

Development of a New Gene-Network Estimation System

Shigeyuki Mitsui

smitsui@tkb.mss.co.jp

Tuyoshi Takada

takada@tkb.mss.co.jp

Yoshiki Miura

ymiura@tkb.mss.co.jp

Hiroya Nobori

nobori@tkb.mss.co.jp

Kunihiko Kishida

kkishida@tkb.mss.co.jp

Hiroshi Ikawa

ikawa@tkb.mss.co.jp

Mitsubishi Space Software Co., Ltd., 1-6-1 Takezono, Tsukuba, Ibaraki 305-0032, Japan

Keywords: gene network, gene expression

1 Introduction

Gene-Network represents interactions between genes. Many techniques to obtain the gene expression data are developed, and the accuracy of the data and the number of available data have increased in recent years. At the same time, techniques of analyzing experimental results have been studied, but difficulty follows on the estimation of Gene-Network. Several trials are carried out with using GA (Genetic Algorithm) [1, 2], but the number of the genes which can be estimated at once is restricted to about five pieces, and additionally, many computer resources are needed. Our software is capable of estimation for more genes with more speed and more few computer resources, without using the heuristic optimization techniques, such as GA and SA (Simulated Annealing). Our software would help users well, and speed up the progress of the studies about Gene-Network.

2 System Overview

Our estimation technique consists of two stages. The first stage includes static state analysis to classify genes into some groups. The following stage contains the dynamic model analysis carried with time course data to estimate interactions in a gene group classified with static state analysis. It is universally said that the 10,000 kinds of genes exist in the living body. It will classify into about 100 gene groups according to static state analysis, and dynamic model analysis will be presented with them.

In static state analysis, the static states of a wild type and the deletion mutant of one gene are used. The difference of two strains implies the influence of the deleted gene, and the genes that express at different level will be in the deleted gene's downstream of the network directly or indirectly. Genes are classified into some groups according to several sets of static state analyses. The coefficient of a differential equation is computed in the dynamic model analysis carried with time course data. The differential equation can choose a liner model and S-System, taking computer resources, the number of genes, etc. into consideration. Genetic Algorithm or Simulated Annealing does not use it for the search technique, but fitting of the time course, a least-squares method, etc. are used.

Visualization is important in order to use an analysis result effectively. Drawing of a network is no longer an easy problem, if the number of genes increases from dozens of pieces. The network-drawing

tool computes the position of the gene on a network figure from the simulation result of a virtual dynamics model. The system overview is shown in Figure 1, and, a network-drawing sample is shown in Figure 2 & 3.

3 Future Plans

We examine some actual gene expression data. The software will be evaluated with more samples and will be brushed up for details and user interface.

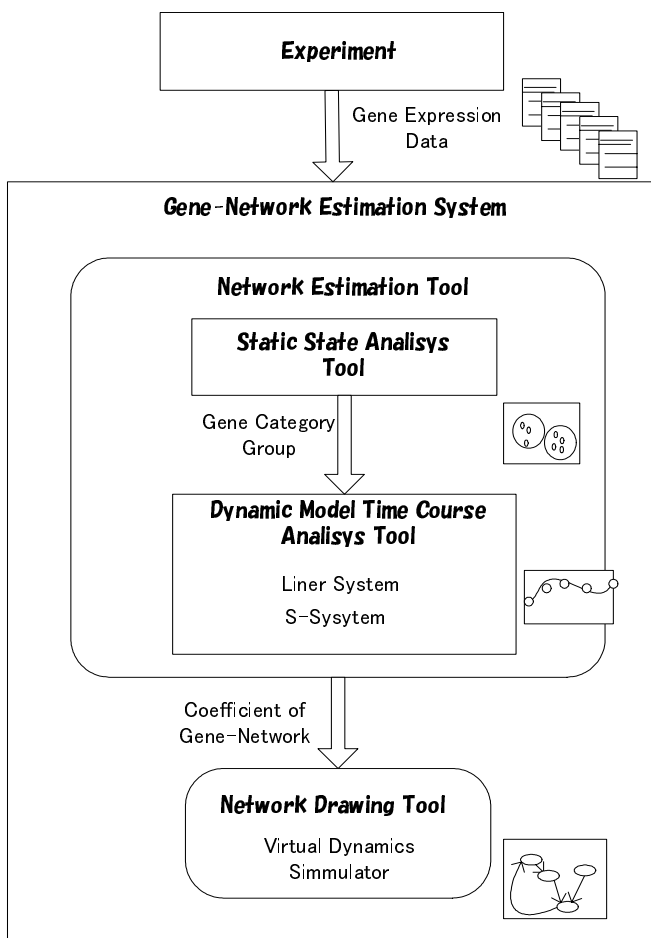


Figure 1: Gene-Network Estimation System.

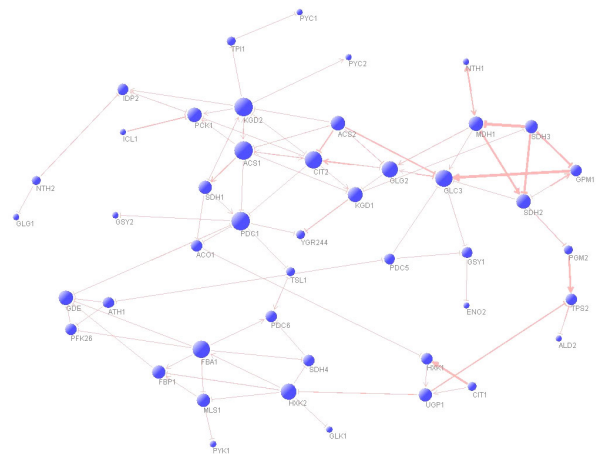


Figure 2: Example of Estimated Gene-Network View (2D).

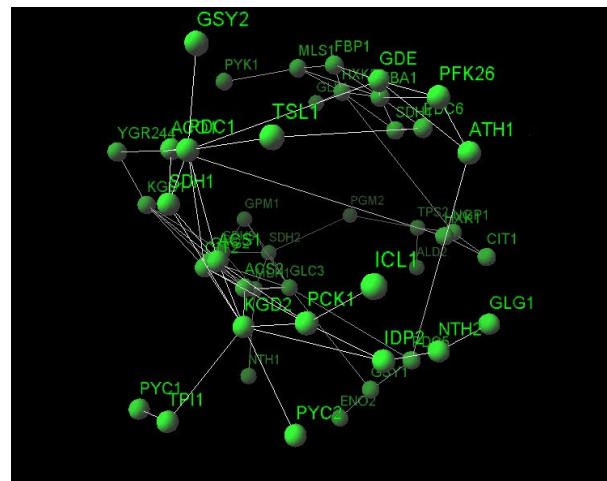


Figure 3: Example of Estimated Gene-Network View (3D).

References

- [1] Tominaga, D., Okamoto, M., Maki, Y., Watanabe, S., and Eguchi, Y., Nonlinear numerical optimization technique based on a genetic algorithm for inverse problems: towards the inference of genetic networks, *Proc. German Conference on Bioinformatics*, 127–140, 1999.
- [2] Voit, E.O. and Almeida, J., *Dynamic Profiling and Canonical Modeling: Powerful Partners in Metabolic Pathway Identification*, Kluwer Academic Publishing, Dordrecht, The Netherlands, 2003.