Multivariate Mutual Information Measures for Discovering Biological Networks

Tho Hoan Pham*, Tu Bao Ho†, Quynh Diep Nguyen*, Dang Hung Tran* and Van Hoang Nguyen‡
*Hanoi National University of Education
136 Xuan-Thuy, Cau-Giay district, Hanoi, Vietnam
Email: {hoanpt, diepnq, hungtd}@hnue.edu.vn
†Japan Advanced Institute of Science and Technology
1-1 Asahidai, Nomi-city, Ishikawa, 1211-1292, Japan Email: bao@jaist.ac.jp
‡ Hanoi University of Agriculture
Trau-Quy, Gia-Lam district, Hanoi
Email: hoangnv@hua.edu.vn

***Abstract*—Most studies on biological networks until now focus only on pairwise interactions/relationships. However interactions/relationships participated by more than two molecules are popular in biology. In this paper, we introduce multivariate mutual information measures to reconstruct multivariate interactions/relationships in biological networks.**

I. Introduction

Biological network models consist of nodes and edges where nodes represents biological molecules and edges represent some kinds of relationships between them. Examples include gene regulatory networks that consist of genes as nodes and transcriptional regulations between them as edges; signal transduction networks that consist of proteins as nodes and activations/inactivations across a set of genes as edges; and metabolic networks that consist of metabolites as nodes and reactions as edges. Thanks to high-throughput technologies that can produce simultaneous measurements of the concentration/expression of thousands of molecular species in a biological system such as genes, proteins and metabolites, we can understand a holistic view of complex interactions by systems theory. Reconstruction of biological networks from the high-throughput profiling data is one of the most challenges and the first step on the road to the ultimate understanding of complex biological systems.

Computational methods for reverse engineering biological networks can be divided into three broad categories. The first one includes information theoretic methods [1], [2], [3] that find statistical dependences between two molecules (dependence here refers to any situation in which random variables do not satisfy a mathematical condition of probabilistic independence). These methods rely on two basic measures of information theory: mutual information and correlation coefficient. The former is more preferred since it takes account any type of dependence while the later is responsive for linear dependence only (see [4] for a comparison). The second consists of Bayesian and graphical networks [5], [6] that maximize a scoring function over some alternative network models to find the best one fitting the data. Since the network model space is often very huge, these graphical methods use some searching heuristic strategies from specific initial networks, in consequence they might get stuck in local optima. The last reconstruction approach includes differential and difference equations [7], [8] that explain the data by a system of mathematical differential and difference equations. Due to the computational complexity, only linear or simple functions are considered in the mathematical equation models. Some excellent reviews on different aspects of biological network reconstructing methods can be found in [9], [10], [11].

Biological networks are often too large; they might consist of thousands of nodes. Since the information theoretic approach has advantages of simplicity and low computational costs, it is superior to the two other approaches in reconstructing biological networks [3]. Graphical models or mathematical equations often incorporate with some information theoretic measures to reduce their search space so that they can work on large biological networks [12], [13]. Therefore information theoretic measures are the key to the success of most biological network reconstruction approaches.

Information theoretic approach assumes that interactions/relationships of molecules (for example, transcriptional regulations, activations/inactivations, reactions) will exhibit some dependencies among them, and therefore such statistical dependencies among variables can be used to reconstruct the original interactions/relationships. Most network reconstruction methods until now mainly focused on pairwise relationships between two variables. However, relationships in real biological networks are often complicated and might consist of more than two variables (molecules), for example, proteins or genes often interact with others to carry out their functions; substrates and products often mutually interact in a reaction. These multi-variable dependencies could not be reducible/represented by pairwise dependencies because the reduction will loss information.

There have been efforts to reconstruct statistical multivariable interactions by using some extensions of mutual information for multiple variables [14], [15]. However, the extension of mutual information from two to multiple variables is not trivial, even for the simplest case of three variables [16],

[17]. There has been little work on such higher-order mutual information, partly due to the fact that the notion of multi-variable dependence itself remains imprecisely defined. There have been two multiple mutual information measures that extend from the pairwise one: total mutual information and interaction information [18], [19], [20], [21]. While the first one is clearly understood and interpreted, there have been a lot of controversy on the interpretation of the second [22], [17]. Forever and most importantly, there has been no research work that systematically identifies different types of dependences existing among multiple variables and provides appropriate measures to capture these multi-variable dependencies. In the previous work [23], we have propose a general mutual information of multi variables that has many variants, each quantifies a type of dependences.

In this work, we want to confirm that our previously proposed mutual information quantities are suitable to reconstruct multi variable relationships in biological networks.

II. Methods

A. Mutual Information of Two Variables

Mutual information of two random variables is a quantity that measures the mutual dependence between the two random variables. The mutual information of two discrete random variables X and Y can be defined as

$$MI(X,Y) = \sum_{x\in X}\sum_{y\in Y} p_{X,Y}(x,y)\log\frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)} \quad (1)$$

where $p_{X,Y}(x,y)$ is the joint probability distribution function of X and Y, and $p_X(x)$ and $p_Y(y)$ are the marginal probability distribution functions of X and Y respectively. In the case of continuous random variables, the above formula of mutual information can be written as:

$$MI(X,Y) = \int\int p_{X,Y}(x,y)\log\frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)}dxdy \quad (2)$$

where $p_{X,Y}(x,y)$ is now the joint probability density function of X and Y, $p_X(x)$ and $p_Y(y)$ are the marginal probability density functions of X and Y respectively.

If X and Y are independent, the mutual information $MI(X,Y) = 0$; if they are perfectly dependent, $MI(X,Y)$ approaches infinity.

The mutual information $MI(X,Y)$ can also be interpreted in terms of information entropy [24] as

$$MI(X,Y) = H(X) + H(Y) - H(X,Y) \quad (3)$$

In [23] we have provided a novel interpretation of the mutual information that bases on the physical view of entropy. That is, mutual information of two variables is the difference between entropy of the joint probability distribution (or density) $p_{X,Y}(x,y)$ of X and Y and that of the fictitious distribution (or density) with the same marginal distributions (or densities) $p_X(x)$ and $p_Y(y)$ but no dependence introduced; i.e.

$$MI(X,Y) = H(p_X * p_Y) - H(p_{X,Y}) \quad (4)$$

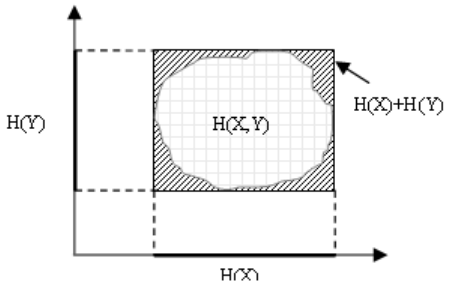

Fig. 1. Mutual information of two variables

Since $H(p_X * p_Y) = H(p_X) + H(p_Y) = H(X) + H(Y)$, the Eq. 4 is equivalent to the Eq. 3.

With this new interpretation, we can visualize the mutual information as the difference of the rectangle that represents $p_X * p_Y$ and the shape that represents the joint distribution (Fig. 1).

B. Mutual Information Measures for Multiple Variables

Generalization of the mutual information for multiple variables is not trivial [22], [14]. In the previous work [23], we have analysed that there are many types of mutual information existing in more than two variables $\{X_1,\ldots,X_n\}$. Each type of mutual information corresponds to a marginal probability distribution or corresponds to a partition $\{D_1,\ldots,D_k\}$ of $\{X_1,\ldots,X_n\}$.

$$\{X_1,\ldots,X_n\} = D_1 \oplus \ldots \oplus D_k \quad (5)$$

The mutual information of $\{X_1, X_2, \ldots, X_n\}$ respective to an interested schedule $\{D_1, D_2, \ldots, D_k\}$ can be defined as follows:

$$MI_{\{D_1,\ldots,D_k\}}(X_1,\ldots,X_n) = H(p_{D_1} * \cdots * p_{D_k}) - H(X_1,\ldots,X_n) \quad (6)$$

or equivalent to

$$MI_{\{D_1,\ldots,D_k\}}(X_1,\ldots,X_n) = H(D_1) + \ldots + H(D_k) - H(X_1,\ldots,X_n) \quad (7)$$

where p_{D_i} is a marginal probability distribution of the joint probability distribution $p(X_1,\ldots,X_n)$ on a subset D_i of variables

For example, with three variables, the general mutual information (Eq. 7 has following variants:

$$MI_{\{X,Y,Z\}}(X,Y,Z) = H(X)+H(Y)+H(Z)-H(X,Y,Z) \quad (8)$$

$$MI_{\{X,<Y,Z>\}}(X,Y,Z) = H(X)+H(Y,Z)-H(X,Y,Z) \quad (9)$$

$$MI_{\{Y,<Z,X>\}}(X,Y,Z) = H(Y)+H(Z,X)-H(X,Y,Z) \quad (10)$$

$$MI_{\{Z,<X,Y>\}}(X,Y,Z) = H(Z)+H(X,Y)-H(X,Y,Z) \quad (11)$$

The Eq. 8 is the *total mutual information* that has been previously defined in [19] as an extension of mutual information. The total mutual information can be visualized as the divergence between dashed shape and the parallelepiped in Fig. 2.

The Equations 9, 10, 11 are mutual information between a single variable and the two others. The mutual information respective to the schedule $(Z, < X, Y >)$ can be visualized as

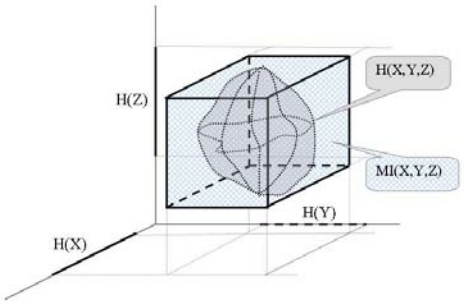

Fig. 2. Total mutual information of three variables

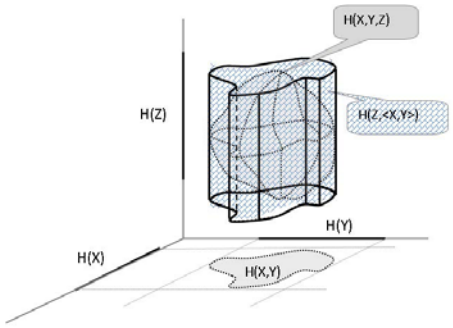

Fig. 3. Mutual information of three variables respective to the schedule $(Z, < X, Y >)$

the divergence between the cylinder (representing $p_Z * p_{X,Y}$) and the dashed shape (representing $p_{X,Y,Z}$) in Fig. 3. We refer these mutual information quantities to *cylinder mutual information measures*.

Each mutual information quantity provides a dependent/independent aspect of a multivariate relationship. We should estimate all mutual information quantities to understand full multivariate interactions/relationships [23].

III. Results and Discussion

A. Reconstructing Pairwise and Triple Interactions from Synthetic Data

In this section, we will illustrated the performance of mutual information (MI) measures, particularly cylinder MI and total MI (see Section II-B, to reconstruct pairwise and triple interactions from some synthetic datasets.

The first dataset consists of 4 binary data points $\{(0,0,0);(0,1,1);(1,0,1);(1,1,0)\}$ (see Fig. 4). We can see that there is a triple relationship *X xor Y xor Z =0*, but there isn't any pairwise relationship in three binary variables X, Y and Z.

As expectation, the pairwise mutual informations of any two of three variables are all zero (see Table I). This demonstrates that there isn't pairwise dependence in any pair of these three variables. However, cylinder mutual informations as well as total mutual information are all 1. This proves that there is a relationship concerning simultaneously all three variables. In this example, multivariate mutual information quantities have been useful to detect multivariate relationships that might be missed if we use only pairwise mutual information measures.

The absolute value of cylinder mutual informations and total mutual information also provide us a complete understanding of relationships among three variables. Another example is the dataset on Fig. 5 where the role of X, Y and Z are different. Table II shows mutual information measures that

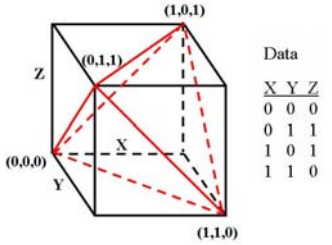

Fig. 4. A synthetic dataset where there is a three-variate interaction $XxorYxorZ = 0$, but it does not show any pairwise relationship between two of three binary variables

TABLE I
Relationships reconstructed from the dataset in Fig. 4 by mutual information measures

Mutual information quantity	Reconstructed relationships
Pairwise MI	
$MI(X,Y) = 0$ $MI(Y,Z) = 0$ $MI(Z,X) = 0$	
Cylinder MI	
$MI(X, < Y, Z >) = 1$	X and $< Y, Z >$
$MI(Y, < Z, X >) = 1$	Y and $< Z, X >$
$MI(Z, < X, Y >) = 1$	Z and $< X, Y >$
Total MI	
$MI(X,Y,Z) = 1$	(X,Y,Z)

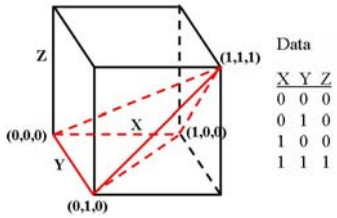

Fig. 5. Another synthetic dataset of three binary variables

TABLE II
Reconstruction of relationships from the dataset in Fig. 5 by mutual information measures

Mutual information quantity	Reconstructed relationships
Pairwise MI	
$MI(X,Y) = 0$	
$MI(Y,Z) = 0.36$	Y and Z
$MI(Z,X) = 0.36$	Z and X
Cylinder MI	
$MI(X, < Y, Z >) = 0.45$	X and $< Y, Z >$
$MI(Y, < Z, X >) = 0.45$	Y and $< Z, X >$
$MI(Z, < X, Y >) = 0.81$	Z and $< X, Y >$
Total MI	
$MI(X,Y,Z) = 0.81$	(X,Y,Z)

indicate the strength of reconstructed relationships. Among relationships found, we can see the relationship between Z and $< X, Y >$ is the strongest. That agrees with the dataset in Fig. 5. Clearly, different mutual information measures provide us different aspects of the relationships.

B. Reconstructing Metabolic Networks

In this section, we aim to experimentally verify the performance of the mutual information quantities on reconstruction of metabolic reactions, which are interactions of multi substrates/products, from metabolome data. Due to the time complexity of mutual information estimation, in this study, we only use the three variate mutual information measures (MIs): cylinder MIs and total three variate MI as in the previous section III-A.

In the first experiment, we use *in silico* metabolome data of red blood cell metabolism (RBC) published by [25]. The dataset is a 1000×39-matrix that are concentrates of 39 metabolites in the RBC model at 1000 time series points. The datasets can be downloaded at *http://menem.com/∼ilya/wiki/index.php/RBC_Metabolic_Network*. The RBC model consisting of 39 metabolites and 44 reactions will be used to validate relationships reconstructed by the mutual information measures. In this work, we use the numerical estimation method of mutual information quantities using B-spline functions as described in [26].

Table III shows 50 top three-variate relationships with the highest total MI. These three-variate relationships are often substrates or products of the same reaction or adjacent reactions in the RBC models. For example, {PG3, PG2, PEP} concern with two adjacent reactions pgm (PG3, PG2) and en (PG2, PEP); {G6P, F6P, GO6P} are three substrates/products concerning glucose synthesis; {G6P, F6P, X5P} concern with two adjacent reactionns pgi (G6P F6P) and $tkii$ (X5P, E4P, GAP, F6P); etc. Note that. the notation $ru5pi$ (RU5P, R5P) describes that the reaction $ru5pi$ consists of two metabolites RU5P and R5P. We can look at different mutual information quantities to understand the full dependent/independent pictures among found three-variable relationships.

The second experiment we aim to evaluate the performance of the total three-variable mutual information (Eq. 8) on the reconstruction ability of three-variable interactions/relationships based on the receiver operating characteristic (ROC) criterion. We first use a small metabolic model (consisting of 10 metabolites and 5 reactions) and generate a time series dataset using *Matlab*. We use the total three-variable mutual information on this dataset to reconstruct three-variable interactions and then use the original model to validate reconstructed interactions. We use different thresholds of the total three-variable mutual information to generate ROC. Figure 6 shows the ROC curve of the total three-variable mutual information-based classifier on this three-variable interaction reconstruction. The area under the curve is quite high (0.85).

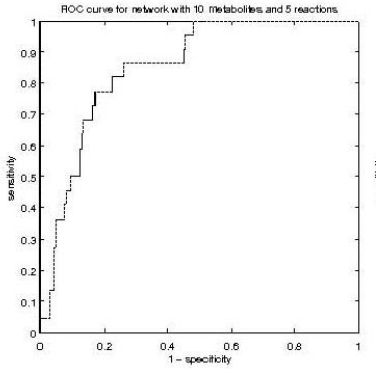

Fig. 6. ROC curve of the total mutual information-based classifier on the reconstruction of three-variable interactions

IV. Conclusion

We have introduced a general mutual information to capture multivariate interactions/relationships. The general formulate has many variants that provide different dependent/independent aspects of the multivariate interactions/relations. We have experimentally confirmed that three-variate mutual information quantities are appropriate measures to reconstruct multivariate interactions/relationships in both synthetic networks as well as metabolic networks.

In this work, we just considered the three-variable mutual information and found that many three-variable interactions/relationships cannot be detected if we use pairwise mutual information quantities only. More generally, many n-variable interactions/relationships could be reconstructed only if we use n-variable mutual information. Therefore, multivariate mutual information measures are important to reconstruct complicated interactions/relationships in systems biology.

Acknowledgment

References

[1] R. Steuer, J. Kurths, C. Daub, J. Weise, and J. Selbig, "The Mutual Information: Detecting and Evaluating Dependencies between Variables." *Bioinformatics*, vol. 2002, no. 18, Suppl 2, pp. S231–S240, 2002.

[2] A. Butte and I. Kohane, "Mutual information relevance networks: Functional genomic clustering using pairwise entropy measurements." *Pacific Symposium on Biocomputing*, pp. 418–429, 2000.

[3] A. Margolin, I. Nemenman, K. Basso, C. Wiggins, G. Stolovitzky, R. Favera, and A. Califano, "ARACNE: an Algorithm for the Reconstruction of Gene Regulatory Networks in a Mammalian Cellular Context." *BMC Bioinformatics*, vol. 7 Sppl. 1:S7, 2006.

[4] D. Camach, P. Licona, P. Mendes, and R. Laubenbacher, "Comparison of Reverse-Engineering Methods Using an in Silico Network." *Ann. N.Y. Acad. Sci.*, vol. doi: 10.1196/annals.1407.006, 2007.

[5] N. Friedman, M. Linial, I. Nachman, and D. Peer, "Using Bayesian Networks to Analyze Expression Data." *J. Comput. Biol.*, vol. 7, pp. 601–620, 2000.

[6] A. Werhli and D. Husmeier, "Reconstructing Gene Regulatory Networks with Bayesian Networks by Combining Expression Data with Multiple Sources of Prior Knowledge." *Statistical Applications in Genetics and Molecular Biology*, vol. Vol. 6 : Iss. 1, Article 15, 2007.

[7] T. Gardner, D. Bernardo, D. Lorenz, and J. Collins, "Inferring Genetic Networks and Identifying Compound Mode of Action via Expression Profiling ," *Science*, vol. 301, pp. 102–105, 2003.

[8] M. Bansal, G. Gatta, and D. Bernardo, "Inference of gene regulatory networks and compound mode of action from time course gene expression profiles." *Bioinformatics*, vol. 22(7), pp. 815–822, 2006.

TABLE III
RECONSTRUCTION OF METABOLIC NETWORKS BY MUTUAL INFORMATION QUANTITIES

Relationships (X,Y,Z)	MI(X,Y)	MI(Y,Z)	MI(Z,X)	MI($X, < Y, Z >$)	MI($Y, < Z, X >$)	MI($Z, < X, Y >$)	MI(X,Y,Z)
PG3, PG2, PEP	0.54	0.54	0.54	0.68	0.68	0.68	1.22
G6P, F6P, GO6P	0.49	0.48	0.48	0.62	0.62	0.62	1.11
G6P, F6P, X5P	0.49	0.48	0.48	0.62	0.62	0.61	1.10
G6P, F6P, RU5P	0.49	0.48	0.48	0.62	0.62	0.61	1.10
G6P, F6P, R5P	0.49	0.48	0.48	0.62	0.62	0.61	1.10
G6P, GO6P, X5P	0.48	0.47	0.48	0.62	0.61	0.61	1.09
F6P, GO6P, X5P	0.48	0.47	0.48	0.62	0.61	0.61	1.09
G6P, GO6P, RU5P	0.48	0.47	0.48	0.62	0.61	0.61	1.09
F6P, GO6P, RU5P	0.48	0.47	0.48	0.62	0.61	0.61	1.09
G6P, GO6P, R5P	0.48	0.47	0.48	0.62	0.61	0.61	1.09
F6P, GO6P, R5P	0.48	0.47	0.48	0.62	0.61	0.61	1.09
G6P, F6P, R1P	0.49	0.47	0.47	0.62	0.62	0.60	1.09
G6P, RU5P, X5P	0.48	0.48	0.48	0.61	0.61	0.61	1.09
F6P, RU5P, X5P	0.48	0.48	0.48	0.61	0.61	0.61	1.09
G6P, R5P, X5P	0.48	0.48	0.48	0.61	0.61	0.61	1.09
G6P, RU5P, R5P	0.48	0.48	0.48	0.61	0.61	0.61	1.09
F6P, R5P, X5P	0.48	0.48	0.48	0.61	0.61	0.61	1.09
F6P, RU5P, R5P	0.48	0.48	0.48	0.61	0.61	0.61	1.09
RU5P, R5P, X5P	0.48	0.48	0.48	0.61	0.61	0.61	1.09
G6P, GO6P, R1P	0.48	0.46	0.47	0.62	0.61	0.60	1.08
F6P, GO6P, R1P	0.48	0.46	0.47	0.62	0.61	0.60	1.08
G6P, RU5P, R1P	0.48	0.47	0.47	0.61	0.61	0.60	1.08
G6P, X5P, R1P	0.48	0.47	0.47	0.61	0.61	0.60	1.08
F6P, RU5P, R1P	0.48	0.47	0.47	0.61	0.61	0.60	1.08
F6P, X5P, R1P	0.48	0.47	0.47	0.61	0.61	0.60	1.08
GO6P, RU5P, X5P	0.47	0.48	0.47	0.61	0.61	0.61	1.08
G6P, R5P, R1P	0.48	0.47	0.47	0.61	0.61	0.60	1.08
F6P, R5P, R1P	0.48	0.47	0.47	0.61	0.61	0.60	1.08
GO6P, R5P, X5P	0.47	0.48	0.47	0.61	0.61	0.61	1.08
GO6P, RU5P, R5P	0.47	0.48	0.47	0.61	0.61	0.61	1.08
RU5P, X5P, R1P	0.48	0.47	0.47	0.61	0.61	0.60	1.08
RU5P, R5P, R1P	0.48	0.47	0.47	0.61	0.61	0.60	1.08
R5P, X5P, R1P	0.48	0.47	0.47	0.61	0.61	0.60	1.08
GO6P, RU5P, R1P	0.47	0.47	0.46	0.60	0.61	0.60	1.07
GO6P, X5P, R1P	0.47	0.47	0.46	0.60	0.61	0.60	1.07
DPG13, PG2, PEP	0.44	0.54	0.44	0.54	0.65	0.65	1.08
GO6P, R5P, R1P	0.47	0.47	0.46	0.60	0.61	0.60	1.07
DPG13, PG3, PEP	0.44	0.54	0.44	0.54	0.64	0.64	1.08
DPG13, PG3, PG2	0.44	0.54	0.44	0.54	0.64	0.64	1.08
G6P, F6P, PRPP	0.49	0.46	0.46	0.62	0.62	0.59	1.07
RU5P, X5P, PRPP	0.48	0.46	0.46	0.61	0.61	0.60	1.08
RU5P, R5P, PRPP	0.48	0.46	0.46	0.61	0.61	0.60	1.07
R5P, X5P, PRPP	0.48	0.46	0.46	0.61	0.61	0.60	1.07
G6P, RU5P, PRPP	0.48	0.46	0.46	0.61	0.62	0.59	1.07
F6P, RU5P, PRPP	0.48	0.46	0.46	0.61	0.62	0.59	1.07
G6P, X5P, PRPP	0.48	0.46	0.46	0.61	0.62	0.59	1.07
F6P, X5P, PRPP	0.48	0.46	0.46	0.61	0.62	0.59	1.07
G6P, R5P, PRPP	0.48	0.46	0.46	0.61	0.61	0.59	1.07
F6P, R5P, PRPP	0.48	0.46	0.46	0.61	0.61	0.59	1.07
RU5P, PRPP, R1P	0.46	0.46	0.47	0.61	0.60	0.60	1.07

[9] M. Styczynski and G. Stephanopoulos, "Overview of computational methods for the inference of gene regulatory networks." *Computers & Chemical Engineering*, vol. 29, no. 3, pp. 519–534, 2005.

[10] K. Cho, S. Choo, S. Jung, J. Kim, H. Choi, and J. Kim, "Reverse engineering of gene regulatory networks." *IET Syst Biol.*, vol. 1, no. 3, pp. 149–163, 2007.

[11] M. Heckera, S. Lambecka, S. Toepferb, E. Somerenc, and R. Guthke, "Gene regulatory network inference: Data integration in dynamic models - A review." *Biosystems*, vol. 96, no. 1, pp. 86–103, 2009.

[12] C. Aliferis, I. Tsamardinos, A. Statnikov, and L. Brown, "Causal Explorer: A Probabilistic Network Learning Toolkit for Biomedical Discovery." *in Proceedings of International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences (METMBS), Las Vegas, Nevada, USA*, pp. 371–376, 2003.

[13] I. Tsamardinos, L. Brown, and C. Aliferis, "The max-min hill-climbing Bayesian network structure learning algorithm." *Machine Learning*, vol. 65, no. 1, pp. DOI10.1007/s10 994–006–6889–7, 2006.

[14] D. Anastassiou, "Computational Analysis of the Synergy among Multiple Interacting Genes." *Molecular Systems Biology*, pp. 3:83, doi:10.1038/msb4 100 124, 2007.

[15] A. Margolin, K. Wang, A. Califano, and I. Nemenman, "Multivariate dependence and genetic networks inference." *IET Syst. Biol.*, vol. 4(6):428, 2010.

[16] A. Jakulin and I. Bratko, "Quantifying and Visualizing Attribute Interactions: An Approach Based on Entropy." *CoRR*, vol. cs.AI/0308002, http://arxiv.org/abs/cs.AI/0308002, 2004.

[17] L. Leydesdorff, "Interaction Information: Linear and Nonlinear Interpretations." *Int. J. General Systems*, vol. 36, no. 6, pp. 681–685, 2009.

[18] W. McGill, "Multivariate information transmission." *Psychometrika*, vol. 19, no. 2, pp. 97–116, 1954.

[19] S. Watanabe, "Information Theoretical Analysis of Multivariate Correlation." *IBM Journal of Research and Development*, pp. 66–82, 1960.

[20] R. Fano, "Transmission of Information." *New York NY, USA: MIT press*, 1961.

[21] T. Han, "Multiple mutual information and multiple interactions in frequency data." *Information and Control*, vol. 46, p. 2645, 1980.
[22] A. Jakulin, "Machine learning based on attribute interactions." *University of Ljubljana*, vol. available at http://stat.columbia.edu/ jakulin/Int/jakulin05phd.pdf, 2005.
[23] Q. Nguyen, T. Pham, T. Ho, H. Nguyen, T. Nguyen, and D. Tran, "Extension and Visualization of Mutual Information for Multiple Variables." *National Conference on Fundamental and Applied IT Research, DongNai, Vietnam*, 2011.
[24] T. Cover and J. Thomas, "Elements of Information Theory (Second edition)." *Molecular Systems Biology*, vol. Wiley-Interscience, A John wiley & Sons, Inc., Publication, 2006.
[25] I. Nemenman, G. Escola, W. Hlavacek, P. Unkefer, C. Unkefer, and M. Wall, "Reconstruction of Metabolic Networks from High-throughput Metabolite Profiling Data: in silico Analysis of Red Blood Cell Metabolism." *Ann N. Y. Acad Sci.*, vol. 1115, pp. 102–115, doi: 10.1196/annals.1407.013, 2007.
[26] C. Daub, R. Steuer, J. Selbig, and S. Kloska, "Estimating Mutual Information using B-spline Functions–an Improved Similarity Measure for Analysing Gene Expression Data." *BMC Bioinformatics*, vol. 5:118, doi:10.1186/1471-2105-5-118, 2004.