Guest Editors Introduction to the Special Section on Bioinformatics Research and Applications

Zhipeng Cai, Oliver Eulenstein, and Cynthia Gibas

◆

THIS special section consists of a selection of papers presented at the Ninth International Symposium on Bioinformatics Research and Applications (ISBRA 2013), which was held at the University of North Carolina at Charlotte, Charlotte, NC, on May 20-22, 2013. The ISBRA symposium provides a forum for the exchange of ideas and results among researchers, developers, and practitioners working on all aspects of Bioinformatics and computational biology and their applications. In 2013, 104 papers were submitted to the conference, including 46 regular submissions (up to 12 pages) and 58 short abstracts (up to four pages), among which 25 papers appeared in the ISBRA proceeding published as volume 7875 of Springer Verlag's Lecture Notes in the Bioinformatics series.

A small number of authors were invited to submit extended versions of their symposium papers to this special section. Following a rigorous review process, four papers were selected for publication. A broad range of Bioinformatics topics have been covered by this special section, including biological networks partition method, semantic similarity of Gene Ontology (GO) terms measurement, incomplete lineage sorting (ILS) method, and protein complexes detection model.

"Partitioning Biological Networks into Highly Connected Clusters with Maximum Edge Coverage" by Falk Hüffner, Christian Komusiewicz, Adrian Liebtrau, and Rolf Niedermeier introduces a combinatorial optimization problem, the HIGHLY CONNECTED DELETION problem. The goal of this optimization problem is to remove as few edges as possible from a graph such that the remaining graph has highly connected components. The authors first demonstrate that HIGHLY CONNECTED DELETION is NP-hard even on four-regular graphs. Subsequently, both exact and heuristic approaches are proposed based on the polynomial time data reduction rules and integer linear programming with column generation.

"Measure the Semantic Similarity of GO Terms Using Aggregate Information Content" by Xuebo Song, Lin Li, Pradip K. Srimani, Philip S. Yu, and James Z. Wang proposes a novel method for the semantic similarity of GO terms measurement. Nowadays, more and more diverse biomedical dataset is annotated by GO terms. With such rapid development of GO, it is very challenging to compute functional or structural similarity of biomedical entities. Song et al. propose a similarity measurement method considering the aggregate information content of all ancestor terms in a graph for each GO term. Note that, aggregate information content not only can implicitly reflect the GO term's location, but also can represent how human being use this GO term. Experimental results indicate that the proposed method outperforms other state-of-the-art methods in terms of correlation with gene expression data.

"Effect of Incomplete Lineage Sorting On Tree-Reconciliation-Based Inference of Gene Duplication" by Yu Zheng and Louxin Zhang analyzes the effect of ILS on gene duplication inference in a species. ILS incurs larger stochastic variation in the topology of a gene tree and it likely introduces false duplication events when a tree reconciliation method is used. The authors investigate the relationship between the effect of ILS on duplication inference in a species tree and its topological parameters. The results indicate the ILS-induced bias should be considered cautiously when gene duplication is inferred via tree reconciliation.

"Detecting Protein Complexes Based on Uncertain Graph Model" by Bihai Zhao, Jianxin Wang, Min Li, Fang-Xiang Wu, and Yi Pan proposes a protein complexes detection algorithm for protein-protein interaction (PPI), called detecting complex based on uncertain graph model (DCU). To access the reliability of high-throughput protein interactions, many computational approaches based on unrealistic graph model, deterministic graphs, have been proposed. This paper intends to investigate the protein complexes detection problem in a more realistic uncertain graph model. The authors propose DCU to predict complexes from a PPI network. The experimental results demonstrate the DCU algorithm outperforms many prior computational methods.

We would like to thank the Program Committee members and external reviewers for volunteering their time to review and discuss symposium papers. We would like to extend special thanks to the Steering and General Chairs of the symposium for their leadership, and to the Finance, Publication, Publicity, and Local Organization Chairs for their hard work in making ISBRA 2013 a successful event. Furthermore, we would like to thank the Editor-in-Chief, Dr. Ying Xu and Associate Editor-in-Chief, Dr. Dong Xu for providing us with the opportunity to showcase some of the exciting research presented at ISBRA 2013 in the *IEEE/ACM*

- *Z. Cai is with the Department of Computer Science, Georgia State University, Atlanta, GA 30303. E-mail: zcai@gsu.edu.*
- *O. Eulenstein is with the Department of Computer Science, Iowa State University, Ames, IA 50011. E-mail: oeulenst@iastate.edu.*
- *C. Gibas is with the Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, Charlotte, NC 28223. E-mail: cgibas@uncc.edu.*

Transactions on Computational Biology and Bioinformatics. Last but not the least, we would like to thank all the ISBRA 2013 authors. The symposium could not continue to thrive without their high quality contributions.

Zhipeng Cai
Oliver Eulenstein
Cynthia Gibas
Guest Editors

Zhipeng Cai is currently an assistant professor in the Department of Computer Science at Georgia State University. He received the BS degree in computer science from the Beijing Institute of Technology, China, and the MS and PhD degrees in computing science from the University of Alberta, Canada. His research interests include wireless networks and optimization theory. Dr. Cai received a number of awards and honors, including an US NSF CAREER Award and a NSERC Postdoc Fellowship.

Oliver Eulenstein studied computer science at the University of Paderborn (Germany) and the University of Bonn (Germany) and received the doctoral degree in 1998. From 1998 to 2000, he held a postdoctoral position at the University of California, Davis, and became an assistant professor in the Department of Computer Science at Iowa State University in 2000. He has been at Iowa State University since then and is currently an associate professor. His research interest is in developing mathematical and algorithmic solutions for problems arising in molecular biology, in particular evolutionary biology.

Cynthia Gibas is a professor in the Department of Bioinformatics and Genomics at the University of North Carolina at Charlotte. She received the PhD degree in biophysics and computational biology from the University of Illinois at UrbanaChampaign in 1996, working with Shankar Subramaniam. She subsequently worked as a bioinformatics analyst at the National Center for Supercomputing Applications and as an assistant professor in the Department of Biology at Virginia Tech. Her research interests are 1) development of visual analytics and business intelligence-inspired methods for comparative analysis of large bacterial genome and transcriptome data sets and 2) understanding the impact of biophysical properties of nucleic acids on the outcomes of high-throughput genomics protocols. She also collaborates with an international consortium in the study of genomics and transcriptomics of Vibrio vulnificus and has a broad interest in improving methods for comparative genomics. Dr. Gibas is also the faculty director of UNC Charlotte's Professional Science Masters in Bioinformatics.

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