



BIBM 2012 Program Schedule

Hyatt Regency Philadelphia

PA, USA

Oct 4-7, 2012

Program

- [October 03, 2012](#)
- [October 04, 2012](#)
- [October 05, 2012](#)
- [October 06, 2012](#)
- [October 07, 2012](#)

Keynote Lecture: 60 minutes (about 45 minutes for talk and 15 minutes for Q and A)

Invited Talk: 40 minutes (about 30 minutes for talk and 10 minutes for Q and A)

Main conference regular Paper: 25 minutes (about 20 minutes for talk and 5 minutes for Q and A)

Main conference short Paper: 20 minutes (about 16 minutes for talk and 4 minutes for Q and A)

03-Oct								
16:00-20:00	Registration: Main Lobby							
04-Oct								
08:00-18:00	Registration: GrandBall Room Foyer (Floor 2) Venue: Floor2: GrandBall Rooms (Gb) A, BC, D, Columbus(Co) A, B, C, Floor3: Washington(Wa) A, B							
08:30-12:30	Tutorial 1	<u>Workshop 2</u>	<u>Workshop 4</u>	<u>Workshop 5</u>	<u>Workshop 13</u>	<u>Workshop 11</u>	<u>Workshop 8</u>	
Session Chairs	Jun Huan	Jing He A. Shehu N. Hapel Brian Chen	Young Cho Pietro Guzzi	D. Cameron A. Sheth O. Bodenreider	Kaizhi Tang Xiong Liu	Illhoi Yoo	Rui Jiang Jane Zheng Zhongming Zhao	
Venue:	Wa-A	Gb-A	Gb-BC	Gb-D	Co-A	Co-B	Co-C	
	Coffee: 10:00-10:20 Floor 2 Foyer							
12:30-13:30	Lunch at your own							
<i>Set up Poster Session (the poster authors could start their poster set-up and the posters will be displayed until Oct 7)</i>								
13:30-18:00	Industrial Session 1	<u>Workshop 2</u>	<u>Workshop 3</u>	<u>Workshop 7</u>	<u>Workshop 10</u>	<u>Workshop 11</u>	<u>Workshop 8</u>	Tutorial 2
Session Chairs	Anastasia Christianson Michael Liebman	Jing He A. Shehu N. Hapel Brian Chen	S. Kim L. Chen K. Horimoto	Dong Xu Guozheng Li Z. Liang	Dongxiao Zhu Steve Qin	Illhoi Yoo	Rui Jiang Jane Zheng Zhongming Zhao	Jalel Akaichi
Venue:	Wa-A	Gb-A	Gb-BC	Gb-D	Co-A	Co-B	Co-C	Wa-B
	Coffee: 16:00-16:20 Floor 2 Foyer							
05-Oct								
08:00-18:00	Registration: GrandBall Room Foyer (Floor 2) Venue: Floor2: GrandBall Rooms (Gb) A, BC, D, Floor3: Riverview (Rv) A, B, Washington(Wa) A, B							
08:00-08:15	Opening and Welcoming Speech Conference co-Chairs: Lyle Ungar (University of Pennsylvania, USA) Cathy Wu (University of Delaware, USA) Program co-Chairs: Jean Gao (University of Texas at Arlington, USA) Reda Alhajj (University of Calgary, Canada) Werner Dubitzky (University of Ulster, UK) Industry Program co-Chairs: Anastasia Christianson (AstraZenica, USA) Michael Lieban (Strategic Medicine Inc., USA) BIBM Steering Committee Chair: Xiaohua Tony Hu (Drexel University)							
Venue:	Gb-ABCD							
08:15-09:15	Session Chair: Lyle Ungar <i>Keynote Lecture 1: The CellOrganizer Project: An Open Source System to Learn Image-derived Models of Subcellular Organization over Time and Space</i> Dr. Robert Murphy, Lane Center for Computational Biology and Department of Biological Sciences, Carnegie Mellon University, USA							
Venue:	Gb-ABCD							
09:15-09:30	Coffee Break, Floor 2 Foyer							
09:30-11:20	Session 1	Session 2	Session 3	Industrial Workshop 1	Tutorial 3			
Session Chair	Huzefa Rangwala	Nurit Haspel	Perry Evans	David Fenske Michael Liebman	Illhoi Yoo			
Venue	Gb-A	Gb-BC	GB-D	Wa-A	Wa-B			
11:20-12:20	Lunch (provided by the conference, OmisSoft presentation at Gb-ABCD from 11:35-11:55am) Poster Session Floor 2 Foyer							

12:20-13:20	Session Chair: Cathy Wu <i>Keynote Lecture 2: Protein 3D structure from genomic sequences and application to cancer genomics</i>				
	Dr. Chris Sander, Chair, Computational Biology Program, Sloan-Kettering Institute				
<i>Venue:</i>	<i>Gb-ABCD</i>				
13:20-13:35	Coffee Floor 2 Foyer				
13:35-15:30	Session 4	Session 5	Session 6	Industrial Session 2	<u>Workshop 11</u>
Session Chair	Jimmy Huang	Michael Ochs	Brian Chen	Anastasia Christianson Michael Liebman	Illhoi Yoo
<i>Venue:</i>	<i>Gb-A</i>	<i>Gb-BC</i>	<i>Gb-D</i>	<i>Wa-A</i>	<i>Wa-B</i>
15:30-17:00	Session 7	Session 8	Session 9	Industrial Session 2	<u>Workshop 11</u>
Session Chair	Fahad Saeed	Ahmet Sacan	Tamer Kahveci	Anastasia Christianson Michael Liebman	Illhoi Yoo
<i>Venue:</i>	<i>Gb-A</i>	<i>Gb-BC</i>	<i>Gb-D</i>	<i>Wa-A</i>	<i>Wa-B</i>

06-Oct					
08:00-18:00	Registration: GrandBall Room Foyer (Floor 2)				
<i>Venue:</i>	<i>Floor2: GrandBall Rooms (Gb) A, BC, D, Floor3: Riverview (Rv) A, B, Washington(Wa) A, B</i>				
08:30-09:30	Session Chair: Jean Gao <i>Keynote Lecture 3: Protein Structure Determination on Demand</i>				
	Prof. Ming Li, Canada Research Chair in Bioinformatics, University of Waterloo				
<i>Venue:</i>	<i>Gb-ABCD</i>				
9:30-9:50	Coffee Floor 2				
9:50-12:00	Session 10	Session 11	Session 12	<u>Workshop 9</u>	Industrial Workshop 2
Session Chair	Nikolay Balov	Taghi Khoshgoftaar	Reda Alhajj	Fangxiang Wu Jianxing Wang	Sabrina Molinaro Michael Liebman
<i>Venue:</i>	<i>Ga-A</i>	<i>Gb-BC</i>	<i>Gb-D</i>	<i>Rv-A</i>	<i>Rv-B</i>
12:00-13:15	Buffet Lunch (Gb-ABCD) Poster Session Floor 2 foyer				
13:15-14:15	Session Chair: Reda Alhajj <i>Keynote Lecture 4: Obstacles and Options for Big-Data Applications in Biomedicine: The role of standards and normalizations</i>				
	Dr. Christopher G. Chute, MD, PH, Professor of Medical Informatics, Mayo Clinic College of Medicine, and IHI Fellow, University of Minnesota, MN				
<i>Venue:</i>	<i>Gb-ABCD</i>				
14:15-14:35	Coffee Floor 2 Foyer				
14:35-15:15	Invited Talk1	Invited Talk 2	Invited Talk 3	Invited Talk 4	Industrial Workshop 3
Speaker:	Michael Ochs	Jimmy Huang	Xue-Wen Chen	Luis Rueda	
Session Chair:	Yidong Chen	Reda Alhajj	Marie-Jean Meurs	Fangxiang Wu	Aaron Kamauu Scott DuVall
<i>Venue:</i>	<i>Gb-A</i>	<i>Gb-BC</i>	<i>Gb-D</i>	<i>Rv-A</i>	<i>Rv-B</i>
15:20-17:30	Session 13	Session 14	Session 15	<u>Workshop 9</u>	Industrial Workshop 3
Session Chair	Yidong Chen	Jun Huan	Marie-Jean Meurs	Fangxiang Wu Jianxing Wang	Aaron Kamauu Scott DuVall
<i>Venue:</i>	<i>Gb-A</i>	<i>Gb-BC</i>	<i>Gb-D</i>	<i>Rv-A</i>	<i>Rv-B</i>
18:30-22:00	Banquet (Best Paper Award, Best Student paper Award, Best Poster Award) (Gb-ABCD)				

07-Oct					
08:00-15:00	Registration: GrandBall Room Foyer (Floor 2)				
<i>Venue:</i>	<i>Floor2: GrandBall Rooms (Gb) A, BC, D, Floor3: Riverview (Rv) A, B, Washington(Wa) A, B</i>				
8:15-10:10	Session 16	Session 17	Session 18	<u>Workshop 6</u>	<u>Workshop 12</u>
Session Chair	Maryam Panahiazar	Anand Mehta	Jean Gao	Ye Duan Yu-Ping Wang	Taesung Park
<i>Venue:</i>	<i>Gb-A</i>	<i>Gb-BC</i>	<i>Gb-D</i>	<i>Rv-A</i>	<i>Rv-B</i>
10:10-10:30	Coffee				
10:30-12:25	Session 19	Session 20	Session 21	<u>Workshop 6</u>	
Session Chair	Zoran Obradovic	Tony Hu	John Rose	Ye Duan Yu-Ping Wang	
<i>Venue</i>	<i>Gb-A</i>	<i>Gb-BC</i>	<i>Gb-D</i>	<i>Rv-A</i>	
12:25-13:30	Lunch at your own				

13:30-15:25	Session 22	Session 23	Session 24
Session Chair	Ye Duan	Hamidreza Chitsaz	Luis Rueda
Venue:	Gb-A	Gb-BC	Gb-D

Keynote Lectures: 4

Keynote 1:

Title: The CellOrganizer Project: An Open Source System to Learn Image-derived Models of Subcellular Organization over Time and Space

Speaker:

Dr. Robert Murphy, Lane Center for Computational Biology and Department of Biological Sciences, Carnegie Mellon University, USA

Abstract:

The CellOrganizer project (<http://cellorganizer.org>) provides open source tools for learning generative models of cell organization directly from images and for synthesizing cell images (or other representations) from one or more of those models. Model learning captures variation among cells in a collection of images. Images used for model learning and instances synthesized from models can be two- or three-dimensional static images or movies. Current components of CellOrganizer can learn models of cell shape, nuclear shape, chromatin texture, vesicular organelle number, size, shape and position, and microtubule distribution. These models can be conditional upon each other: for example, for a given synthesized cell instance, organelle position will be dependent upon the cell and nuclear shape of that instance. The models can be parametric, in which a choice is made about an explicit form to represent a particular structure, or non-parametric, in which distributions are learned empirically. One of the main uses of the system is in support of cell simulations: models learned from separate experiments can be combined into one or more synthetic cell instances that are output in a form compatible with cell simulation engines such as MCell, Virtual Cell and Smoldyn. Another important application of the system is in comparison of target patterns and perturbation effects in high content screening and analysis. This is currently done using numerical features, but these are difficult to compare across different microscope systems or cell types since features can be affected by changes in more than one aspect of cell organization. More robust comparisons can be made using generative model parameters, since these can distinguish effects on cell size or shape from effects on organelle pattern. Ultimately, it is anticipated that collaborative efforts by many groups will enable creation of image-derived generative models that permit accurate modeling of cell behaviors, and that can be used to drive experimentation to improve them through active learning.

Short Bio:

Robert F. Murphy is the Ray and Stephanie Lane Professor of Computational Biology and Professor of Biological Sciences, Biomedical Engineering, and Machine Learning at Carnegie Mellon University, and Director (Department Head) of the Lane Center for Computational Biology in the School of Computer Science. He is also Honorary Professor of Biology at the Albert Ludwig University of Freiburg, Germany, a Fellow of the American Institute for Medical and Biological Engineering, and the recipient of an Alexander von Humboldt Foundation Senior Research Award. He is Past-President of the International Society for Advancement of Cytometry, and is a member of the National Advisory General Medical Sciences Council and the NIH Council of Councils. He has published over 190 research papers in the areas of cell and computational biology. Dr. Murphy's career has centered on combining fluorescence-based cell measurement methods with quantitative and computational methods. In the mid 1990's, his group pioneered the application of machine learning methods to high-resolution fluorescence microscope images depicting subcellular location patterns. His current research interests include image-derived models of cell organization and active machine learning approaches to experimental biology.

Keynote 2:

Title: Protein 3D structure from genomic sequences and application to cancer genomics

Speaker:

Dr. Chris Sander, Computational and Systems Biology Memorial Sloan-Kettering Cancer Center

Abstract:

Amino acid covariation in proteins, extracted from the evolutionary sequence record, can be used to fold proteins, including transmembrane proteins. Addressing a fundamental challenge in computational molecular biology, a new prediction method (EVold) applies a maximum entropy approach to infer evolutionary couplings between sequence positions from correlated mutations in the multiple sequence alignment of a protein family. When translated to distance constraints, such residue-residue couplings are sufficient to generate good all-atom models of proteins from different fold classes, ranging in size from 50 to more than 300 residues. We use the technique to predict previously unknown 3D structures of large transmembrane proteins of biomedical interest, from their sequences alone. We show how the method can plausibly predict oligomerization, functional sites, and conformational changes in transmembrane proteins. Project co-leader: Debora Marks, Harvard Medical School; co-authors (alphabetical): Lucy Colwell, Thomas Hopf, Andrea Pagnani, Burkhard Rost, Robert Sheridan, Riccardo Zecchina. See <http://bit.ly/tob48p> (PDF) and www.evold.org. The discovered evolutionary couplings provide insight into essential interactions constraining protein evolution and, with the rapid rise in large-scale sequencing, are likely to facilitate a comprehensive survey of the universe of protein structures by a combination computational and experimental technology. Applications to cancer genomics relate to the interpretation of the functional impact of cancer-related mutations and the design of targeted therapeutics.

Short Bio:

Chris Sander is acknowledged as an initiating leader in the field of computational biology, an interdisciplinary field that aims to solve important problems in biology using techniques of mathematics, physics, engineering, and computer science. He is Head of the Computational Biology Center at Memorial Sloan Kettering Cancer Center and Tri-Institutional professor at Rockefeller and Cornell Universities. Sander's current research interests are in computational genomics and systems biology, with a focus on network pharmacology and the development of targeted combinatorial therapy in cancer. His group uses the results of high-throughput sequencing to compute protein 3D structures and functional sites; and studies the regulation of gene expression by small RNAs. In 2012, he is active in the International Cancer Genomics Consortium, the NIH Cancer Genome Atlas Project, the NCI Integrative Cancer Biology Program and a leader in the bioPAX and PathwayCommons community efforts to create an

open-source information resource for biological pathways. He has published more than 250 peer-reviewed articles in physics and biology (<http://bit.ly/Uk990K>) with an h-index of 100. Previously, Sander co-founded the research section of the European Bioinformatics Institute in Cambridge, England, and was founding chair of the department of Biocomputing at the European Molecular Biology Laboratory in Heidelberg. He is a Fellow of the International Society for Computational Biology.

Keynote 3:

Title: Protein Structure Determination on Demand

Speaker:

Prof. Ming Li, Canada Research Chair in Bioinformatics, School of Computer Science, University of Waterloo

Abstract:

Protein structure prediction by computers at best may serve as a screening method, and the current high-throughput protein structure determination methods are costly and will never exhaust all proteins. A complementary approach is "protein structure determination on demand", say in a week. We will discuss two approaches that would realize this goal: automatic protein structure determination using NMR data and mass spectrometry data.

Short Bio:

Ming Li is a Canada Research Chair in Bioinformatics and a University Professor at the University of Waterloo. He is a fellow of the Royal Society of Canada, ACM, and IEEE. He is a recipient of E.W.R. Steacie Fellowship Award in 1996, the 2001 Killam Fellowship, and the 2010 Killam Prize. Together with Paul Vitanyi they have co-authored the book "An Introduction to Kolmogorov Complexity and Its Applications". He is a co-managing editor of Journal of Bioinformatics and Computational Biology.

Keynote 4:

Title: Obstacles and Options for Big-Data Applications in Biomedicine: The role of standards and normalizations

Speaker:

Dr. Christopher G. Chute, MD, PH, Professor of Medical Informatics, Mayo Clinic College of Medicine, and IHI Fellow, University of Minnesota, MN

Abstract:

Advances in computing capabilities are palpably evident throughout many industries manifest by unprecedented, large-scale data integration and inferencing. Branded as "big-data" in many cases, the question of whether such techniques can leverage advances in biomedicine and clinical practice are obvious. High-throughput clinical analytics, synthesizing genomic and clinical attributes of a particular patient, portends predictive models that can directly influence clinical care decisions. However, to make this widely shared vision practical and scalable, barriers attributable to data heterogeneity dominate. Methods and strategies to increase the comparability and consistency of healthcare related data will be discussed.

Short Bio:

Dr. Chute received his undergraduate and medical training at Brown University, internal medicine residency at Dartmouth, and doctoral training in Epidemiology at Harvard. He is Board Certified in Internal Medicine, and a Fellow of the American College of Physicians, the American College of Epidemiology, and the American College of Medical Informatics. He became founding Chair of Biomedical Informatics at Mayo in 1988, stepping down after 20 years in that role. He is now Professor of Medical Informatics and Section Head. He is PI on a large portfolio of research including the HHS/Office of the National Coordinator (ONC) SHARP (Strategic Health IT Advanced Research Projects) on Secondary EHR Data Use, the ONC Beacon Community (Co-PI), the LexGrid projects, Mayo's CTSA Informatics, and several NIH grants including one of the eMERGE centers from NCHRI, which focus upon genome wide association studies against shared phenotypes derived from electronic medical records. Dr. Chute serves as Vice Chair of the Mayo Clinic Data Governance for Health Information Technology Standards, and on Mayo's enterprise IT Oversight Committee. He is presently Chair, ISO Health Informatics Technical Committee (ISO TC215) and Chairs the World Health Organization (WHO) ICD-11 Revision. He also serves on the Health Information Technology Standards Committee for the Office of the National Coordinator in the US DHHS, and the HL7 Advisory Board. Recently held positions include Chair of the Biomedical Computing and Health Informatics study section at NIH, Chair of the Board of the HL7/FDA/NCI/CDISC BRIDG project, on the Board of the Clinical Data Interchange Standards Consortium (CDISC), ANSI Health Information Standards Technology Panel (HITSP) Board member, Chair of the US delegation to ISO TC215 for Health Informatics, Convener of Healthcare Concept Representation WG3 within the (TC215), Co-chair of the HL7 Vocabulary Committee, Chair of the International Medical Informatics Association (IMIA) WG6 on Medical Concept Representation, American Medical Informatics Association (AMIA) Board member, and multiple other NIH biomedical informatics study sections as chair or member

Invited Talks: 4

Invited Talk 1:

Title: Identifying Deregulated Signaling Pathways and Targeting Aberrant Proteins in Cancer

Speaker:

Prof. Michael Ochs, Division of Biostatistics and Bioinformatics, School of Medicine, The Johns Hopkins University

Abstract:

For cancer to develop, a number of cellular regulatory programs must be circumvented. Many of these programs are controlled by signaling networks, and identification of deregulated pathways and aberrant protein activity promises to provide targets for therapeutic development. First we present our work on the development of machine learning techniques for identifying transcription factor activity from transcriptomic data to deduce deregulated signaling processes as well as their response to targeted therapy. Second we present methods that learn signaling deregulation through integration of diverse omics data. We show that application to omics data from tumor samples leads to identification of unanticipated responses to targeted therapy and to new pathways and targets in specific cancer types.

Short Bio:

Dr. Michael Ochs is Associate Professor of Oncology in the Division of Oncology Biostatistics and Bioinformatics at the Johns

Hopkins School of Medicine. He holds a Ph.D. in Physics from Brandeis University with a focus on radio astronomy and the astrophysics of quasars. He migrated into medical research through in vivo spectroscopic imaging research, and directed the Bioinformatics Group at the Fox Chase Cancer Center until September 2006, when he joined the faculty at Johns Hopkins. Dr. Ochs has served as director of the Bioinformatics Shared Resource at Johns Hopkins since 2009.

Invited Talk 2:

Title: Promoting Ranking Diversity for Biomedical Information Retrieval Using Wikipedia

Speaker:

Prof. Jimmy Huang, Information Retrieval and Knowledge Management Research Lab, School of Information Technology, York University, Toronto, Canada

Abstract:

Traditional Information Retrieval models assume that the relevance of a document is independent of the relevance of other documents. However, in reality, this assumption may not hold. The usefulness of retrieving a document usually depends on previous ranked documents, since a user may want to see the top ranked documents concerning different aspects of his/her information need instead of reading relevant documents that only deliver redundant information. In this talk, I will discuss how to find relevant documents that can deliver more different aspects of a query. In particular, I will discuss how to use Wikipedia to detect aspects covered by retrieved documents. I will discuss new models derived from the survival analysis theory for measuring aspect novelty. An aspect filter based on a two-stage model will be introduced and a new re-ranking method that combines the novelty and the relevance of a retrieved document at the aspect level will also be presented. Through extensive experiments on standard large-scale TREC biomedical collections, I will show that the proposed models and methods are effective in promoting ranking diversity for biomedical information retrieval.

Short Bio:

Jimmy Huang is the Director & a full Professor at the School of Information Technology and the founding director of Information Retrieval & Knowledge Management Research Lab at the York University. He joined York University as an Assistant Professor in July 2003. Previously, he was a Post Doctoral Fellow at the School of Computer Science, University of Waterloo. He did his PhD in Information Science at City University in London. He also worked in the financial industry in Canada, where he was awarded a CIO Achievement Award. Since 2003, he has published more than 150 refereed papers in top-tier journals (such as JASIST, IPM, IEEE TKDE, Information Sciences, IR, BMC Bioinformatics and BMC Genomics), book chapters and international conference proceedings (such as ACM SIGIR, ACM CIKM, COLING and IEEE ICDM). He received the Dean's Award for Outstanding Research in 2006, an Early Researcher Award, formerly the Premier's Research Excellence Awards in 2007, the Petro Canada Young Innovators Award in 2008, the SHARCNET Research Fellowship Award in 2009 and the Best Paper Award at the 32nd European Conference on Information Retrieval in 2010. He was the General Conference Chair for the 19th International ACM CIKM Conference and the General Program Chair for IEEE/ACM International Joint Conferences on Web Intelligence & Intelligent Agent Technology in 2010. His research focuses on the area of information retrieval and its applications to medical healthcare and Web.

Invited Talk 3:

Title: Systems Biology of Human Diseases

Speaker:

Prof. Xue-Wen Chen, Professor and Chair, Department of Computer Science, Wayne State University.

Abstract:

While much of molecular biology research has led to a wealth of knowledge about individual cellular components and their functions, it has become increasingly clear that most cellular functions are carried out by complex networks of interconnected components, and that the characterization of isolated cellular components is not sufficient to understand the cell's complexity. In recent years, the development of high-throughput technologies has provided the scientific community with exciting new opportunities for systematically studying biological networks on a whole-genome scale. In this talk, I will highlight some examples of how systems biology provides a framework for impacting medicine and healthcare research, particularly for understanding human diseases with a network view of the factors that cause diseases. I will also discuss the challenges and opportunities.

Short Bio:

Xue-wen Chen is currently Professor and Chair in the Department of Computer Science at Wayne State University. He is a senior IEEE member, the chair of IEEE Computer Society Technical Committee on Computational Life Sciences, the chair of task force in Systems Biology in the IEEE Computational Intelligence Society Bioinformatics and Biomedicine Technical Committee, steering committee chair of the IEEE Computer Society Healthcare Informatics, Imaging, and Systems Biology, and a member in the IEEE Computer Society Bioinformatics and Biomedicine Steering Committee. He serves in the editorial board in several international journals. He is also a member of the leadership team in the IEEE Life Science Initiative. Dr. Chen is the conference chair for ACM Conference on Information and Knowledge Management (CIKM) 2012, which will be held in October 29 in Maui, Hawaii.

Dr. Chen received his PhD degree from Carnegie Mellon University, Pittsburgh, USA in 2001. He was a recipient of the prestigious NSF CAREER Award. He served as conference chair (and co-chairs) for the IEEE international Conference on Healthcare Informatics, Imaging, and Systems Biology, 2011, the IEEE International Conference on Bioinformatics and Biomedicine (BIBM) in 2009 and the International Conference in Machine Learning and Applications in 2011 and 2008. He also served as a program committee member in numerous conferences such as KDD, CIKM, BIBM, ICDM, and CEC. His research interest includes machine learning, data mining, bioinformatics, systems biology, and healthcare informatics.

Invited Talk 4:

Title: The Role of Sequence and Structural Domains in Mediating Protein-protein Interactions

Speaker:

Prof. Luis Rueda, Associate Professor, School of Computer Science, University of Windsor, Canada

Abstract:

Protein-protein interactions (PPIs) play crucial roles in most biological processes in living cells, including cell regulation, basic biochemical reactions, and many others. An important aspect that has been studied in PPIs is the identification of different types of protein complexes. Physicochemical, geometric and conservation-based properties have been used successfully in

prediction of PPIs. Among these, characterization of PPIs in terms of smaller units, namely sequence, structural and functional domains, and their interactions – domain-domain interactions or DDIs – have been shown beneficial as well. Some limitations exist on the number of structures available, compared to the millions of proteins known by sequence, and the difficulties inherent in calculating some of these relevant properties. On the other hand, motifs, and in particular short, linear motifs – SLiMs – have been shown to play an important role in mediating protein interactions with other proteins and specific functional domains, and in important metabolic processes, such as hijacking viruses or determining functional units in disordered proteins. A recent study has shown that the use of SLiMs can characterize stable interactions, predicting obligate and nonobligate complexes with almost perfect accuracy – more than 99%. The role of physicochemical properties, DDIs and SLiMs in stability of complexes will be discussed, as well as their implications in prediction and analysis.

Short Bio:

Luis Rueda is an associate professor in the School of Computer Science and director of the Pattern Recognition and Bioinformatics Lab at the University of Windsor in Canada. He joined the School of Computer Science in 2002, after completing his Master's and Ph.D. degrees in computer science from Carleton University in Canada. His research mostly focuses on finding biomarkers in transcriptomic data and in understanding stability of complexes in protein-protein interactions. He holds three patents on data encryption and more than 100 publications in prestigious journals and conferences, mostly in pattern recognition and bioinformatics. He is a Senior Member of the IEEE and a Member of the IAPR - International Association for Pattern Recognition.

Conference Paper Presentations

Session 1

Regular	B211 "Manifold learning reveals nonlinear structure in metagenomic profiles" Xingpeng Jiang, Xiaohua Hu, Huiyu Shen, and Tingting He
Regular	B293 "LSH-Div: Species Diversity Estimation using Locality Sensitive Hashing" Zeehasham Rasheed, Huzefa Rangwala, and Daniel Barbara
Short	B207 "Extending the Ball-Histogram Method with Continuous Distributions and an Application to Prediction of DNA-Binding Proteins" Ondrej Kuzelka, Andrea Szaboova, and Filip Zelezny
Short	B218 "Pre-miRNA Classification via Combinatorial Feature Mining and Boosting" Jason Wang, Ling Zhong, Dongrong Wen, and Bruce Shapiro
Short	B249 "OWA-PSSM - A Position Specific Scoring Matrix based Method Integrated with OWA Weights for HLA-DR Peptide Binding Prediction" Wen-Jun Shen and Hau-San Wong

Session 2

Regular	B307 "Efficient Basin Hopping in the Protein Energy Surface" Brian Olson and Amarda Shehu
Regular	B366 "Improving Interacting Residue Prediction Using Long-Distance Information in Hidden Markov Models" Colin Kern, Alvaro Gonzalez, Li Liao, and Vijay Shanker
Short	B286 "Rotation Crossover and K-Site Move Mutation for Evolutionary Protein Folding in 3D FCC HP Model" Shih-Chieh Su and Jyh-Jong Tsay
Short	B288 "Identifying essential proteins via integration of protein interaction and gene expression data" Xi-wei Tang, Jianxin Wang, and Yi Pan
Short	B291 "Combining Homolog and Motif Similarity Data with Gene Ontology Relationships for Protein Function Prediction" Hafeez ur Rehman, Alfredo Benso, Stefano Di Carlo, Gianfranco Politano, Alessandro Savino, and Prashanth Suravajhala

Session 3

Regular	B481 "Estimating a gene's mutation burden by the number of observed synonymous base substitutions" Perry Evans and Michael Krauthammer
Regular	B498 "De-noise Biological Network from Heterogeneous Sources via Link Propagation" Nan Du, Jing Gao, Vishrawas Gopalakrishna, and Aidong Zhang
Short	B295 "Identifying enterotype in human microbiome by decomposing probabilistic topics into components" Xingpeng Jiang, Dushoff Jonathan, Xin Chen, and Xiaohua Hu
Short	B316 "A Basin Hopping Algorithm for Protein-protein Docking" Irina Hashmi and Amarda Shehu
Short	B330 "Not All Protein Complexes Exhibit Dense Structures in <i>S. cerevisiae</i> PPI Network" Bolin Chen, Jinhong Shi, and Fang-Xiang Wu

Session 4

Regular	B225 "Efficient Filtration for Similarity Search with Spaced k-mer Neighbors" Weiming Li, Bin Ma, and Kaizhong Zhang
Regular	B263 "A Log-Linear Graphical Model for Inferring Genetic Networks from High-Throughput Sequencing Data" Genevera Allen and Zhandong Liu
Regular	B285 "Prediction of Novel Systems Components in Cell Cycle Regulation in Malaria Parasite by Subnetwork Alignments" Hong Cai, Changjin Hong, Jianying Gu, Timothy Lilburn, Rui Kuang, and Yufeng Wang
Short	B240 "Drug-Target Network in Myocardial Infarction: a Structural Analysis" Haiying Wang, Huiru Zheng, Francisco Azuaje, and Xing-Ming Zhao
Short	B247 "A Model of Cellular Decision Making in Photodynamic Therapy of Cancer" Ioannis Gkigkitzis and Xin Hua Hu

Session 5	
Regular	B298 "Significance Analysis by Minimizing False Discovery Rate" Yuanzhe Bei and Pengyu Hong
Regular	B299 "SiS: Significant Subnetworks in Massive Number of Network Topologies" Md Mahmudul Hasan, Yusuf Kavurucu, and Tamer Kahveci
Regular	B310 "Identifying Context-Specific Transcription Factor Targets from Prior Knowledge and Gene Expression Data" Elana J Fertig, Alexander V Favorov, and Michael F Ochs
Short	B252 "Similarity Analysis of Feature Ranking Techniques on Imbalanced DNA Microarray Datasets" David Dittman, Taghi Khoshgoftaar, Randall Wald, and Amri apolitano
Short	B277 "The Effect of Measurement Approach and Noise Level on Gene Selection Stability" Randall Wald, Taghi Khoshgoftaar, and Ahmad Abu Shanab

Session 6	
Regular	B346 "Aligning protein-protein interaction networks using random neural networks" Hang Phan, Michael Sternberg, and Erol Gelenbe
Regular	B381 "Computation of biochemical pathway fluctuations beyond the linear noise approximation using iNA" Philipp Thomas, Hannes Matuschek, and Ramon Grima
Short	B353 "A Comparison Study on Protein-protein Interaction Network Models" Mingyu Shao, Yi Yang, Jihong Guan, and Shuigeng Zhou
Short	B296 "Modelling Non-Stationary Gene Regulatory Process with Hidden Markov Dynamic Bayesian Network" Shijia Zhu and Yadong Wang
Short	B478 "Aligning Ligand Binding Cavities by Optimizing Superposed Volume" Ruobing Chen, Katya Scheinberg, and Brian Chen

Session 7	
Regular	B417 "Combining gene expression and function in a spatially localized approach" Evangelia Zacharaki, Angeliki Skoura, Li An, Desmond Smith, Scott Faro, and Vasileios Megalooikonomou
Regular	B440 "An Advanced Method for Identifying Protein-Protein Interaction by Analyzing TAP/MS Data" Xiaoyun Sun, Pengyu Hong, Meghana Kulkarni, Young Kwon, and Norbert Perrimon
Regular	B459 "A two-variable model for stochastic modelling of chemical events with multi-step reactions" Qianqian Wu, Kate Smith-Miles, and Tianhai Tian
Short	B338 "Accurate detection of SNPs using base-specific cleavage and mass spectrometry" Ruimin Sun, Xiang Gao, Nanyu Han, Qiong Wu, Yuguang Mu, Kai Tang, and Xin Chen
Short	B374 "An Efficient Algorithm for Clustering of Large-Scale Mass Spectrometry Data" Fahad Saeed, Trairak Pisitkun, Mark Knepper, and Jason Hoffert

Session 8	
Regular	B463 "Identifying Protein Binding Functionality of Protein Family Sequences by Aligned Pattern Clusters" En-Shiun Annie Lee and Andrew K. C. Wong
Regular	B467 "A Weighted Hypergeometric Statistic for the Enrichment of Gene Sets" Rehman Qureshi and Ahmet Saçan
Regular	B468 "An adaptive feature selection method for microarray data analysis" Jie Cheng, Joel Greshock, Leming Shi, Jeff Painter, Xiwu Lin, Kwan Lee, Shu Zheng, Richard Wooster, Lajos Pusztai, and Alan Menius
Short	B377 "Robust RFCM Algorithm for Identification of Co-expressed miRNAs" Sushmita Paul and Pradipta Maji
Short	B415 "Using Gene Sets to Identify Putative Drugs for Breast Cancer" Tzu-Hung Hsiao, Hung-I Harry Chen, Yidong Chen, Yu-Heng Chen, and Eric Y. Chuang

Session 9	
Regular	B475 "Fast Sparse Representation Approaches for the Classification of High-Dimensional Biological Data" Yifeng Li and Alioune Ngom
Regular	B484 "A Discrete Bayesian Network Framework for Discrimination of Gene Expression Profiles" Nikolay Balov
Regular	B499 "A Link Prediction based Unsupervised Rank Aggregation Algorithm for Informative Gene Selection" Kang Li, Nan Du, and Aidong Zhang
Short	B419 "On the Design of Advanced Filters for Biological Networks using Graph Theoretic Properties" Kathryn Dempsey, Tzu-Yi Chen, Sanjukta Bhowmick, and Hesham Ali
Short	B464 "Stress induces biphasic-rewiring and modularization patterns in the metabolomic networks of Escherichia coli" Fayez Aziz, Philemon Chan, Johan Osorio, Bushra Minhas, Vaisak Parekatt, and Gustavo Caetano-Anolles

Session 10	
Regular	B486 "A random walk based approach for improving protein-protein interaction network and protein complex prediction" Chengwei Lei and Jianhua Ruan
Regular	B492 "Unsupervised Kernel Parameter Estimation by Constrained Nonlinear Optimization for Clustering Nonlinear Biological Data" Hyokyeong Lee and Rahul Singh
Short	B470 "Inferring Fuzzy Cognitive Map Models for Gene Regulatory Networks from Gene Expression Data" Ye Chen, Lawrence Mazlack, and Long Lu

Short	B306 "Uncertain interactions affect degree distribution of biological networks" Andrei Todor, Alin Dobra, and Tamer Kahveci
Short	B494 "Drug-Drug Interaction Analysis Using Heterogeneous Biological Information Network" Kyubum Lee, Sunwon Lee, Minji Jeon, Jaehoon Choi, and Jaewoo Kang

Session 11

Regular	B385 "Data Driven Knowledge Acquisition Method for Domain Knowledge Enrichment in the Healthcare" Sujan Perera, Cory Henson, Krishnaprasad Thirunarayan, Amit Sheth, and Suhas Nair
Short	B441 "MDAsim: a Multiple Displacement Amplification Simulator" Zeinab Taghavi and Sorin Draghici
Short	B255 "Bridging Encounter Forms and Electronic Medical Record Databases: Annotation" Yuan An, Ritu Khare, Xiaohua Hu, and Il-Yeol Song
Short	B356 "Formalization of clinical trial eligibility criteria: Evaluation of a pattern-based approach" Krystyna Milian, Anca Bucur, and Annette ten Teije
Short	B379 "Skeleton Timed Up and Go" Okko Lohmann, Thomas Luhmann, and Andreas Hein

Session 12

Regular	B258 "Integrating Protein Networks for Identifying Cooperative miRNA Activity in Disease Gene Signatures" Mohammed Alshalalfa and Reda Alhadj
Regular	B260 "PPIExtractor: A Protein-Protein Interaction Extractor for Biomedical Literature" Zhihao Yang, Zhehuan Zhao, Yanpeng Li, Yuncui Hu, and Hongfei Lin
Regular	B264 "Drug-induced QT Prolongation Prediction Using Co-regularized Multi-view Learning" Jintao Zhang and Jun Huan
Short	B458 "ENISI Visual, an agent-based simulator for modeling gut immunity" Yongguo Mei, Raquel Hontecillas, Xiaoying Zhang, Keith Bisset, Stephen Eubank, Stefan Hoops, Madhav Marathe, and Josep Bassaganya-Riera
Short	B372 "The e!DAL JAVA-API: Store, Share and Cite Primary Data in Life Sciences" Daniel Arend, Matthias Lange, Christian Colmsee, Steffen Flemming, Jinbo Chen, and Uwe Scholz

Session 13

Regular	B268 "MultiFacTV: Finding Modules from Higher-order Gene Expression Profiles with Time Dimension" Xutao Li, Yunming Ye, Qingyao Wu, and Michael Ng
Regular	B278 "Modeling Semantic Influence for Biomedical Research Topics using MeSH Hierarchy" Dan He
Regular	B297 "M-Finder: functional association mining from protein interaction networks weighted by semantic similarity" Young-Rae Cho, Tak Chien Chiam, and Yanxin Lu
Short	B411 "Building a Classifier for Identifying Sentences Pertaining to Disease-Drug Relationships in Tardive Dyskinesia" Xia Bi, Hongzhan Huang, Sherri Matis-Mitchell, Peter McGarvey, Manabu Torii, Hagit Shatkay, and Cathy Wu
Short	B439 "Composition of Bioinformatics Model Federations using Communication Aspects" Keith Lee and David Stotts

Session 14

Regular	B314 "Integration of Multiple Annotators by Aggregating Experts and Filtering Novices" Ping Zhang and Zoran Obradovic
Regular	B332 "An Efficient Sequential Pattern Mining Algorithm for Motifs with Gap Constraints" Chiang-Chi Liao and Ming-Syan Chen
Regular	B339 "Parallel Simulation of Apoptotic Receptor-Clustering on GPGPU Many-Core Architectures" Claus Braun, Markus Daub, Alexander Schöll, Guido Schneider, and Hans-Joachim Wunderlich
Short	B371 "CTGR-Span: Efficient Mining of Cross-Timepoint Gene Regulation Sequential Patterns from Microarray Datasets" Chun-Pei Cheng, Yi-Lin Tsai, and Vincent S. Tseng
Short	B496 "Placement of Unique Restriction Sites in Synthetic Genomes using Multi-Objective Optimization" Mahfuza Sharmin and Sohel Rahman

Session 15

Regular	B388 "Identifying Protein Complexes Based on Local Fitness Method" Jun Ren, Jianxin Wang, and Min Li
Regular	B392 "iSimp: A Sentence Simplification System for Biomedical Text" Yifan Peng, Catalina O. Tudor, Manabu Torii, Cathy H. Wu, and K. Vijay-Shanker
Short	B311 "Linking and Querying Genomic Datasets Using Natural Language" Bobby McKnight and Budak Arpinar
Short	B362 "Personalized Semantic Assistance for the Curation of Biochemical Literature" Fedor Bakalov, Marie-Jean Meurs, Birgitta König-Ries, Bahar Sateli, René Witte, Greg Butler, and Adrian Tsang

Session 16

Regular	B422 "An Accurate Scalable Template-based Alignment Algorithm" David Gardner, Weijia Xu, Jamie Cannone, Daniel Miranker, Stuart Ozer, and Robin Gutell
Regular	B430 "PhylOnt: A Domain-Specific Ontology for Phylogeny Analysis" Maryam Panahiazar, Ajith Ranabahu, Vahid Taslimitehrani, Hima Yalamanchili, Arlin Stoltzfus, Jim Leebens_Mack,

	and Amit Sheth
Regular	B447 "A Novel Dynamic Graph-Based Computational Model for Predicting Salivary Gland Branching Morphogenesis" Nimit Dhulekar, Lauren Bange, Abhirami Baskaran, Daniel Yuan, Basak Oztan, Bulent Yener, Shayoni Ray, and Melinda Larsen
Short	B290 "Incorporating Semantic Similarity into Clustering Process for Identifying Protein Complexes from Affinity Purification/Mass Spectrometry Data" Bingjing Cai, Haiying Wang, Huiru Zheng, and Hui Wang
Short	B460 "A Neural Network Approach to the Identification of b-/y-ions in MS/MS Spectra" James Cleveland and John Rose

Session 17

Regular	B471 "Labor Contraction Prediction via Demographic and Obstetrical Information Analysis" Zifang Huang, Mei-Ling Shyu, James Tien, David Birnbach, and Michael Vigoda
Regular	B236 "Predicting distant metastasis in breast cancer using ensemble classifier based on context-specific miRNA regulation modules" Xionghui Zhou, Juan Liu, and Jianghui Xiong
Short	B394 "Improved biomarker performance for the detection of hepatocellular carcinoma by inclusion of clinical parameters" Mengjun Wang, Timothy Block, Jorge Marrero, Adrian Di Bisceglie, Karthik Devarajan, and Anand Mehta
Short	B271 "Using microenvironments to identify allosteric binding sites" Christopher Foley, Sana AlAzwar, Mark Duffton, and John Wilson
Short	B275 "The Effect of Unhealthy β -cells in Synchronized Insulin Secretion" Yang Pu, Saangho Lee, David Samuels, Layne Watson, and Yang Cao

Session 18

Regular	B479 "CWT-PLSR for Quantitative Analysis of Raman Spectrum" Shuo Li, Jean Gao, James Nyagilo, and Digant Dave
Short	B304 "A New Method for Computational Drug Repositioning Using Drug Pairwise Similarity" Jiao Li and Zhiyong Lu
Short	B202 "Keyword Annotation of Biomedical Documents with Graph-based Similarity Methods" Shuguang Wang and Milos Hauskrecht
Short	B242 "Immune System Simulation: Modeling the Mast Cell" Charles da Silva, Alcione Oliveira, Maurilio Possi, Fabio Cerqueira, Andreia Gomes, Rodrigo

Session 19

Regular	B378 "Improving Health Records Search Using Multiple Query Expansion Collections" Dongqing Zhu and Ben Carterette
Regular	B276 "Epileptic EEG Signal Analysis and Identification Based On Nonlinear Features" Yuedong Song
Short	B483 "Predicting Viral Infection by Selecting Informative Biomarkers From Temporal High-Dimensional Gene Expression Data" Qiang Lou and Zoran Obradovic
Short	B270 "Automatic Analysis Method of Protein Expression images based on Generalized Data Field" Shuliang Wang, Ying Li, Wenchen Tu, and Peng Wang
Short	B489 "ForeTell - Facilitating doctor-patient conversation through interactive information visualization of risk prediction index" Pratibha Bhaskaran, Milind Kaduskar, Preeti Saluja, Smitha Tallimani, Sandip Bhaumik, and Sanghyun Yoo

Session 20

Regular	B284 "A Monte Carlo Approach to Biomedical Time Series Search" Jonathan Woodbridge, Bobak Mortazavi, Majid Sarrafzadeh, and Alex Bui
Regular	B466 "Comprehensive human membrane protein database" Min-sung Kim and Gwan-Su Yi
Regular	B303 "Early Classification of Multivariate Time Series Using a Hybrid HMM/SVM model" Mohamed Ghalwash, Dusan Ramljak, and Zoran Obradovic
Short	B251 "A Semi-Supervised Learning Method for Names of Traditional Chinese Prescriptions and Drugs Recognition" Dongfeng Cai, Changlin Ding, Junjun Zuo, and Yu Bai
Short	B395 "Extracting BI-RADS Features from Portuguese Clinical Texts" Houssam Nassif, Filipe Cunha, Inês Moreira, Ricardo Cruz-Correia, Eliana Sousa, David Page, Elizabeth Burnside, and Inês Dutra

Session 21

Regular	B319 "Network-based Inferring Drug-Disease Associations from Chemical, Genomic and Phenotype Data" Yu-Fen Huang, Hsiang-Yuan Yeh, and Von-Wun Soo
Regular	B320 "A Data Mining Approach for Optimization of Acute Inflammation Therapy" Vladan Radosavljevic, Kosta Ristovski, and Zoran Obradovic
Regular	B364 "Prediction of Human Immunodeficiency Virus Type 1 Drug Resistance: Representation of Target Sequence Mutational Patterns via an n-Grams Approach" Majid Masso
Short	B456 "A Bayesian-based Prediction Model for Personalized Medical Health Care" Jiashu Zhao, Jimmy Xiangji Huang, Xiaohua Hu, Joseph Kurian, and William Melek

Short	B491 "Automated Wound Identification System Based on Image Segmentation and Artificial Neural Networks" Bo Song and Ahmet Sacan
-------	--

Session 22

Regular	B233 "Robust Segmentation of Biomedical Figures Toward an Image-based Document Retrieval" Luis Lopez, Jingyi Yu, Catalina Tudor, Cecilia Arighi, Hongzhan Huang, K Vijay-Shanker, and Cathy Wu
Regular	B407 "Classification of Multicolor Fluorescence In-Situ Hybridization (M-FISH) Image Using Structure Based Sparse Representation Model" Jinyao Li and Yu-Ping Wang
Regular	B410 "Biomarker Identification for Diagnosis of Schizophrenia with Integrated Analysis of fMRI and SNPs" Hongbao Cao, Yu-Ping Wang, Dongdong Lin, Junbo Duan, and Vince Calhoun
Regular	B477 "Novel Features for Categorizing Biomedical Images" Jianqiang Sheng, Songhua Xu, Weicai Deng, and Xiaonan Luo
Short	B351 "An Automated Blood Vessel Extraction Algorithm in Fundus Images" Marwan Saleh and C. Eswaran

Session 23

Regular	B446 "A Multi-Objective Program for Quantitative Subtyping of Clinically Relevant Phenotypes" Jiangwen Sun, Jinbo Bi, and Henry Kranzler
Short	B476 "Discovering Distal Regulatory Elements by Integrating Multiple Types of Chromatin State Maps" Li Teng and Kai Tan
Short	B342 "Systematic Measurement of Mismatch Effect for Designing Inter-Species Microarray" Fukuzaki Mutsumi, Masa-aki Yoshida, Atsushi Ogura, and Jun Sese
Short	B352 "Finding Genomic Features from Enriched Regions in ChIP-Seq Data" Iman Rezaeian and Luis Rueda
Short	B358 "Reconstructing Isoform Graphs from RNA-Seq data" Stefano Beretta, Paola Bonizzoni, Gianluca Della Vedova, and Raffaella Rizzi

Session 24

Short	B437 "Maps, Rates, and Fuzzy Mountains: Generating Meaningful Risk Maps" Tamara Jimenez, Chetan Tiwari, Armin Mikler, and Marty O'Neill II
Short	B393 "The Human Imprintome v1.0: Over 120 Imprinted Genes in the Human Genome Impose a Major Review on Previous Censuses" Samara Silva-Santiago, Ana Carolina Landim-Pacheco, Elton José Rosas-Vasconcelos, Samyra Maria Vieira-Brasil, Tereza Cristina L Rocha, Mônica Mor
Short	B427 "De Novo Co-Assembly of Bacterial Genomes from Multiple Single Cells" Narjes Movahedi, Elmirasadat Forouzmand, and Hamidreza Chitsaz
Short	B409 "A Model to Predict and Analyze Protein-protein Interaction Types Using Electrostatic Energies" Gokul Vasudev and Luis Rueda
Short	B282 "Recognizing Drosha processing sites by a two-step prediction model with structure and sequence information" Xingchi Hu, Chuang Ma, and Yanhong Zhou

Workshops

1. Workshop on Multi-Level and Multi-Scale Modeling, Analysis, and Simulation in Biology (Merged with Workshop 6)

2. The 2012 Computational Structural Bioinformatics Workshop

8:30 – 10:10 Session 1 Session Chair: Jing He
Ibrahim Al-Blawi, Marc Vaisset, Thierry Siméon, and Juan Cortés, <i>Coarse-grained elastic networks, normal mode analysis and robotics-inspired methods for modeling protein conformational transitions</i>
Bahar Akbal-Delibas and Nurit Haspel, <i>Refining Multimeric Protein Complexes Using Conservation, Electrostatics and Probabilistic Selection</i>
Thomas Evangelidis, Li Xie, Philip Bourne, and Lei Xie, <i>An integrated workflow for proteome-wide off-target identification and polypharmacology drug design</i>
Jim Havrilla and Ahmet Saçan, <i>Meta-analysis of Protein Structural Alignment (short)</i>
Lauro Galvão, Luiz Nunes, Pablo Moscato, and Heitor Lopes, <i>A New Greedy Heuristic for 3DHP Protein Structure Prediction With Side Chain (short)</i>
10:10-10:20 Coffee Break, poster setup
10:20-11:50 Session 2 Session Chair: Brian Chen Session Co-chair: Brian Godshall
Ankur Dhanik, John McMurray, and Lydia Kavrakı, <i>AutoDock-based incremental docking protocol to improve docking of large ligands</i>
Andrew McKnight, Kamal Al Nasr, Dong Si, Andrey Chernikov, Nikos chrisochoides, and Jing He, <i>CryoEM Skeleton Length Estimation using a Decimate Curve</i>
Majid Masso, <i>Knowledge-Based Scoring Function Derived from Atomic Tessellation of Macromolecular Structures for Prediction of Protein-Ligand Binding AffinityX</i>
Masood Zamani and Stefan C. Kremer, <i>Protein Secondary Structure Prediction Using Support Vector Machines and a Codon Encoding Scheme (short)</i>
11:40-12:30 poster session 1, 12:30-1:30 Lunch (on your own)
1:30-3:30 Session 3 Session Chair: Amarda Shehu Session Co-chair: Irina Hashmi
Keynote Speaker: Roland Dunbrack, Fox Chase Cancer Center "Structural bioinformatics of proteins and protein complexes."

Yajia Zhang and Kris Hauser, <i>Unbiased, Scalable Sampling of Constrained Kinematic Loops</i>
Brian Godshall and Brian Chen, <i>Improving Accuracy in Binding Site Comparison with Homology Modeling</i>
Sameh Saleh, Brian Olson, and Amarda Shehu, <i>A Population-based Evolutionary Algorithm for Sampling Minima in the Protein Energy Surface</i>
3:30-3:50 poster session (coffee: 3:30-3:50)
3:50-5:20 Session 4 Session Chair: Nurit Haspel Session Co-chair: Bahar Akbal-Delibas.
Boyu Zhang, Daniel Yehdego, Kyle Johnson, Ming-Ying Leung, and Michela Taufer, <i>A modularized MapReduce framework to support RNA secondary structure prediction and analysis workflows</i>
Kevin Molloy and Amarda Shehu, <i>A Robotics-inspired Method to Sample Conformational Paths Connecting Known Functionally-relevant Structures in Protein Systems</i>
Weiwei Han and Lei Xie, <i>Structural Basis of Polypharmacological Effects of Metformin (short)</i>
Bahar Akbal-Delibas, Filip Jagodzinski, and Nurit Haspel, <i>Towards a Hybrid Method for Detecting Critical Protein Residues (short)</i>

3. Pharmaco-Informatics for Drug Discovery

4. The Fifth International Workshop on Biomolecular Network Analysis (IBNA)

Session 1 (8:30am – 10:15am)

8:30am - 8:35am Opening remarks.

8:35am – 9:00am *Discovering Maximal Cohesive Subgraphs and Patterns from Attributed Biological Networks.*

Saeed Salem, Rami Alroobi, Syed Ahmed, and Mohammad Hossain.

9:00am – 9:25am *Relevance Judgment Algorithm for Detecting Protein Complexes from Protein Interaction Networks.*

Huaxiong Yao, Xiaolong Li, and Yan Yang.

9:25am – 9:50am *Mining hub-based protein complexes in massive biological networks.*

Lin Zhijie.

9:50am – 10:15am *Improvements of graph entropy approach to detect protein complexes by ontological analysis of PPIs.*

Frank Pena and Young-Rae Cho.

Break (10:15am – 10:25am)

Session 2 (10:25am – 12:10pm)

10:25am – 10:50am *Predator prey modeling at the molecular level: Analytical considerations about the coupling of biochemical events during photo-chemotherapy and cell survival probability.*

Ioannis Gkigkitzis.

10:50am – 11:15am *A Computational Model for Genetic and Epigenetic Signals in Colon Cancer.*

Irina-Afrodita Roznovat and Heather J. Ruskin.

11:15am – 11:40am *Effective Parameter Optimization for S-system Model-based Genetic Networks Using Hybrid Soft Computing.*

Yuan-Ming Yeh and Wei-Chang Yeh.

11:40am – 12:05pm *Align-MCL: Comparative Analysis of Protein Interaction Networks through Markov Clustering.*

Marco Mina and Pietro Hiram Guzzi

5. The First International Workshop on the role of Semantic Web in Literature-Based Discovery

08:15 – 10:00 Session 1 Session Chair: Delroy Cameron	
08:15-09:00	Invited Talk: Trevor Cohen <i>Semantic Predications in Vector Space: New Representations for Literature-based Discovery</i>
09:00-09:30	J. Caleb Goodwin, Trevor Cohen and Thomas Rindflesch <i>Discovery by scent: Closed literature-based discovery system based on the Information Foraging Theory</i>
09:30-10:00	Sanmitra Bhattacharya and Padmini Srinivasan <i>A Semantic Approach to involve Twitter in LBD Efforts</i>
10:00-10:20 Coffee Break, poster setup	
10:20-12:30 Session 2 Session Chair: Delroy Cameron	
10:20-10:50	Trevor Cohen, Dominic Widdows, Roger Schvaneveldt and Thomas Rindflesch <i>Discovery at a Distance: Farther Journeys in Predication Space</i>
10:50-11:20	Lubomir Stanchev <i>Building Semantic Corpus from WordNet</i>
11:20-11:50	Delroy Cameron, Varun Bhagwan and Amit P. Sheth <i>Towards Comprehensive Longitudinal Healthcare Data Capture</i>
Closing Remarks	

6. International Workshop on Multiscale Biomedical Imaging Analysis(MBIA)

8:30 – 10:10 Session 1	
Keynote Speak: <i>Moving towards Omni-tomography</i> by Ge Wang, Samuel Reynolds Pritchard Professor and Director of Biomedical Imaging Division, Virginia Tech	
Murat Seckin Ayhan, Ryan Benton, Vijay Raghavan, and Suresh Choubey, <i>Utilization of Domain-Knowledge for Simplicity and Comprehensibility in Predictive Modeling of Alzheimer's Disease</i>	
Chen Geng, Jian Yang, Tong Li, and Yongtian Wang, <i>A Novel Non-contact Interactive Medical Image Viewing System</i>	
10:30-12:30 Session 2	
Dongdong Lin, Jigang Zhang, Jingyao Li, Vince Calhoun, and Yu-Ping Wang, <i>Correspondence between fMRI and SNP Data by Group Sparse Canonical Correlation Analysis</i>	
Mohammad Majid al-Rifaie, Ahmed Aber, and Ahmed Majid Oudah, <i>Utilizing Stochastic Diffusion Search to Identify Metastasis in Bone Scans and Microcalcifications on Mammographs</i>	
Saurabh Shirgaonkar, Dong Hyun Jeong, Toan Huynh, and Soo-Yeon Ji, <i>Designing a Robust Bleeding Detection Method for Brain</i>	

<i>CT Image Analysis</i>
Zhen Liu and Yang Cao, <i>Approximation of a Multi-Scale Model Based on Multiple Site Phosphorylation</i>
Yongzhi Ong and Winfried Kurth, <i>A Graph Model and Grammar for Multi-scale Modeling using XL</i>
Ratchadaporn Kanawong, Tayo Obafemi-Ajayi, Jun Yu, Dong Xu, Shao Li, and Ye Duan, <i>Classification in Traditional Chinese Medicine based on Modified Specular-free Tongue Images</i>

7. The Third International Workshop on Information Technology for Chinese Medicine (ITCM2012)

13:30 – 13:40 Opening Remarks
13:40-15:25 Session 1 (Invited Talk) Session Chair: Dong Xu/ Zhaohui, Liang
Ziping Li, <i>Distal Point Acupuncture for Cervical Spondylosis with Radiculopathy Based on Flow of Meridians</i>
Stefan Jaeger, <i>An Information-Theoretic Neural Model Based on Concepts in Chinese Medicine</i>
Zhaohui Liang, <i>Application of machine learning in Chinese medicine: current & future studies</i>
15:25 – 15:50 Coffee Break
15:50-17:30 Session 2 (Oral Presentations) Session Chair: Dong Xu/ Zhaohui, Liang
Lei Zhang, Wei Mao, Yin Li, Guobin Su, and Xusheng Liu, <i>Gifts from Chinese medicine for diabetic nephropathy: ancient formulas in modern times</i>
Yuan-fang Chen, Yu Chen, Tao Yu, Zhaohui Liang, and Song Wei, <i>Theory and Practice: a Discourse-based Translation Strategy of Chinese Medicine</i>
Gang Zhang, Xiangyang Shu, Zhaohui Liang, Yunting Liang, Shuyi Chen, Jian Yin, <i>Multi-Instance Learning for Skin Biopsy Image Features Recognition</i>
Yuntao Liu, Qing Liu, Yi Luo, Danwen Zheng, Qingping Deng, Zhanpeng Tan, Xiaoyang Jin, Wei Huang, Yaling Lei, and Jian Yin, <i>Application and Improvement Discussion about Apriori Algorithm of Association Rules Mining in Cases Mining of Influenza Treated by Contemporary Famous Old Chinese Medicine</i>
Da Guo, Jian Li, Jun Liu, Gang Zhang, Weixiang Lu, and Shaojian Xu, <i>Research on Optimal Traditional Chinese Medicine Treatment of Knee Osteoarthritis with Data Mining Algorithms</i>
17:30-17:45 Closing Remarks

8. The Third Integrative Data Analysis in Systems Biology (IDASB 2012)

Session 1: Pathway and Network Applications (Chair: Zhongming Zhao)

9.10 - 9.35	Using Expression Data to help Pathway Curation	Keith Decker, Preeti Anday, Liang Sun, and Carl Schmidt
9.35 - 10.00	Identification and Analysis of Gene Clusters in Biological Data	Liliana Marcela Olarte Mesa, Liliana López-Kleine, and Luis Fernando Niño Vásquez
10.00 -10.20	Coffee break	
10.20 - 10.45	Pathway Curation: Application of Text-Mining Tools eGIFT and RLIMS-P	Carl Schmidt, Liang Sun, Cecilia Arighi, Keith Decker, K. Vijay-Shanker, Manabu Torii, Catalina O. Tudor, Cathy Wu, and Peter D'Eustachio
10.45 - 11.10	A Network and Functional Investigation of Illicit Drugs and their Targets	Ravi Atreya, Jingchun Sun, and Zhongming Zhao
11.10- 11. 35	Connecting Clusters of Patient to Drug Responses of Cell Lines to Suggest Personalized Therapeutics for Breast Cancer	Simon Gordonov and Avi Ma'ayan
11.45-13.30		

Session 2: Integrative Systems Biology Approaches (Chair: Jane Zheng)

13.30 - 13.55	Enriching miRNA binding site specificity with sequence profile based filtering of 3'-UTR region of mRNA	Jasjit Banwait, Hesham Ali, and Dhundy Bastola
13.55 - 14.20	Constraint-driven optimization of the plant defense model parameters	Dragana Miljkovic, Matjaž Depolli, Tjaša Stare, Igor Mozetič, Marko Petek, Kristina Gruden, and Nada Lavrač
14.20 - 14.45	Spatial Index and Visualization of Protein Structure	Hui Li and Chunmei Liu
14.45 - 15.10	Predicting Protein-Protein Interactions Using Full Bayesian	Hui Li and Chunmei Liu

9. International Workshop on Computational Proteomics

9:50 – 12:00 Session 1
Opening Remarks
Fahad Saeed, Trairak Pisitkun, Jason Hoffert, Guanghui Wang, Marjan Gucek, and Mark Knepper, "An Efficient Dynamic Programming Algorithm for Phosphorylation Site Assignment of Large-Scale Mass Spectrometry Data"
Yanhua Wang, Zhihao Yang, Hongfei Lin, and Yanpeng Li, "A Syntactic Rule-based Method for Automatic Pathway Information Extraction from Biomedical Literature"
Wei Xiong, Hui Liu, Jihong Guan, and Shuigeng Zhou, "Protein Function Prediction by Collective Classification with Explicit and Implicit Edges in Protein-protein Interaction Networks"
Soyoung Ryu, David Goodlett, William Noble, and Vladimir Minin, "A statistical approach to peptide identification from clustered tandem mass spectrometry data"
Gang Chen and Jianxin Wang, "Identifying functional modules in tissue specific protein interaction networks"
Xijun Liang, Zhonghang Xia, Xinnan Niu, Andrew Link, Liping Pang, Fang-Xiang Wu, and Hongwei Zhang, "A Fuzzy Cluster-based Algorithm for Peptide Identification"
12:00-13:15 Lunch (Provided by the conference)
15:20-17:30 Session 2

Andrea Szaboova, Ondrej Kuzelka, and Filip Zelezny, "Prediction of Antimicrobial Activity of Peptides using Relational Machine Learning"
Murtada Elfaki, Jianxin Wang, and Fang-Xiang Wu, "A Hybrid Approach of Support Vector Machines with Logistic Regression for Beta-turn Prediction"
Jian Wang, Qian Xu, Hongfei Lin, Zhihao Yang, and Yanpeng Li, "Combining Labeled and Unlabeled Data For Biomedical Event Extraction"
Bo Xu, Kavishwar B. Waghlikar, Hongfei Lin, Zhihao Yang, and Hongfang Liu, "Classifying Protein Complexes from Candidate Subgraphs using Fuzzy Machine Learning Model"
Mohammad R Nezami Ranjbar, Yi Zhao, Mahlet Tadesse, Yue Wang, and Habtom Resson, "Evaluation of Normalization Methods for Analysis of LC-MS Data"
Closing Remarks

10. Data-mining of Next-Generation Sequencing

8:30 – 10:00 Session 1
Nikolaos Alachiotis, Simon Berger, Tomas Flouri, Solon Pissis, and Alexandros Stamatakis, " <i>libgapmis: an ultrafast library for short-read single-gap alignment</i> "
Shuai Yuan and Zhaohui Qin, " <i>Read-mapping using personalized diploid reference genome for RNA sequencing data reduced bias for detecting allele-specific expression</i> "
Anurag Nagar and Michael Hahsler, " <i>A Novel Quasi-Alignment-Based Method for Discovering Conserved Regions in Genetic Sequences</i> "
10:00-10:20 Coffee Break, poster setup
10:20-11:50 Session 2
Oliver Bonham-Carter, Hesham Ali, and Dhundy Bastola, " <i>A Meta-genome Sequencing and Assembly Preprocessing Algorithm Inspired by Restriction Site Base Composition</i> "
Julia Warnke and Hesham Ali, " <i>An Efficient Overlap Graph Coarsening Approach for Modeling Short Reads</i> "
Mohammad Ashraf, Girija Chetty, Dharmendra Sharma, and Dat Tran, " <i>Feature Selection Techniques on Thyroid, Hepatitis, and Breast Cancer Datasets</i> "
11:50-12:30 poster session 1
1:30-2:50 Session 3
Po-Yen Wu, John Phan, and May Wang, " <i>The Effect of Human Genome Annotation Complexity on RNA-Seq Gene Expression Quantification</i> "
Zhe Zhang, Jeremy Leipzig, Ariella Sasson, Juan Carlos Perin, Michael Xie, Mahdi Sarmady, Patrick Warren, and Peter White, " <i>Bamchop: A bioinformatics utility to summarize and visualize exome and other types of targeted resequencing data</i> "
Avindra Fernando, Jun Huan, Justin Blumenstiel, Jin Lin, Xue-wen Chen, and Bo Luo, " <i>Identification of Transposable Elements of the Giant Panda (<i>Ailuropoda Melanoleuca</i>) Genome</i> "
3:00-3:50 poster session (coffee: 3:30-3:50)
3:50-5:00 Session 4
Maristela Terto Holanda, Renato de Paula, Sérgio Lifschitz, and Maria Emilia M. T. Walter, " <i>Managing Data Provenance in Genome Project Workflows</i> "
Ali Alatabbi, Carl Barton, and Costas Iliopoulos, " <i>On the Repetitive Collection Indexing Problem</i> "

11. The 2012 International Workshop on Biomedical and Health Informatics (BHI 2012)

10/4 (Thursday): Room Co-B

10:00 am	Welcome Illhoi Yoo (Workshop Chair)
10:10	José Ferrão, Mónica Oliveira, Filipe Janela, and Henrique Martins, Clinical Coding Support Based on Structured Data Stored in Electronic Health Records
10:40	Sharmili Roy, Michael Brown, and George Shih, Extracting Volumetric Information from Standard Two-Dimensional Radiological Annotations within the Clinical Workflow
11:10	[S] Tripti Swarnkar and Pabitra Mitra, Graph Based Unsupervised Feature Selection for Microarray Data
11:30	Min Song, Chris Yang, Xuning Tang, and Wook-shin Han, Mapping the Field of Bioinformatics with a Content and Co-authorship Analysis
1:30 pm	Ishwor Thapa, Sanjukta Bhowmick, and Dhundy Bastola, A Comparison between Hierarchical Clustering and Community Detection Method in the Collection of Gene Targets for Molecular Identification of Pathogenic Fungi
2:00	[S] Dengyuan Wu, Li Sheng, Eric Xu, Kai Xing, and Dechang Chen, Analysis of an Ensemble Algorithm for Clustering Cancer Data
2:20	Vivek Datla, King-Ip Lin, and Max Louwerse, Capturing disease-symptom relations using higher-order co-occurrence algorithms
2:50	[S] Soubhik Paul and Sandip Bhaumik, A Context and Knowledge based Health Event Storage for Efficient Clinical Information Retrieval
3:10	Break
3:30	[S] Kyungtae Kang, Min-Young Nam, Jaemyoun Lee, Juyoung Park, Homin Yoo, and Lui Sha, Model-Based Design of a Wireless Telemetry System and QoS Assessment using AADL
3:50	Shuihua Wang and Yingli Tian, Detecting Stairs and Pedestrian Crosswalks for the Blind by RGBD Camera
4:20	Chenyang Zhang, Mohsin Shabbir, Yingli Tian, and Despina Stylianou, Computer Vision-based Mathematics Learning Enhancement for Children with Visual Impairments
4:50	[S] Francisco Rocha Pirolla, Marilde T. P. Santos, Joaquim C. Felipe, and Marcela Xavier Ribeiro, Dimensionality Reduction to Improve Content-based Image Retrieval - A Clustering Approach

10/5 (Friday): Room Wa-B

2:20 pm	Alberto De la Rosa Algarin, Steve Demurjian, Solomon Berhe, and Jaime Pavlich-Mariscal, A Security Framework for XML Schemas and Documents for Healthcare
2:50	[S] Kotoe Katayama, Rui Yamaguchi, Seiya Imoto, Keiko Matsuura, Kenji Watanabe, and Satoru Miyano, Analysis of questionnaire for Traditional Medical and develop decision support system
3:10	Timoteus Ziminski, Alberto De la Rosa Algarin, Rishi Saripalle, Steve Demurjian, and Eric Jackson, SMARTSync: Towards Patient-Driven Medication Reconciliation
3:40	Michael Blechner, Rishi Kanth Saripalle, and Steve Demurjian, A Proposed Star Schema and Extraction Process to Enhance the Collection of Contextual & Semantic Information for Clinical Research Data Warehouses
4:30	Jingshan Huang, Jiangbo Dang, Xingyu Lu, Dejing Dou, Judith Blake, William Gerthoffer, and Ming Tan, An Ontology-Based MicroRNA Knowledge Sharing and Acquisition Framework
5:00	[S] Pablo Vasquez and Bengt Mandersson, Frequency tracking of resonant-like sounds from audio recordings of Arterio-Venous Fistula Stenosis
5:30	Ehsan Nasiri Khoozani, Jamshaid Ashraf, and George Fink, Stress Knowledge Base: An Ontology-Based Information Repository for the Stress Domain
6:00	[S] Patrizia Vizza, Giuseppe Tradigo, Antonio Curcio, Ciro Indolfi, and Pierangelo Veltri, Intracavitary Signal Analysis for Atrial Fibrillation Prediction

12. The Third Workshop on Data mining from genomic variants and its application to genome-wide analysis

[Session 1] Chair: Prof. Taesung Park, Workshop Chair

08:15-08:55 Invited Talk, Youngchul Kim, "Personalized chemotherapy for ovarian cancer by integrating genomic data with clinical data "

08:55-09:20 SC206 Seo-Jin Bang, Yong-Gang Kim, and Taesung Park, "Joint Selection of SNPs for Improving Prediction in Genome-wide Association Studies"

09:20-09:45 SC205 Jungsoo Gim and Taesung Park, "Identifying differential expression for RNA-seq data"

09:45-10:10 SC204 Nagabhushan S. N., TaeJin Ahn, Srikanth M. R., Taesung Park, Ajit Bopardikar, and Rangavittal Narayanan, "A Data Aggregation Framework for Cancer Subtype Discovery"

10:10-10:30: Break Time

[Session 2] Chair: Prof. Youngchul Kim

10:30-11:10 Invited Talk, Chris Yoon, "Probe-wise analysis of copy number variation from population data"

11:10-11:35 SC202 Amin Assareh, L. Gwenn Volkert, and Jing Li, "Feature Selections Using AdaBoost: Application in Gene-Gene Interaction Detection"

11:35-12:00 SC203 Kyounga Kim, Min-Seok Kwon, Sung Young Lee, Junghyun Namkung, Ming D. Li, and Taesung Park, "GxG-Viztool: a program for visualizing gene-gene interactions in genetic association analysis"

12:00-12:25 B355 Xijia Fei and Dechang Xu, "Integrated analysis of hela cell microRNAs' microarray data and anti-cancer pathway prediction of β -sitosterol"

13. Workshop on Nanoinformatics for Biomedicine

8:35-9:00	Michele Ostraat, Karmann Mills, and Kimberly Guzan , <i>The Nanomaterial Registry: Opportunities and Challenges in Informatics,</i>
9:00-9:25	James Luo , <i>Nanomaterial Registry: A resource for biological and environmental interactions of nanomaterials,</i>
9:25-9:50	Dennis G Thomas, Satish Chikkagoudar, Alan R Chappell, and Nathan A Baker, <i>Annotating the structure and components of a nanoparticle formulation using computable string expressions</i>
9:50-10:15	Kaizhi Tang, Xiong Liu, Stacey Harper, Jeffery Steevens, and Roger Xu , <i>NEIMiner: A Model Driven Data Mining System for Studying Environmental Impact of Nanomaterials</i>
10:15-10:40	Xiong Liu, Kaizhi Tang, Stacey Harper, Bryan Harper, Jeffery Steevens, and Roger Xu, <i>Predictive Modeling of Nanomaterial Biological Effects</i>
10:45-11:15	Christie Sayes and Ivan Ivanov , <i>Grouping of Colloidal Metal Nanoparticles Based on Their Measurable Characteristics: A Proposed Framework</i>
11:15-11:40	Hyunsoo Kim and Markus Bredel , <i>Predicting Survival by Cancer Pathway Gene Expression Profiles in the TCGA</i>
11:40-12:05	Sushmita Paul and Pradipta Maji , <i>Rough Sets and Support Vector Machine for Selecting Differentially Expressed miRNAs</i>
12:05-12:30	Stanley Saamoah Moffatt, Richard Cristiano, and Rose Boyle , <i>Combined Formulation of Doxorubicin-Arg-Gly-Asp (RGD) and Modified PEGylated PLGA-encapsulated Nanocarrier improves Anti-tumor Activity</i>

14. Bioinformatics Strategies for Integrative Cancer Genomics (Merged with Workshop 12)

Industrial Track

Industrial Session 1:

Discovering breast cancer prognostic biomarkers using a novel feature selection tool, *Jie Cheng, Glaxo SmithKline*

A high-throughput analysis pipeline for large next generation DNA sequencing studies, *Zayed Albertyn and Jörg Hakenberg, Roche*

In Silico Target Portal: An Integrated Oncology Target Discovery Web Portal, *Ying Li, Roche*

Approach to Information Management in an Externalized Business Environment, *Bob O'Hara, ResultWorks*

Industrial Session 2:

Clinical Case: Enhancing medical monitoring with visualization and analytics, *Michael Farnum, Johnson and Johnson*

Click-ON: Methodologies and infrastructure supporting personalized medicine, *Sabrina Molinaro, Institute for Clinical Physiology, National Research Council of Italy*

Systematic Drug Repositioning: A new paradigm in Drug Discovery, *Vinod Kumar, Glaxo SmithKline*

New approaches for better decisions in Pharma R&D, *Anastasia Christianson, AstraZeneca*

Data-Driven Methods for Leveraging Electronic Health Records in support of Clinical Trials: Protocol Optimization, Site Selection and Patient Recruitment, *Aaron Kamauu, Anolinx LLC*

Industrial Workshop 1:

Title: Hindsight, Insight and Foresight - How to Optimize Industry Research by Measuring and Predicting Change Effects in Bioinformatics Applications

Workshop Organizers:

- David Fenske, Dean, iSchool, Drexel University
- Michael N. Liebman, PhD (Managing Director, Strategic Medicine, Inc)

Background:

This session will include perspectives from industry on collaborative efforts to bring next generation visual and decision support tools and techniques to enable decision makers in the bioinformatics industry to improve fundamentally the way an organization's information is analyzed, interpreted and used.

Objectives:

- Dean David Fenske and Dr Chaomei Chen will outline some of the ground-breaking research of the National Science Foundation Center for Visual and Decision Informatics, awarded this year to the *iSchool* at Drexel and the University of Louisiana at Lafayette. Industry and academic perspectives in this effort are integrated to probe research applications and deliver novel approaches to optimize industry research while cutting time and costs. Dr Fenske will describe examples of the Center's projects
- Dirk Husselman and Dr. Liebman will present examples of solutions into complex problem-solving from the user perspective back to the technology application

Program Outline:

1. National Science Foundation Center for Visual and Decision Informatics , an integration of industrial and academic perspectives on research needs(David Fenske, Dean, iSchool, Drexel University)
2. Tracking and modeling exploratory searches in a structure-function space, with application to drug discovery and compound optimization (Chaomei Chen, Professor, iSchool, Drexel University)
3. Risk, Challenges and Opportunities in Healthcare using Cloud-based solutions(Dirk Husselman, President, 2lemetry, Inc)
4. Optimizing Clinical Trial Success by Evaluating and Refining the Hypothesis (Michael Liebman, Managing Director, IPQ Analytics, LLC)
5. Panel discussion with a question and answer session.

Industrial Workshop 2:

Title: Data-mining: Issues and Challenges in Commercial Applications

Workshop Organizers:

- Sabrina Molinaro, PhD (Head, Division of Epidemiology and Health Research, National Research Council of Italy, Institute for Clinical Physiology, Pisa)
- Michael N. Liebman, PhD (Managing Director, Strategic Medicine, Inc)

Background:

Biomedicine is confronted with the challenges and opportunities presented by the evolving access to large data sets and the need to identify and create value from their contents. This workshop will focus on the development and application of methods for data-mining to resources ranging from genomic, proteomic and molecular data to clinical observations and data.

Objectives:

- To empower researchers with an understanding of the challenges and issues around data-mining in large, complex data repositories, virtual or integrated
- To familiarize attendees with available methods and tools.

Description:

The focus of the workshop will be to introduce issues and challenges in data-mining from virtual and/or integrated heterogeneous data and describe methodologies that are being used to develop value for commercial clients.

Program Outline:

1. Next generation in Technology for Advances in Biological data Analysis (Sultan Meghji, EVP, Appistry, Inc)
2. Fostering Computational Reasoning for Big Data by Non-mathematicians (Peter Schad, CSO, Digital Infuzion, Inc)
3. Applying Social Network Analysis for Data-mining in Complex Medical Data (Sabrina Molinaro,Head, Epidemiology and Health Research, ICP, CNR)

4. Network analysis of high throughput data for prediction of disease-related genes and identification of drug targets (Marina Bessarabova, Director Computational Biology & Bioinformatics ,Thomson Reuters)
5. (Ilya Mazo, Elsevier)
6. Panel discussion with a question and answer session.

Industrial Workshop 3:

Title: An Introduction to Natural Language Processing (NLP) Methods in Clinical Research

Workshop Presenters:

- Aaron Kamauu MD, MS, MPH (President, Anolinx LLC)
- Scott DuVall PhD (Associate Director of VINCI, VA Salt Lake City Health Care System)

Background:

As the use of natural language processing methods in preparing data for research continues to increase, researchers should understand the benefits and limitations of such a tool. While NLP is not a “solved” science, there are many tasks that NLP can do reliably. Extracting concepts (symptoms, diseases, medications) and values (lab values, vital signs) that are stored in the text is one example. More complex tasks, such as determining what caused an event of interest or why a patient discontinued a medication can also be addressed using the right tools. This workshop will introduce researchers to NLP and explore ways that NLP can support ongoing research.

Objectives:

- To empower researchers with knowledge of strengths and limitations of NLP methods.
- To familiarize attendees with available methods and tools.

Description:

The focus of the workshop will be to introduce the methodology of natural language processing (NLP) and demonstrate how it can be used in pharmacoepidemiological research.

Program Outline:

1. Explanation of what NLP is and discussion of the practical role it can play in pharmacoepidemiology (Aaron Kamauu)
2. Discussion of the different approaches for processing clinical text, including symbolic and statistical (Scott DuVall)
3. Practical demonstration of the interplay between clinical and technical teams to determine the relevant clinical concepts for training the NLP system (Aaron Kamauu and Scott DuVall)
4. Interactive walk-through of NLP tools on sample data (Scott DuVall and Aaron Kamauu)
5. Summary: Open discussion with a question and answer session.

Tutorials

Tutorial 1:

Title: Big Data in Drug Discovery: Opportunities and Challenges (4 hours)

Tutorial Organizer: Dr. Jun Huan, University of Kansas, USA
jhuan@ku.edu

Description:

A number of universities recently started drug discovery centers. Among the list include Duke, University of North Carolina, Emory, Vanderbilt, Broad Institute, John Hopkins and many others. NIH has maintained its intramural drug discovery programs for a number of years. Recently NIH has started ambitious extramural funding programs to support academic-based drug discovery programs (with an estimated funding of ~\$600M in the next a few years). With these new initiatives, huge volumes of data have been collected in an open and collaborative environment and such data are shared with public free of charge.

In this tutorial, the presenter will review big data analytics in the drug discovery process, with an emphasize on machine learning and data mining techniques, including but not limited to active learning, multitask learning, multiview learning, transfer learning, and multi-instance learning among others. The targeted audience of the tutorial is bioinformatics researchers and practitioners who are interested in developing and applying advanced learning techniques for the exciting application area of (broadly defined) drug discovery. No domain knowledge of drug discovery is required. General knowledge of statistics is assumed.

About the instructor:

Dr. Jun (Luke) Huan is an associate professor in the Electrical Engineering and Computer Science department at the University of Kansas. He directs the Bioinformatics and Computational Life Sciences Laboratory at the Information and Telecommunication Technology Center (ITTC) at KU. He also holds courtesy appointments at the KU Bioinformatics Center and KU Bioengineering Program. Dr. Huan received his Ph.D. in Computer Science from the University of North Carolina at Chapel Hill in 2006 (with Wei Wang, Jan Prins, and Alexander Tropsha). Before joining KU in 2006, he worked at the Argonne National Laboratory (with Ross Overbeek) and the GlaxoSmithKline plc. (with Nicolas Guex). Dr. Huan was a recipient of the National Science Foundation Faculty Early Career Development Award in 2009. He has published more than 70 peer reviewed papers in leading conferences and journals. His group won the best student paper award in ICDM'11 and the best paper award (runner-up) in CIKM'09. Dr. Huan served on the program committees of leading international conferences including ACM SIGKDD, IEEE ICDE, ACM CIKM, and IEEE ICDM.

Tutorial 2:

Title: Mobility data analysis to understand unknown diseases behavior: the case of facial paralysis (2 hours)

Tutorial Organizer: Dr. Jalel Akaichi, University of Sciences and Technologies of Lille, France
j.akaichi@gmail.com

Description:

Daily actions performed by or on health care actors leave obviously digital traces in modern health information systems. This is due to the frequent use of different kind of positioning technologies and mobile context aware devices in various health care activities. Through these digital traces, it is feasible to sense, make use of, and analyze mobility data resulting from health care applications related to specific diseases. Consequently, their potential value, for medical decision making, is high because of their increasing volume, pervasiveness and positioning accuracy.

In fact, the analysis of mobility data raises opportunities for discovering behavioral patterns to be exploited in applications like mobile hospitals, mobile physicians, facial paralysis, heart disease, etc. Online analytical processing and data mining techniques can be employed in order to convert the vast amount of raw data, resulting from digital traces, into useful knowledge. The tutorial aims to show how mobility data modeling and analysis help to understand unknown diseases behavior through the exploitation of digital traces. As illustration, it takes care of, in detail, the case of facial paralysis disease.

Tutorial Outline:

The proposed tutorial is composed of the followings parts:

- Health care pervasive systems part, in which we present concepts and applications examples.
- Mobility data modeling part, in which we describe a new conceptual approach.
- Mobility data analysis part, in which we present an extended multidimensional approach.
- Bell's palsy case study part, in which we show how the above approaches help to understand the disease behavior.

About the instructor:

Jalel Akaichi received his PhD in Computer Science from the University of Sciences and Technologies of Lille (France) and then his Habilitation degree from the University of Tunis (Tunisia) where he is currently an Associate Professor in the Computer Science Department. Jalel Akaichi has published in many international journals and conferences, and has served on the program committees of several international conferences and journals. He is currently the Chair of the Master Science in Business Intelligence, and the President of the Business Intelligence Association of Tunisia. Jalel Akaichi visited and taught in many institutions such as the State University of New York at Potsdam, Worcester Polytechnic Institute, National Institute of Applied Sciences in Lyon, University of Blaise Pascal at Clermont Ferrand, University of Sciences and Technologies of Lille, University of Toulouse 1 -Capitole, etc.

Tutorial 3:

Title: Introduction to biomedical vocabularies and ontologies (2 hours)

Tutorial Organizer: Dr. Illhoi Yoo, University of Missouri School of Medicine, USA
yooil@health.missouri.edu

Description:

Recently, researchers have realized that using domain knowledge in biomedical vocabularies and ontologies can significantly improve the performance and usefulness of the systems. At the same time, biomedical vocabularies and ontologies can considerably enhance the interoperability of diverse healthcare systems. They are the only way to overcome terminological variations in biomedicine/healthcare. For this reason, most information retrieval and data/text mining approaches adopt biomedical vocabularies and ontologies to improve their performance in terms of recall/precision/accuracy.

Tutorial Outline:

- Introduction to controlled vocabularies and ontologies
 - This section is basically the summary of ANSI/NISO's document (Z39.19-2005) titled Guidelines for the construction, Format, and Management of Monolingual Controlled Vocabularies
 - For each guideline and standard, many biomedical examples will be given especially from MeSH and UMLS.
- Medical Subject Heading (MeSH)
 - Introduction to MeSH
 - Internal structure of XML MeSH
 - PubMed searches with MeSH
- Unified Medical Language System (UMLS)
 - Introduction to UMLS
 - Metathesaurus (internal structure)
 - Semantic network

About the instructor:

Dr. Illhoi Yoo is an Association professor of health informatics at the University of Missouri School of Medicine and a core faculty member in University of Missouri Informatics Institute. He received his PhD in Information Science and Technology, Drexel University *iSchool*. He has used biomedical vocabularies and ontologies such as MeSH and UMLS in his research for around 8 years so that he can provide a lot of hands-on experience. He has published 40 peer-reviewed papers in leading conferences and journals. His papers have been cited nearly 200 times. The primer tutorial is based on his graduate course, HMI 8441 (Theory and Application of Controlled Terminologies, aka Biomedical and Health Ontologies and Applications). The student course evaluation scores of the course are 4.2 (2011) and 4.0 (2010).

Posters

- P205 Ioannis Gkigkitzis, Monotonicity Functional for a Transient Mathematical Model of Oxygen Depletion during Photodynamic Therapy
- P207 Sagar Patel, Hetalkumar Panchal, and Kalpesh Anjaria, Phylogenetic analysis of some Leguminous trees using CLUSTALW2 Bioinformatics Tool
- P208 Sagar Patel, Hetalkumar Panchal, and Kalpesh Anjaria, DNA Sequence analysis by ORF Finder & GENOMATIX Tool : Bioinformatics Analysis of some tree species of Leguminosae Family
- P210 Daniel Arend, Matthias Lange, Christian Colmsee, Steffen Flemming, Jinbo Chen, and Uwe Scholz, The e!DAL JAVA-API: Store, Share and Cite Primary Data in Life Sciences
- P212 WOONGHEE LEE and JOHN MARKLEY, PACSY database, a relational database management system for protein structure and nuclear magnetic resonance chemical shift analysis
- P214 Sameh Saleh, Brian Olson, and Amarda Shehu, An Evolutionary Framework to Sample Near-native Protein
- P215 Teng Liu, Yu Zhang, and Jonathan Dennis, Joint Clustering and Alignment for Nucleosome Occupancy Analysis
- P216 Yu Zhang, Xiuwen Liu, and Jonathan Dennis, Quantitative Models for Statistical Nucleosome Occupancy Prediction
- P218 Bahar Akbal-Delibas, Filip Jagodzinski, and Nurit Haspel, Combining Conservation and Rigidity Methods to Detect Critical Residues in Proteins
- P219 Bahar Akbal-Delibas and Nurit Haspel, An Evolutionary-guided Iterative Refinement Approach for Protein Multimers
- P220 Daniel Veltri and Amarda Shehu, Physico-chemical Features for Recognition of Antimicrobial Peptides
- P221 Xiuwen Liu, Yu Zhang, Teng Liu, and Jonathan Dennis, One-base-pair Localization of Nucleosomes
- P222 Brian Olson and Amarda Shehu, Jumping Low, jumping High: Controlling Hopping in the Protein Energy Surface
- P224 Irina Hashmi and Amarda Shehu, Sampling Low-energy Protein-protein Configurations with Basin Hopping
- P225 Zahra Hajihashemi and Mihail Popescu, Predicting Health Patterns Using Sensor Sequence Similarity and NLP
- P227 Aldrin Montana, Douglas Brandt, Bob Somers, Alex Dekhtyar, Chris Lupo, Michael Black, Anya Goodman, and Chris Kitts, Pyroprinting Sensitivity Analysis on the GPU
- P229 Byung-Cheol Kim and Sunghoon Kim, Assemblable Project Formulation for Drug Target Discovery
- P230 Ana Paula Bezerra, Samara Silva-Santiago, Elton José Vasconcelos, Ana Carolina Pacheco, Mônica Silva, and Diana Magalhães de Oliveira, In Silico Analyses of 24 Human Imprinted Genes Associated to Diabesin Triad (diabetes, obesity and inflammation)
- P232 Sai Gouravajhala, David Wang, Lunal Khuon, and Forrest Bao, EpSMART: Epileptic Seizure Monitoring with Alerts in Real Time
- P233 Sungchul Ji, Larry Cheng, Weronika Szafran, and Richard Carmona, Microarray Therapeutic Index: A novel method to combine molecular and clinical data from cancer patients.
- P235 zeehasham rasheed, Huzefa Rangwala, and Patrick Gillevet, Classification and Clustering in Metagenomics with Unified Data Management and Computational Framework
- P238 En-Shiun Annie Lee and Andrew K. C. Wong, Identifying Protein Binding Functionality of Protein Family Sequences by Aligned Pattern Clusters
- P239 Kevin Molloy and Amarda Shehu, Mapping Conformational Pathways Between Known Functional Protein States
- P240 Eric Juhn and John Springer, A Usability Study of the Komen Virtual Tissue Bank
- P241 Ananda Mondal and Jianjun Hu, Mutation Analysis of Disease Causing Proteins
- P242 Kevin Molloy and Amarda Shehu, A Tree-based Search to Bias Sampling of Protein Decoy Conformations