The 12th International Symposium on Operations Research and Its Applications
The 9th International Conference on Systems Biology
(ISORA/ISB 2015)

Organizers

August 21-24, 2015
Luoyang, China

ISORA/ISB 2015 Schedule

August 21 Friday	15:00-22:00	Registration (hotel lobby at Dongshan Hotel)	
	18:00-20:00	Dinner	
August 22 Saturday	08:30-08:50	Opening Session (Chair: Xiang-Sun Zhang)	
	08:50-10:30	ISORA/ISB Plenary Session P1 Speakers: Hiro Ito, Sunghoon Kim Chair: Xiang-Sun Zhang	
	10:30-10:50	Coffee break	
	10:50-12:30	ISORA Session A1 Optimization Paper ID: a4, a11, a15, a28, a43 Chair: Tatsuo Oyama	ISB Session B1 Highlight I Paper ID: b3, b114, b101, b103, b104 Chair: Zhaolei Zhang
	12:30-14:00	Lunch	
	14:00-15:40	ISORA Session A2 OR Applications Paper ID: a9, a18, a25, a38, a51 Chair: Chuangyin Dang	ISB Session B2 Invited Talks Speakers: Ting Lu, Zhaolei Zhang Chair: Sunghoon Kim
	15:40-16:00	Coffee break	
	16:00-18:00	ISORA Session A3 Combinatorics and Graph Theory Paper ID: a5, a10, a12, a16, a21, a36 Chair: Ping Ji	ISB Session B3 Highlight II Paper ID: b111, b106, b107, b108, b109, b110 Chair: Jinzhi Lei
	18:30-20:00	Reception	
	20:00-21:30	Board member meeting of ORSC-CSB	

ISORA/ISB 2015 Schedule

August 23 Sunday	07:00-12:30	Half day excursion in Longmen Grottoes	
	12:30-14:00	Lunch	
	14:00-15:40	ISORA Session A4 Bioinformatics and Complex Networks Paper ID: a27, a30, a32, a41, a47 Chair: Zhenping Li	ISB Session B4 Bioinformatics Paper ID: b7, b12, b115, b116, b20 Chair: Fengfeng Zhou
	15:40-16:00	Coffee break	
	16:00-18:00	ISORA Session A5 Logistics and Supply Chain Management Paper ID: a7, a20, a24, a42, a44, a8 Chair: Yu Song	ISB Session B5 Highlight III Paper ID: b4, b13, b102, b105, b112, b113 Chair: Shihua Zhang
	18:30-20:00	Banquet	
August 24 Monday	8:30-10:10	ISORA/ISB Plenary Session P2 Speakers: Jin Wang, Chao-Nan Qian Chair: Luonan Chen	
	10:10-10:30	Coffee break	
	10:30-12:10	ISORA Session A6 OR in Management Paper ID: a13, a26, a39, a45 Chair: Wuyi Yue	ISB Session B6 Systems Biology Paper ID: b5, b8, b10, b18, b16 Chair: Zhiping Liu
	12:30-14:00	Lunch	
	14:00-15:40	ISORA Session A7 OR Applications Paper ID: a19, a34, a35, a46 Chair: Jihong Zhang	ISB Session B7 Systems Biology Paper ID: b6, b9, b14, b15, b17 Chair: Lin Gao

ISORA/ISB 2015 Program
August 21-24, Luoyang, China

August 21 (Friday) Registration

15:00-22:00 Registration
Participants arrive in Luoyang, check in Dongshan Hotel, and registration package pick up (Hotel lobby at Dongshan Hotel).

18:00-20:00 Dinner

Locations:

Hotel lobby: Building 5
Conference rooms: Floor 1 of Building 6
Big Multi-functional Room: Opening Session, Plenary Sessions, Sessions A1-A6
Small Multi-functional Room: Sessions B1-B6, ORSC-CSB Board Meeting

August 22 (Saturday) Technical sessions

08:30-11:30 Registration for late arrivals (Floor 1 of Building 6)

08:30-08:50 Opening Session
Chair: Xiang-Sun Zhang

8:50-10:30 ISORA/ISB Plenary Session P1
Chair: Xiang-Sun Zhang

8:50-09:40 *Generalized shogi and chess are constant-time testable*
Hiro Ito
The University of Electro-Communications, Japan

09:40-10:30 *Functional expansion of ancient protein synthesis enzymes for system control*
Sunghoon Kim
Seoul National University, Korea

10:30-10:50 Coffee break

10:50-12:30 ISORA Session A1
Topic: Optimization
Chair: Tatsuo Oyama

10:50-11:10 *An efficient algorithm for linear semi-infinite programming over positive polynomials*
Meiling Xu and Zongwei Luo
South University of Science and Technology of China, China
Paper ID: a4

11:10-11:30 *Computational complexity of inverse word search problem*
Hiro Ito and Shinnosuke Seki
University of Electro-Communications, Japan
Paper ID: a11

11:30-11:50 *A Method for Calculating Probability of Scores for Men's Team Competition in Artistic Gymnastics*
Nobuyoshi Hirotsu, Mutsumi Harada and Minoru Kano
Juntendo University, Japan
Paper ID: a15

11:50-12:10 *Projection Method for Support Vector Machines with Indefinite Kernels*
Hao Jiang, Wai-Ki Ching, Yushan Qiu and Xiaoqing Cheng
Renmin University of China, China
Paper ID: a28

12:10-12:30 *Sufficient Optimal Conditions for Unconstrained Quadratic Binary Problems*
Liu Liu, Chunli Liu and Qiuling Xie
Shanghai University of Finance and Economics, China
Paper ID: a43

10:50-12:30 ISB Session B1
Topic: Highlight I
Chair: Zhaolei Zhang

10:50-11:10 *Online learning approach to prediction of protein-protein interaction strengths*
Morihiro Hayashida, Mayumi Kamada and Hitoshi Koyano
Kyoto University, Japan
Paper ID: b3

11:10-11:30 *MeSHLabeler: improving the accuracy of large-scale MeSH indexing by integrating diverse evidence*
Shanfeng Zhu
Fudan University, China
Paper ID: b114

11:30-11:50 *GPU-parallelized constrained feature selection algorithm improves the breast cancer subtyping model*
Ren Zhong, Guoqin Mai and Fengfeng Zhou
Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China
Paper ID: b101

11:50-12:10 *Integrated omics analysis in tumor genesis of Hepatocellular Carcinoma (HCC)*
Lei Liu
Fudan University, China
Paper ID: b103

12:10-12:30 *Stem Cell Regeneration as a multi-scale dynamical system: From modeling to simulation*
Jinzhi Lei
Tsinghua University, China
Paper ID: b104

12:30-14:00 Lunch break

14:00-15:40 ISORA Session A2
Topic: OR Applications
Chair: Chuangyin Dang

14:00-14:20 *Research on the task assignment problem of warehouse robots in the smart warehouse*
Zhenping Li, Wenyu Li and Lulu Jiang
Beijing Wuzi University, China
Paper ID: a9

14:20-14:40 *Election of Best Nine in Nippon Professional Baseball Based on the DEA Model*
Tohru Ueda and Hirofumi Amatatsu
Seikei University, Japan
Paper ID: a18

14:40-15:00 *Statistical data analyses of recent national elections in Japan*
Akm Abul Kalam Azad, Md. Arifur Rahman and Tatsuo Oyama
National Graduate Institute for Policy Studies, Japan
Paper ID: a25

15:00-15:20 *Employing post-dea method in budget management of health care sectors*
Xiaoya Li
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China
Paper ID: a38

15:20-15:40 *Enrichment Set Cover Problem*
Yinliang Liu, Xiang-Sun Zhang, Ling-Yun Wu
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China
Paper ID: a51

14:00-15:40 ISB Session B2
Topic: Invited Talks
Chair: Sunghoon Kim

14:00-14:50 *Engineering Bacteria from Single Cells to Multispecies Communities*
Ting Lu
University of Illinois Urbana-Champaign, USA
14:50-15:40 *New Computational Methods in Analyzing MicroRNA Regulations in Cancer*
Zhaolei Zhang
University of Toronto Faculty of Medicine, Canada

15:40-16:00 Coffee break

16:00-18:00 ISORA Session A3
Topic: Combinatorics and Graph Theory
Chair: Ping Ji

16:00-16:20 *Replacement first and last for a parallel system with constant and random units*
Xufeng Zhao, Cunhua Qian, Syouji Nakamura and Toshio Nakagawa
Qatar University, Qatar
Paper ID: a5
16:20-16:40 *A method for computing a sequence of circumscribing polygons and its analysis*
Kensuke Onishi
Tokai University, Japan
Paper ID: a10
16:40-17:00 *A polynomial-space exact algorithm for tsp in degree-5 graphs*
A polynomial-space exact algorithm for tsp in degree-5 graphs
Technical University of Malaysia Malacca, Malaysia
Paper ID: a12
17:00-17:20 *A method for generating colorings over graph automorphism*
Fei He and Hiroshi Nagamochi
Kyoto University, Japan
Paper ID: a16
17:20-17:40 *The Equal-Subsets Problem Is in the Class PPAD*
Chuangyin Dang
City University of Hong Kong, Hong Kong
Paper ID: a21
17:40-18:00 *An improved algorithm for the machine scheduling problem with job delivery coordination*
Yuzhong Zhang, Qiongyi Zheng, Jianfeng Ren and Long Zhang
Qufu Normal University, China
Paper ID: a36

16:00-18:00 ISB Session B3
    Topic: Highlight II
    Chair: Jinzhi Lei

16:00-16:20 *Tunable Sensitivity and Its Biological Insights*
    Guanyu Wang
    South University of Science and Technology of China, China
    Paper ID: b111

16:20-16:40 *Combining genetic and environmental variations to infer genes for phenotypes in crops*
    Jingdong Liu
    Monsanto Company, USA
    Paper ID: b106

16:40-17:00 *ARCS: Assemble short-reads by using combinatorial optimization in scaffolding*
    Dongbo Bu
    Institute of Computing Technology, Chinese Academy of Sciences, China
    Paper ID: b107

17:00-17:20 *Reconstructing phylogenetic trees for ultra-large unaligned DNA sequences via with Hadoop*
    Quan Zou
    Tianjin University, China
    Paper ID: b108

17:20-17:40 *RNA structurome reveals the second layer of genetic information*
    Qiangfeng Zhang
    Tsinghua University, China
    Paper ID: b109

17:40-18:00 *The human microbe-disease network*
    Qinghua Cui
    Peking University, Chin
    Paper ID: b110

18:30-20:00 Welcome Reception

20:00-21:30 Board member meeting for Computational Systems Biology Society of ORSC

August 23 (Sunday) Technical sessions

07:00-12:00   Half day excursion in Longmen Grottoes

12:30-14:00 Lunch break

14:00-15:40 ISORA Session A4
Topic: Bioinformatics and Complex Networks
Chair: Zhenping Li

14:00-14:20 *Fast community detection algorithm based on relationship strength coupling in social networks*
Ye Li and Hui-Jia Li
Central University of Finance and Economics, China
Paper ID: a27

14:20-14:40 *On Classification of Biological Data Using Outlier Detection*
Yushan Qiu, Xiaoqing Cheng, Wenpin Hou and Wai-Ki Ching
The University of Hong Kong, Hong Kong
Paper ID: a30

14:40-15:00 *A Multiple Linear Regression Model for Structure of N-linked Oligosaccharides*
Xiaoqing Cheng, Wai-Ki Ching, Wenpin Hou and Kiyoko Aoki-Kinoshita
The University of Hong Kong, Hong Kong
Paper ID: a32

15:00-15:20 *Meteorological station network and its application to forecast infestation of tropilaelaps mites*
Jin-Shan Li
Beijing Forestry University, China
Paper ID: a41

15:20-15:40 *Melanocytic globules detection in skin lesion images*
Leszek Nowak, Kasia Grzesiak-Kopec and Maciej Ogorzalek
Jagiellonian University, Poland
Paper ID: a47

14:00-15:40 ISB Session B4
Topic: Bioinformatics
Chair: Fengfeng Zhou

14:00-14:20 *A Tool: Identification of Feature Genes from Genomics Data Using Factor Analysis*
Changhe Fu, Su Deng, Jinghua Wu, Xianqiong Wu and Zhi-Hui Fu
Shenyang Normal University, China
Paper ID: b7

14:20-14:40 *Detecting disease genes of non-small lung cancer based on consistently differential interactions*
Xiaoping Liu, Qianqian Shi and Luonan Chen
Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, China
Paper ID: b12

14:40-15:00 *Discrimination of recurrent CNVs from individual ones from multisample aCGH by jointly constrained minimization*
Hongmin Cai
South China University of Technology, China
Paper ID: b115

15:00-15:20 *WGX-50, A Drug Candidate from Sichuan Pepper and Its Potential Role Against AD And in Anti-aging*
Dong-Qing Wei
Shanghai Jiaotong University, China
Paper ID: b116

15:20-15:40 *Intra-residue atom contacts outperform other features in determining protein-protein interface residues*
Yongxiao Yang and Xinqi Gong
Renmin University of China, China
Paper ID: b20

15:40-16:00 Coffee break

16:00-18:00 ISORA Session A5
Topic: Logistics and Supply Chain Management
Chair: Yu Song

16:00-16:20 *A model and method for supply chain management problems*
Guirong Pan and Hongchun Sun
Linyi University, China
Paper ID: a7

16:20-16:40 *An inventory routing problem with soft time windows*
Zhenping Li, Chongyu Jiang and Lulu Jiang
Beijing Wuzi University, China
Paper ID: a20

16:40-17:00 *Analysis of a Markovian Queue with Two Heterogenous Servers and a Threshold Assignment Policy*
Dequan Yue, Hui Li, Guoxi Zhao and Wuyi Yue
Yanshan University, China
Paper ID: a24

17:00-17:20 *The pricing and copyright's payment strategy in online movie distribution*
Ming-Xi Zhu, Ji-Hong Zhang and Xiao-Shuang Han
Beijing Foreign Studies University, China
Paper ID: a42

17:20-17:40 *Solve the vehicle routing problem via a discrete particle swarm optimization*
Hongwei Liu
Beijing Wuzi University, China
Paper ID: a44

17:40-18:00 *Study of multi-vehicle routing problem with time window*
Yufeng Bai, Xiaoguang Zhou, Yuxiang Zhang and Menggu Yang
Beijing University of Posts and Telecommunications, China
Paper ID: a8

16:00-18:00 ISB Session B5
Topic: Highlight III
Chair: Shihua Zhang

16:00-16:20 *Systematic identification of transcriptional and post-transcriptional regulations in human respiratory epithelial cells during influenza A virus infection*
Zhi-Ping Liu
Shandong University, China
Paper ID: b4

16:20-16:40 *Biomarker discovery in biomedical big data*
Tao Zeng and Luonan Chen
Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, China
Paper ID: b13

16:40-17:00 *lncRNA-MFDL: identification of human long non-coding RNAs by fusing multiple features and using deep learning*
Xiao-Nan Fan and Shao-Wu Zhang
Northwestern Polytechnical University, China
Paper ID: b102

17:00-17:20 *CCLasso: Correlation Inference for Compositional Data through Lasso*
Huaying Fang, Chengcheng Huang, Hongyu Zhao, and Minghua Deng
Peking University, China.
Paper ID: b105

17:20-17:40 *Inferring Direct PPI Networks from AP-MS Data*
Zengyou He
Dalian University of Technology, China
Paper ID: b112

17:40-18:00 *Organelle-focused interactome of rice*
Ming Chen
Zhejiang University, China
Paper ID: b113

18:30-20:00 Banquet

August 24 (Monday) Technical Sessions

8:30-10:10 ISORA/ISB Plenary Session P2
Chair: Luonan Chen

8:30-09:20 *The landscape and flux theory for biological networks*
Jin Wang
Stony Brook University, USA

09:20-10:10 *Identifying the key molecules promoting metastasis of nasopharyngeal carcinoma cells*
Chao-Nan Qian
Sun Yat-sen University Cancer Center, China

10:10-10:30 Coffee break

10:30-12:10 ISORA Session A6
Topic: OR in Management
Chair: Wuyi Yue

10:30-10:50 *A stackelberg game theoretic approach to competitive product portfolio management*
Xiaojie Liu, Gang Du and Yi Xia
Tianjin University, China
Paper ID: a13

10:50-11:10 *The Study of Evaluation of Regional Innovation Capability of the High-tech Industry*
Ren Lin and Guohua Zhou
Hunan city University, China
Paper ID: a26

11:10-11:30 *Mining online product reviews to identify consumers' fine-grained concerns*
Jian Jin and Ping Ji
Beijing Normal University, China
Paper ID: a39

11:30-11:50 *Optimization of Artificial Neural Network for Stock Market Return Prediction*
Mingyue Qiu and Yu Song
Fukuoka Institute of Technology, Japan
Paper ID: a45

10:30-12:10 ISB Session B6
Topic: Systems Biology
Chair: Zhiping Liu

10:30-10:50 *Identifying early-warning signals of critical transitions by distribution embedding*
Rui Liu, Pei Chen, Kazuyuki Aihara and Luonan Chen
South China University of Technology, China
Paper ID: b5

10:50-11:10 *A reflected stochastic differential equation model for biochemical reaction systems*
Yuanling Niu
Shanghai Institutes for Biological Science, Chinese Academy of Sciences, China
Paper ID: b10

11:10-11:30 *PDCD5 interacts with p53 and functions as a regulator of p53 dynamics in DNA damage response*
Changjing Zhuge, Xiaojuan Sun, Yingyu Chen and Jinzhi Lei
Bejing Forestry University, China
Paper ID: b8

11:30-11:50 *Critical stage with early-warning signals during chronic inflammation to hepatocellular carcinoma*
Meiyi Li, Chen Li, Rong Zeng and Luonan Chen
Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, China
Paper ID: b18

11:50-12:10 *CHiP-seq data analysis of DAF-16 transcription factor binding site*
Lu Wang and Jiarui Wu
Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, China
Paper ID: b16

12:30-14:00 Lunch

14:00-15:40 ISORA Session A7
Topic: OR in Management
Chair: Jihong Zhang

14:00-14:20 *The Multiple Knapsack Problem with Compatible Bipartite Graphs*
Jianping Li, Weidong Li and Hao Wang
Yunnan University, China
Paper ID: a19

14:20-14:40 *MPS: a web server for Metabolism Pathway Synthesis on E.coli*
Anjun Zhu
University of Science and Technology of China, China
Paper ID: a34

14:40-15:00 *A Signaling Pathway Analysis Method Based on Information Divergence*
Hang Wei and Haoran Zheng
University of Science and Technology of China, China
Paper ID: a35

15:00-15:20 *Medical treatment capability analysis based on m/m/s/gd/c/∞ model*
Juyun Wang, Jie Sui, Mei Jia Li and Hua Yu
Communication University of China, China
Paper ID: a46

14:00-15:40 ISB Session B7
Topic: OR in Management
Chair: Lin Gao

14:00-14:20 *Singularity Analysis of Grb10 signaling pathway quantitatively reveals Grb10 promotes lipolysis and thermogenesis*
Hao Kang, Guanyu Wang and Luonan Chen
ShanghaiTech University, China
Paper ID: b6

14:20-14:40 *Hopf bifurcation in a prey-predator model with two time delays*
Yiming Ding and Yan Meng

University of Science and Technology Beijing, China
Paper ID: b9

14:40-15:00 *Identifying progressive biomarkers utilizing time-course information*
Jingxue Xin, Tao Zeng, Yong Wang and Luonan Chen
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China
Paper ID: b14

15:00-15:20 *Identifying specific and consistent dysfunctional modules from multiple cancer sites*
Chuan-Chao Zhang, Qianqian Shi, Tao Zeng, Juan Liu and Luonan Chen
Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, China
Paper ID: b15

15:20-15:40 *Quantifying direct associations in networks by sample-based data*
Juan Zhao and Luonan Chen
Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, China
Paper ID: b17

*The above program subjects to revision based on further information and Ad Hoc presentation requests.

Book of Abstracts

ISORA/ISB Plenary Sessions

Generalized shogi and chess are constant-time testable

Hiro Ito

The University of Electro-Communications, Japan

We present constant-time testing algorithms for the generalized shogi (Japanese chess) and chess. These problems are known to be EXPTIME-complete. A testing algorithm (or a tester) for a property accepts an input if it has the property and rejects it if it's far from having the property in high probability (e.g., at least 2/3) by reading only a constant part of the input. A property is said to be testable if there is a tester. The generalized shogi (and chess) problem is, given any position on $\sqrt{n} \times \sqrt{n}$ board with $O(n)$ pieces, for testing the property ``the player who moves first has a winning strategy.'' We present that this property is testable for both shogi and chess. The shogi tester is one-sided-error and the chess tester is surprisingly no-error! In the last decade, many problems have been found to be testable. However, almost all of such problems are in class NP. This is the first result on constant-time testability of EXPTIME-complete problems.

Functional expansion of ancient protein synthesis enzymes for system control

Sunghoon Kim

Seoul National University, Korea

Aminoacyl-tRNA synthetases (ARSs) are multi-functional proteinsplayers, not only playing catalytic roles in protein synthesis as catalysts, but also mediating regulatory roles in many signal pathways. Being an ancient enzymes decoding genetic information to proteins, they have evolved to acquire additional functions through various structural modifications such as new domain addition, post-translational modifications, peptide cleavage and alternative splicing. Among many interesting protein complexes mediated by ARSs, systematic interactome studies have been conducted to a macromolecular protein complex consisting of nine different ARSs and three auxiliary factors named AIMP1-3. The results suggest insights into the structural organization and functional implications of ARS-mediated functional complexes.

The landscape and flux theory for biological networks

Jin Wang

Stony Brook University, USA

We developed a landscape and flux theory for biological networks [1]. We identify the two global driving forces for biological network. One is the gradient of the underlying landscape and the other is from the curl flux. The underlying landscape is linked to the steady state probability distribution and provides a global picture for describing the networks. We found that the landscape can be used to quantify the global stability and robustness of the system. The non-zero flux breaks the detailed balance and therefore gives a quantitative measure of how far away the system is from the equilibrium state, reflecting the degree of the energy input to the system. Our decomposition of the driving forces of the complex systems into landscape gradient and curl flux establishes the link between the dynamics and the underlying thermodynamic non-equilibrium natures. We applied our theory to several biological systems such as cell cycle[2], stem cell differentiation and reprograming [3,4], cancer [5]. For cell cycle oscillations [2], we found the underlying landscape has a Mexican hat ring shape topology. The height of the center island Mexican hat determines the global stability. The landscape gradient attracts the system down to the oscillation ring. The curl flux is the driving force for coherent oscillation on the ring. Along the cell cycle oscillation ring there are a few basins of attractions which can be identified and quantitatively described as the G1, SG2 and M phases. The barriers between these local basins become the check points of the cell cycle. The speed of the cycle is determined by the flux originated from nutrition and the barriers at the check points between the basins. Global sensitivity analysis on these two global factors gives us information on key genes and regulations determining the function. From this, new anti-cancer strategy can be designed aiming at reducing the speed of the cell cycle. We also applied our landscape and flux theory to stem cell differentiation and development. We quantify the Waddington landscape for differentiation and identify the stem cell and differentiation basins of attractions. We quantify the differentiation and reprograming path for a human embryonic stem cell network and identify the key genes and regulations [3,4]. We also constructed a cancer gene network and quantify the landscape for cancer where we identify the normal, cancer and apoptosis basins of attractions. We quantify the major pathways of cancerization and identify the key genes and regulations responsible. [5]

1. J. Wang, L. Xu, E. K. Wang. Proc. Natl. Acad. Sci. USA, 105: 12271-12276. (2008).
2. C.H. Li, J. Wang. Proc. Natl. Acad. Sci. USA. 111(39), 14130-14135 (2014)
3. J. Wang, K. Zhang, L. Xu, E.K. Wang. Proc. Natl. Acad. Sci. USA. 108(20):8257-8262(2011).
4. C.H. Li, J. Wang. PLoS Comput Biol., 9(8): e1003165 (2013).
5. C.H. Li, J. Wang. J. R. Soc. Interface. 11: 20140774. (2014).

Indentifying the key molecules promoting metastasis of nasopharyngeal carcinoma cells

Chao-Nan Qian

Sun Yat-sen University Cancer Center, China

Nasopharyngeal carcinoma (NPC) is a common malignancy in southern China and Southeast Asia with the highest metastasis rate among head and neck cancers. In order to explore the molecular mechanisms underlying NPC metastasis, we have established the cellular and animal models. By high through-put gene expression profiling and following functional studies, we have so far identified several key molecules promoting NPC metastasis, including serglyin, interleukin 8 (IL-8), and WNT5A. Among them, IL-8 is the most recognized molecule as an inflammatory and pro-angiogenic factor for years. We have found that IL-8 strongly promotes metastasis of NPC cells via autocrine and paracrine means, involving activation of AKT signaling and inducing EMT in NPC cells. IL-8 can also be secreted into circulation. Our simultaneously detection of serum IL-8 together with 38 other circulating cytokines found that IL-8 was an important prognostic marker in multiple malignancies. In summary, targeting the key molecules promoting NPC metastasis is a promising therapeutic strategy, and the strategies to target different categories of these key molecules will be discussed.

ISB Invited Sessions

Engineering Bacteria from Single Cells to Multispecies Communities

Ting Lu

Department of Bioengineering, University of Illinois at Urbana-Champaign, USA

My research focuses on the analysis, construction and exploitation of bacterial gene networks for functionality programming. In this talk, I will report a pathway engineering platform for lactic acid bacteria that we recently developed, and illustrate its applications in bacteriocin overproduction and pathway optimization. I will also discuss a multiscale biophysical framework for modeling bacterial communities, and illustrate its applications in understanding the spatiotemporal patterns of bacterial populations.

New Computational Methods in Analyzing MicroRNA Regulations in Cancer

Zhaolei Zhang

Donnelly Centre for Cellular and Biomolecular Research, University of Toronto Faculty of Medicine, Canada

microRNAs are a class of abundant small RNAs that can repress the translation of thousands of genes. They have critical regulatory functions in many essential cellular pathways, and perturbations in microRNA pathways often contribute to human diseases such as cancer. In this talk I will describe some of the computational approaches that we have recently developed, which include (i) a target prediction method that can integrate sequence information with expression data, (ii) a framework that model the competition between mRNAs for common microRNA regulators, and (iii) a pan-cancer analysis of microRNAs in the TCGA dataset. I will further describe a new approach that can integrate multiple omics data sets to model the transcriptional regulation of human genes; I will use acute myeloid leukemia (AML) as an example to illustrate the performance of this approach.

ISORA Sessions

Paper ID: a4

An efficient algorithm for linear semi-infinite programming over positive polynomials

Meiling Xu and Zongwei Luo

South University of Science and Technology of China, China

This paper describes an ecient implementation of a form of linear semi-innite programming (LSIP). We look at maximizing (minimizing) a linear function over a set of constraints formed by positive trigonometric polynomials. Previous studies about LSIP are formulated using semi-denite programming (SDP), this is typically done by using the Kalman Yakubovich Popov (KYP) lemma or using a trace operation involving a Grammian matrix, which can be computationally expensive. The proposed algorithm is based on simplex method that directly solves the LSIP without any parameterization. Numerical results show that the proposed LISP algorithm is signicantly more ecient than existing SDP solvers using KYP lemma and Grammian matrix, in both execution time and memory.

Paper ID: a5

Replacement first and last for a parallel system with constant and random units

Xufeng Zhao, Cunhua Qian, Syouji Nakamura and Toshio Nakagawa
Qatar University, Qatar

This paper observes optimal replacement times for a parallel system with n units, when it is operating for successive jobs with a random working cycle. The classical approach of whichever occurs first and the newly proposed approach of whichever occurs last are respectively employed for replacements scheduled at time T and at working cycle Y , whose policies are called replacement rst and replacement last. Two cases when the number of units for this parallel system is a given constant value and a random variable with estimated distribution are considered into modelings. We obtain their expected replacement cost rates and optimal solutions analytically. Further, comparisons of replacement rst and replacement last are described in detail to determine which policy could save more replacement cost rate.

Paper ID: a7

A model and method for supply chain management problems

Guirong Pan and Hongchun Sun
Linyi University, China

In this paper, a mathematical model on supply chain management problem is given. To present optimal decision for the problem, we propose a new type of algorithm, and the global convergence of algorithm is also established under milder conditions. Furthermore, we prove that the method has R-linear convergence rate with the underlying mapping being P-uniform function and Lipschitz continuous.

Paper ID: a8

Study of multi-vehicle routing problem with time window

Yufeng Bai, Xiaoguang Zhou, Yuxiang Zhang and Menggu Yang
Beijing University of Posts and Telecommunications, China

The vehicle routing problem (VRP) is an attractive topic in logistics research work. Multi-vehicle routing problem with time window (MVRTW) is a variant of VRP, which accommodates realistic system specifics such as capacity of multi-vehicle, time constraint and network constraint (one-way, banning of turning movement etc.). To solve the MVRPTW, an improved approach combining geographical information system (GIS) with parallel genetic algorithm (PGA) is proposed. Shortest paths could be calculated by spatial analysis module and topology construction of road network in GIS. In order to strengthen the search ability, an adaptive generation mechanism of initial population and evolutionary operators are used in PGA. The suggested approach proved to be efficient by a practical case of Changchun City.

Paper ID: a9

Research on the task assignment problem of warehouse robots in the smart warehouse

Zhenping Li, Wenyu Li and Lulu Jiang
Beijing Wuzi University, China

The task assignment problem of warehouse robots in the smart warehouse based on cargo-to-person is investigated. Firstly, given the sites of warehouse robots and the order picking tasks, the task assignment problem of warehouse robots is formulated into a mathematical model to mini-mize the total operation cost. Then the heuristic algorithm is designed to solve the task assignment problem. Finally, simulations are done by using the orders data of online bookstore A. The feasibility and validity of the model and algorithm are verified. The model and algorithm in this paper provide a theoretical basis to make the task assign-ment schemes of warehouse robots in the smart warehouse.

Paper ID: a10

A method for computing a sequence of circumscribing polygons and its analysis

Kensuke Onishi
Tokai University, Japan

The computational complexity of the two methods is O(n^2) time and O(n) space for a simple polygon with n vertices. We applied the methods to sets of simple polygons. We measured the number of intersection checks of two line segments and the execution time of each methods.

Paper ID: a11

Computational complexity of inverse word search problem

Hiro Ito and Shinnosuke Seki

University of Electro-Communications, Japan

Word search is a classical puzzle to search for all given words on a given assignment of letters to a rectangular grid (matrix). This problem is clearly in P. The inverse of this problem is more difficult, which asks to assign letters in a given alphabet to a matrix of given size so that every word in a given wordset can be found horizontally, vertically, or diagonally. This problem is in NP; it admits a trivial polynomial-size certificate. We prove its NP-hardness. It turns out to be so even under the following restrictions: 1) the alphabet size is 2 (binary) and 2) all the words to be found are of length at most 2. These results are optimal in the sense that decreasing these bounds 2 to 1 makes the problem be trivially in P.

Paper ID: a12

A polynomial-space exact algorithm for tsp in degree-5 graphs

Norhazwani Yunos, Aleksandar Shurbevski and Hiroshi Nagamochi

Technical University of Malaysia Malacca, Malaysia

The Traveling Salesman Problem (TSP) is one of the most well-known NP-hard optimization problems. Following a recent trend of research which focuses on developing algorithms for special types of TSP instances, namely graphs of limited degree, and thus alleviating a part of the time and space complexity, we present a polynomial-space branching algorithm for the TSP in graphs with degree at most~5, and show that it has a running time of $O^*(2.4531^n)$. To the best of our knowledge, this is the first exact algorithm specialized to graphs of such high degree. While the base of the exponent in the running time bound is greater than two, our algorithm uses space merely polynomial in an input instance size, and thus by far outperforms Gurevich and Shelah's~$O^*(4^n n^{\log n})$ polynomial-space exact algorithm for the general TSP (Siam Journal of Computation, Vol. 16, No. 3, pp. 486-502, 1987). In the analysis of the running time, we use the measure-and-conquer method, and we develop a set of branching rules which foster the analysis of the running time.

Paper ID: a13

A stackelberg game theoretic approach to competitive product portfolio management

Xiaojie Liu, Gang Du and Yi Xia
Tianjin University, China

We are concerned with a product portfolio management problem in which a firm wants to enter a competitive market by offering new products where there are existing products belonging to a competitor. The objective of the new entrant firm is to find out an optimal product portfolio that maximizes its expected shared surplus. The market incumbent firm can react by adjusting its existing product portfolio through offering new products or closing old ones with the aim of maximizing its own expected shared surplus. We formulate a bilevel zero-one integer nonlinear programming model based on the Stackelberg leader-follower game where the new entrant firm is the leader and the incumbent firm is the follower. In the absence of a closed-form solution, the new model is illustrated through a numerical case under different market scenarios and calculated results are compared with Nash equilibrium.

Paper ID: a15

A Method for Calculating Probability of Scores for Men's Team Competition in Artistic Gymnastics

Nobuyoshi Hirotsu, Mutsumi Harada and Minoru Kano
Juntendo University, Japan

In this paper, we propose a method for calculating probability of scores for men's team competition in artistic gymnastics. We here assume each gymnast's score as a normal distribution and provide a mathematical formulation for 6-5-4 format in the team competition. By setting the different mean and standard deviation (SD) of the normal distribution, we calculate the distributions of the team score, and obtain the relationship between the mean and SD of each gymnast and the expected number of the team score. We demonstrate an application of this method in the selection of five gymnasts to compete in each event.

Paper ID: a16

A method for generating colorings over graph automorphism

Fei He and Hiroshi Nagamochi
Kyoto University, Japan

Given a graph $G=(V,E)$ with a set $W\subseteq V$ of vertices, enumerating colorings to W such that for every two enumerated colorings c and c' the corresponding colored graphs (G,c) and (G,c') are not isomorphic has an important application in the study of isomers of chemical graphs such as generation of benzen isomers from a tree-like chemical graph structure.

The number of such colorings can be computed efficiently based on Polya's theorem. However, enumerating each from the set of these colorings without using a large space is a challenging problem in general. In this paper, we propose a method for enumerating these colorings when the automorphism of G is determined by two axial symmetries, and show that our algorithm can be implemented to run in polynomial delay and polynomial space.

Paper ID: a17

An extended bailey–welch rule with no-show and walk-in for outpatient appointment scheduling

Zhihuang Mei, Ronghuan Huang and Jiping Tao
Department of Automation, Xiamen University, China

In clinic appointment scheduling system no-shows have been a significant and confirmed issue with a bad influence on patient accessibility and clinic efficiency. Over-booking is a common and important approach to deal with no-show. The problem of walk-in has often been seen as the opposite of no-show problem. In this work we revisit a walk-in admitting based approach to mitigate the bad influence of no-show. First we establish a model which utilizes marginal benefit objective function to balance the interests of the clinic, the patient and the doctor and prove that no-show and walk-in cancels out each other straightly has a bad property. Then we propose a new rule which is an extension of the well-known Bailey–Welch rule, the simulation results show that our rule has an improvement comparing with the common rule that cancels them out straightly.

Paper ID: a18

Election of Best Nine in Nippon Professional Baseball Based on the DEA Model

Tohru Ueda and Hirofumi Amatatsu
Seikei University, Japan

The Best Nine is a post-season All-Star honor given out in Nippon Professional Baseball. Every year one player per league is voted as the best at each position by baseball writers. However some players who were not elected may have gotten superior results than players who were elected. Thus we try election of another Best Nine based only on official records. We elect players who have high efficiency scores which are derived by methods based on Data Envelopment Analysis. At first candidates were limited to players with more than 220 plate appearances (PA). As a result there were many cases choosing players with low values of PA. Secondly we limited to players with more PA than 350 PA. Since importance of items presenting batting ability depends on the batting order, we showed also a new Best Nine considering the batting order.

Paper ID: a19

The Multiple Knapsack Problem with Compatible Bipartite Graphs

Jianping Li, Weidong Li and Hao Wang
Yunnan University, China

The multiple knapsack problem is to pack some items into given knapsacks, such that the sum of the knapsack profits is maximized. This paper is concerned with a variant of the multiple knapsack problem, called the multiple knapsack problem with compatible bipartite graph (MKPCBG), where two items can be packed into the same knapsack only if their corresponding vertices are adjacent in the given com- patible bipartite graph. Under two different objectives, we prove that the MKPCBG problem is strongly NP-hard, design some 1/2-approximation algorithms, and de-sign two optimal algorithms for the special case where all knapsacks have the same capacity.

Paper ID: a20

An inventory routing problem with soft time windows

Zhenping Li, Chongyu Jiang and Lulu Jiang
School of information, Beijing Wuzi University, China

Inventory routing problem (IRP) is a typical NP-hard Problem integrated with Inventory Control and Vehicle Routing Problem, and also is the key of Inventory and distribution coordinated optimization under the vendor managed inventory (VMI). We consider an IRP of petrol secondary delivery system in a large petroleum and petrochemical enterprise group. A homogeneous set of vehicles with capacity constraint is used for fuel distribution from one depot to a set of petrol stations that have deterministic fuel consumption. To minimize the total cost of the system, we provide a mixed integer programming model of the inventory routing problem with soft time window (IRPRTW). The model was tested on an example and the result is presented by Lingo program. Based on solutions with different weight coefficients we analyse the impact on the total cost by vehicle assignment cost and travel cost. A theoretical basis of solution to the problem of petrol secondary distribution problem of multiple terminal stations is provided.

Paper ID: a21

The Equal-Subsets Problem Is in the Class PPAD

Chuangyin Dang
City University of Hong Kong, Hong Kong

By introducing an integer labeling rule in an augmented set and applying a cubic triangulation of the Euclidean space, it is shown in this paper that the equal-subsets problem is in the class PPAD.

Paper ID: a24

Analysis of a Markovian Queue with Two Heterogenous Servers and a Threshold Assignment Policy

Dequan Yue, Hui Li, Guoxi Zhao and Wuyi Yue
Yanshan University, China

This paper considers a parallel queuing system with two heterogeneous servers where the task is dispatched to the two servers. The threshold assignment policy dispatches the tasks according to the number of customers in server 1. Firstly, we obtain the stationary condition of the system. Secondly, we give the stationary performance indices by using a matrix-geometric solution theory. Finally, we develop the average cost function and analyze the effect of the parameters on the average cost function by using numerical examples.

Paper ID: a25

Statistical data analyses of recent national elections in Japan

Akm Abul Kalam Azad, Md. Arifur Rahman and Tatsuo Oyama
National Graduate Institute for Policy Studies, Japan

This paper investigates the recent national elections in Japan quantitatively by applying various statistical methods and mathematical models. Firstly, we examine the regional characteristics that affect voter turnouts and the efficiency of the political parties of Japan. We find that the high income and most urbanized prefectures are categorized in the same cluster regarding voter turnouts. The voting efficiency according to prefectures and different parties also shows certain patterns. In addition, the relationship between the vote share and the seat share for all political parties was examined by applying mathematical models with both polynomial and exponential functions. Actual data obtained from Japan's recent national elections for both the House of Councilors and the House of Representatives from 2005 to 2014 have been used. We find that an exponential model can approximate the seat shares as a function of the vote shares obtained by the parties of the respective elections. We thus conclude that an exponential model can be used to predict the seat shares for the contesting parties in an election where many parties take part in the election process. Additionally, we investigate pass votes and fail votes to determine the efficiency of the political parties obtained in these elections.

Paper ID: a26

The Study of Evaluation of Regional Innovation Capability of the High-tech Industry

Ren Lin and Guohua Zhou
College of Mathematics and Computer Science, Hunan city University, China

This paper establishes the comprehensive evaluation index system which can reflect the high-tech

industry innovation capability. Factor analysis is carried out using statistical data of high-tech industry innovation ability, constructing the corresponding comprehensive evaluation model. Through 30 provinces (cities and autonomous regions), quantitative analysis of statistical data found that there exists anunbalanced development in the high-tech industry between Eastern and Western parts of China; the eastern parts show obvious advantages and a steadily increasing trend. Finally, according to the method of factor scores by cluster, these 30 provinces are divided into three categories, and then provide constructive suggestions of how to accelerate the developmental level of the high-tech industry.

Paper ID: a27

Fast community detection algorithm based on relationship strength coupling in social networks

Ye Li and Hui-Jia Li

School of Management Science and Engineering, Central University of Finance and Economics, China

Community structure detection is one of the most interesting issues in the study of social networks. However, there are seldom polynomial time algorithms which able to uncover community structure accurately. Inspired by ideas of famous Modularity optimization, in this paper, we proposed a novel k-strength relationship which naturally represents the coupling distance between two nodes. Community structure detection algorithm is presented using a generalized Modularity measure based on the k-strength matrix. To obtain the optimal number of communities, we then propose a new parameter-free framework using the eigenvalue gap of specific transition matrix. Finally, we apply our algorithm on both benchmark network and real networks. Theoretical analysis and experiments show that the algorithm is able to uncover communities fast and accurately, which can be easily extended to large scale real networks.

Paper ID: a28

Projection Method for Support Vector Machines with Indefinite Kernels

Hao Jiang, Wai-Ki Ching, Yushan Qiu and Xiaoqing Cheng
Renmin University of China, China

In this paper, we tackle with indefinite kernels by introducing projection matrix to formulate a positive semidefinite kernel. The projection matrix has a nice property of sharing the same set of eigenvectors with the original kernel. The proposed model can be regarded as a generalized version of spectrum method(denoising method and flipping method) by varying parameter λ. The problem of selecting optimal λ for optimizing the prediction performance is also considered. Using the Bregman matrix divergence theory, one can realize kernel learning by using unconstrained optimization. And our suggested λ in projection matrix helps to exhibit optimal performance for different values of λ.

Paper ID: a30

On Classification of Biological Data Using Outlier Detection

Yushan Qiu, Xiaoqing Cheng, Wenpin Hou and Wai-Ki Ching
The University of Hong Kong, Hong Kong

With the rapid development of information technology, the number of datasets, as well as their complexity and dimension,have been growing dramatically. This dramatic growth of biology data and non-biological commercial databases becomes a challenging issue in data mining. Classification technique is one of the major tools in the captured research area. However, the performance of classification may be degraded when there exists noise in the captured databases. Therefore, outlier detection becomes an urgent need and the issue of how to integrate outlier detection method and classification techniques is an important and challenging issue. In this paper, we proposed a novel and effective approach based on k-means clustering to identify outliers in the databases. In particular, we employed one of famous classification techniques, Support Vector Machine (SVM), owing to its ability to handle high-dimensional data set. We also compare the classification results with the multivariate outlier detection method. Numerical results on two different data sets indicate that the classification results after removing the outliers by our proposed method are much better than the multivariate outlier detection method.

Paper ID: a32

A Multiple Linear Regression Model for Structure of N-linked Oligosaccharides

Xiaoqing Cheng, Wai-Ki Ching, Wenpin Hou and Kiyoko Aoki-Kinoshita
The University of Hong Kong, Hong Kong

It is well-known that carbohydrate sugar chains, or glycans, play various well in cellular processes, including cancer, but the elucidation of glycans is difficult because of their complex structure. Both computational methods and mathematical models are necessary to integrate and analyze the information of glycomics data so as to efficiently detect glycan structures. In this paper, we propose a new model to predict the structure of N-glycans, which are the most common type of glycans. Our proposed prediction method is based on a Multiple Linear Regression (MLR) model. The coefficients of our proposed model are solved by using experimental data. We obtain our data from High Performance Liquid Chromatography (HPLC) experiments. Three sources of our data are adopted and they are divided into two parts: elution value on an Amide column and elution value on an OctaDecylSilane (ODS) column. After pre-processing the data, we then construct our proposed MLR model. The obtained correlation coefficients are 0.9680 for the Amide data and 0.9263 for the ODS data. We have also tested the correctness of the model statistically. The model test and correlation coefficients demonstrate both the accuracy and efficiency of our proposed model.

Paper ID: a34

MPS: a web server for Metabolism Pathway Synthesis on E.coli

Anjun Zhu

University of Science and Technology of China, China

An integrated system MPS has been developed which includes three modules: MRSD (Metabolic Route Search and Design), GD (Gene Design) and OD (Oligonucleotide Design). It can complete the process of from designing and searching a route to designing DNA oligonucleotides (oligos) for de novo gene synthesis. And each module can also be accessed and used individually. MRSD submodule searches and designs routes based on data from KEGG. GD submodule reversely translates multiple catalytic enzymes contained in chosen pathway into DNA sequence combining with operons for expression. OD submodule cuts resultant multiple DNA sequences into a series of oligonucleotide sequences characterized by highly homogeneous melting temperatures and a minimized tendency for secondary structure formation and mis-hybridization. The ideal software developed for biologists to de novo sequences design synthesizes new metabolic route for target compound and provides a good interactive interface. Availability: http://bioinfo.ustc.edu.cn/softwares/MPS/.

Paper ID: a35

A Signaling Pathway Analysis Method Based on Information Divergence

Hang Wei and Haoran Zheng

University of Science and Technology of China, China

Abnormal regulation of signaling pathways is the key factor causing disease. For better understanding disease mechanisms, many methods have been proposed to identify the significantly differential pathways between diseases and normal individuals via microarray gene expression datasets. Unlike previous common analysis processes, which is focused on merging gene difference into difference of pathway indirectly. In this paper, the idea of information divergence is introduced and a novel signaling pathway analysis method from a holistic view is presented to improve the detection results. We identify significantly differential pathways directly via computing the KL divergence between real and simulated probability distributions of gene-gene regulatory ability. We test our method on four human microarray expression datasets. The results illustrate that the capability of our approach in detecting significantly differential pathways between two sample groups is superior to other three classical pathway analysis methods.

Paper ID: a36

An improved algorithm for the machine scheduling problem with job delivery coordination

Yuzhong Zhang, Qiongyi Zheng, Jianfeng Ren and Long Zhang

Academy of Management, Qufu Normal University, China

This paper considers a two-stage supply chain scheduling problem in which the first stage is job production and the second stage is job delivery. The focus is on the study of the integration of production scheduling with delivery of finished products to customers. In our considered model each job can be processed on either of two identical machines, and then delivered by a vehicle to a customer location. We present an improved algorithm which the worst-case performance ratio is close to 14/9, which improves the known upper bound of 5/3.

Paper ID: a38
Employing post-dea method in budget management of health care sectors

Xiaoya Li

Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China

An increasing attention has been paid to efficiency analysis in the health care area. Among the existing efficiency assessment techniques, data envelopment analysis (DEA) plays an important role in a wide range of applications as for measuring the relative efficiency of different health care sectors. This paper focuses on a second stage of the analysis, which is operated after efficiency evaluation and called as post-DEA stage, and mainly cope with the budget management problems such as budget allocation and budget prediction. In the post-DEA stage, a comprehensive DEA based techniques are adopted to make the budget prediction and allocation based on the outcome of the first stage. A framework of budget management from macro aspect is built, together with corresponding resource allocation and prediction models proposed from micro aspect. The effect of the post-DEA method is illustrated by a numerical example with considering 10 hospitals, and with considering elasticity in post-DEA method, the outcome leads to an efficiency incentive effect in budget management.

Paper ID: a39
Mining online product reviews to identify consumers' fine-grained concerns

Jian Jin and Ping Ji
Beijing Normal University, China

Online product reviews contain valuable information about customer requirements (CRs). Intelligent analysis of a large volume of online CRs attracts interest from re-searchers in different fields. However, many research studies only concern sentiment polarity in different level and designers still need to read these reviews to absorb comprehensive CRs. In this research, online reviews are analyzed to obtain consumers' fine-grained concerns. Specifically, aspects of product features and detailed reasons of consumers are extracted from online reviews. This re-search starts from the identification of product features and the sentiment analysis with the help of pros and cons reviews. Next, the approach of conditional random fields is employed to detect aspects of product features and detailed reasons jointly. In addition, a co-clustering algorithm is devised to group

similar aspects and reasons to provide concise descriptions about CRs. Finally, with hundreds of customer reviews of six mobiles in Amazon.com, a case study is presented to illustrate how the proposed approaches benefit product designers in the elicitation of CRs by the analysis of online opinion data.

Paper ID: a41

Meteorological station network and its application to forecast infestation of tropilaelaps mites

Jin-Shan Li

College of Science, Beijing Forestry University, China

Human activities, particularly modern transportations, are becoming main contribution to quick spreading of pests or some epidemic disease when meteorological condition is suitable to their living. Tropilaelaps mites are ectopara-sites of Apis honey bees. With the aid of bee-keeper mov-ing bee colonies, Tropilaelaps mites spread quickly in most area of China when meteorology is appropriate for them to live. For aim of finding dynamic pattern of the outbreak of these animals, geographical coordinates of meteorological stations are applied to build up a meteoro-logical station network connecting all the meteorological stations with edges, and their edge distances could be able to guarantee two connected stations as similar meteorolo-gy as possible, and rightly reflect the movement of bee-keepers whose aims are for commercial benefit. To reach these aims, a multi-objective programming is used to look for an edge distance threshold attempting to connect as many stations (nodes) as possible into network and to use as less edges as possible to connect them. After building the meteorological station network, inner product between means of selected meteorological variables of specimen places of Tropilaelaps mites and that of every station in a month is computed to distinguish whether the area around this station has high probability of infestation of Tropi-laelaps mites in this month, and these stations with high probability connect together if two of them are neighbours in the built meteorological station network. And thus monthly dynamics of infestation of Tropilaelaps mites can be viewed on the built meteorological station network, at the same time, some knowledge about infestation of Tropilaelaps mites also can be known.

Paper ID: a42

The pricing and copyright's payment strategy in online movie distribution

Ming-Xi Zhu, Ji-Hong Zhang and Xiao-Shuang Han

International Business School, Beijing Foreign Studies University, China

This paper discusses the online movie distribution model. Under this model, the movie would be released through both theatrical channels and online channels. The producer has to bargain with the online distributor in order to determine the pricing and the payment method of copyright. Varied cases are considered, including the formation of the market players and the substitution rate between

theatrical version and online version. This paper further gives out the prevailing payment method under different scenarios and analyses how the price and the demand would change in different cases.

Paper ID: a43

Sufficient Optimal Conditions for Unconstrained Quadratic Binary Problems

Liu Liu, Chunli Liu and Qiuling Xie

School of Mathematics, Shanghai University of Finance and Economics, China;

In this article, we present several sufficient optimal conditions for unconstrained quadratic binary problems, which can be applied in algorithms combining with SDP relaxations in branch-and-bound approaches for the primal problem. These optimal conditions can work for many situations when the Lagrangian duality gap is zero.

Paper ID: a44

Solve the vehicle routing problem via a discrete particle swarm optimization

Hongwei Liu

School of Informations, Beijing Wuzi University, China

This paper introduces a new hybrid algorithmic approach based on discrete Particle Swarm Optimization (PSO) and optimal splitting procedure for solving one of the most popular supply chain management problems, the Vehicle Routing Problem (VRP). The VRP is a well known NP-hard problem in which a vehicle with finite capacity leaves from the depot with full load and has to serve a set of customers whose demands are known only when the vehicle arrives to them. Experiment results show that the proposed algorithm is an efficient algorithm in solving VRP. This paper also shows that the proposed algorithm algorithm can find the feasible solution effectively and almost can find the global optimal solution for small instance. For larger instances, if the size of populations in the proposed algorithm increases, the possibility of finding the global optimal solution will increase.

Paper ID: a45

Optimization of Artificial Neural Network for Stock Market Return Prediction

Mingyue Qiu and Yu Song

Fukuoka Institute of Technology, Japan

The application of artificial neural network for stock market return prediction has been examined for many years. Back propagation algorithm is a classical learning algorithm for the neural network, and had been widely applied by many prior studies. In view of the limitations of gradient search of

back propagation technique, global search techniques are proposed for training neural networks. In this study, we proposed hybrid genetic algorithm and simulated annealing for weight and bias optimization of artificial neural network to overcome the shortcoming of back propagation algorithm. Due to the complexity of stock market data, we selected the effective input variables from the data of 71 variables that covered financial and economic information of Japanese stock market by the fuzzy surfaces. To verify the prediction ability of the selected input variables, we use the monthly data of Japanese stock market index to empirically compare the hybrid approach based on genetic algorithms and simulated annealing with back propagation for training neural networks.

Paper ID: a46

Medical treatment capability analysis based on m/m/s/gd/c/∞ model

Juyun Wang, Jie Sui, Mei Jia Li and Hua Yu
Communication University of China, China

When biochemical terrorist attacks happen, a large number of victims will flock to the nearest hospital in a short period of time. It is necessary to analyze these situations of both victims and the hospital real time so that the victims can get medical treatment as soon as possible. We proposed a method to analyze medical treatment capability based on Queuing Theory, validated the model and algorithm using the data of Japanese Tokyo Sarin Event in 1995, calculated the arrival status of victims and the medical treatment capability. The method can provide the emergency managers with decision support.

Paper ID: a47

Melanocytic globules detection in skin lesion images

Leszek Nowak, Kasia Grzesiak-Kopec and Maciej Ogorzalek
Jagiellonian University, Poland

In this paper a method is presented for detection of mela-nin globules often present in melanocytic skin lesions im-ages. The detection is done by performing image analysis similar to the one used in clinical evaluation. The method uses multi-stage image filtering to extract objects present in the dermoscopic image that match globule structure pattern. Classification of the found objects is made based on shape and size of globule structure. The classification is problematic task due to color and scale differences be-tween dermatologic images and is related to differences between image acquisition equipment used in der-matoscopy. First we describe characteristic of globule structure needed for correct classification, along with method for calcu-lating those characteristic. Next we presented a method for globules detection that is a part of computer-aided diagnostic process of melanocytic skin lesions. Evaluation of such lesions is a basis for early detection of malignant lesions.

Enrichment Set Cover Problem

Yinliang Liu, Xiang-Sun Zhang, Ling-Yun Wu

Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China

Enrichment analysis, as a common task in the fields of bioinformatics, is to investigate the functional association between a gene list of interest which often derived from biological experiments, and specific gene sets in a large database. The core problem of enrichment analysis can be characterized as a two-objective optimization problem. In this paper, we formulated the multiple gene sets enrichment problem into a variant of the set cover problem, named enrichment set cover problem, and designed four approximation algorithms to solve the NP-hard enrichment set cover problem. The performance of proposed approximation algorithms were analyzed theoretically and evaluated on the data sets from real biological databases.

ISB Sessions

Paper ID: b3

Online learning approach to prediction of protein-protein interaction strengths

Morihiro Hayashida, Mayumi Kamada and Hitoshi Koyano

Bioinformatics Center, Institute for Chemical Research, Kyoto University, Japan

Protein-protein interactions take various important roles in living cells. Many researchers have analyzed interactions of proteins, and have developed its prediction methods. As well as protein-protein interactions, interaction strengths provide useful knowledge to understand complicated cellular networks, and several prediction methods have been developed. In our previous study, we proposed new feature space mappings based on protein domains, and employed support vector regression and relevance vector machine. The combination of the mapping and the supervised regression method outperformed the existing prediction methods based on domains.

Paper ID: b4

Systematic identification of transcriptional and post-transcriptional regulations in human respiratory epithelial cells during influenza A virus infection

Zhi-Ping Liu

Shandong University, China

Revealing genome-wide transcriptional and post-transcriptional regulatory relationships can further advance our understanding of the molecular mechanisms of
airway epithelial cell responses to viral infection, which motivates the development of novel and more efficient computational methods to simultaneously infer the transcriptional and post-transcriptional regulatory networks. In this work, we propose a novel framework to investigate the interactions among transcription factors (TFs), microRNAs (miRNAs) and target genes. Our simulation studies on networks of different sizes suggest that the proposed framework can effectively determine the genuine regulations among TFs, miRNAs and target genes; also, we compare our framework with several selected state-of-the-art algorithms to further evaluate its performance. By applying the proposed framework to mRNA and miRNA expression data generated from human lung epithelial A549 cells in response to A/Mexico/InDRE4487/2009 (H1N1) virus infection, we are able to detect the activated transcriptional and post-transcriptional regulatory relationships as well as the significant regulatory motifs. The results obtained for human respiratory epithelial cells suggest the importance of the transcriptional, post-transcriptional regulations as well as their synergies in the innate immune responses against influenza A virus infection.

Paper ID: b5

Identifying early-warning signals of critical transitions by distribution embedding

Rui Liu, Pei Chen, Kazuyuki Aihara and Luonan Chen
South China University of Technology, China

Identifying early-warning signals of a critical transition for a complex system is a difficult task, especially when this system is constantly perturbed by big noise, which makes the traditional methods fail due to the strong fluctuations on the observed data. In this work, we show that the critical transition is not traditional state-transition but probability distribution-transition when the noise is not sufficiently small, which, however, is a general case in real systems. We present a new model-free computational method based on the observed time-series data to detect the warning signals just before such a critical transition. The key idea behind this method is a new strategy: ``making big noise smaller" by distribution embedding scheme, which transforms the data from the observed state-variables with big noise to their distribution-variables with small noise and thus makes the traditional criteria effective because of the significantly reduced fluctuations. Specifically, increasing the dimension of the observed data by moment expansion that changes the system from state-dynamics to probability distribution-dynamics, we can derive new data in a higher-dimensional space but with much smaller noise. Then, we develop a criterion based on the dynamical network marker (DNM) to detect the early-warning signal of a critical transition in this transformed higher-dimensional data. From both theoretical and computational aspects, we demonstrate that our work can effectively detect early-warning signals for many real-world problems by using observed data, e.g., a pre-transition state of acute lung injury, a tipping point of an eutrophic lake state, and a catastrophic phenomenon of financial markets.

Paper ID: b6

Singularity Analysis of Grb10 signaling pathway quantitatively reveals Grb10 promotes lipolysis and thermogenesis

Hao Kang, Guanyu Wang and Luonan Chen
School of Life Science and Technology, ShanghaiTech University, China

Identification of connections between key regulators of lipid metabolism and thermogenesis has important therapeutic implications for obesity and other metabolic diseases. Here we develop a mathematical model of lipolysis and thermogenesis response network, centering on the key regulator Grb10 which was discovered by Feng Liu [Cell Metab. 2014 Jun 3; 19(6):967-80] and plays important roles in mTOR and AKT signaling. The model is in the form of coupled differential equations. By applying singularity and bifurcation analysis, we show that the phosphorylation of Grb10 exhibits a periodic bang-bang response in response to external stimulus, while mTOR activity oscillates with a phase lag. The periodic oscillations are modulated by cold signals, demonstrating cells' different lipolysis and thermogenesis strategies in response to different cold exposures. Our analysis reveals system-level mechanisms in regulating biological functions, which may provide valuable insights into therapeutic interventions.

Paper ID: b7

A Tool: Identification of Feature Genes from Genomics Data Using Factor Analysis

Changhe Fu, Su Deng, Jinghua Wu, Xianqiong Wu and Zhi-Hui Fu
Shenyang Normal University, China

In this article, a software tool (IFGFA) for identification of feature genes from genomic data based on factor analysis is present. Most computation methods and statistic models have been successfully applied in this study. They are, however, only appropriate for some special gene expression profiling data. We have designed a computation framework to predict colon cancer related genes, which also works effectively for other diseases. The framework is based on a well-established latent factor model, Bayesian factor and regression model (BFRM) and prior gene knowledge of some cancer in Online Mendelian Inheritance in Man (OMIM). Furthermore, the framework validated the identified genes by checking their somatic mutations of the same samples from DNA sequencing data. Here, we realize the framework by visual programming method facilitating users' operation.The software tool is freely available at http://www.fupage.org/downloads/ifgfa.zip.

Paper ID: b8

PDCD5 interacts with p53 and functions as a regulator of p53 dynamics in DNA damage response

Changjing Zhuge, Xiaojuan Sun, Yingyu Chen and Jinzhi Lei
Bejing Forestry University, China

Background: The tumor suppressor p53 plays a central role in cell fate decisions after DNA damage. Programmed Cell Death 5 (PDCD5) is known to interact with the p53 pathway to promote cell apoptosis. Recombinant human PDCD5 can significantly sensitize different cancers to chemotherapies.

Method: We construct a computational model that includes PDCD5 interactions in the p53 signaling network and study the effects of PDCD5 on p53-mediated cell fate decisions during the DNA damage response.

Results: Our results revealed that PDCD5 functions as a co-activator of p53 that regulates p53-dependent cell fate decisions via the mediation of p53 dynamics. The effects of PDCD5 are dose-dependent such that p53 can display either sustained or pulsed dynamics at different PDCD5 levels. Moreover, PDCD5 regulates caspase-3 activation via two mechanisms during the two phases of sustained and pulsed p53 dynamics.

Conclusion: This study provides insights regarding how PDCD5 functions as a regulator of the p53 pathway and might be helpful for increasing our understanding of the molecular mechanisms by which PDCD5 can be used to treat cancers.

Paper ID: b9

Hopf bifurcation in a prey-predator model with two time delays

Yiming Ding and Yan Meng

University of Science and Technology Beijing, China

In this paper, we study a modified prey-predator model with two unequal time delays.We show that the positive equilibrium is asymptotically stable in the absence of time delay,but loses its stability via the Hopf bifurcation when one of the two time delay increases beyond a threshold.Furthermore,we study the relationship between this threshold and the other time delay.Numerical calculations are preformed to illustrate our theoretical results.

Paper ID: b10

A reflected stochastic differential equation model for biochemical reaction systems

Yuanling Niu

Key Laboratory of Systems Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Science, Chinese Academy of Sciences, China

In this paper, we gave a framework of modelling and simulation of the biochemical reaction systems by stochastic differential equations with reflection not in a heuristic way but in a mathematical way. The model is computationally efficient compared with the discrete-state Markov chain approach, and ensures that its analytic and numerical solutions remain in a biologically realistic region. The domain D where the species numbers should lie in was given and it is much closer to the physical reality than that was given before. Our domain D is obtained according to the structure of the corresponding chemical Langevin equations, that is to say the boundary is inherent in the biochemical reaction system. A variant of projection method was employed to solve the reflected stochastic differential equations model. Euler-Maruyama method was applied to the equations first and the only further calculation is to check whether the point lies within the domain D and if not perform an orthogonal projection. It is found that the projection onto the closure $\bar{D}$ is the solution to a convex quadratic programming problem. Then existing methods for the convex quadratic programming problem can be employed for the orthogonal projection map. Numerical testing on several important problems in systems biology confirms the efficiency and accuracy of this approach.

Paper ID: b12

Detecting disease genes of non-small lung cancer based on consistently differential interactions

Xiaoping Liu, Qianqian Shi and Luonan Chen

Key Laboratory of Systems Biology, SIBS-Novo Nordisk Translational Research Centre for PreDiabetes, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, China

Systematic identification of causal disease genes can shed light on the mechanisms underlying complex diseases and provide crucial information to develop efficient biomarkers or design suitable therapies. The present paper describes a novel approach to detect potential disease genes for lung cancer, based on consistently differential interaction (CDI) scheme from heterogeneous disease datasets. In particular, reliable disordered regulations in disease states were discovered by identifying the CDIs, from which the disease genes were further detected based on their topological structures on the network. As an application of the CDI-based method, the RNA-seq data of two subtypes of non-small lung cancer were used to identify CDIs from normal to cancer onset. The results of analysis well agree with the prior knowledge as well as the experiments, thereby implying the predictive power of the CDI-based method. The comparison with other approaches also indicated the superiority of the CDIbased method in terms of accuracy and effectiveness on detecting disease-specific genes for lung cancer and metastasis. In contrast to conventional molecular biomarkers, the identified CDIs as novel network biomarkers or edge biomarkers can be applied to predict patient survival for both subtypes of lung cancers, and the interactions among CDIs can be further used as new edgetic targets for network drug design. In addition, a potential molecular mechanism was developed to explain the key roles of the identified CDIs in lung cancer and metastasis from a network perspective.

Paper ID: b13

Biomarker discovery in biomedical big data

Tao Zeng and Luonan Chen

Key Laboratory of Systems Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, China

The sources and structures of biomedical big data have been widely accessible in public for biomarker and disease study. The biomedical big data is actually "Large p, Small n", i.e., small samples but with high dimensions on features (e.g., omics data) for each individual, in contrast to traditional big data as "Large n, Small p" in other fields, i.e., big samples but with low dimensions on features.

Big-data based edge biomarker is a new concept to characterize disease features based on biomedical big data in a dynamical and network manner, which also provides alternative strategies to indicate disease status in single samples. In particular, in contrast to using the information of the common molecules or edges (i.e. molecule-pairs) across a population in traditional biomarkers including network and edge biomarkers, big-data based edge biomarkers are specific for each individual and thus can accurately evaluate the disease state by considering the individual heterogeneity. Therefore, the measurement of big data, i.e., high dimensional data is required not only in the learning process but also in the diagnosing or predicting process of the tested individual. As a model of big-data based edge biomarker, module network rewiring-analysis (MNR) has been proposed to use the module network rewiring model to characterize functional reorganization of a complex biological system, especially applied to systematically study dynamical drug sensitivity and resistance during drug treatment. The dry-experiments of MNR have been carried on two

scenarios: (i) the expression data of patients with Hepatitis C virus infection receiving Interferon therapy; (ii) and the temporal expression data from a malaria vaccine trial. All results demonstrated that consistent module genes derived by MNR could be directly used to reveal new genotypes relevant to drug sensitivity, unlike the other differential analyses of gene expressions; the functional connections and reconnections among consistent modules bridged by biological paths were necessary for achieving effective responses of a drug; and the hierarchical structures of the temporal module network can be considered as spatio-temporal biomarkers to monitor the efficacy, efficiency, toxicity, and resistance of the therapy.

Besides, to quantitatively illustrate the efficiency of different categories of biomarkers in the context of dynamic drug sensitivity and resistance, a number of state-of-the-art methods have also been applied to the case studies. The integration of high-confidence prior knowledge is expected to provide additional information on the discovery of edge biomarkers or big-data biomarkers. And the future systematical evaluation of biomarkers in different application scenarios including the study of drug sensitivity and resistance, will inspire more powerful models and technologies in biomarker discovery and application.

Paper ID: b14

Identifying progressive biomarkers utilizing time-course information

Jingxue Xin, Tao Zeng, Yong Wang and Luonan Chen
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China

Biomarker study has become an important task in biological and medical research with the rapid development of high-throughput sequencing technologies today. However, accurately predicting and diagnosing complex diseases is still challenging since people have limited knowledge about the molecular mechanism of these complex diseases. We noticed that large amount of time course or disease stage specific data in expression and medical records are available. In addition, conventional methods for biomarker discovery are limited in interpretability by using the same differentially expressed gene set as biomarkers to distinguish different disease classes. In this study, we proposed to utilize the valuable information from time dependence among samples and consider the dynamics of biomarker set along with time or stage. Specifically, we developed a novel linear programming method to simultaneously identify disease stage specific and stage common biomarkers from a multiple-class classification problem. We identify various clusters with time-course property and the progress of disease development in biomarker study is considered. As a result, our model identifies different biomarkers for disease stage, and then utilizes these stage-specific biomarkers to make prediction. Our optimization model ensures to identify a small number of biomarkers with low redundancy and high accuracy. Computational experiment demonstrates the effectiveness and efficiency of our method, which can improve the classification accuracy compared with the existing biomarker identification method. Furthermore, we could use these common and specific biomarkers among different stages to analyze the development of complex disease.

Paper ID: b15

Identifying specific and consistent dysfunctional modules from multiple cancer sites

Chuan-Chao Zhang, Qianqian Shi, Tao Zeng, Juan Liu and Luonan Chen
Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, China

Motivation: There are many methods proposed to identify disease genes or module biomarkers of complex diseases based on the gene expression data and the related molecular network. However, for complex diseases with multiple subtypes (e.g. cancer), it is still unclear about the specific and consistent gene modules corresponding to particular disease subtypes, and their associations.

Method: In this paper, we present a novel network-based approach to identify the specific and consistent network modules from multiple disease subtypes simultaneously, by integrating gene expression data and protein-protein interaction network together. We first make full use of the specific and conservative information (e.g. the differential gene expression/integrations between the disease and normal samples) among multiple disease subtypes, and then divide the integrated protein-protein interaction network into various modules based on an enhanced Markov Clustering algorithm. Next, we applied a multi-class classification model to identify the module combinations, which reveals the specific and consistent disease modules, and their compositions.

Results: As a proof-of-concept study to real data, we applied our approach to six types of cancer, including breast cancer, colon cancer, liver cancer, lung cancer, oesophagus cancer and thyroid cancer; and identified several specific and consistent network modules for characterizing different cancer sites. Functional enrichment analysis and independent validations on three additional gene expression datasets (i.e. breast cancer, lung cancer and esophageal squamous cancer) both verified site-specific and site-common roles of dysfunctional modules in cancer.

Paper ID: b16

CHiP-seq data analysis of DAF-16 transcription factor binding site

Lu Wang and Jiarui Wu
Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, China

Background: The activity of DAF-16 transcription factor is closely related to cell apoptosis and animal lifespan. The higher the activity of DAF-16, the longer the worm lives.

In MAPK pathway, JKK-1(JUN kinase kinase) regulate the downstream protein JK-1 to activate transcription factor DAF-16. When miR-83(miRNA) binds to specific sites, those sites are resisted to DAF-16(transcription factor) binding. The sequence of those sites can be deal with gene knock-out method by collaborating wet lab.

Because the mainly transcription factor binding site databases (e.g. TRANSFAC、TRRD) has been outdated, we want to find the specific binding site for miR-83 in the ChIP-seq database (Encode).

Methods: Through LiftOver tool in UCSC, we got a reference genome response to the ChIP-seq data. The raw data was first mapped to the respective reference genome (Worm S220 reference),

then the resulting coordinate site was translated into the reference site. This mapping process is called peak-calling. Based on the global similarity of transcription factor binding sites, we set the main standard using PCC (Pearson Correlation Coefficient) to eliminate the noise in the result of peak-calling.

Results: Using Bowtie to map the sequence of ChIP-seq data to the genome of the Worm S220 reference, we got the result of peak-calling. According to experience in recently study, we estimated the effective genome size in mapping is going to be 5.6*107 bp in the genome of worm (8*107 bp). To find the more specific binding sites, we set the higher MFOLD value to be 5.

Conclusion: By checking the bed files resulted from peak-calling, we found the miR-83 related sequence site (7845152-7847135). With this specific binding site, the knock-out experiment would be performed accordingly. We are waiting for the experiment result of the other side to carry out the next step of cooperation.

Paper ID: b17

Quantifying direct associations in networks by sample-based data

Juan Zhao and Luonan Chen

Shanghai Institutes for Biological Sciences Chinese Academy of Sciences, China

Quantitatively identifying direct dependencies between variables is an important task in data analysis, in particular, for reconstructing various types of networks and causal relations in science and engineering. One of the most widely used criteria is partial correlation, however, it can only measure linearly direct dependencies and miss nonlinear associations. In this work, we propose a new concept and measure, 'Part Mutual Information (PMI)', to overcome the problem. Specifically, we first defined PMI to measure nonlinearly direct dependencies between variables, and then derived its relations with both mutual information (MI) and conditional mutual information (CMI). Finally, we used a number of simulated data (benchmark examples) to numerically demonstrate PMI features, and further real gene expression data from E.coli and yeast to reconstruct gene regulatory networks, which all validated advantages and effectiveness of PMI for accurately quantifying nonlinearly direct dependencies.

Paper ID: b18

Critical stage with early-warning signals during chronic inflammation to hepatocellular carcinoma

Meiyi Li, Chen Li, Rong Zeng and Luonan Chen

Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, China

Recently increasing studies have supported that chronic inflammation contributes to development and progression of most of cancers, including hepatocellular carcinoma (HCC). To identify the critical period and discover underlying mechanisms from chronic inflammation to HCC, here we introduced the c-myc tumor-prone transgenic mouse model infected by woodchuck hepatitis virus

(WHV) and applied our mathematical model based on dynamical network biomarker (DNB). Analyzing the well-designed time-series proteomic data of WHV/c-myc mice and age-matched wt-C57BL/6 mice, we identified the 5-month as a critical transition period of cancer initiation, which is consistent with the clinical symptoms during the progression of the c-myc-induced HCC. Meanwhile, according to functional analysis of inversing DNB-centered subnet at the previous and the latter periods, we discovered that dysfunction of PLA2G6 and CYP2C44-associated arachidonic acid metabolism actives the upstream of inflammatory response through the inflammatory mediators of transient receptor potential (TRP) channels, which may further lead to impairments of liver detoxification. Especially, inflammation-related dysfunction appears at the previous stage but angiogenesis-associated after the critical period. These potential mechanisms can help us understand the pathogeneses of c-myc-induced hepatocarcinogenesis even the human-suffered HCC and provide some new insight into intervention strategies for preventing from malignancy of chronic hepatitis.

Paper ID: b20

Intra-residue atom contacts outperform other features in determining protein-protein interface residues

Yongxiao Yang and Xinqi Gong
Renmin University of China, China

Protein-protein interactions are of great significance in many biological processes. Since it is very difficult and expensive to determine protein complex structures in experiment, the computational prediction has become more and more important. The information of interface residues plays a key role in the accurate prediction of protein complex. Here, we characterized the surface residues of protein monomer by nine simple descriptors and employed these features to discriminate the interface residues from the non-interface residues. We found that the interior contact area of the surface residue is the best one when only a single feature is adopted for identifying the interface residues. For more than 70% of the monomers corresponding to the dimmers in Docking Benchmark 4.0, there exists at least one interface residue in the three surface residues with the biggest interior contact areas. In order to investigate the effects of all kinds of combinations of the nine features, we used backward propagated neural network to train models in the set which contains all the surface residues of the monomers corresponding to the dimmers in Docking Benchmark 3.0, then tested these models in the set of surface residues of the monomers corresponding to the dimmers updated in Docking Benchmark 4.0. For the trained models, the percentage mentioned above in the test set can be more than 75%. Additionally, the number of features is more, the result is not better, which suggests that the discovery of the key impact factors of protein-protein interactions is very important for interface residue prediction. These results should be beneficial for the understanding of the mechanism of protein-protein interactions and the solving of protein-protein interface residue prediction problem.

Paper ID: b21

Analysis on acute metabolic disease by dynamical network biomarker

Zhonglin Jiang, Xiaoping Liu and Luonan Chen

Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, China

Metabolic diseases, such as non-alcoholic fatty liver disease (NAFLD) and obesity, have become the most common and emerging form of chronic human disease. However, the regulatory processes of these diseases are still not clear due to complex interplay of many internal and external factors. With the rapid advances of high-throughput technologies, such as microarray and next-generation sequencing, big data generated from the technologies provide unprecedented opportunities to study the molecular mechanisms of the metabolic diseases from a system-wide perspective. In this work, we combined dynamical network biomarker (DNB) and biological network to detect the early signal of these complex diseases and analyze the potential disease associated pathway based on a time series RNA-Seq data from the model mouses with NAFLD. Finally, we have identified not only a DNB in the critical stage but also the potential regulation relationship for the DNB on pathway. It not only indicates the emergence of the disease state but also provides the insights into the driving factors of the disease.

Paper ID: b22

Diagnosing Different Stages of Alzheimer's Disease from resting-state Functional MRI Data

Yiwei Zhou

Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, China

Alzheimer's disease (AD), also known in medical literature as Alzheimer disease, is the most common form of dementia. There is no cure for the disease, which worsens as it progresses, and eventually leads to death. Most often, AD is diagnosed in people over 65 years of age, although the less-prevalent early-onset Alzheimer's can occur much earlier. To effectively diagnose AD, the brain image technologies become popular in the brain research due to much less damage to human bodies and also repeatable. In addition, the dynamic information of brain can be achieved by brain image technologies. In this work, we define 6 stages of all samples from normal people to severe Alzheimer's disease patients, which are NC (Normal Controlled), SCI (Subjective Cognitive Impairment), MCI (Mild Cognition Impairment), MCIP (post Mild Cognition Impairment), ADM (mild Alzheimer's Disease), AD (severe Alzheimer's Disease). MCIP is defined as those patients who are in MCI but will definitely change to AD. The patients in SCI and MCI can be treated and become normal, but the patients in MCIP, ADM and AD are difficult to be treated by today's medical technique. We first adopt Lasso Elastic Net Method to identify biomarkers for classifying samples, which achieve over 80% accuracy. Then, we use the Dynamic Network Biomarker (DNB) theory to detect combinations of brain regions as biomarkers, which indicate the early warning signals of the disease. The results show that the identified DNBs have the power to classify the NC samples and SCI samples, or the NC samples and MCI samples.

Paper ID: b101

GPU-parallelized constrained feature selection algorithm improves the breast cancer subtyping model

Ren Zhong, Guoqin Mai and Fengfeng Zhou

Shenzhen Institutes of Advanced Technology, and Key Lab for Health Informatics, Chinese Academy of Sciences, China

Breast cancer is the top one factor for the caner induced death for women across the world. One of the reasons is that breast cancer has four major subtypes, i.e. luminal A, luminal B, Her2 and Basallike, and the four subtypes have varied relapse rates and treatment options. PAM50 is a 50-gene panel designed to classify these four subtypes of breast cancers versus the normal samples, and is supposed to reflect the recurring patterns. But this panel does not include two breast cancer biomarkers, i.e. BRCA1 and BRCA2. This study hypothesizes that a biomarker set consisting of BRCA1/BRCA2 may perform better than PAM50. The experimental data shows that a 33-feature panel performs better in classification accuracy than the PAM50 panel, while using much fewer features than the 112 features mapped to the 50 genes in PAM50. By expanding the feature screening space, the biomarker detection algorithm requires much more computing power and is parallelized by the GPU CUDA architecture.

Paper ID: b102

lncRNA-MFDL: identification of human long non-coding RNAs by fusing multiple features and using deep learning

Xiao-Nan Fan and Shao-Wu Zhang

Key Laboratory of Information Fusion Technology of Ministry of Education, School of Automation, Northwestern Polytechnical University, China

Long noncoding RNAs (lncRNAs) are emerging as a novel class of noncoding RNAs and potent gene regulators, which play an important and varied role in cellular functions. lncRNAs are closely related with the occurrence and development of some diseases. High-throughput RNA-sequencing techniques combined with de novo assembly have identified a large number of novel transcripts. The discovery of large and 'hidden' transcriptomes urgently requires the development of effective computational methods that can rapidly distinguish between coding and long noncoding RNAs. In this study, we developed a powerful predictor (named as lncRNA-MFDL) to identify lncRNAs by fusing multiple features of the open reading frame, k-mer, the secondary structure and the most-like coding domain sequence and using deep learning classification algorithms. Using the same human training dataset and a 10-fold cross validation test, lncRNA-MFDL can achieve 97.1% prediction accuracy which is 5.7, 3.7, and 3.4% higher than that of CPC, CNCI and lncRNA-FMFSVM

predictors, respectively. Compared with CPC and CNCI predictors in other species (e.g., anole lizard, zebrafish, chicken, gorilla, macaque, mouse, lamprey, orangutan, xenopus and C. elegans) testing datasets, the new lncRNA-MFDL predictor is also much more effective and robust. These results show that lncRNA-MFDL is a powerful tool for identifying lncRNAs. The lncRNA-MFDL software package is freely available at http://compgenomics.utsa.edu/lncRNA_MDFL/ for academic users.

Paper ID: b103

Integrated omics analysis in tumor genesis of Hepatocellular Carcinoma (HCC)

Lei Liu

Fudan University, China

Tumor cells with stemness (stem-cell) features contribute to initiation and progression of hepatocellular carcinoma (HCC), but the involvement of long noncoding RNAs (lncRNAs) remains largely unclear. Genome-wide analyses were applied to identify tumor-associated lncRNA-DANCR. The DANCR expression level and the prognostic values of DANCR were assayed in two HCC cohorts (China and Korea, n = 135 and 223). Artificial modulation of DANCR (down- and over-expression) was done to explore the role of DANCR in tumorigenesis and colonization and tumor-bearing mice were used to determine the therapeutic effects. We found that lncRNA-DANCR is over-expressed in stem-like HCC cells and this can serve as a prognostic biomarker for HCC patients. Experiments showed that DANCR markedly increased stemness features of HCC cells to promote tumorigenesis and intra-/extra-hepatic tumor colonization. Conversely, DANCR knock-down attenuated the stem-cell properties and in vivo interference with DANCR action led to decreased tumor cell vitality, tumor shrinkage and improved mouse survival. Additionally, we found that the role of DANCR relied largely on association with and regulation of CTNNB1. Association of DANCR with CTNNB1 blocked the repressing effect of miR-214, miR-320a and miR-199a on CTNNB1. This observation was confirmed in vivo, suggesting a novel mechanism of tumorigenesis involving lncRNAs, mRNAs and microRNAs.

Paper ID: b104

Stem Cell Regeneration as a multi-scale dynamical system: From modeling to simulation

Jinzhi Lei

Tsinghua University, China

This talk presents resent progress about the modeling of stem cell regeneration through the multi-scale model with cross talk between genetic and epigenetic regulation. Mathematical models include the long-term dynamics from inflammation to cancer, controls of stem cell regeneration, p53 dynamics in DNA damage responses, and multiple cellular simulations. We also introduce the method of GPU simulation, which is a powerful technique in the study of long-term regeneration of a group of heterogeneous stem cells.

Paper ID: b105

CCLasso: Correlation Inference for Compositional Data through Lasso

Huaying Fang, Chengcheng Huang, Hongyu Zhao, and Minghua Deng
LMAM, School of Mathematical Sciences, Peking University, China.

Direct analysis of microbial communities in the environment and human body has become more convenient and reliable owing to the advancements of high-throughput sequencing techniques for 16S rRNA gene profiling. Inferring the correlation relationship among members of microbial communities is of fundamental importance for genomic survey study. Traditional Pearson correlation analysis treating the observed data as absolute abundances of the microbes may lead to spurious results because the data only represent relative abundances. Special care and appropriate methods are required prior to correlation analysis for these compositional data. In this article, we first discuss the correlation definition of latent variables for compositional data. We then propose a novel method called CCLasso based on least squares with $\ell 1$ penalty to infer the correlation network for latent variables of compositional data from metagenomic data. An effective alternating direction algorithm from augmented Lagrangian method is used to solve the optimization problem. The simulation results show that CCLasso outperforms existing methods, e.g. SparCC, in edge recovery for compositional data. It also compares well with SparCC in estimating correlation network of microbe species from the Human Microbiome Project.
Availability: Under GNU LGPL v3, CCLasso is open source and freely available from https://github.com/huayingfang/CCLasso.

Paper ID: b106

Combining genetic and environmental variations to infer genes for phenotypes in crops

Jingdong Liu
Monsanto Company, USA

Understanding how gene-level changes generate phenotypic variations has been a central problem in biology that has great impact for crop improvement. In recent years, predictive modeling has been increasingly applied to address diverse biological problems. In this talk, applications of predictive modeling will be illustrated with examples in predicting plant phenotypes using a diverse range of gene-level information representing system variations introduced by both genetic and environmental factors, with the aim of identifying important genes for crop improvement.

Paper ID: b107

ARCS: Assemble short-reads by using combinatorial optimization in scaffolding

Dongbo Bu

BioInformatics Lab., Institute of Computing Technology, Chinese Academy of Sciences, China

In genome sequence assembly, the de Bruijn graph-based approaches gained great successes, yet at least two challenges remain as the obstacles: 1) repetitive regions, especially inexact repeats, entangle the de Bruijn graph, making it difficult to assemble; 2) chimeric paired-end reads, although not large in quantity, form erroneous links that can mis-lead scaffolding process. Existing assembly programs generally adopt simple heuristics or greedy strategies to solve these problems. In this study, we present a more rigorous and novel approach to these problems and implement it into an assembler, called ARCS. ARCS first estimates the copy numbers of contigs, and assembles and positions the identified non-repeat contigs optimally through a linear programming to unravel relative distance information among them. Then ARCS decomposes inexact repeats into a collection of exact repeats interspersed with unique sequences that distinguish one repeat from others. These unique sequences serve as bridges to connect neighboring contigs and local scaffolds into longer scaffolds. ARCS solves the chimeric paired-end reads problem by using a statistical model to iteratively filter out chimera read pairs. Finally, ARCS uses a gap filling method to merge contigs within each scaffold. Experimental results demonstrate that on the real sequencing data of the E. Coli genome, ARCS generates scaffolds with an N50 of 132kbp, and outperforms the state-of-art assembly tool SOAP-denovo2 with N50 scaffold of 95kbp. On two other genome sequencing data from D9 and D12 genomes, ARCS also assembles longer scaffolds than SOAP-denovo. Through case studies we demonstrate that the performance improvement can be explained by the accurate identification of chimeric read pairs, and the assembly of inexact repeats.

Paper ID: b108

Reconstructing phylogenetic trees for ultra-large unaligned DNA sequences via with Hadoop

Quan Zou

School of Computer Science and Technology, Tianjin University, China

Building the phylogenetic trees for ultra-large DNA sequence datasets is of vital importance and is a great challenge, especially for DNA sequences without alignment. We have employed the Hadoop parallel platform and develop a Java-based phylogenetic tree reconstruction software tool for ultra-large unaligned DNA datasets. Clustering and multiple-sequence alignment were executed in parallel, and the basic phylogenetic trees were built using the neighbour-joining model. I would also introduce some algorithm skills in the multiple sequence alignment process. Experiments on two large datasets, both more than 1 GB, show that our software tool can outperform other common phylogenetic reconstruction tools.

The software tool along with its codes and datasets are accessible at http://datamining.xmu.edu.cn/software/Phylogenetic_tree/ and http://datamining.xmu.edu.cn/software/halign/.

Paper ID: b109
RNA structurome reveals the second layer of genetic information

Qiangfeng Zhang

MOE Key Laboratory of Bioinformatics, Tsinghua-Peking Center for Life Sciences, School of Life Sciences, Tsinghua University, China

The most remarkable findings in the past two decades in biology include the mammalian genome is largely transcribed and how versatile functions RNA molecules can have. RNA structure may play a critical role in defining its function and regulation. However, due to lack of information, our knowledge on RNA structural language is very limited. In this talk, I will describe our recent effort in using new chemistry and deep sequencing techniques to probe RNA structures within a tube and also inside a cell, on a genome-wide scale. The study provides both the landscape and also the variation of human and mouse structural transcriptome. Analysis reveals structure features behind many important biological processes including translation, RNA methylation, and RNA-protein interaction etc. Our results highlight the potentially broad contribution of RNA structure and its variation to gene regulation.

Reference: Spitale R*, Flynn RA*, Zhang QC*, Pete Crisalli, Byron Lee, Jong-Wha Jung, Hannes Y. Kuchelmeister, Pedro J. Batista, Eduardo A. Torre, Eric T. Kool, and Chang HY. (2015) Structural imprints in vivo decode mechanisms of RNA regulation. Nature. 519 (7544): 486-90. (*Co-first authorship)

Paper ID: b110
The human microbe-disease network

Qinghua Cui
Peking University, China

The microbiota living in the human body has critical impact on our health and disease, but a systems understanding of its relationships with disease remains limited. Here, we use a large-scale manually-curated microbe-disease association dataset to construct a microbe-based human disease network and investigate the relationships between microbes and disease genes, symptoms, chemical fragments, and drugs. We reveal that the microbe-based disease loops are significantly coherent. Microbe-based disease connections have strong cross talks with those constructed by disease genes, symptoms, chemical fragments, and drugs. Moreover, we confirm that the microbe-based disease analysis is able to predict novel connections and mechanisms for disease, microbes, genes, and drugs. The presented network, methods and findings can be a resource helpful for addressing some important issues of medicine, for example, the discovery of bench knowledge and bedside clinical solutions for disease mechanism understanding, diagnosis, and therapy.

Paper ID: b111

Tunable Sensitivity and Its Biological Insights

Guangyu Wang

South University of Science and Technology of China, China

Compared with cooperative sensitivity, zero-order ultrasensitivity is tuning friendly. Using a mathematical model of regulated Goldbeter-Koshland kinetics, we find the tuning of ultrasensitivity can be decomposed into two orthogonal modes, which confer a remarkable separation of biological functions. These discoveries provide valuable insights into biological processes such as tissue development and cell signaling. Cell cycle is such a process, in which both irreversibility (e.g., the one-way cycling per se) and reversibility (e.g., cell cycle arrest) are important. We show that sensitivity and its tuning play essential roles to guarantee both properties. These discussions reveal the intricate relationship between circuit structure and biological function, and lead to the discovery of subtle but essential difference between positive feedback and double negative feedback.

Paper ID: b112

Inferring Direct PPI Networks from AP-MS Data

Zengyou He

School of Software, Dalian University of Technology, China

Affinity purification-mass spectrometry (AP-MS) is a mainstream experimental method for identifying PPIs in a highthroughput manner. In each AP-MS experiment, a tagged protein (bait) is first selectively purified along with its potential interacting partners (preys) from a cell or tissue lysate. Then, MS is used to identify and quantify these affinity purified proteins. Such purification experiments are repeated many times with different bait proteins. The set of bait-prey pairs from all purifications, termed the AP-MS data, is used to infer the underlying protein-protein interaction network structure. Ideally, one bait protein should have a real and direct interaction relationship with each associated prey protein. However, there are a large number of false positive interactions in the AP-MS data, where the prey protein can be a non-specific contaminant. In addition, some prey proteins don't interact with the bait protein directly, which connect to the bait protein via other intermediate proteins. Despite of recent algorithmic advances on inferring PPI networks, there are still several challenging problems that remain unsolved. In this report, we focus on one of such questions: Can we accurately infer the direct PPI network from AP-MS data? From the viewpoint of data analysis, we first discuss the algorithmic nature of such inference problem and then present several algorithms that are capable of tackling this problem. Finally, we present some empirical results and finish this report with some future research issues.

Paper ID: b113

Organelle-focused interactome of rice

Ming Chen

Department of Bioinformatics, College of Life Sciences, Zhejiang University, China

Proteomic analysis (proteomics) refers to the systematic identification and quantification of the complete complement of proteins (the proteome) of a biological system (cell, tissue, organ, biological fluid, or organism). To better understand the interactions of proteins in rice, we developed PRIN, a predicted rice interactome network. The protein-protein interactions data of PRIN are based on interologs of six model organisms where the large-scale protein-protein interaction experiments are applied. An example showed that proteins functional complex and biological pathways could be effectively expanded in our predicted network. Protein subcellular localization has been a long-standing key problem in investigating proteins' function, which provides important clues for revealing their functions and aids in understanding their interactions with other biomolecules at the cellular level. We presented a novel integrative approach (PSI) that derives the wisdom of multiple specialized predictors via a joint-approach of group decision making strategy and machine learning methods to give an integrated best result. We systematically defined the organelle-focused proteomes and interactomes in rice. A total of 83.42% of the whole rice proteome obtained their subcellular localizations based on manual annotation, manual adjustment and the prediction results of PSI. We illustrated the cross talk bias between different organelles and the function organization accounting for nine organelles. Motif analysis illustrated the protein interaction bias in different organelles to implement certain biology functions.

Paper ID: b114

MeSHLabeler: improving the accuracy of large-scale MeSH indexing by integrating diverse evidence

Shanfeng Zhu

Fudan University, China

Medical Subject Headings (MeSHs) are used by National Library of Medicine (NLM) to index almost all citations in MEDLINE, which greatly facilitates the applications of biomedical information retrieval and text mining. To reduce the time and financial cost of manual annotation, NLM has developed a software package, Medical Text Indexer (MTI), for assisting MeSH annotation, which uses k-nearest neighbors (KNN), pattern matching and indexing rules. Other types of information, such as prediction by MeSH classifiers (trained separately), can also be used for automatic MeSH annotation. However, existing methods cannot effectively integrate multiple evidence for MeSH annotation. We propose a novel framework, MeSHLabeler, to integrate multiple evidence for accurate MeSH annotation by using 'learning to rank'. Evidence comes form multiple sources, such as numerous predictions from MeSH classifiers, KNN, pattern matching and MTI. MeSHLabeler won the first place in Task 2A of 2014 BioASQ challenge, and its performaance is around 10% higher than the one achieved by MTI.

Paper ID: b115

Discrimination of recurrent CNVs from individual ones from multisample aCGH by jointly constrained minimization

Hongmin Cai

South China University of Technology, China

Copy number variations (CNVs) are associated with complex diseases and particular tumor types; thus, reliable CNV identification has substantial potential value. Several different high-throughput technologies can be used to identify CNV sites. One commonly used approach to detect CNVs is array-based comparative genomic hybridization (aCGH). Recent advances in sequencing technology make it affordable to obtain aCGH data for multiple samples, and an increasing number of methods have become available for detecting recurrent CNV regions across samples. However, copy number is highly dynamic in cancer cells and thus individually specific. In contrast, researchers anticipate that detecting recurrent CNVs in samples is an indication that the tumors share the same origin and thus possibly also have common oncogene drivers and tumor insurgence. Therefore, accurate discrimination of recurrent from individual CNVs is vital to explain various phenotype differences and genetic diseases. To address this problem, we present a general model to identify and discriminate two types of CNVs, namely sample-wised individual and group-wised recurrent CNVs from multi-sample aCGH profiles. We first imposed general assumptions on the sample-wised and group-wised CNVs. Then, we detected CNVs by proposing a convex optimization with multi-constraints to distinguish the two CNV types. An efficient numerical algorithm was then presented to solve the problem. We demonstrated the performance of the proposed method by comparing the results with those of popular alternative methods on both simulated and empirical breast cancer datasets. The experimental results demonstrate that the proposed method outperformed its peers by successfully detecting CNV patterns and accurately discriminating their differences.

PaperID: b116

WGX-50, A Drug Candidate from Sichuan Pepper and Its Potential Role Against AD And in Anti-aging

Dong-Qing Wei

State Key Lab. Of Microbial Metabolism and College of Life Science and Biotechnology, Shanghai Jiaotong University, China

Abstract: Based on structure of membrane protein targets acquired by bioinformatics tools, and database of molecules extracted from the Traditional Chinese medicines, various cheminformatics procedures were performed to screen for potential active compounds. Many promising molecules were obtained, for example, agaritine and wgx-50 was singled out through similarity search, molecular docking and molecular dynamics simulations. In vivo, the pharmacological kinetics

profiles of wgx-50 were obtained to evaluate the effect of wgx-50 on cognition ability of scopolamine induced acute memory damage mice and Amyloid Precursor Protein transgenic (APP-Tg) mice by Morris water maze testing (MWM), and the β-amyloid (Aβ) oligomerics in the brain of APP-Tg mice by immunohistochemisty. In vitro, the direct effect of wgx-50 on Aβ oligomerics was observed by the atomic force microscope; the neurocyte protection by Western blot and cell apoptosis assay; and measured intracellular calcium current ([Ca2+]i) by Laser confcal microscopy. Experiments in vivo showed that wgx-50 could penetrate the blood brain barrier and improve the cognition ability and decrease the Aβ oligomeric accumulation on cerebral cortex. Results in vitro displayed that wgx-50 could disassemble the Aβ oligomerics, inhibit Aβ-induced neurocyte apoptosis as well as apoptotic gene expression, and relieve neuron calcium intoxication. These results strongly suggest that wgx-50 possess biologic functions against AD. Discoveries were also made in its potential role in anti-aging.

References
1)   R.-X. Gu, H. Gu, Z.-Y. Xie, J.-F. Wang, H.R. Arias, D.-Q. Wei* and K. Ch. Chou, "Possible Drug Candidates for Alzheimer's Disease Deduced from Studying their Binding Interactions with α7 Nicotinic Acetylcholine Receptor", Medicinal Chemistry, 5, 250-262(2009).
2)   Shigao Chen, Ruoxu Gu* and Dong-Qing Wei*, "Virtual Screening with Respect to alpha7 Nicotinic Acetylcholine Receptor for Possible Leads against Alzheimer's disease", Journal of Molecular Graphics and Modeling, 39, 98-107(2013).
3)   H. R. Arias, Ruo-Xu Gu, Dominik Feuerbach, Bao-Bao Guo, Yong Ye, and D.Q. Wei*, "Novel Positive Allosteric Modulators of the Human α7 Nicotinic   Acetylcholine Receptor", Biochemistry, 50, 5263–5278(2011).
4)   Ruo-Xu Gu, Limin Angela Liu*, and Dong-Qing Wei*, " Equilibrium of four binding states of anti-viral drug rimantadine in M2-lipid bilayer system", J. Am. Chem. Soc., 133 (28), 10817–10825(2011).
5)   H. R. Arias, Ruo-Xu Gu, Dominik Feuerbach, Bao-Bao Guo, Yong Ye, and D.Q. Wei*, "Novel Positive Allosteric Modulators of the Human α7 Nicotinic   Acetylcholine Receptor", Biochemistry, 50, 5263–5278(2011).
6)   Maoping Tang, Zhaoxia Wang, Ying Zhou, Wangjie Xu, Shengtian Li,Lianyun Wang, Dong-Qing Wei*, Zhongdong Qiao*, "A novel drug candidate for Alzheimer disease treatment - gx-50 derived from Zanthoxylum Bungeanum", J. Alzheimer's Disease, 34, 203–213 (2013).