Integrating Epigenetic Prior in Dynamic Bayesian Network for Gene Regulatory Network Inference

Haifen Chen[1#], D.A.K. Maduranga[1#], Piyushkumar A Mundra[1], Jie Zheng[1,2*]
[1] Bioinformatics Research Centre, School of Computer Engineering, Nanyang Technological University, Singapore 639798
[2] Genome Institute of Singapore, A*STAR (Agency for Science, Technology, and Research), Biopolis, Singapore 138672
{hchen009, kasun1, pamundra, zhengjie}@ntu.edu.sg

[#] The two authors made equal contribution to this paper
* Corresponding author

***Abstract*—Gene regulatory network (GRN) inference from high throughput biological data has drawn a lot of research interest in the last decade. However, due to the complexity of gene regulation and lack of sufficient data, GRN inference still has much space to improve. One way to improve the inference of GRN is by developing methods to accurately combine various types of data. Here we apply dynamic Bayesian network (DBN) to infer GRN from time-series gene expression data where the Bayesian prior is derived from epigenetic data of histone modifications. We propose several kinds of prior from histone modification data, and use both real and synthetic data to compare their performance. Parameters of prior integration are also studied to achieve better results. Experiments on gene expression data of yeast cell cycle show that our methods increase the accuracy of GRN inference significantly.**

Keywords— Gene regulatory network; dynamic Bayesian network; epigenetics; histone modification; gene expression; yeast cell cycle

I. INTRODUCTION

Reconstructing gene regulatory network from gene expression data has been intensely studied for decades. Many approaches were proposed to build models for GRN. These approaches can be categorized into four types: Boolean networks, differential equations, information theory models and Bayesian networks [1]. They each have their own advantages and limitations. Boolean networks [2] are efficient and can model dynamic events. Nevertheless, they require data discretization which may cause loss of essential information. Differential equations [3] are quantitative and flexible, thus can model complex relations. But they are limited by mathematical difficulty and only applicable to small networks. Information theory models [4] are very simple and fast, and are suitable for large-scale networks. Yet they are static and cannot model multiple regulations. Bayesian networks [5-7] simulate expression values of genes as random variables and regulatory relations as conditional probabilities. Among these four types of approaches, Bayesian networks are most widely used because of at least four advantages: 1) they can model randomness and noise, which are ubiquitous in biological processes; 2) they are flexible to integrate various data sources and prior knowledge; 3) Bayesian networks are able to learn causal relationships; 4) they can handle incomplete data sets and can avoid over-fitting of data [5]. However, Bayesian networks are computationally expensive, and cannot handle feedback loops because directed acyclic graphs (DAG) are employed as model structures. Dynamic Bayesian networks [8] were proposed to cope with the problem of feedback loops in GRN. Many heuristic and stochastic algorithms have been applied to reduce the computational cost of Bayesian learning, such as genetic algorithms, simulated annealing, Markov Chain Monte Carlo (MCMC) [9].

In spite of intense studies, GRN reconstruction still suffers from poor accuracy, which is partly because of the inherent complexity of biological processes, lack of gold standard networks, limited availability of data samples and noise or bias in the data. On the other hand, with recent advancement in sequencing technologies, new datasets such as epigenetic data are gathered which could help achieve accurate inference of GRN. Epigenetics is the study of phenotype changes (such as gene expression) caused by some functional modifications to the genome rather than inherent changes of DNA sequences. Such modifications include histone modifications, DNA methylation, and so on.

The relations between epigenetic features and gene expression have been studied in recent years. Yu et al. built a Bayesian network to infer causal and combinatorial relationships between gene expression and histone modifications [10]. Su et al. constructed an epigenetic interaction network of histone modifications, DNA methylation and gene expression using partial correlation and Pearson correlation [11]. Various machine learning techniques were also adopted to study the relationship between chromatin features and gene expression. Xu et al. applied two regression models to estimate gene expression using histone modifications [12]. Cheng et al. used support vector machine (SVM) to predict gene expression levels based on histone modifications [13]. All these papers focus on the "global" effect of epigenetic features on gene expression. How the epigenetic factors affect regulatory relations between individual genes is still far from clear. Zheng et al. [14] integrated histone modifications data as

prior in Bayesian network to improve the performance of GRN reconstruction from yeast cell cycle expression data. Their results showed that the integration of epigenetic data with gene-expression data improve the accuracy of GRN inference significantly. However, their experiments were restricted to static Bayesian network which cannot capture the feedback regulations between genes. The effect of such prior on dynamic Bayesian network based GRN inference is untested. Besides, they used the average enrichment of each histone modification across different loci of a gene, which may lose some important information, and hence, may not be optimal.

In this paper, we extend work of [14] and integrate epigenetic prior in dynamic Bayesian network. Instead of taking average HM enrichment across all loci for each gene, we treat two different kinds of regions separately: intragenic and intergenic regions. Based on their distinct inherent characteristics, we propose several new priors and analyze their effect on GRN reconstruction using DBN. We first confirm the finding from [14] that connected gene-pairs have stronger correlation of histone modification profiles than random gene-pairs in the global transcriptional regulatory network of yeast. Then by applying DBN on both real and synthetic expression data of yeast, we demonstrate that our priors can improve the performance (F-measure) of GRN inference by up to 6%. Furthermore, the weight of prior on Bayesian learning is studied by adjusting the parameter of prior probability distribution. Our methods are described in section II of this paper, which is followed by the experiments and results analysis. Lastly, we discuss and conclude our work in section IV.

II. METHODS

A. *Dynamic Bayesian Network*

Bayesian networks are defined by three components namely, a conditional probability distribution, its parameters and graphical structure [22]. The structure of a Bayesian network consists of a set of edges which are directed and nodes. The edges in the structure indicate conditional dependence relationships among the nodes and this is also called as parent child relationship. Bayesian network decomposes the joint probability of set of random variable $(a_1, a_2, ..., a_n)$ into a product of conditional probabilities as follows [22]:

$$p(a_1, a_2, ..., a_n) = \prod_{i=1}^{n} p(a_i \mid b_i) \qquad (1)$$

where b_i denotes the set of parents of i^{th} variable (a_i).

Dynamic Bayesian network expands the scope of the Bayesian network model to capture the dynamic properties of data and represent feedback loops. Note that vector $x = (x_i(t))_{t=1}^{T}$, contains expression data of gene i that is collected over *T* time points. Suppose q and s represent the parameters and the structure of the networks respectively. Let's assume that each gene has L number of levels. Then by applying the decomposability property, as explained by (1), likelihood of gene expression is given by:

$$p(x \mid s, q) = \prod_{i=1}^{I} \prod_{j=1}^{J} \prod_{l=1}^{L} q_{ijl}^{M_{ijl}} \qquad (2)$$

where $q_{ijl} = p(x_i(t) = l \mid b_i(t-1) = j)$ and M_{ijl} shows the number of times that x_i obtains the state l in t^{th} time sample when b_i (x_i 's parents) being j at *(t-1)*th time sample. The GRN structure s^* obtained by maximizing posterior probability $p(s \mid x)$ as follows [22]:

$$s^* = \arg\max_{s} p(s \mid x). \qquad (3)$$

According to the Bayes rule of the probability, we can write posterior probability as a multiplication of prior probability and marginal likelihood,

$$p(s \mid x) = \frac{p(x \mid s)\, p(s)}{\sum_{s'} p(x \mid s')\, p(s')} \qquad (4)$$

where $p(s)$ is the prior probability, $\sum_{s'} p(x \mid s')\, p(s')$ denotes the normalization factor, s' denotes the all possible network structures and $p(x \mid s)$ denotes the marginal likelihood of the network structure and it is obtain by integrating (2) with respect to parameters:

$$p(x \mid s) \propto \int p(x \mid s, q) p(q \mid s) dq \qquad (5)$$

The network structure s^* is obtained by maximizing posterior probability. But the number of possible gene regulatory network structures for a given number of genes (nodes) increases super exponentially with the increase in number of genes (nodes) in the network [22]. So it is impractical to compute the value of the denominator ($\sum_{s'} p(x \mid s')\, p(s')$) in (4). Because of such problem, it is hard to obtain straightforward solution to (3). Therefore, we need to sample network structures from a posterior probability distribution and apply the Markov chain Monte Carlo (MCMC) [9] simulation.

In this paper, we have implemented the MCMC simulation which is associated with the Metropolis-Hastings criteria [15]. Let's take s_{old} as the current regulatory network, and s_{new} as the propose network. This new regulatory network structure is accepted if it satisfies Metropolis-Hastings criteria which is given by below equation,

$$\min\left\{1, \frac{p(s_{new}\mid x)}{p(s_{old}\mid x)} \times \frac{Q(s_{old}\mid s_{new})}{Q(s_{new}\mid s_{old})} \times \frac{p(s_{new})}{p(s_{old})}\right\}, \quad (6)$$

where Metropolis-Hasting acceptance ratio is:

$$\frac{p(s_{new}\mid x)}{p(s_{old}\mid x)} \times \frac{Q(s_{old}\mid s_{new})}{Q(s_{new}\mid s_{old})}, \quad (7)$$

where $Q(s_{new}\mid s_{old})$ denotes the proposal probability and $p(s\mid x)$ is obtained from the (4). The most important advantage of the Metropolis-Hasting acceptance ratio is that it cancels out the intractable denominator of the posterior probability as it uses the ratio. In this paper, initial structure is obtained by using mutual information among the gene expression data. Then elementary operations (deleting, reversing, or adding an edge) are applied to current structure in each iteration to obtain a new network structure. The second part of the (7) is called Hastings ratio and it is computed by taking the ratio between sizes of neighborhood of s_{old} and s_{new} [22].

B. Prior intergration

The process of epigenetic prior integration is similar to [14], which is based on the framework described in [16]. Using the epigenetic data of histone modification (HM) profiles from [17], we first construct a prior matrix for DBN inference. Instead of taking the average HM enrichment of all locations for each gene, we average the HM enrichment across different locations from intragenic and intergenic regions separately.

To estimate the epigenetic association between gene i and gene j, we calculate Pearson correlation coefficient (PCC) ρ_{ij} between their HM vectors. Since PCC ranges between -1 and 1, while the Bayesian prior need to be between 0 and 1, we scale to [0, 1] by taking the absolute value of PCC. Let $\mathrm{b}_{i,j}$ be the prior value of gene i and gene j, then $\mathrm{b}_{i,j} = |\rho_{ij}|$. We take the absolute PCC because high correlation suggests connection in regulatory network, regardless of positive or negative. In the following part of this paper, we use 'correlations' as short for 'absolute correlations'. And we use 'prior' as short for the prior matrix $\{\mathrm{b}_{i,j}\}_{n\times n}$ where n is the number of genes. Similarly, 'intragenic prior' and 'intergenic prior' represent the prior matrixes generated from HM profiles of intragenic regions and intergenic regions respectively. And 'epigenetic prior' means the prior matrix generated from epigenetic data, which are HM data here.

After that, we use Gibbs distribution to obtain the prior probability $P(s)$ for a given network structure s, using the following equation:

$$P(s) = \frac{1}{Z} e^{-\beta E(\mathrm{s})} \quad (8)$$

where Z is a normalizing factor also referred as normalizing partition function in [14]. The effect of Z is automatically cancels out since we are using ratio between prior probabilities in (6). The $E(\mathrm{s})$ in (8) is given by

$$E(\mathrm{s}) = \sum_{i,j=1}^{n} |\mathrm{b}_{i,j} - \mathrm{c}_{i,j}|, \quad (9)$$

where $\mathrm{b}_{i,j}$ is element of the prior matrix generated from the HM profiles, and $\mathrm{c}_{i,j}$ is from the connectivity matrix corresponding to the structure s. For more details, please refer to [14]. In (8), the parameter β is a hyperparameter which indicates the influence of prior evidence to DBN inference process relative the data. When $\beta \to 0$, prior probability given by (8) is flat and not provide sufficient information about the network structure but when β is increasing it provide more information about the structure.

We find that the prior matrix from HM profiles of intragenic regions has overall higher correlation values than those from intergenic regions. However, for intergenic regions, correlations of truly connected gene-pairs are more likely higher than those of random pairs. The box plots of such comparison are shown in Fig. 1. As seen from Fig. 1, correlations of HM profiles from the intragenic regions (the left two boxes in Fig. 1) are generally higher than those from the intergenic regions (the right two boxes in Fig. 1). On the other hand, in intergenic regions, the correlations of the confirmed gene-pairs are higher than random pairs, which is not the case for intragenic regions. The higher correlations of confirmed gene-pairs than random pairs can serve as prior knowledge for GRN inference.

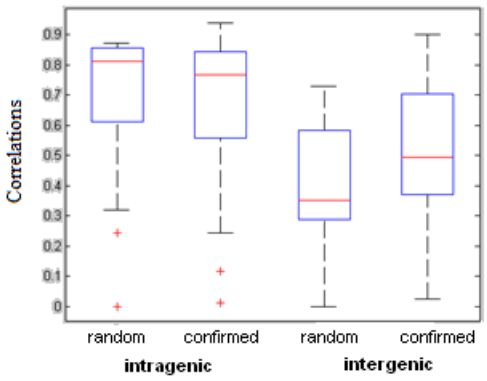


Figure 1. Comparison of correlations of yeast HM profiles between random gene-pairs and confirmed gene-pairs in intragenic and intergenic regions respectively.

To take advantage of both the high correlations in intragenic regions and the discriminative power of intergenic regions, we combine the intragenic prior with intergenic prior. To avoid the low correlations of intergenic prior jeopardizing the performance of the combined prior, we first normalize the intergenic prior in the following way: first randomly pick N (here N=5000) elements $e_1, e_2, \ldots, e_N$ from the matrix of intergenic prior to form a random

distribution $RD = \{e_1, e_2, ..., e_N\}$, then each value v in the matrix of intergenic prior is re-calculated as

$$v_{new} = \frac{\left|\{e_i \mid e_i \le v, e_i \in RD\}\right|}{N}. \quad (10)$$

where |…| denotes the number of elements of the set "…". After normalization, the overall correlations in the matrix of intergenic prior are increased and its discriminative power is enhanced, see Fig. 2.

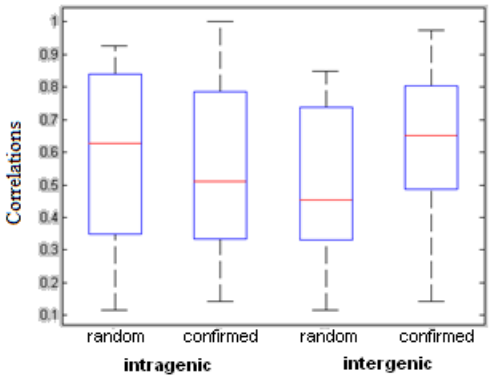


Figure 2. Comparison of normalized correlations of yeast HM profiles between random gene-pairs and confirmed gene-pairs in in intragenic and intergenic regions respectively.

Although the intergenic prior has increased values after normalization, its correlations are still not high enough. We need to combine it with intragenic prior, which has high correlation values. Let G be the matrix of intragenic prior, and In be the matrix of normalized intergenic prior, the new combined prior GIn is obtained by:

$$GIn(i, j) = (G(i, j) + In(i, j)) / 2, \quad (11)$$

where $G(i, j)$ is the prior value between gene i and gene j. Similar meaning for $In(i, j)$ and $GIn(i, j)$.

Therefore, we propose several priors as follows, and we will compare their performance in GRN reconstruction using DBN.

1) A: the prior used in [14];

2) An: the normalized version of A;

3) G: the prior generated from HM profiles in intragenic regions;

4) Gn: the normalized version of G;

5) I: the prior generated from HM profiles in intergenic regions;

6) In: the normalized version of I;

7) GIn: the combination of G and In, using (11).

C. Synthetic expression data generation

Time-series gene expression data were simulated for a given network topology using first-order multivariate vector autoregressive (MVAR) [18, 19] model. Here we use the 9-gene yeast regulatory network in [14] as network topology to generate the expression data. Let vector $x(t) = (x_i(t))_{i=1}^n$ represent expression of a gene network which has n number of genes at time t. According to first order MVAR model, gene expression values of all the genes at time t obtained by,

$$x(t) = x(t-1) * S_{weight} + \varepsilon(t), \quad (12)$$

where $S_{weight} = \{S_{weight}\}_{n \times n}$ denotes the connectivity matrix of the gene network with randomly assigned weights and $\varepsilon(t)$ denotes the added Gaussian random noise to the gene expression at time t. Random weights in S_{weight} were allocated to all connections by obtaining values from a uniform distribution on the interval [-1, -0.5] and [0.5, 1]. When there is no edge between two genes in the connectivity matrix, their weights were set to zero. This random allocation of weights helps to maintain the number of positive and negative weights approximately equal [19]. The initial gene expression vector $x(t)_{t=0}$ was constructed by sampling from the uniform distribution on the interval [0, 1]. Then subsequent time points are generated using (12). Gene expression values were represented as continuous data in generated data set. Then these data were converted to three discrete values {+1, 0, -1} by applying equal width interval binning method. Each of these discrete levels indicates up regulation, no regulation and down regulation respectively.

III. Experiments and Results

First we confirmed that the gene-pairs in a regulatory network have significantly higher correlations of histone modification profiles than random gene-pairs. We confirmed this finding in the global transcriptional regulatory network of yeast Saccharomyces cerevisiae [20], which contains 4441 genes and 12873 regulatory interactions. This network was assembled from biochemical, genetic and ChIP-chip experiments [20]. The data of histone modification profiles are from [17]. We can see from Fig. 3 that the gene-pairs in the global transcriptional regulatory network of yeast have significant higher correlations of HM profiles than random gene-pairs, which supports our method of using HM data as prior knowledge in GRN inference.

Next we compare the performance of all the seven priors we have introduced earlier using both real data and synthetic data. The benchmark network we use is the same as in [14]. The real data are expression data of yeast cell cycle (24 time points of cdc-15 cell cycle arrest) from [21]. And the synthetic data are generated by MVAR model as described in the Methods section. Six synthetic datasets have been generated, which contain 10, 30, 50, 70, 100, 200 time points respectively. Each dataset consists of 20 time series and the

final results take the average of the performance metrics. We use precision, recall, and F-measure as performance metrics. Fig. 4 shows the comparison of F-measure between the seven different priors, where the error bars represent the standard deviation. As seen, the *GIn* prior has the highest performance among all the proposed priors in real data and most synthetic data. For detailed comparison, please refer to Table I, where 'Precision', 'Recall' and 'F-measure' are calculated as follows:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN}$$

$$F - measure = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision+Recall}},$$

where *TP*, *FP* and *FN* denote the number of true positive, false positive and false negative respectively.

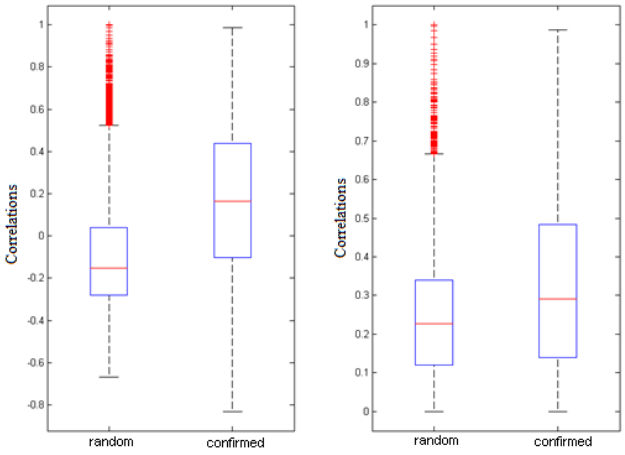


Figure 3. Comparison between HM profiles correlations of random gene-pairs vs. confirmed gene-pairs in the global gene regulatory network of yeast, where the left figure is using the original correlation and the right one is using absolute correlations. (Both Wilcoxon test p-value < 0.001)

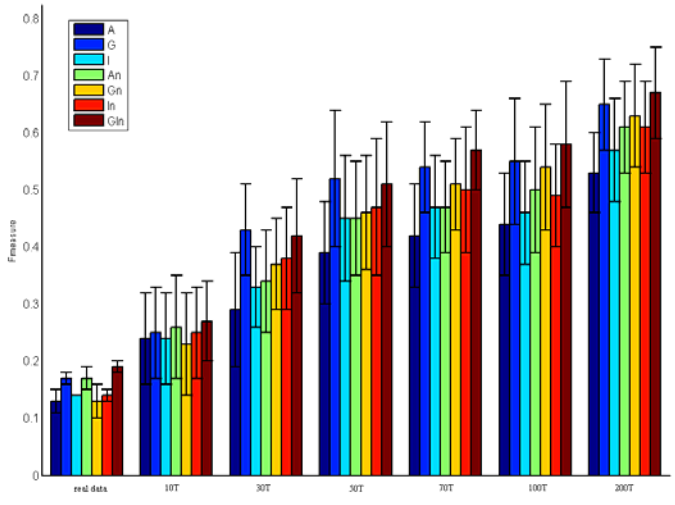


Figure 4. Performance (F-measure) comparison between priors in GRN inference on both real data and synthetic data (10~200 time points). β =1

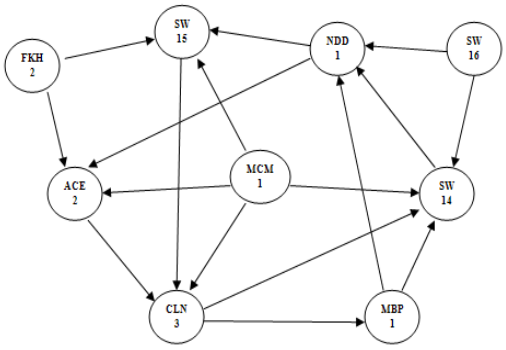
TABLE I. Effect of prior types and number of time points on DBN performance when β =1

Prior	Data	Precision	Recall	F-measure
No prior	Real	0.21	0.13	0.16
	10	0.29	0.18	0.22
	30	0.48	0.28	0.35
	50	0.68	0.37	0.47
	100	0.91	0.42	0.57
A	Real	0.15	0.11	0.13
	10	0.29	0.20	0.24
	30	0.41	0.23	0.29
	50	0.63	0.29	0.39
	100	0.86	0.29	0.44
An	Real	0.17	0.16	0.17
	10	0.30	0.23	0.26
	30	0.44	0.29	0.34
	50	0.63	0.36	0.45
	100	0.83	0.36	0.50
G	Real	0.17	0.18	0.17
	10	0.28	0.23	0.25
	30	0.49	0.39	0.43
	50	0.68	0.43	0.52
	100	0.85	0.43	0.55
Gn	Real	0.14	0.13	0.13
	10	0.28	0.20	0.23
	30	0.46	0.31	0.37
	50	0.65	0.37	0.46
	100	0.84	0.40	0.54
I	Real	0.18	0.12	0.14
	10	0.30	0.20	0.24
	30	0.46	0.26	0.33
	50	0.66	0.35	0.45
	100	0.88	0.32	0.46
In	Real	0.16	0.12	0.14
	10	0.30	0.22	0.25
	30	0.49	0.31	0.38
	50	0.63	0.38	0.47
	100	0.87	0.34	0.49
GIn	Real	0.26	0.15	0.19
	10	0.32	0.24	0.27
	30	0.51	0.37	0.42
	50	0.66	0.43	0.51
	100	0.87	0.44	0.58

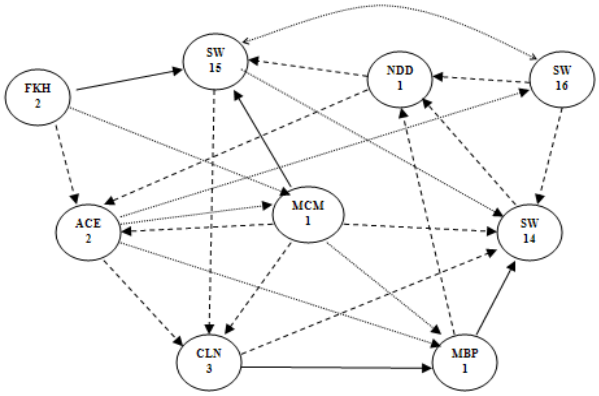
To look closely at the predicted networks by DBN with the *GIn* prior compared to DBN without prior, we give an example which is shown in Fig. 5, where solid line denotes true positive (TP), square dot line represents false positive (FP), and dash line means false negative (FN). As seen in Fig. 5, large number of FP and FN are obtained when prior information is not incorporated in DBN. On the other hand, *GIn* prior can bring more TP while keep the number of FP and FN under control.

We also study the effect of Gibbs prior parameter β on the performance of different priors. In previous experiments, this parameter was set as 1. The parameter β controls the weight of prior on GRN inference. That is to say, if one is very confident about the prior, a large value to β could be set. Fig. 6 plots the values of F-measure of different priors with the variation of β from 1 to 10. As can be seen from Fig. 6, most priors have increasing F-measure as β is raised from 1 to 5. But from 6 to 10, the increment of β does not bring any obvious improvement. To see this more clearly,

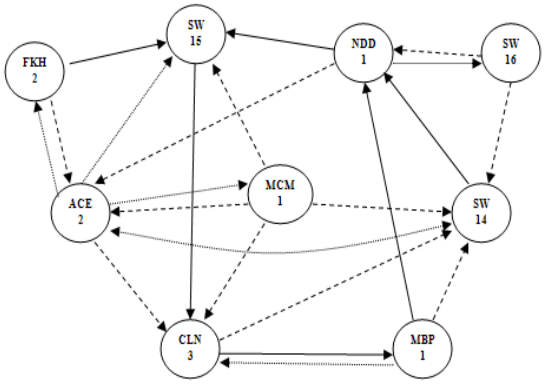
please refer to Fig. 7, which describes the variation of F-measure when β changes from 1 to 10, using benchmark network as prior. This extreme example shows that even when the prior matrix is very confident, β does not need to be too large. The results shown in Fig. 6 and 7 are from experiments on synthetic data with 30 time points. Experiments on other datasets including real data also have similar results (not shown here).



(a)



(b)



(c)

Figure 5: (a) benchmark network; (b) and (c) are predicted networks by DBN without prior and with *GIn* prior respectively.

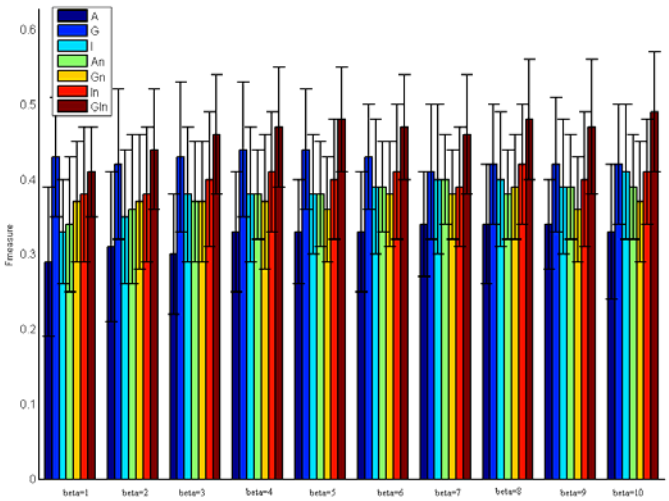


Figure 6. F-measure variation of GRN inference using different priors with beta from 1 to 10.

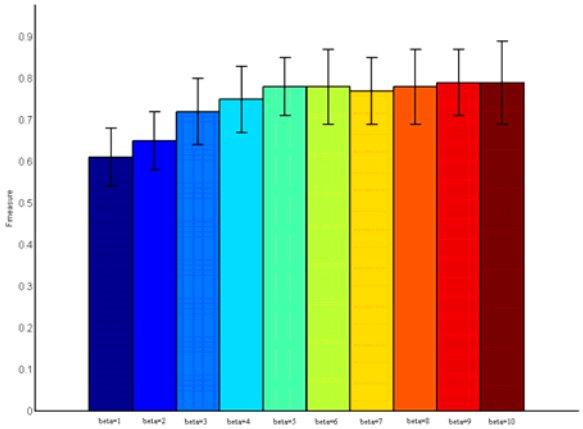


Figure 7. F-measure variation of GRN inference using benchmark as prior with beta from 1 to 10.

IV. DISCUSSIONS AND CONCLUSIONS

In this paper, we integrated epigenetic priors into dynamic Bayesian network for GRN reconstruction. We proposed several new priors from the same epigenetic profile by treating the intragenic and intergenic regions separately and using the technique of random normalization. Experiments on both real data and synthetic data of yeast show that epigenetic prior can improve the performance of GRN reconstruction significantly. According to the results, F-measure performance was improved averagely by 4% by intragenic prior (G) and 6% by combined prior of intragenic and normalized intergenic prior (GIn) for synthetic data. For real data only GIn prior provide improvement in F-measure (3%) when compare with other priors. Such improvement was obtained by increasing the number of true positives while decreasing the number of false negatives, because the higher correlations of HM profiles of confirmed gene-pairs than random pairs can serve as prior knowledge for GRN inference.

The contributions of this paper are as follows: 1) extend the work of [14] by integrating epigenetic prior in dynamic Bayesian network; 2) propose several new priors and improve the accuracy of GRN inference; 3) study the effect of weighting of prior integration for improving GRN inference.

Several priors have been proposed here. But their performance has been tested on just one benchmark network. It would be interesting to test them on other benchmark networks of yeast and other species. Here we generate our priors based on different genomic regions (i.e. intragenic and intergenic). New priors could also be generated by selecting the types of HMs. Finally, other epigenetic data especially DNA methylation are potential good priors for GRN inference.

ACKNOWLEDGMENT

This project is partially supported by Startup Grant COE_SUG/RSS_1FEB11_1/8 from Nanyang Technological University, and AcRF Tier 2 MOE2010-T2-1-056 (ARC 09/10), Ministry of Education, Singapore.

REFERENCES

[1] M. Hecker, S. Lambeck, S. Toepfer, E. van Someren, and R. Guthke, "Gene regulatory network inference: data integration in dynamic models-a review," *Biosystems,* vol. 96, pp. 86-103, Apr 2009.

[2] S. Bornholdt, "Boolean network models of cellular regulation: prospects and limitations," *Journal of the Royal Society Interface,* vol. 5, pp. S85-S94, 2008.

[3] T. S. Gardner, D. di Bernardo, D. Lorenz, and J. J. Collins, "Inferring genetic networks and identifying compound mode of action via expression profiling," *Science Signalling,* vol. 301, p. 102, 2003.

[4] J. M. Stuart, E. Segal, D. Koller, and S. K. Kim, "A gene-coexpression network for global discovery of conserved genetic modules," *Science,* vol. 302, pp. 249-255, 2003.

[5] D. Heckerman, "A tutorial on learning with Bayesian networks," *Innovations in Bayesian Networks,* pp. 33-82, 2008.

[6] N. Friedman, M. Linial, I. Nachman, and D. Pe'er, "Using Bayesian networks to analyze expression data," *Journal of computational biology,* vol. 7, pp. 601-620, 2000.

[7] C. J. Needham, J. R. Bradford, A. J. Bulpitt, and D. R. Westhead, "A primer on learning in Bayesian networks for computational biology," *PLoS computational biology,* vol. 3, p. e129, 2007.

[8] K. Murphy and S. Mian, "Modelling gene expression data using dynamic Bayesian networks," Technical report, Computer Science Division, University of California, Berkeley, CA1999.

[9] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika,* vol. 57, pp. 97-109, 1970.

[10] H. Yu, S. Zhu, B. Zhou, H. Xue, and J. D. J. Han, "Inferring causal relationships among different histone modifications and gene expression," *Genome research,* vol. 18, pp. 1314-1324, 2008.

[11] J. Su, Y. Qi, S. Liu, X. Wu, J. Lv, H. Liu, R. Zhang, and Y. Zhang, "Revealing epigenetic patterns in gene regulation through integrative analysis of epigenetic interaction network," *Mol Biol Rep,* vol. 39, pp. 1701-12, Feb 2012.

[12] X. Xu, S. Hoang, M. W. Mayo, and S. Bekiranov, "Application of machine learning methods to histone methylation ChIP-Seq data reveals H4R3me2 globally represses gene expression," *BMC Bioinformatics,* vol. 11, p. 396, 2010.

[13] C. Cheng, K. K. Yan, K. Y. Yip, J. Rozowsky, R. Alexander, C. Shou, and M. Gerstein, "A statistical framework for modeling gene expression using chromatin features and application to modENCODE datasets," *Genome Biol,* vol. 12, p. R15, 2011.

[14] J. Zheng, I. Chaturvedi, and J. Rajapakse, "Integration of Epigenetic Data in Bayesian Network Modeling of Gene Regulatory Network," *Pattern Recognition in Bioinformatics,* pp. 87-96, 2011.

[15] S. Chib and E. Greenberg, "Understanding the metropolis-hastings algorithm," *The American Statistician,* vol. 49, pp. 327-335, 1995.

[16] D. Husmeier and A. V. Werhli, "Bayesian integration of biological prior knowledge into the reconstruction of gene regulatory networks with Bayesian networks," in *Comput Syst Bioinformatics Conf*, 2007, pp. 85-95.

[17] D. K. Pokholok, C. T. Harbison, S. Levine, M. Cole, N. M. Hannett, T. I. Lee, G. W. Bell, K. Walker, P. A. Rolfe, and E. Herbolsheimer, "Genome-wide map of nucleosome acetylation and methylation in yeast," *Cell,* vol. 122, pp. 517-527, 2005.

[18] A. Fujita, J. R. Sato, H. M. Garay-Malpartida, R. Yamaguchi, S. Miyano, M. C. Sogayar, and C. E. Ferreira, "Modeling gene expression regulatory networks with the sparse vector autoregressive model," *BMC Systems Biology,* vol. 1, p. 39, 2007.

[19] J. C. Rajapakse and P. A. Mundra, "Stability of building gene regulatory networks with sparse autoregressive models," *BMC Bioinformatics,* vol. 12, p. S17, 2011.

[20] S. Balaji, M. M. Babu, L. M. Iyer, N. M. Luscombe, and L. Aravind, "Comprehensive analysis of combinatorial regulation using the transcriptional regulatory network of yeast," *J Mol Biol,* vol. 360, pp. 213-27, Jun 30 2006.

[21] P. T. Spellman, G. Sherlock, M. Q. Zhang, V. R. Iyer, K. Anders, M. B. Eisen, P. O. Brown, D. Botstein, and B. Futcher, "Comprehensive identification of cell cycle–regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization," *Molecular biology of the cell,* vol. 9, pp. 3273-3297, 1998.

[22] D. Husmeier. "Sensitivity and specificity of inferring genetic regulatory interactions from microarray experiments with dynamic Bayesian networks," *Bioinformatics,* vol. 19, pp. 2271-2282, 2003