

Guest Editorial for Special Section on BIBM 2014

Illhoi Yoo and Amarda Shehu

THIS special section includes a selection of six papers presented at the Eighth Annual IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2014) held in Belfast, United Kingdom, on November 2-5, 2014. BIBM is a premier research conference that covers all aspects of bioinformatics, biomedicine, and healthcare informatics. The scientific program highlights five research themes: (1) genomics and molecular structure, function, and evolution; (2) computational systems biology; (3) medical informatics and translational bioinformatics; (4) cross-cutting computational methods and bioinformatics infrastructure; and (5) healthcare informatics.

BIBM 2014 had an international program committee of 288 members. All 291 research papers submitted to the conference were reviewed by at least three independent reviewers. After the review process, 111 papers were accepted, of which 56 (19.2 percent) papers were regular research papers and 55 (18.9 percent) were short papers. The authors of six high quality papers were invited to submit an extended version to this special section. Following a rigorous review process, these six papers were selected for publication.

The paper “Discovering Protein-DNA Binding Cores by Aligned Pattern Clustering” by En-Shiun Annie Lee, Ho-Yin (Anotnio) Sze-To, Man-Hon Wong, Kwong-Sak Leung, Terrence Chi-Kong Lau, and Andrew K. C. Wong addresses the problem of modeling protein-DNA binding cores to improve our understanding of protein-DNA (TF-TFBS) binding and gene regulation. A new algorithm is proposed that incorporates residue/nucleotide variations in sequence data. The algorithm additionally makes use of structural models to verify that conserved regions are binding cores. The algorithm is shown to outperform related methods on both precision and running time.

The paper “Multi-View Clustering of Microbiome Samples by Robust Similarity Network Fusion and Spectral Clustering” by Yong Zhang, Xiaohua Hu, and Xingpeng Jiang address the problem of clustering microbiome samples in the presence of multi-view information on a microbiome in the form of genes, functions, and taxonomic assignments. A novel approach is proposed that exploits the ability of similarity

network fusion to efficiently integrate similarities built from each view of data into a unique network representing the full spectrum of the data. The Robust Similarity Network Fusion (RSNF) approach, which combines the strength of random forest and the advantage of SNF at data aggregation, is shown to substantially improve the clustering performance compared to several state-of-the-art methods on several datasets.

The paper “Optimizing Analytical Depth and Cost Efficiency of IEF-LC/MS Proteomics” by Ilona Kifer, Rui M. Branca, Amir Ben-Dor, Linhui Zhai, Ping Xu, Janne Lehtiö, and Zohar Yakhini addresses the cost and analysis time in the IEF LC-MS/MS analytical method used in mass spectroscopy-based protein identification. The proposed approach selects a subset of IEF fractions for LC-MS/MS analysis that is highly informative in the context of a group of proteins of interest without compromising coverage. Algorithmic techniques and heuristics are developed to optimize the selection of the IEF fractions for LC-MS/MS. The benefits of the approach are demonstrated on experimental data obtained from IEF LC-MS/MS analysis of yeast and human samples.

The paper “Analysis of Organization of the Interactome Using Dominating Sets: a Case Study on Cell Cycle Interaction Networks” by Huiru Zheng, Chaoyang Wang, and Haiying Wang addresses the problem of identifying critical and redundant proteins in a protein interaction network. A new method based on the minimum dominating set based approach is developed and implemented as a Cytoscape plugin. A clear distinction is observed between critical and redundant proteins when examining topological parameters, suggesting the central role of critical proteins in the control of a network. The method reveals statistically significant enrichments of proteins that govern human diseases, including cancer-related and virus-targeted genes in the corresponding set of critical proteins.

The paper “Muscle Tissue Labeling of Human Lower Limb in Multi-Channel mDixon MR Imaging: Concepts and Applications” by Matthias Becker and Nadia Magnenat-Thalmann shows how multi-channel images can be used for robust air masking and reliable muscle tissue detection in the human lower limb. The authors exploit additional channels usually discarded in clinical routine and use the common mDixon acquisition protocol for MR imaging. A fully automated approach is demonstrated on four subjects and compared with manual labeling. The approach is shown to provide advanced and intuitive visualization, quantification of tissue types, pose estimation, and initialization of further segmentation methods, making it promising to be useful in clinical environments.

The paper “Improving Recognition of Antimicrobial Peptides and Target Selectivity through Machine Learning

- I. Yoo is with the Department of Health Management and Informatics, University of Missouri School of Medicine, CE718 Clinical Support and Education Building (DC006.00), Five Hospital Dr., Columbia, MO 65212. E-mail: yooil@health.missouri.edu.
- A. Shehu is with the Department of Computer Science, George Mason University, 4400 University Drive, MSN 4A5, Fairfax, VA, 23030. E-mail: amarda.shehu@gmail.com.

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.2016.2567998

and Genetic Programming” by Daniel Veltri, Uday Kamath, and Amarda Shehu seeks to identify the biological signals or features that govern activity in naturally-occurring antimicrobial peptides as a way to assist wet-laboratory studies in designing novel antimicrobial-based therapeutics to combat antibiotic resistant bacteria. The method proposes a novel representation of a biological signature that integrates distal sequence information and develops a genetic programming algorithm to evolve such features towards those that most effectively discriminate between antimicrobial peptides and other peptides of no known antimicrobial activity. The approach is demonstrated capable not only of improving antimicrobial peptide recognition over state-of-the-art methods but also of predicting for the first time target selectivity based on models of activity against only Gram-positive, only Gram-negative, or both types of bacteria.

Ilhoi Yoo
Amarda Shehu
Guest Editors



Ilhoi Yoo received the BS degree in biomedical engineering from Inje University, South Korea, the MS degree in information science from the University of Pittsburgh, and the PhD degree in information science & technology from Drexel University. He is an associate professor of health informatics, University of Missouri School of Medicine, and a core faculty member of the University of Missouri Informatics Institute. His research areas are biomedical information retrieval (PubMed), biomedical/healthcare data mining, biomedical literature mining, and healthcare smartphone applications. He has published 17 journal articles, 6 book chapters, and 24 conference papers (excluding abstracts) in biomedical and health informatics. The total number of the citations of the articles published for the last five years, in which he was either a first or corresponding author is more than 480, according to Google Scholar (as of 5/11/2016).



Amarda Shehu received the BS degree in computer science and mathematics from Clarkson University in Potsdam, NY, in 2002 and the PhD degree in computer science from Rice University in Houston, TX, in 2008, where she was an NIH fellow of the Nanobiology Training Program of the Gulf Coast Consortia. She is an associate professor in the Department of Computer Science at George Mason University. She holds affiliated appointments in the School of Systems Biology and the Department of Bioengineering at George Mason University. Her research contributions are in computational structural biology, biophysics, and bioinformatics with a focus on issues concerning the relationship between sequence, structure, dynamics, and function in biological molecules. Her research on probabilistic search and optimization algorithms for protein structure modeling is supported by various US National Science Foundation (NSF) programs, including Intelligent Information Systems, Computing Core Foundations, and Software Infrastructure. She also received a NSF CAREER award in 2012.

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