Bioinformatics Center
- Biological Information Networks -

http://www.bic.kyoto-u.ac.jp/takutsu/index.html

Prof
AKUTSU, Tatsuya
(D Eng)

Assist Prof
HAYASHIDA, Morihiro
(D Inf)

Assist Prof
TAMURA, Takeyuki
(D Inf)

Program-Specific Res
KATO, Yuki
(D Eng)

Guest Res Assoc
SONG, Jiangning
(Ph D)

Students

MOURI, Kazunari (D3)
BROWN, John (D2)
POOLSAP, Unyanee (D1)
NARITA, Yuki (M1)

Visitor

TAN, Hao Monash University, Australia, 7 December 2007–1 March 2008

Scope of Research

Due to rapid progress of the genome projects, whole genome sequences of organisms ranging from bacteria to human have become available. In order to understand the meaning behind the genetic code, we have been developing algorithms and software tools for analyzing biological data based on advanced information technologies such as theory of algorithms, artificial intelligence, and machine learning. We are recently studying the following topics: systems biology, scale-free networks, protein structure prediction, inference of biological networks, chemo-informatics, discrete and stochastic methods for bioinformatics.

Research Activities (Year 2008)

Publications

Hayashida M, Tamura T, Akutsu T, Zhang SQ, Ching WK: Algorithms and Complexity Analyses for Control of Singleton Attractors in Boolean Networks, *EURASIP J Bioinform Syst Biol*, **2008**, [521407-1]-[521407-16] (2008).

Fujiwara H, Wang J, Zhao L, Nagamochi H, Akutsu T: Enumerating Treelike Chemical Graphs with Given Path Frequency, *J. Chem. Inf. Model*, **48**, 1345-1357 (2008).

Song J, Tan H, Takemoto K, Akutsu T: HSEpred: Predict Half-sphere Exposure from Protein Sequences, *Bioinformatics*, **24**, 1489-1497 (2008).

Presentations

An Improved Algorithm for Detecting a Singleton Attractor in a Boolean Network Consisting of AND/OR Nodes, Tamura T, Akutsu T, 3rd International Conference on Algebraic Biology, 31 July 2008.

Prediction of Protein Beta-Sheets: Dynamic Programming versus Grammatical Approach, Kato Y, Akutsu T, Seki H (Nara Institute of Science and Technology), 3rd IAPR International Conference on Pattern Recognition in Bioinformatics (PRIB2008), 15 October 2008.

On Distribution and Enumeration of Attractors in Probabilistic Boolean Networks, Hayashida M, Tamura T, Akutsu T, Ching WK (The University of Hong Kong), The 2nd International Symposium on Optimization and Systems Biology, 1 November 2008.

Grants

Akutsu T, Goto S, Mochizuki A, Tokita K, Mathematical Analysis of Structure and Dynamics of Biological Information Networks, Grant-in-Aid for Scientific Research on Priority Areas, 1 April 2005–31 March 2010.

Akutsu T, Kawabata T, Nagamochi H, Hayashida M, A Novel Approach to Computational Drug Design Based on Graph Theory and Kernel Methods, Grant-in-Aid for Scientific Research(A), 1 April 2007–31 March 2011.

Akutsu T, Data Compression Based Approach to Elucidation of Principles of Complex Biological Systems, Grant-in-Aid for Exploratory Research, 1 April 2007–31 March 2009.

Topics

Emergence of Scale-free Distribution in Protein-protein Interaction Networks based on Random Selection of Interacting Domain Pairs

Recent researches for biological and artificial networks have uncovered common network architecture, called scale-free topology. The origin of the scale-free topology has been explained using growth and preferential attachment mechanisms. In a cell, proteins are the most important carriers of functions, and are contain domains as elemental units responsible for the physical interaction between protein pairs.

We propose a model for protein-protein interaction networks that reveals the emergence of two possible topologies. We show that depending on the number of randomly selected interacting domain pairs, the connectivity distribution follows either a scale-free distribution, even in the absence of the preferential attachment, or a normal distribution. This new approach only requires an evolutionary model of proteins (nodes) but not for the interactions (edges). The edges are added by means of random interaction of domain pairs. As a result, this model offers a new mechanistic explanation for understanding complex networks with a direct biological interpretation because only protein structures and their functions evolved through genetic modifications of amino acid sequences. These findings are supported by numerical simulations using *H. sapiens* protein domain data from UniProt database.

Nacher JC, Hayashida M, Akutsu T: *BioSystems*, in press.

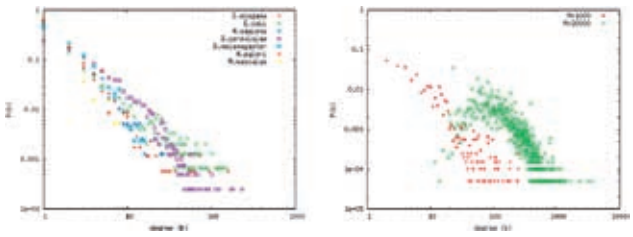

Figure 1. Degree distribution $P(k)$ of PPI networks (Left) real data for several organisms (Right) simulated data using *H. sapiens* data for the number of randomly selected interacting domain pairs, N =1000, 30000. A power-law distribution and a normal distribution were respectively observed.

Prediction of RNA Secondary Structure with Pseudoknots Using Integer Programming

RNA secondary structure prediction is one major task in bioinformatics, and various computational methods have been proposed so far. Pseudoknot is one of the typical substructures appearing in several RNAs, and plays an important role in a number of RNA functions such as ribosomal frameshifting and splicing. Prediction of RNA secondary structure with pseudoknots is still challenging since the problem is NP-hard when arbitrary pseudoknots are taken into consideration.

We propose a new method of predicting RNA secondary structure with pseudoknots based on integer programming. In our formulation, we aim at minimizing the value of the objective function that reflects free energy of a folding structure of an input RNA sequence since many single-stranded RNAs are considered to fold back on themselves to be thermodynamically stable. We focus on a practical class of pseudoknots by setting constraints appropriately. Experimental results for a set of real RNA sequences show that our proposed method outperforms several existing methods in sensitivity. Furthermore, for a set of sequences of small length, our approach achieved good performance in both sensitivity and specificity. Our integer programming-based approach to RNA secondary structure prediction is flexible and extensible enough to describe various types of secondary structures.

Poolsap U, Kato Y, Akutsu T: *BMC Bioinformatics*, 10 (Suppl 1) (2009).

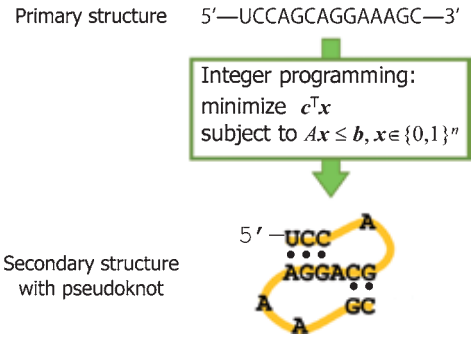

Figure 2. Prediction of RNA pseudoknotted structure.

Awards

Kato Y, IPSJ Best Paper Award, RNA Pseudoknotted Structure Prediction Using Stochastic Multiple Context-Free Grammar, Information Processing Society of Japan, 30 May 2008.

Kato Y, SIGBIO Best Paper Award, RNA Pseudoknotted Structure Prediction Using Stochastic Multiple Context-Free Grammar, Information Processing Society of Japan (Bioinformatics and Genomics), 17 December 2008.