Computational methods for the Inference of Gene Regulatory Networks

Seema More†, Accaamma V††

†Faculty, Department of Computer Science & Engineering, M S Ramaiah Institute of Technology, Bangalore, India
††Faculty, Department of Computer Science & Engineering, Atria Institute of Technology, Bangalore, India

Summary
In the last 10 years, Inference of Gene Regulatory Networks from Microarray data has become an important research area in Bioinformatics. Several algorithms have been proposed and Regulatory networks have been inferred with good accuracy for Saccromycin cervisea, Escherichia coli, Hela cells etc. A large amount of knowledge on various biological systems, e.g. gene regulation, metabolic regulations, and signal transduction are being continually accumulated over the years, though there remains a large portion that is not well understood. This paper is a survey of different techniques used for the inference of Gene regulatory networks.

Key words:
Data mining, Bioinformatics, Gene Regulatory Network, Stochastic Model, Differential Equations, Association rule mining, Bayesian Network

1. Introduction

The input to the Gene network inference problem is Microarray data. If the Microarray data is Time series, the edges of the inferred network indicate the potential regulation relationships between genes. Thus a directed edge (i, j) would imply that gene i regulates gene j. The idea is to consider time-course gene expression experiments and correlate sustained positive and negative changes in the expression levels while incorporating biological considerations.

Various models have been proposed in literature to represent and simulate the behavior of Gene Regulatory Networks. Boolean, Bayesian Networks, Differential Equations, Weight Matrices, S-System are some of the prominent ones. The actual choice of a modeling formalism for a gene network will depend on the type and amount of data available, prior knowledge about the interactions in the network, nature of the questions one needs answered, area of formal training of the modeler, experimental and computational resources, and possibly other study- or organism-specific factors.

One important factor in determining a good algorithm is the number of time points required to infer the network. Most of the algorithms require huge datasets for effective determination of the regulatory networks. Another factor is the noise in the available gene expression data. Some algorithms perform badly in the case of noisy data.

The models can be classified as deterministic and stochastic models.

In deterministic models the expression states of the genes are either given by a formula or belong to a specific class. Measured at two different times or places, while keeping all other parameters the same, a gene's expression would be the same. The precision of the observed expression values, then, depends solely on the experimental setup and technological precision, and can be refined indefinitely with technological advances. The edges stand for relationships, which, like the node states, are also deterministic Eg: S-systems, Differential equations.

Stochastic models, on the other hand, start from the assumption that gene expression values are described by random variables, which follow probability distributions. The difference with the deterministic models is fundamental: randomness is modeled to be intrinsic to the observed processes, and thus all things being equal, a gene's expression on two different occasions may be different. Stochastic edges indicate probabilistic dependencies, and their absence may indicate independencies between nodes. They are not easy to interpret in practice. Eg: Bayesian networks

2. Differential Equations

This section reviews the modeling of GRNs using Differential equations.

Differential equations (DE) are the starting point for quantitative modeling of complex systems. DEs are continuous and deterministic modeling formalisms, capable of describing non-linear and emerging phenomena of complex dynamical systems.

DE models of gene networks are based on rate equations, quantifying the rate of change of gene expression as a function of the expressions of other genes. In this class of models, a system of differential equations is used to model the behavior of the gene network. This model is best suited for inferring structure from data collected from biological processes like developmental processes, cell cycle processes…etc.

Manuscript received March 5, 2011
Manuscript revised March 20, 2011

Rate equations have the mathematical form

$$\frac{dx_i}{dt} = f_i(x), 1 \le i \le n$$

Where Each fi(·) quantifies the combined effect of its arguments, or regulators, on xi, and it subsumes all the biochemical effects of molecular interactions and degradation.

Given observed gene expression data, the first step to identifying the gene network is to guess or approximate the fi(·)'s. Since the identification process depends solely on the form of the fi(·)'s, the functions are typically linearized. The reason for linearizing the original system is to turn it into a linear differential equation system, in which the parameters can easily be fitted to the data using linear algebra methods. In practice, the gene expression matrix is very long and narrow, that is, there are typically many more genes than experiments,

To infer linear ordinary differential equation systems there are two main approaches: (i) Singular Value Decomposition (SVD) based methods that calculate a solution for the interaction matrix by imposing additional mathematical constraints, and (ii) methods to identify sparse interaction matrices by combinatorial search strategies

2.1 Linear Differential Equations

The simplest interesting form that the fi(·)'s can take are linear additive functions

$$\frac{dx_i(t)}{dt} = ext_i(t) + w_{i1}x_1(t) + ... + w_{in}x_n(t)$$

With possibly some additional linear terms on the right hand side, indicating the degradation rate of gene i's mRNA or environmental effects, which can all be incorporated in the wij parameters, assuming their influence on xi is linear. The term exti(t) indicates a possible controlled external influence on gene i, like a perturbation for example, and is directly observable.

This model takes advantage of the continuous aspect of the data and is therefore suitable for genes with periodic expression, such as are important in the cell cycle, and all genes considered in the model are assumed to show this kind of expression pattern. Disadvantage of this model is that very fine sampling times need to be used to approximate them. [4], [5] and [31] used linear models to infer regulatory networks and claim very good results.

In the Non-linear Differential Equation model, the regulatory effect is represented as a non-linear function. [6] Used sigmoid function on the regulators

$$\frac{dz}{dt} = \frac{k_1}{1+\exp(-wy+b)} - k_2 z$$

Where the constant k2 represents the rate constant of degradation of the target gene product, and k1 is its maximal rate of expression; y is approximated with a polynomial of degree n

$$y \approx a_0 + a_1 t + a_2 t^2 + ... + a_n t^n$$

Coefficients $\{a_0, \ldots, a_n\}$ were computed from the experimental gene expression profile using a least squares minimization procedure.

This technique was applied on eukaryotic cell cycle dataset published by Spellman et al by [6]. When compared to linear model this model gave markedly better results both in the sense correct identification of regulators and the goodness of fit of the computed target gene expression profile.

Second Order Differential Equation is given as

$$\ddot{x}_i(t) + a_i \dot{x}_i(t) + b_i x_i(t) = G_i(t) + \varepsilon_i(t)$$

Where Gi(t) is the upstream regulatory function to influence the expression profile Xi(t) of the i-th gene while ai, and bi are the parameters that characterize the dynamics inherent property of the gene like degradation and oscillation, and εi(t) is the noise of current microarray data or the residue of the model. The regulatory pathway is constructed by tracing back Ri regulatory genes from the identified regulatory function of the target gene as the following kinetic relationship,

$$\hat{G}_i(t) = c_{i0} + \sum_{j \in R_i} c_{ij} \tilde{X}_j(t - \tau_j) + e_i(t)$$

After the combination of above equations the whole regulatory pathway is obtained as

$$\dot{X}_i(t) = -a_i X_i(t) - b_{i_i} X_i(t) + c_{i0} + \sum c_{ij} \tilde{X}_j(t - \tau_j)$$

[7] Applied this technique to successfully infer the circadian regulatory pathway in Arabidopsis thaliana and metabolic shift pathway from fermentation to respiration in yeast Saccharomyces cerevisiae, It was used in [8] to reconstruct the Dynamic Gene regulatory network of Cancer cell cycle.

Stochastic Differential Equation model is given as

$$\Delta X_t = \left[c_0 + \sum_{i=1}^{n} c_i f_i(x_{it}) \right] \Delta t + \varepsilon_{t,\Delta t}$$

where Xt and Xit are expression level of target gene and i-th regulator gene, ci is contribution of i-th regulator, fi is sigmoid function depending on i-th regulator and εt,Δt , is a random error generated from the normal distribution, i.e. N(0, σ2Δt)

[9] Suggested and tried to estimate the S.O.S DNA Repair network of E.coli with proposed model. The results obtained were good but more regulations were liable to be detected than expected. [10] coupled the model with log-likelihood and Akaike Information Criterion evaluated to estimate the goodness-of-fit. An important feature in the

SDE model is that dynamics arise as a natural consequence of time course in the duration of the cell cycle process. The SDE model takes time course into account for addressing continuous-time random error in Brownian motion. The fitted curves adequately depict differently shaped expression patterns while keeping the model parameters as few as possible. [11] used enhanced the SDE of [10] by using AIC strategy for the selection of the best fitting combination of the pool of regulators. With the addition of beta sigmoid pattern, the SDE model renders good prediction results even in the case of the previously worst fitted genes obtained by [10]

Any of the above Differential equation models can be combined with various Parameter estimation techniques to estimate the parameters part of the equation.

Maximum Likelihood Estimation is one of the widely used techniques to estimate the values. For the fixed set of data and underlying probability model, this method picks the values of the model parameters that make the data "more likely" than any other values of the parameters would make them. It was used by [7][8]coupled with the Second order differential equation technique.

1.1.1 Least squares minimization technique tries to fit the model into the data by estimating the parameters. The predicted data and actual data are compared and the error is minimized. A related method is the least mean squares (LMS) method. It occurs when the number of measured data is 1 and the gradient descent method is used to minimize the squared residual. LMS is known to minimize the expectation of the squared residual, with the smallest number of operations per iteration). However, it requires a large number of iterations to converge

AdaBoost was used coupled with SDE in [32] and with experiments shown that it was better than any combination of SDE, Linear or Weaver model coupled with AIC or BIC.

3. Bayesian Networks

A Bayesian network consists of an annotated directed acyclic graph G (X, E), where the nodes, $x_i \in X$, are random variables representing gene expressions and the edges indicate the dependencies between the nodes. The random variables are drawn from conditional probability distributions $P(x_i|Pa(x_i))$, where $Pa(x_i)$ is the set of parents for each node. A Bayesian network implicitly encodes the Markov Assumption that given its parents; each variable is independent of its non-descendants. With this assumption each Bayesian network uniquely specifies a decomposition of the joint distribution over all variables down to the conditional distributions of the nodes

The approach of Bayesian networks were first applied to the problem of reverse engineering genetic networks from microarray expression data by (Friedman et al. 2000)

While a static BN is restricted to be acyclic, a dynamic Bayesian network (DBN) can be used to infer cyclic phenomena such as feedback loops that are prevalent in biological systems. (Murphy and Mian, 2002) are to be credited with first employing DBN for modeling time-series expression data.

3.1 Structural Learning of Bayesian Networks

Structural learning is an unsupervised learning problem which can be stated as follows: given a dataset D = {d1; d2; :::; dN} of independent observations, find the structure that best matches D. Each datapoint dl of the N samples of dataset D is a n-dimensional vector dl = {dl1; :::; dln}.

The process of learning Bayesian networks from the data is essentially two-fold

The first part is model selection: Given observed data find the best graph (or model) G of relationships between the variables. For model selection, heuristics are used to efficiently search the space of models by looking at neighboring graphs around a given graph, by adding and deleting edges, or reversing directions of edges. Following are some of the techniques employed for model selection.

[33] Used hill-climbing algorithm that at each step performs the local change that results in the maximal gain, until it reaches a local maximum

Comparative study of Greedy algorithm, Simulated annealing and Genetic algorithm done by [34] proves that all three methods return identical networks with high recall and precision. However, greedy algorithm performs better than the two techniques.

The second step of Learning Bayesian networks is parameter fitting: Given a graph G and observed data find the best conditional probabilities for each node. Parameter fitting is the easier of the two in general. Given a graph model G, good candidates for the best conditional probability distributions can be found by various techniques

The expectation-maximization (EM) algorithm is a commonly used method to cope with missing data. [17] Used a learning algorithm based on the EM algorithm and on the Maximum likelihood maximization. [35] Proposed a new DBN model embedded with structural expectation maximization (SEM), which is capable of efficiently dealing with missing data.

The Bayesian Learning techniques use different scoring techniques to evaluate the goodness of the estimation done.

Studies in [34] show that The Bayesian Information Criterion (BIC) is a good approximation to the full posterior Bayesian Dirichlet equivalent (BDe) score and is faster to compute; however, it is known to over-penalize

with small amounts of data. BIC penalizes complexity more than Akaike Information Criterion (AIC)
Minimizing the (Minimum Description Length) MDL-score is equivalent to maximizing BIC, which is exactly the negation of MDL but with a completely different origin

4. Evolutionary Algorithms

The mathematical modeling of regulatory systems raises two problems:
1. Solutions those appear optimal under the objective function, but which do not correspond to the true model. These exist, because the system is underdetermined.
2. Suboptimal solutions, to which the optimization methods may converge.

Evolutionary algorithms are stochastic optimization techniques that mimic the natural evolution process of repeated mutation and selection as proposed by Charles Darwin. They have proven to be a powerful tool for solving complex optimization problems.

In the evolutionary algorithms technique, there are two aspects.

1. Representation of the dynamics of the Gene regulation as a model.
2. Optimization of the parameters of the model using evolutionary techniques.

Numerous criteria. e.g., mean squared error, Minimum Description Length, Bayesian, Information Criteria, and Akaike's Information Criteria, in correspondence with the various models that has been employed:

Two popular models used in this technique are S-Systems and Linear weight matrices

S-systems allows for capturing the non-linearity and general dynamics of the gene regulation. They are a type of power-law formalism and can be described by a set of nonlinear differential equations:

In Linear weight matrices, the regulative interactions between the genes are represented by a weight matrix, W, where each row of W represents all the regulatory inputs for a specific gene.

Evolved target GRNs is more reliably reconstructed by evolutionary algorithms than are 'random' target GRNs, and there is often no correlation between the best fit expression vector and recovery of the target GRN. Therefore, EA methods for biological-GRN reverse engineering are favored even if other methods more closely match the target expression vector.

5. Association Rule Mining

Previous works on gene expression association rule mining are mostly based on the 'support–confidence' framework. In Creighton and Hanash (2003), the Apriori algorithm was adopted with some additional criteria, such as extracting frequent itemsets larger than size of seven, to narrow the search space of candidate itemsets. Even so, tens of thousands of frequent itemsets were extracted out, many of which were redundant, and it is still very time-and-memory consuming to generate rules from such large number of itemsets. A manual search was done with the itemsets that seemed to be closed (itemsets that were not subsets of some larger itemsets), based on which rules were finally extracted. Kotala et al. (2001) adopted the Peano Count Tree (P-tree) to efficiently calculate the support and confidence by a high-order bit first and a single attribute first approach. Those methods of setting additional criteria to prune the itemsets before and after applying Apriori helps to narrow the vast majority of frequent itemsets to some extent; however, since the relations of gene expression data are very complicated and there is little Apriori knowledge about the gene network, it is a great challenge for researchers to set the proper criteria.

6. Inference of cancer-specific gene regulatory networks

Inference of gene regulatory network helps in extracting lot of information especially in disease related expression data. Perturbations of gene regulatory networks are essentially responsible for oncogenesis. Therefore, inferring the gene regulatory networks is a key step to overcoming cancer. Yeh et. al.[39] presented a computational method for inferring genetic regulatory networks from micorarray data automatically with transcription factor analysis and conditional independence testing to explore the potential significant gene regulatory networks that are correlated with cancer, tumor grade and stage in the prostate cancer. A computational framework was provided to reconstruct the genetic regulatory network from the microarray data using biological knowledge and constraint-based inferences. They predicted not only individual gene related to cancer but also discovered significant gene regulation networks. [40] Proposes a method for inferring directed gene regulatory networks based on soft computing rules, which can identify important cause-effect regulatory relations of gene expression. First, important genes associated with a

specific cancer are identified (colon cancer) using a supervised learning approach. Next, the gene regulatory networks are reconstructed by inferring the regulatory relations among the identified genes and their regulated relations by other genes within the genome. There were two meaningful findings. One is that upregulated genes are regulated by more genes than downregulated ones, while downregulated genes regulate more genes than upregulated ones. The other one is that tumor suppressors suppress tumor activators and activate other tumor suppressors strongly, while tumor activators activate other tumor activators and suppress tumor suppressors weakly, indicating the robustness of biological systems. These findings provide valuable insights into the pathogenesis of cancer.

The complicated molecular mechanism underlying cancer lies in the perturbations of gene-interaction networks at some level. Therefore, identifying cancer genes and the pathways they control through the networks is a key step toward overcoming cancer. Generally speaking, directed gene regulatory networks reflect the gene interactions more genuinely than undirected gene co-expression networks in that the principal cause-effect relations between genes can be disclosed in directed gene regulatory networks. [40] aims at inferring directed gene regulatory networks under specific disease conditions using formalized rules, which facilitate the interpretability of the inference model. First the genes that are relevant to a specific disease are identified by supervised learning algorithms, and then the regulatory relations among the identified genes and their regulated relations by all other genes are inferred. Wang et. al's approach for inferring regulation networks is based on soft computing rules. The reliability of inferred regulation relations depends on the confidence of corresponding rules, which is governed by the controllable parameter α. To ensure sufficiently high reliabilities of the inferred relations, a high threshold is set for α. When analyzing the properties of inferred networks, a network was selected induced with a rational value of α, which contain substantial and reliable regulatory relations.

7. Selection of Regulators

A drawback of all published algorithms for inference of transcriptional regulatory networks is that the candidate regulators are selected from the pool of potential regulators defined independently. If the regulator is not identified, it inevitably escapes identification by the modeling approach. The less characterized the genome of an organism is, the higher the probability of this type of error.

Clustering is one of the techniques used as a preprocessing step. If the research covers a large amount of genes then the technique adapted is to find clusters of genes, which have similar expression patterns, Arguments used, may be that co-expressed genes are probable to have related functions.

Several techniques for clustering expression data time-series have been proposed in the literature, based on measures like Euclidian distance, mutual information, linear correlation, and rank correlation. The use of clustering algorithms is motivated by the idea that two genes exhibiting similar expression patterns over time may regulate each other or be co regulated by a third gene. Additional analyses may permit one to extract more information on putative regulatory connections between co-expressed genes in the graph, such as the analysis of time lags [19]

8. Results

Most of the algorithms have been tried on artificial datasets. Some GRNs inferred from biological microarray data has been summarized in table 1.

Table 1: Summary GRN's inferred from microarray data

Dataset	***Size***	***DBN***	***Differential***	***Evolutionary***
Saccharomyces cerevisiae Cell cycle	800 genes		[6][10]	[37]
Metabolic shift pathway from fermentation to respiration in yeast Saccharomyces cerevisiae	6400		[7]	
S.O.S. DNA Repair network of the Escherichia coli bacterium	9 genes	[17]	[36][9]	
Hela Cancer cell cycle	775 genes		[8]	

9. Conclusion

In this paper, we provided a survey of different algorithms that exist for the inference of Gene Regulatory Networks. The Differential equations methods are faster than Bayesian and Evolutionary methods. But, their drawback is that they can be strongly affected by noise. The evolutionary techniques always yield optimal solutions. But, they have huge computation times. Therefore, evolutionary computation with High performance computing resources is a promising direction.

References

[1] Patrik D'Haeseleer, S. Liang, and R. Somogyi. Genetic network inference: From co-expression clustering to reverse engineering. Bioinformatics, 16(8):707--726, 2000.

[2] Zeke S. H. Chan, Nikola K. Kasabov, Lesley Collins: A two-stage methodology for gene regulatory network extraction from time-course gene expression data. Expert Syst. Appl. 30(1): 59-63 (2006)

[3] Christian Spieth, Felix Streichert, Nora Speer, Andreas Zell: Clustering-based approach to identify solutions for the inference of regulatory networks. Congress on Evolutionary Computation 2005: 660-667

[4] T. Chen, H. L. He, and G. M. Church. Modeling gene expression with differential equations. In Pacific Symposium Biocomputing '99, pages 29--40, 1999.

[5] M.J.L. de Hoon, S. Imoto, K. Kobayashi, N. Ogasawara and S. Miyano. Inferring gene regulatory networks from time-ordered gene expression data of Bacillus subtilis using differential equations. Pacific Symposium on Biocomputing, 8, 17-28, 2003

[6] Tra Thi Vu and Jiri Vohradsky, Nonlinear differential equation model for quantification of transcriptional regulation applied to microarray data of Saccharomyces cerevisiae, Nucleic Acids Research, 2007, Vol. 35, No. 1 279–287

[7] Wen-Chieh Chang, Chang-Wei Li, Bor-Sen Chen: Quantitative inference of dynamic regulatory pathways via microarray data. BMC Bioinformatics 6: 44 (2005)

[8] Cheng-Wei Li, Yung-Hsiang Chu and Bor-Sen Chen, Construction and Clarifi cation of Dynamic Gene Regulatory Network of Cancer Cell Cycle via Microarray Data, Cancer Informatics 2006: 2 223–241

[9] Shinya Nabatame, Hitoshi Iba, Estimation of Gene Regulatory Network Using Stochastic Differential Equation Model

[10] Kuang-Chi Chen, Tse-Yi Wang, Huei-Hun Tseng, Chi-Ying F. Huang, Cheng-Yan Kao, A stochastic differential equation model for quantifying transcriptional regulatory network in Saccharomyces cerevisiae. BIOINFORMATICS ORIGINAL PAPER Vol. 21 no. 12 2005, pages 2883–2890

[11] Adriana Climescu-Haulica and Michelle D Quirk, A stochastic differential equation model for transcriptional regulatory networks, BMC Bioinformatics 2007, 8(Suppl 5):S4

[12] Friedman, N., Linial, M., Nachman, I., and Pe'er, D, Using Bayesian networks to analyze expression data.J. Comput. Biol. 7, 601–620, 2000

[13] D. Heckerman, D. Geiger, and D. Chickering. Learning bayesian networks: The combination of knowledge and statistical data. Mach. Learn., 20:197–243, 1995.

[14] Tie-Fei Liu, Wing-Kin Sung: Learning Gene Network Using Conditional Dependence. ICTAI 2006: 800-804

[15] T. F. Liu, W. K. Sung, and A. Mittal. Learning Multi-Time Delay Gene Network Using Bayesian Network Framework. In The 16th IEEE International Conference on Tools with Artificial Intelligence, pages 640.645, 2004.

[16] Linus Göransson, Timo Koski, Using a Dynamic Bayesian Network to Learn Genetic Interactions

[17] Bruno-Edouard Perrin, Liva Ralaivola, Aurélien Mazurie, Samuele Bottani, Jacques Mallet, Florence d'Alché-Buc: Gene networks inference using dynamic Bayesian networks. ECCB 2003: 138-148

[18] Mathäus Dejori, Martin Stetter: Bayesian Inference of Genetic Networks from Gene-Expression-Data: Convergence and Reliability. IC-AI 2003: 323-327

[19] Min Zou, Suzanne D. Conzen: A new dynamic Bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data. Bioinformatics 21(1): 71-79 (2005)

[20] Eisen, M.B., Spellman, P.T., Brown, P.O., Botstein, D.: "Cluster analysis and display of genome-wide expression patterns", Proc. Natl. Acad. Sci. USA 95: 14863-14868, 1998.

[21] Shinichi Kikuchi, Daisuke Tominaga, Masanori Arita, Katsutoshi Takahashi, Masaru Tomita: Dynamic modeling of genetic networks using genetic algorithm and S-system. Bioinformatics 19(5): 643-650 (2003)

[22] M. Swain, T. Hunniford, Johannes J. Mandel, Niall M. Palfreyman, Werner Dubitzky: Modeling gene-regulatory networks using evolutionary algorithms and distributed computing. CCGRID 2005: 512-519

[23] M.K.S. Yeung, J. Tegn˜ar, and J.J. Collins. Reverse engineering gene networks using singular value decomposition and robust regression. Proc. Natl. Acad. Sci. USA, 99:6163–6168, 2002.

[24] Inferring a System of Differential Equations for a Gene Regulatory Network by using Genetic Programming, Erina Sakamoto, Hitoshi Iba, 2001

[25] Estimation of Gene Regulatory Network by Genetic Algorithm and Pairwise Correlation Analysis, Shin Ando, Hitoshi Iha, 2003

[26] On the Reconstruction of Gene Regulatory Networks from Noisy Expression Profiles, Nasimul Noman and Hitoshi Iba, 2006

[27] Investigating Issues in the Reconstructability of Genetic Regulatory Networks, David Corne, Carey Pridgeon, 2004

[28] Clustering-based Approach to Identify Solutions for the Inference of Regulatory Networks, Christian Spieth, Felix, Streichert, Nora Speer, and Andreas Zell

[29] Feedback Memetic Algorithms for Modeling Gene Regulatory Networks,C. Spieth, F. Streichert, J. Supper, N. Speer, and A. Zell,2006

[30] Inferring Gene Regulatory Networks using Differential Evolution with Local Search Heuristics Nasimul Noman and Hitoshi Iba, Member, IEEE, 2006

[31] Tabus;, C. D. Giurcaneanu, and J. Astola, "Genetic networks inferred from time series of gene expression data," in Proc. First International Symposium on Control, Communications and Signal Processing (ISCCSP '04), pp. 755-758, Hammamet, Tunisia, March 2004.

[32] Shinya Nabatame, Hitoshi Iba, Integrative Estimation of Gene Regulatory Network by Means of AdaBoost,

[33] N. Friedman, M. Linial, I. Nachman, and D. Peer. Using Bayesian Networks to Analyze Expression Data. In International Conference on Research in Computational Molecular Biology (RECOMB), pages 127.135, 2000.

[34] Yu, J., Smith, V., Wang, P., Hartemink, A., & Jarvis, E. (2002) "Using Bayesian Network Inference Algorithms to Recover Molecular Genetic Regulatory Networks." International Conference on Systems Biology 2002 (ICSB02), December 2002.

[35] Yu Zhang, Zhidong Deng, Hongshan Jiang, Peifa Jia: Gene Regulatory Network Construction Using Dynamic Bayesian Network (DBN) with Structure Expectation Maximization (SEM). RSKT 2006: 402-407
[36] Christian Spieth, Felix Streichert, Nora Speer, and Andreas Zell, A Memetic Inference Method for Gene Regulatory Networks Based on S-Systems, 2004 IEEE
[37] Nasimul Noman and Hitoshi Iba, Inferring Gene Regulatory Networks using Differential Evolution with Local Search Heuristics, IEEE Transactions on Computational Biology and Bioinformatics.
[38] Mukesh Bansal, Giusy Della Gatta and Diego di Bernardo, Inference of gene regulatory networks and compound mode of action from time course gene expression profiles, Bioinformatics Original Paper, Vol. 22 no. 7 2006
[39] Yeh HY, Cheng SW, Lin YC, Yeh CY, Lin SF, Soo VW. Identifying significant genetic regulatory networks in the prostate cancer from microarray data based on transcription factor analysis and conditional independency. BMC Med Genomics. 2009 Dec 21; 2:70.
[40] Xiaosheng Wang, Osamu Gotoh Inference of Cancer-specific Gene Regulatory Networks Using Soft Computing Rules, Gene Regul Syst Bio. 2010; 4: 19–34.

Seema More received the B.E. from Gulbarga University, Karnataka, and M.S. degree, from Birla Institute of Technology, Rajasthan. Pursuing research in the field of data mining in Bioinformatics. She is working as an associate professor at M S Ramaiah Institute of Technology Bangalore, India. Her research interest includes acoustic signal processing, sound and data mining and its applications in the field of Bioinformatics. He is a member of IEEE.