Large-scale integration of heterogeneous biological networks

Jian Peng
Department of Computer Science
University of Illinois at Urbana-Champaign
Heterogeneous data
Metabolic network

From wiki
Protein-protein interaction networks generated by String database
Heterogeneous association network
How to make sense of these networks?

"And that's why we need a computer."
Dimensionality reduction

High-dimensional space

Node vector space

Low-dimensional space

Learning

Prediction

Diffusion Component Analysis: Unraveling Functional Topology in Biological Networks
Hyunghoon Cho, Bonnie Berger and Jian Peng
RECOMB, 2015
Mashup: network integration

Diffusion state

Network 1

Network 2

Network K

Probability

Node

Observed $s_i$

Minimize Difference

Model $\hat{s}_i$

Global vector $\uparrow$

Network-specific vector $\uparrow$

Low-dimensional representation

$\hat{s}_{ij}^k \propto \exp\{x_i^T w_j^k\}$

$\dim(x_i) = d \ll n$

$\dim(s_i) = n$
Example: Gene function prediction

Human, GO BP, 31-100

- Integrated
- neighborhood
- fusion
- co-occurrence
- co-expression
- experimental
- database

AUPR

Individual networks
Example: Gene function prediction
A incomplete list of successful applications

- Drug target prediction
- Disease gene identification
- Module detection
- Gene essentiality prediction
- Interaction prediction
- Drug efficacy prediction
- Cancer patient stratification

......
Drug discovery through model organisms
TransposeNet: from yeast to human
Parkinson’s disease network
Mashup: Compact Integration of Multi-Network Topology for Functional Analysis of Genes
Hyunghoon Cho, Bonnie Berger, Jian Peng
Cell Systems, to appear

Genome-scale molecular networks link diverse neurodegenerative disease genes and processes to alpha-synuclein
Cell Systems, to appear

In situ proteome approaches connect alpha-synuclein directly to endocytic trafficking and mRNA metabolism.
Cell Systems, to appear

ProSNet: integrating homology with molecular networks for protein function prediction
Sheng Wang, Meng Qu, Jian Peng
PSB, 2017, to appear

Exploiting Ontology Graph for Predicting Sparsely Annotated Gene Function
ISMB/ECCB, 2015
Bioinformatics, 2015

Diffusion Component Analysis: Unraveling Functional Topology in Biological Networks
Hyunghoon Cho, Bonnie Berger and Jian Peng
RECOMB, 2015
Acknowledgements

UIUC:
Sheng Wang, Yunan Luo, Jiawei Han, Saurabh Sinha

MIT:
Hoon Cho, Chee Yeun Chung, Ernest Fraenkel, Susan Lindquist, Bonnie Berger

Harvard:
Vikram Khurana