International Journal of Artificial Intelligence and Neural Networks
Volume 6 : Issue 1 *[ISSN 2250-3749]*
Publication Date : 31 August, 2016

Reconstruction of Gene Regulatory Network from Gene Perturbation Data, Current Methods and Problems

Winddy, Sharifallilah, Ali and Saifulaman

***Abstract*—The inference of regulators is the core factor in interpreting the actual regulatory conditions in gene regulatory networks (GRNs). Various methods have been developed to reconstruct GRNs with the motivation of improving the accuracy and scalability of network inference. Thus, this study will brief the structure of GRNs, discuss current methods of GRNs reconstruction and problems when dealing with gene perturbation data. Most of the information gathered from bioinformatics and system biology literature. At the end of the study several of GRNs reconstruction methods will be reviewed and identified their problem when dealing with gene perturbation data. This study is useful as a reference to develop more accurate GRNs inference methods particular for gene perturbation data.**

***Keywords*— gene regulatory network, reconstruction of gene regulatory networks (GRNs), gene expression data, gene perturbation, machine learning methods**

I. Introduction

Reconstruction of gene regulatory networks (GRNs) is the core factor in interpreting the actual regulatory conditions in GRNs [1]. It helps us to understand the working mechanisms of the cell in pathophysiological conditions [2]. GRNs describe control at the gene expression level and could be inferred from microRNAs (miRNAs), regulatory motifs, gene expression profiles and interactions between regulatory targets [3]. This provides a clearer blueprint on the relationship between the genes that affect the expression of other genes and adequately describing these effects [4]. The importance of gene interaction research studies has provided several useful applications such as the identification and discovery of potential targets for therapeutic intervention in diseases such as cancer [5]. Thus, a number of techniques have been proposed for GRNs reconstruction. Here we will first brief the structure of GRN, discuss the current methods for GRNs reconstruction, and present some problems when dealing with gene perturbation data.

Winddy Pindah
Universiti Teknologi MARA (UiTM), Malaysia
shendy1015@gmail.com

Sharifallilah Nordin
Universiti Teknologi MARA (UiTM), Malaysia
sharifa@tmsk.uitm.edu.my

Ali Seman
Universiti Teknologi MARA (UiTM), Malaysia
aliseman@tmsk.uitm.edu.my

Mohamed Saifulaman Mohamed Said
Universiti Teknologi MARA (UiTM), Malaysia
drsaifulaman@yahoo.com

II. Structure of Gene Regulation Network (GRN)

GRN is a model in which nodes corresponding to gene-activities and can be represented in mixed graph, $G = (V, U, D)$ over a set V of nodes, with ordered pairs D, the directed edges and unordered pairs U, the indirected edges. A directed edge d_{ij} from v_i to v_j is present if a causal effect run from node v_i to v_j and there exist no nodes or subsets of nodes in V that are intermediating the causal influence (it may be mediated by hidden variables, i.e. variables not in V). An indirected edges u_{ij} between nodes v_i to v_j is present if gene-activities v_i and v_j are associated due to confounding, and there exist no nodes or subsets of nodes in V that explain that association (it is caused by a variable hidden to V).

The nodes in GRNs structure representing the genes, proteins, metabolites, their complexes or even modules while the edges represent direct or indirect interactions between nodes. Proteins and metabolites appear as hidden variables and GRNs are inferred only from gene expression data as observable variables. These hidden variables can model unobserved effects that cannot be measured. Fig. 1 presents the projection of interactions from the space of metabolites and proteins in genes space. Dashed lines represent gene interactions and the full lines represent the interactions among genes, proteins, metabolites and their complexes [6], [7].

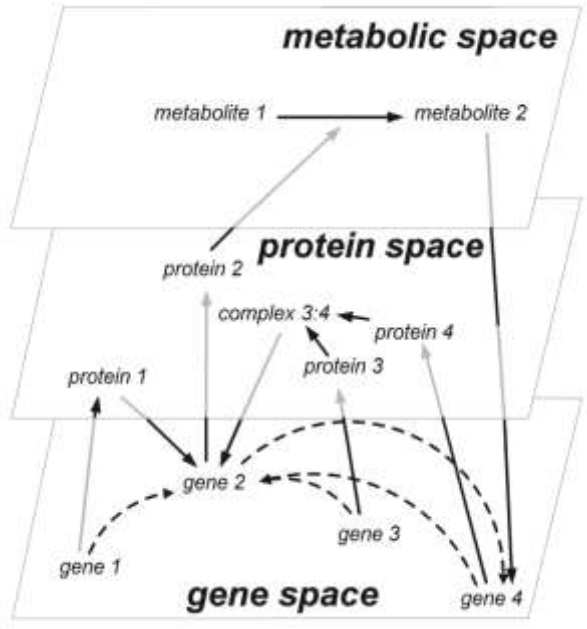


Figure 1. Projection of GRN in different spaces [7].

A gene may directly influence the activity of other target gene or gene product. Influence may be indirect by coding a transcription factor (TF) that in turn regulates another gene. A possible causal relationship in GRN is shown in Fig. 2. Apparently, four different types of causal relationship may be possible in a living cell. Based on the above figure we can derive the following causal relationship [8], [9].

- A gene can enhance the activity of more than one gene (relationship between A, B, C, and D).
- A gene's activity may be influenced by more than one gene (relationship between B, D, and F). Often F is referred as Collider [9], [10].
- Gene can also influence the activity of itself (node B).
- A gene may inhibit the activity of another gene (D inhibits E). Inhibition or negative regulation may also follow above three relationships, i.e., many to one, one-to-many, and self.

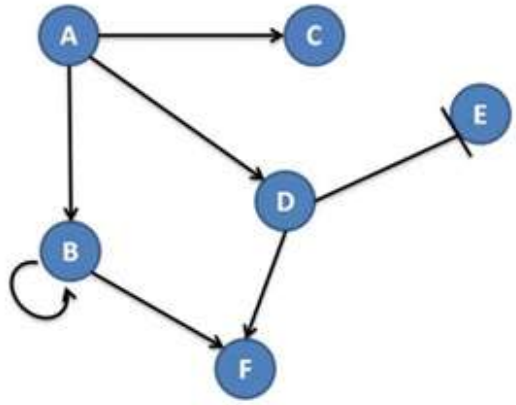


Figure 2. Possible causal dependency in GRN graph [9].

III. Algorithms for Reconstruction of GRN

Recently, various methods have been developed to reconstruct GRNs by producing hypotheses about the presence or absence of interactions among genes, hypotheses that can later be tested by laboratory experiments [4], [11]. In general, these GRNs reconstruction methods fall into two categories, namely model-based methods and machine learning-based methods [12], [13]. For model-based methods, chemical reaction of transcription and translation, as well as other cellular processes are described as linear or nonlinear differential equations, in which the parameters represent the regulation strengths of the regulators. Among the algorithms in this category include multiple linear regression [14], singular value decomposition method [12], [15]–[18], network component analysis [19], structural equation model (SEM) [20], sparse vector autoregressive [19] and linear programming. Adding to the list of methods is hybrid method [21]–[24], in which researchers have developed a novel method by incorporating more than one method.

Regression analysis widely used for prediction and forecasting, where its use has substantial overlap with the field of machine learning [25]. Multiple regression analysis is a very advanced statistical tool and it is extremely powerful when you are trying to develop a "model" for predicting a wide variety of outcomes [26]. Since the nature of GRNs that consists of simultaneous observation and analysis of more than one outcome variable [13], multiple regression analysis wise choice to reconstruct GRNs. There are a number of methods in this category, such as Multiple Linear Regression [27], Principle Component Regression [28], Partial Least Squares [29], Least Absolute Shrinkage [1] and Selection Operator (LASSO) [30] and Canonical Correlation Analysis [31]. While the linear regression model consists of a deterministic part and a random part, generally defined as

$$y = \beta_0 + \beta_1 x + \epsilon \quad (1)$$

The deterministic portion of the model,

$$\beta_0 + \beta_1 x \quad (2)$$

defines as, for any value of the independent variable, x, the population mean of the dependent or response variable, y, is described by the straight-line function $\beta_0 + \beta_1 x$. The linear regression-based method developed by [20] performs better in terms of power of detection, but requires a future work as the false discovery rate decrease in high-noise context, and apply new strategies to handle large-size gene networks.

Structural equation model (SEM) also produces results comparable to the results from any that regression techniques. For the reason, SEM has been applied frequently in the reconstruction of GRNs [13], [32]–[35]. Among the favoured position of using SEMs are allowing feedback loops, differentiating direct relations from indirect relations [13], able to correct for measurement error and falsifiable (able to be proved if the SEM model is wrong) . Maximum Likelihood (ML) is the method to estimate the parameters in SEM. It works iteratively to find the best solution by searching through solutions and testing them until it finds the best one as applied in [36].

For the machine learning-based methods, the network is inferred by measuring the dependencies or causalities between transcriptional factors (TFs) and target genes. Among methods fall in this category include the partial correlation coefficient (PCC) [37], Bayesian network [38], dynamic Bayesian network [39], Boolean network [40], probabilistic Boolean network [41], Ordinary differential Equations [42], mutual information (MI) [43], graph theory [44] and neural network [45]. Besides that, numerous methods for GRNs reconstruction are proposed, such as Collateral-Fuzzy Gene Regulatory Network Reconstruction (CF-GeNe) [46]. The detail descriptions including the advantages and disadvantages of each method reviewed in [6], [47].

Neural network was inspired by animals' central nervous systems. It is a flexible statistical method capable of recognizing input patterns, and modelling any functional relationships and data structure [47], [48]. For the reason, this method among the most popular used in GRNs reconstruction, especially the recurrent neural network (RNN) [17], [49], [50]. It is consists of feedback connections and dynamic memory units [17], [49], [51], [52]. The basic concept RNN, each node of the GRN represents a particular gene, and the wiring between the nodes defines as regulatory interactions. However, RNN is sometimes restricted in performance by the limited number

of samples (or time points) in DNA microarray experiments; their number generally being smaller than the number of attributes (or the parameters) to be estimated [53].

Another the most popular used method in GRNs reconstruction is the correlation based approaches such as Mutual Information (MI). MI proposed to extract genetic networks [54] by computing MI for all gene pairs in a microarray dataset and infer that two genes are biologically related if their MI is above a certain threshold [55]. The MI of discrete random variables X and Y is defined as

$$I(X < Y) = \sum_{x \epsilon X, y \epsilon Y} P(x,y) \log \frac{P(x,y)}{P_X(x) P_Y(y)} \quad (3)$$

where P_X, P_Y are the marginal probability mass functions of X and Y, respectively, and P is the joint probability mass function of (X, Y). Methods such as ARACNE [55], CLR [56], MRNET [57] and PCA-CMI [58] have succeeded in inferring the GRNs using MI-based methods [59]. Even though MI works well with co-expressed or positively regulated patterns, it fails in handling gene profile with negative and mixed patterns [54]. Moreover, MI is incapable of detecting edges directionality [59], therefore it unable to identify the relationships accurately [54] particularly when applied to a limited number of samples [20]. Despite of the limitations, MI is further improved by using Conditional Mutual Information (CMI) which measures conditional dependency between two variables (genes) given other gene(s) [58]. Although the CMI is able to identify the direct regulations, it generally underestimates the regulation strength. To overcome the problems, a novel concept, namely conditional mutual inclusive information (CMI2) proposed by [31], to describe the regulations between genes.

Despite the advantages of CMI2, there is still room to improve it. Firstly, similar to PCA-CMI, CMI2 cannot directly infer edge directionality, which is also a general problem of many other methods, especially for those not working on time series data [55], [31]. Secondly, it is still a challenge task to select the conditional genes in an optimization way. Recently a method incorporated Gaussian Noise Model and Pearson Correlation Coefficient has been proposed by [13]. This method has demonstrated the ability to (1) predict the presence of regulatory interactions between genes, (2) their directionality and (3) their states (activation or suppression). However, this prediction method generated high false positive values because the indirect regulations have been wrongly predicted as true relationships. The list of GRN inference methods can be seen in Table 1.

TABLE 1. LIST OF ALGORITHM FOR GRNs RECONSTRUCTION

Categories	Methods	References
1. Model-based	Multiple linear regression	[14]
	Singular value decomposition	[15]–[18]
	Network component analysis	[19]
	Structural equation model (SEM)	[20]
	Sparse autoregressive vector	[19]
	Linear programming	[60]
	Hybrid method	[21]–[24]
2. Machine Learning-based	Partial correlation coefficient (PCC)	[37]
	Bayesian network	[38]
	Dynamic Bayesian network	[39]
	Boolean network	[40]
	Probabilistic Boolean network	[41]
	Ordinary differential Equations	[42]
	Mutual information (MI)	[43]
	Graph theory	[44]
	Neural network	[45]
	Collateral-Fuzzy Gene Regulatory Network Reconstruction (CF-GeNe)	[46]

IV. Gene Perturbation Data for Reconstruction of GRN

Gene expression data are crucial for GRN construction. Gene expression data allow biologists to observe the expression level of genes on a large scale [61]. There are two types of gene expression data used for GRN construction: time series and gene perturbation experiments. Time series expression data enable biologists to investigate the temporal pattern in biological networks. While, data obtained from experiments that gene perturbation by knockouts or RNA interference contain useful information for addressing the GRN reconstruction problem [33] because it provides information on interactions direction.

However, several factors have impeded the accuracy GRNs reconstruction using gene perturbation data. First, it is needed to identify the two important parameters that are affected by noise: (1) the unaffected genes and (2) the wild-type strain values, which are more difficult to identify when a larger number of genes are involved. Second, though past research has been conducted in reconstructing GRN, only a few researchers applied their methods to real experimental GRN datasets, as an addition to synthetic data. Third, most previous research only focuses on GRN prediction, only a few attentions given to determining the directionality of the genes. Finally, many high false positive (FP) values are generated using the current GRNs reconstruction methods because the indirect regulations have been incorrectly predicted as true relationships.

V. Conclusion

This study discusses several of GRNs reconstruction methods and identified their problem when dealing with gene perturbation data. Since the main motivation of developing GRNs reconstruction methods to improve its accuracy and scalability, this study useful as a reference to develop more accurate GRNs reconstruction methods particular when dealing with gene perturbation data. In future work we are going to develop a method of GRNs reconstruction by attentions given to determining the directionality of the genes and reduce the high false positive (FP) values are generated using real gene perturbation experimental datasets. The experiments that assess the stated features of our method will be covered in our future research studies.

Acknowledgment

As a token of my appreciation, I would like to thank my supervisor and co-supervisors, Dr. Sharifallilah Nordin, Dr. Ali Seman and Prof. Dr. Mohamed Saifulaman Mohamed Said for their continuous guidance and advice, also to all my seniors and course mates for their priceless support and assistances throughout the review process.

References

[1] W. C. Young, A. E. Raftery, and K. Y. Yeung, "Fast Bayesian inference for gene regulatory networks using ScanBMA.," *BMC Syst. Biol.*, vol. 8, p. 47, 2014.

[2] F. M. Alakwaa, N. H. Solouma, and Y. M. Kadah, "Construction of gene regulatory networks using biclustering and bayesian networks," *Theor. Biol. Med. Model.*, vol. 8, no. 1, p. 39, Jan. 2011.

[3] F. Emmert-streib, M. Dehmer, and B. Haibe-kains, "Gene regulatory networks and their applications : understanding biological and medical problems in terms of networks," *Front. cell Dev. Biol.*, vol. 2, no. August, pp. 1–7, 2014.

[4] T. Abdullah, A. Qazlan, A. Hamdi-cherif, and C. Kara-mohamed, "State of the Art of Fuzzy Methods for Gene Regulatory Networks Inference," vol. 2015, 2015.

[5] I. Shmulevich, E. R. Dougherty, and W. Zhang, "Gene perturbation and intervention in probabilistic Boolean Networks," *Bioinformatics*, vol. 18, no. 10, pp. 1319–1331, 2002.

[6] B. Ristevski, "A survey of models for inference of gene regulatory networks," *Nonlinear Anal. Model. Control*, vol. 18, no. 4, pp. 444–465, 2013.

[7] P. Brazhnik, A. De Fuente, P. Mendes, and V. Bioinformatics, "Gene networks: how to put the function in genomics," *Trends Biotechnol.*, vol. 20, no. 11, pp. 467–472, 2002.

[8] S. Roy, D. Das, D. Choudhury, and G. G. Gohain, "Causality Inference Techniques for In-Silico Gene Regulatory Network," in *Springer*, 2013, pp. 432–443.

[9] S. Roy and P. H. Guzzi, "Biological Network Inference from Microarray Data , Current Solutions , and Assessments," *Methods Mol. Biol.*, 2015.

[10] C. Olsen, P. E. Meyer, and G. Bontempi, "Inferring causal relationships using information-theoretic measures," in *5th Benelux Bioinformatics Conference (BBC09)*, 2009, no. December 2009, p. 2009.

[11] P. C. H. Ma and K. C. C. Chan, "A Fuzzy Data Mining Technique for the Reconstruction of Gene Regulatory Networks from Time Series Expression Data," pp. 1–8, 2006.

[12] J. Richer, X. Zhang, K. Liu, Z. Liu, X. Zhao, J. Hao, and L. Chen, "Systems biology NARROMI : a noise and redundancy reduction technique improves accuracy of gene regulatory network inference," *Bioinformatics*, vol. 29, no. 1, pp. 106–113, 2013.

[13] F. H. Mohamed Salleh, S. M. Arif, S. Zainudin, and M. Firdaus-Raih, "Reconstructing gene regulatory networks from knock-out data using Gaussian Noise Model and Pearson Correlation Coefficient," *Comput. Biol. Chem.*, 2015.

[14] F. Gregoretti, V. Belcastro, D. Bernardo, and G. Oliva, "A Parallel Implementation of the Network Identification by Multiple Regression (NIR) Algorithm to Reverse- Engineer Regulatory Gene Networks," *PLoS One*, vol. 5, no. 4, pp. 2–5, 2010.

[15] H. Li, "GENE REGULATORY NETWORK RECONSTRUCTION USING," 2013.

[16] Y. H. Chang and C. Tomlin, "Reconstruction of Gene Regulatory Networks with Hidden Nodes," pp. 1492–1497, 2014.

[17] S. Vineetha, "Reconstruction of Gene Regulatory Network from Expression Profile of Plasma RNA Data of Colorectal Cancer Patients using Soft Computing Techniques," Cochin University of Science and Technology, Kochi-682 022, India, 2012.

[18] V. Filkov, "Identifying Gene Regulatory Networks from Gene Expression Data," in *Handbook of Computational Molecular Biology*, 2005, pp. 1–30.

[19] J. C. Liao, R. Boscolo, Y.-L. Yang, L. M. Tran, C. Sabatti, and V. P. Roychowdhury, "Network component analysis: reconstruction of regulatory signals in biological systems.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 26, pp. 15522–7, Dec. 2003.

[20] Z. Dong, T. Song, and C. Yuan, "Inference of Gene Regulatory Networks from Genetic Perturbations with Linear Regression Model," *PLoS One*, vol. 8, no. 12, 2013.

[21] K. Y. Yip, R. P. Alexander, K. Yan, and M. Gerstein, "Improved Reconstruction of In Silico Gene Regulatory Networks by Integrating Knockout and Perturbation Data," *PLoS One*, vol. 5, no. 1, pp. 1 – 9, 2010.

[22] P. Smolen, D. A. Baxter, and J. H. Byrne, "Modeling Transcriptional Control in Gene Networks — Methods , Recent Results , and Future Directions," *Bull. Math. Biol.*, vol. 62, pp. 247–292, 2000.

[23] J. Pahle, "Biochemical simulations : stochastic , approximate stochastic and hybrid approaches," *Brief. Bioinform.*, vol. 10, no. 1, pp. 53 – 64, 2009.

[24] S. Menz, "Hybrid Stochastic – Deterministic Approaches for Simulation and Analysis of Biochemical Reaction Networks," 2013.

[25] Y. Li, M. Liang, and Z. Zhang, "Regression analysis of combined gene expression regulation in acute myeloid leukemia.," *PLoS Comput. Biol.*, vol. 10, no. 10, pp. 1 –14, 2014.

[26] J. Higgins and D. Ed, "Chapter 4: Introduction to Multiple Regression," in *The Radical Statistician*, 2005, pp. 1–15.N. E. F. Time-course and E. Data, "Network extraction from time-course," no. 1, pp. 8–11, 2004.

[27] X. Chen, C. Wang, A. N. Shajahan, R. B. Riggins, R. Clarke, and J. Xuan, "Reconstruction of Transcription Regulatory Networks by Stability-Based Network Component Analysis," *IEEE/ACM Trans Comput Biol Bioinform*, vol. 10, no. 6, pp. 1347–1358, 2014.

[28] S. C. Chan, H. C. Wu, and K. M. Tsui, "A New Method for Preliminary Identification of Gene Regulatory Networks from Gene Microarray Cancer Data Using Ridge Partial Least Squares With Recursive Feature Elimination and Novel Brier and Occurrence Probability Measures," *IEEE Trans. Syst. Cybern. A Syst. HUMANS*, vol. 42, no. 6, pp. 1514–1528, 2012.

[29] P. Li, "Inferring Gene Regulatory Networks From Time Series Microarray Data," no. August, 2009.

[30] X. Zhang, J. Zhao, J.-K. Hao, X.-M. Zhao, and L. Chen, "Conditional mutual inclusive information enables accurate quantification of associations in gene regulatory networks," *Nucleic Acids Res.*, vol. 43, no. 5, pp. 1–10, 2014.

[31] B. a. Logsdon and J. Mezey, "Gene expression network reconstruction by convex feature selection when incorporating genetic perturbations," *PLoS Comput. Biol.*, vol. 6, no. 12, 2010.

[32] A. Shojaie, A. Jauhiainen, M. Kallitsis, and G. Michailidis, "Inferring regulatory networks by combining perturbation screens and steady state gene expression profiles," *PLoS One*, vol. 9, no. 2, 2014.

[33] M. Vignes, J. Vandel, D. Allouche, N. Ramadan-Alban, C. Cierco-Ayrolles, T. Schiex, B. Mangin, and S. de Givry, "Gene regulatory network reconstruction using bayesian networks, the dantzig selector, the lasso and their meta-analysis," *PLoS One*, vol. 6, no. 12, 2011.

[34] N. Bing, "Statistical Analysis of Gene Expression Profile : Transcription Network Inference and Sample Classification Statistical Analysis of Gene Expression Profile : Transcription Network Inference and Sample Classification," 2004.

[35] A. Madar, A. Greenfield, E. Vanden-eijnden, and R. Bonneau, "DREAM3 : Network Inference Using Dynamic Context Likelihood of Relatedness and the Inferelator," *PLoS One*, vol. 5, no. 3, pp. 1 – 13, 2010.

[36] K. Raza and R. Jaiswal, "RECONSTRUCTION AND ANALYSIS OF CANCER- SPECIFICGENEREGULATORYNETWORKS FROM GENEEXPRESSION PROFILES," *Int. J. Bioinforma. Biosci.*, vol. 3, no. 2, pp. 21–33, 2013.

[37] F. K. Ahmad, S. Deris, and N. H. Othman, "The inference of breast cancer metastasis through gene regulatory networks," *J. Biomed. Inform.*, vol. 45, no. 2, pp. 350–362, 2012.

[38] H. Wu and X. Liu, "Dynamic Bayesian Networks Modeling for Inferring Genetic Regulatory Networks by Search Strategy : Comparison between Greedy Hill Climbing and MCMC Methods," vol. 2, no. 8, pp. 151–161, 2008.

[39] A. Saadatpour and R. Albert, "Boolean modeling of biological regulatory networks: A methodology tutorial," *Methods*, vol. 62, no. 1, pp. 3–12, 2013.

[40] I. Shmulevich, E. R. Dougherty, S. Kim, and W. Zhang, "Probabilistic Boolean Networks: a rule-based uncertainty model for gene regulatory networks.," *Bioinformatics*, vol. 18, no. 2, pp. 261–274, 2002.

[41] J. Cao, X. Qi, and H. Zhao, "Modeling gene regulation networks using ordinary differential equations.," *Methods Mol. Biol.*, vol. 802, pp. 185–97, 2012.

[42] K. Raza and R. Parveen, "Reconstruction of gene regulatory network of colon cancer using information theoretic approach," *4th Int. Conf. Conflu. 2013*, pp. 1–5, 2013.

[43] G. a Pavlopoulos, M. Secrier, C. N. Moschopoulos, T. G. Soldatos, S. Kossida, J. Aerts, R. Schneider, and P. G. Bagos, "Using graph theory

to analyze biological networks.," *BioData Min.*, vol. 4, no. 1, p. 10, 2011.

[44] M. Francesca, I. Casorelli, A. Colosimo, F. Simone, M. Bignami, and A. Giuliani, "A recursive network approach can identify constitutive regulatory circuits in gene expression data," *Physica A*, vol. 348, pp. 349–370, 2005.

[45] M. S. B. Sehgal, I. Gondal, and L. S. Dooley, "CF-GeNe : Fuzzy Framework for Robust Gene Regulatory Network Inference," *J. Comput.*, vol. 1, no. 7, pp. 1–8, 2006.

[46] L. E. Chai, S. K. Loh, S. T. Low, M. Saberi, S. Deris, and Z. Zakaria, "A review on the computational approaches for gene regulatory network construction," *Comput. Biol. Med.*, vol. 48, pp. 55–65, 2014.

[47] J. Vohradsky, C. Republic, and J. Vohradsky, "Neural Model of the Genetic Network *," *J. Biol. Chem.*, vol. 276, no. 39, pp. 36168–36173, 2001.

[48] N. Noman, L. Palafox, and H. Iba, "Reconstruction of gene regulatory networks from gene expression data using decoupled recurrent neural network model," *Nat. Comput. Beyond*, pp. 93–103, 2013.

[49] J. A. Bullinaria, "Recurrent Neural Networks," 2015, pp. 1–20.

[50] R. Xu, G. K. Venayagamoorthy, and D. C. Wunsch, "Modeling of gene regulatory networks with hybrid differential evolution and particle swarm optimization.," *Neural Netw.*, vol. 20, no. 8, pp. 917–927, 2007.

[51] B. Yang, Y. Chen, and M. Jiang, "Reverse engineering of gene regulatory networks using flexible neural tree models," *Neurocomputing*, vol. 99, pp. 458–466, 2013.

[52] R. Xu, D. C. W. Ii, and R. L. Frank, "Inference of Genetic Regulatory Networks with Recurrent Neural Network Models Using Particle Swarm Optimization," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 4, no. 4, pp. 681–692, 2007.

[53] S. Roy, D. K. Bhattacharyya, and J. K. Kalita, "Reconstruction of gene co-expression network from microarray data using local expression patterns.," *BMC Bioinformatics*, vol. 15 Suppl 7, no. Suppl 7, p. S10, Jan. 2014.

[54] A. a Margolin, I. Nemenman, K. Basso, C. Wiggins, G. Stolovitzky, R. Dalla Favera, and A. Califano, "ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context.," *BMC Bioinformatics*, vol. 7 Suppl 1, p. S7, 2006.

[55] J. J. Faith, B. Hayete, J. T. Thaden, I. Mogno, J. Wierzbowski, G. Cottarel, S. Kasif, J. J. Collins, and T. S. Gardner, "Large-Scale Mapping and Validation of Escherichia coli Transcriptional Regulation from a Compendium of Expression Profiles," *PLOS Biol.*, vol. 5, no. 1, pp. 54–66, 2007.

[56] M. J. Brim, L. DeRose, B. P. Miller, R. Olichandran, and P. C. Roth, "MRNET:ASCALABLE INFRASTRUCTURE FORDEVELOPMENT OF PARALLEL TOOLS AND APPLICATIONS Michael," in *Cray User Group 2010 Proceedings*, 2010, pp. 1–9.

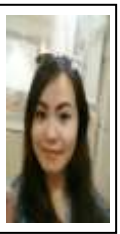
[57] X. Zhang, X. Zhao, K. He, L. Lu, Y. Cao, J. Liu, J. Hao, Z. Liu, and L. Chen, "Inferring gene regulatory networks from gene expression data by path consistency algorithm based on conditional mutual information," *Bioinformatics*, vol. 28, no. 1, pp. 98–104, 2012.

[58] X. Guo, Y. Zhang, W. Hu, H. Tan, and X. Wang, "Inferring Nonlinear Gene Regulatory Networks from Gene Expression Data Based on Distance Correlation," *PLoS One*, vol. 9, no. 2, pp. 1–7, 2014.

[59] Y. Li and S. a. Jackson, "Gene Network Reconstruction by Integration of Prior Biological Knowledge," *G3: Genes|Genomes|Genetics*, vol. 5, no. 6, pp. 1075–1079, 2015.

[60] W. Pindah, S. Nordin, A. Seman, and M. S. M. Said, "Review of Dimensionality Reduction Techniques using Clustering Algorithm in Reconstruction of Gene Regulatory Networks," in *IEEE-2015 International Conference on Computer, Communication, and Control Technology (I4CT 2015)*, 2015, no. I4ct, pp. 172–176.

About Author (s):



Winddy Pindah is a Ph.D candidate in Universiti Teknologi MARA (UiTM), Malaysia. She is interested in doing research in the field of Bioinformatics specifically in the reconstruction of Gene Regulatory Network

Dr. Sharifalillah Nordin is a lecturer in the Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM), Malaysia. She is interested in doing research in the field of Bioinformatics and Artificial Intelligence.

Dr. Ali Seman is a lecturer in the Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM), Malaysia. He is interested in doing research in the field of Bioinformatics and Machine Learning

Prof. Dr. Mohamed Saifulaman Mohamed Said is a lecturer in Faculty of Applied Sciences, Universiti Teknologi MARA (UiTM), Malaysia. He is interested in doing research in the field of Molecular Biology and Bioinformatics.