IEEE Transactions on Parallel and Distributed Systems*, the *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *NAR*, *Bioinformatics*, etc.) and conferences (including SIGMOD, VLDB, ICDE, AAAI, IJCAI, ICCV, CVPR, SIGKDD, SODA, RECOMB, and ISMB, etc.). More information about him can be found at <http://admis.fudan.edu.cn/~sgzhou>.



Yi-Ping Phoebe Chen (SM'07) received the BInfTech degree with first class honors, and the PhD degree in computer science (bioinformatics) from the University of Queensland. She is a professor and chair in the Department of Computer Science and Computer Engineering, La Trobe University, Melbourne, Australia. She is the chief investigator of the ARC Centre of Excellence in Bioinformatics. She is currently working on knowledge discovery technologies and is especially interested in their application to genomics and biomedical science. Her research focus is to find the best solutions for mining, integrating, and analyzing complex data structure and functions for scientific and biomedical applications. She has been working in the area of bioinformatics, health informatics, multimedia databases, query system, and systems biology, has coauthored more than 200 research papers, with many published in top journals and conferences. She is steering committee chair of the Asia-Pacific Bioinformatics Conference (founder) and the International Conference on Multimedia Modelling. She has been on the program committees of more than 100 international conferences, including top ranking conferences such as ICDE, ICPR, ISMB, CIKM, etc. More information about her can be found at <http://homepage.cs.latrobe.edu.au/ypchen/index.htm>.



Hiroshi Mamitsuka received the BS degree in biochemistry and biophysics, the ME degree in information engineering, and the PhD (Doctor of Sciences) degree in information sciences from Tokyo University, in 1988, 1991, and 1999, respectively. He is currently a professor with the Bioinformatics Center, Institute for Chemical Research, Kyoto University, and is jointly appointed as a faculty of the School of Pharmaceutical Sciences at the same university. Before he joined Kyoto University, he worked in the Machine Learning group of NEC Research Laboratories in Japan from 1991 to 2002. His research interests include data mining, machine learning, and their applications to biological problems, particularly those over graphs and networks, such as gene networks, medical pathways, and molecular structures. More information about him can be found at <http://www.bic.kyoto-u.ac.jp/pathway/mami/>

▷ For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.