Technique of Gene Regulatory Networks Reconstruction Based on ARACNE Inference Algorithm

Sergii Babichev[1,2][0000-0001-6797-1467], Bohdan Durnyak[2][0000-0003-1526-9005],
Vsevolod Senkivskyy[2][0000-0002-4510-540X], Oleksandr Sorochynskyi[2][0000-0003-0964-2598],
Mykhailo Kliap[3][0000-0003-1933-6148] and Orest Khamula[2][0000-0003-0926-9156]

[1]Jan Evangelista Purkyne University in Usti nad Labem, Usti nad Labem, Czech Republic
`sergii.babichev@ujep.cz`
[2]Ukrainian Academy of Printing, Lviv, Ukraine
`durnyak@uad.lviv.ua, senk.vm@gmail.com, somsoroka@gmail.com`
[3]Uzhhorod National University, Uzhhorod, Ukraine
`mihaylo.klyap@uzhnu.edu.ua`

Abstract. The paper presents the technique of gene regulatory networks reconstruction based on ARACNE (Algorithm for the Reconstruction of Accurate Cellular Networks) inference algorithm. The research concerning optimisation of gene network topology on the basis of the complex apply of both the topological parameters of network and Harrington desirability index has been presented in the paper. Affymetrix DNA-chips mouse expression array *moe430a* from the database *ArrayExpress* was used as the experimental data during the experiment implementation. This data contains the gene expression profiles of *mus* musculus organism obtained as a result of DNA-microarray experiment implementation. The optimal parameters of ARACNE algorithm was determined during the simulation process. These parameters correspond to maximum value of Harrington desirability index which include as the components the topological parameters of the network.

Keywords: gene regulatory network, gene expression, gene expression profiles, Harrington desirability index, thresholding, reconstruction

1 Introduction

In most cases, current systems of information processing are based on the use of analogies of biological processes functioning which are occurred in living organisms. Such systems are the follows: immune system of organism, neural network, gene network, et.al. Their particularities are high level of complexity, ability of self-learning, ability to information recognizing and decisions making, decentralized parallel information processing. In this reason, development of current artificial intelligent models should be carried out based on the complex approach considering complex use of techniques of molecular biology, mathematics, informatics, physics, etc. Implementation of this approach creates the conditions for better understanding

particularities of operation of the biological system in order to influence to this process.

Reconstruction of gene regulatory networks and further simulation of the obtained models form the basis for investigation and analysis of both the character of molecular systems elements interconnections and influences of these interconnections to functional possibilities of the investigated objects. The complexity of gene networks reconstruction is determined by the follows: the experimental data which are used for the reconstruction process usually does not allows defining the network structure and pattern of genes interconnection in the network. Moreover, large quantity of genes complicates the interpretation of the network elements interconnections. In this case, it is necessary to conduct research concerning evaluation of both the network topology and the pattern of genes interconnection in network with the use of experimental data obtained by the use of both the DNA microchip experiment or RNA-molecules sequencing method. Qualitatively reconstructed gene regulatory network allows investigating the pattern of the biological organism development at the genetic level. It creates the conditions for both making new effective medicines and development of methods of early diagnostics and effective treatment of complex diseases. This fact indicates the actuality of the research in this subject area.

2 Formal Problem Statement

Fig. 1 presents a structure block-diagram of procedure of data processing in order to reconstruct gene networks and to validy the reconstructed models. A matrix of genes expression is used as the initial dataset.
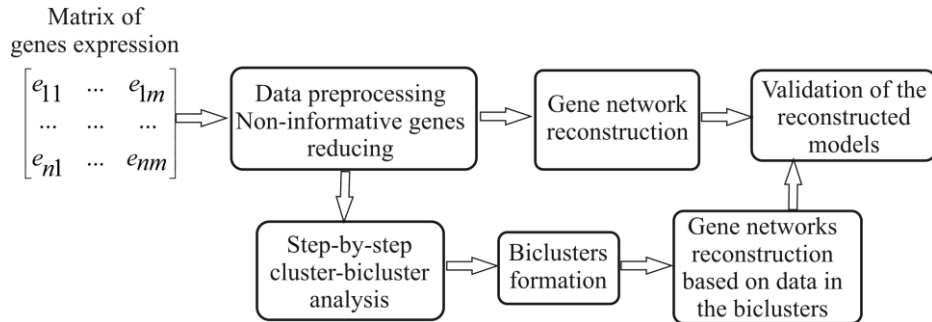

Fig. 1. A block-diagram of procedure of gene networks reconstruction and validation of the reconstructed models

As it can be seen, implementation of this process assumes gene expression profiles reducing in order to remove non-informative genes and further reconstruction of gene regulatory network based on full set of informative genes. Concurrently, the step-by-step cluster-bicluster analysis is performed. The result of this step implementation is a certain set of the biclusters, each from them contains mutually correlated genes and

samples. At the next step it is necessary to reconstruct gene networks based on the data in the biclusters. The final step of this procedure implementation is validation of the reconstructed gene networks by comparison analysis of patterns of genes interconnections in basic network and in networks reconstructed usin the data in the previously obtained biclusters.

The main task which should be solved within the framework of this research is development the technique to optimize the network topology using the network topological parameters. The current research presents one of the ways of this tasks solve.

3 Literature Review

Reconstruction of gene regulatory network is very complicated task which has no unambiguous solution nowadays. The first works concerning the reconstruction of gene regulatory networks based on gene expressions data were published in the end of 90's of the last century [1–4]. Several of approaches concerning reconstruction of gene networks are proposed in these works. Reviews concerning techniques of gene networks reconstruction using the data from DNA microarray experiments are presented in [5–7]. The authors have considered in details the stages of the gene regulatory network reconstruction, they also have done comparison analysis of various methods with allocation of both the advantages and shortcoming appropriate method.

Nowadays, there are sufficient information about the properties of gene regulatory networks of natural biological systems. So, in [8,9] the authors have formulated the rules of gene regulatory networks reconstruction. Their applying allows us significantly limit the dimension of search space of the optimal network topology. The sparseness property is the most common and important property of gene network. This property means that the topology of the network is sparse. Each of the genes has a small number of regulatory inputs [10]. However, it should be noted that there are a small number of genes (master-genes) that can control other genes. This property is used to limit the search space of the optimal topology by limiting the number of regulatory links. In [11, 12] the authors have shown that the frequency distribution of the number of regulating inputs of the gene network nodes of biological systems often is performed based on the stepwise Pareto distribution law. It means that in the case of a without-scale network, most of genes are weakly linked between each other. However, there are several nodes with large quantity of relationships. These nodes are named nodes-hubs and they correspond to genes that perform most of the overall regulation of other nodes in the network. The presence of hubs leads to network localization since all network nodes are connected with hubs by short links, the quantity of which are limited. Moreover, the hubs increase the model's stability to external influences and various types of oscillations, since they bind the network and do not allow the network to be split into separate modules.

The next property of gene networks, which should be considered when the network reconstruct, is modularity. This property means that genes in the network

cannot be considered as separate elements. In the general case, genes can be divided into functional ones that perform the regulating function and genes which function concordantly performing a common function. It is obvious that in this case the genes can be grouped into modules depending on the functional similarity of their expression profiles. The papers [13–16] present reviews concerning reconstruction of various types of gene regulatory networks. According to the authors' research, the models of gene networks can be divided in the following types: logical, continuous, single-molecular and hybrid models. The choice of the appropriate model is determined by both the amount and quality of information regarding gene expression values and appropriate regulatory elements between elements of the gene network.

In [17] the authors have proposed the information technology of gene expression profiles processing for purpose of gene regulatory network reconstruction. Implementation of this technology involves the following stages: gene expression array formation, non-informative genes reducing, step-by-step process of gene expression profiles cluster and bicluster analysis, reconstruction of gene networks and validation of the reconstructed models. The research concerning evaluation of this technology stability in the case of existence of increasing level of noise components is presented in [18].

However, it should be noted that despite the achievements in this subject area, nowadays there is absent an effective technique of gene regulatory network reconstruction, which can allow us with high level of probability to predict the pattern of biological organism development at genetic level. This problem can be solved using modern methods of complex data processing [19] based on ontologies of genes interconnections in network [20].

The aim of the paper is development of the technique of gene regulatory networks reconstruction based on ARACNE inference algorithm.

4 Materials and Methods

4.1 Topological Parameters of Gene Regulatory Network

The implementation of a technique to reconstruct gene regulatory network based on gene expression values assumes the possibility of creating various network topologies that are differed by the number of nodes, the number of arcs connecting the corresponding nodes, and the particularities of the connections between the network nodes. In this reason, there is a necessity to evaluate the network topology using parameters which can allow us to determine the appropriate network topology considering both the type of experimental data and aim of the current task. In general cases, the gene regulatory network can be presented as directed or undirected graph whose arcs can be weighted or unweighted. Therefore, the graph theory can be used to define the parameters of the network topology [21].

The follows topological parameters are used to evaluate the network topology within the current research:

- *Number of network nodes*. This parameter determines the general number of interconnected between each other genes.
- *Degree of a network node* or its *connectivity* is the total weight of connections (arcs) that connect a given node with neighbour nodes:

$$k_i = \sum_{j=1, j \neq i}^{n_i} w_{ij} \tag{1}$$

where n_i is a number of neighbour nodes of i-th node; w_{ij} is the weight of arc which connect the nodes i and j.

- *Average of degree* or *average of connectivity* is determined as an average of degrees of all network nodes:

$$\bar{k} = \frac{1}{n} \sum_{i=1}^{n} k_i \tag{2}$$

- Maximal degree determines the maximum value of elements the connectivity vector of all nodes in the network:

$$k_{max} = \max(k_1, k_2, ..., k_n) \tag{3}$$

The high value of this parameter indicates a high level of complexity, since all network nodes have a large number of connections with their neighbours. This fact complicates the interpretation of the reconstructed network. It is obvious that the minimum value of this parameter is optimal if the number of the network nodes is maximal.

- *Network density* is defined as the ratio of the number of weighted connections between the nodes to the maximum number of connections between the nodes in the network:

$$DS = \frac{2 \cdot \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w_{ij}}{n \cdot (n-1)} \tag{4}$$

The density value is varied from 0 to 1. If $DS = 0$, there is no connection between appropriate nodes, in the case of $DS = 1$ we have a fully connected network.

- The *clustering coefficient* of a node determines the probability that the nearest neighbours of a current node are directly connected between each other. The network clustering coefficient is defined as the average of the clustering coefficients of all network nodes:

$$CL = \frac{1}{n} \sum_{i=1}^{n} \frac{c_i}{0.5 \cdot k_i (k_i - 1)} \tag{5}$$

where n is a number of the network nodes, c_i is a number of connections of i-th node with neighbour nodes, k_i determines the numbers of neighbours of i-th node including this node which can make up a complete cluster. This parameter is a quantitative measure of the network connectivity. If its value is unity, the network is fully connected. In the case of zero values, there are no connections between the neighbours of the network nodes.

- The network *centralization* parameter determines the degree of proximity to the star topology:

$$Centr = \frac{n}{n-2}\left(\frac{k_{\max}}{n-1} - DS\right)$$ (6)

If the network topology looks like a lattice where all nodes are connected equally, the value of this parameter is zero. A higher value of the centralization parameter indicates a higher degree of network similarity to a star topology

- Network *heterogeneity* determines the degree of nonuniformity of the network topology and it is expressed using the variance and mean values of the average degree of nodes as follows:

$$GT = \frac{\sqrt{\mathrm{var}\left(\overline{k}\right)}}{mean\left(\overline{k}\right)}$$ (7)

Homogeneous network has zero heterogeneity value and otherwise, increasing the value of this parameter indicates a greater level of the network nonuniformity.

To make a final decision concerning network topology, it is necessary to calculate the complex topology parameter which should include the private parameters as the components. To solve this task, we propose a Harrington's desirability function [22]. Implementation of the proposed technique assumes transforming topological parameters scales into non-dimensional parameter Y, values of which are varied within the range from -2 to 5. Private values of desirabilities are calculated as follows:

$$dpr = \exp(-\exp(-Y))$$ (8)

The algorithm to calculate the general topological parameter involves the following steps:

1. Transformation of topological parameters scale into non-dimensional parameter Y scale in accordance with following linear equations:

$$\begin{cases} Y_{DS} = a_{DS} - b_{DS} \cdot DS; \\ Y_{CL} = a_{CL} - b_{CL} \cdot CL; \\ Y_{Centr} = a_{Centr} - b_{Centr} \cdot Centr; \\ Y_{GT} = a_{GT} - b_{GT} \cdot GT \end{cases}$$ (9)

The a and b coefficients for appropriate topological parameter are determined empirically considering the appropriate topological parameter boundary values.

2.   Calculation of the non-dimensional parameter Y for each of the topological parameters by equations (9)

3.   Determination of private desirabilities for the used topological parameters by the equation (8).

4.   Calculation of the general topological parameter as geometric average of all private desirabilities:

$$GTP = \sqrt[n]{\prod_{i=1}^{n} dpr_i} \qquad (10)$$

The maximum value of the parameter (10) indicates about optimal topology in terms of the used topological parameters.

4.2 ARACNE Inference Algorithm of Gene Regulatory Network Reconstruction

ARACNE inference algorithm (Algorithm for the Reconstruction of Accurate Cellular Networks) of gene regulatory networks reconstruction forms the links between genes (arcs) based on the analysis of statistical hypotheses concerning presence or absence of appropriate link [23]. As a result, the vector of probabilities is formed for each gene at the stage of gene network reconstruction. Each of the elements of this vector determines both the existence and the force of appropriate link. The estimation of the degree of interconnection between a pair of genes in the case of N of connection variants presence is performed using the Gaussian kernel estimation based on Shannon entropy:

$$I(g_i, g_j) = \frac{1}{N} \sum_{k=1}^{N} \log \frac{H_k(g_i, g_j)}{H_k(g_i) H_k(g_j)} \qquad (11)$$

where $H(g)$ is a Shannon entropy which is calculated for profile of g gene. The principal idea of ARACNE algorithm is a following: if there are different paths of appropriate genes linking in network, each of them is characterized by degree of the appropriated connection $I(g_i, g_j)$, it is selected the connection which satisfied to the condition:

$$I(g_i, g_j) = \min\left[I(g_i, g_s), I(g_s, g_p), \ldots, I(g_h, g_j)\right] \qquad (12)$$

where $g_s, g_p, \ldots, g_h$ are intermediate genes through which are linked the genes g_i and g_j. As a result, we obtain a network of interacting genes, the weight of the linkage between the corresponding genes is determined by the degree of linkage between the genes. The number of relationships in network is also limited by the appropriate thresholding coefficient. It is assumed that if the weight of the corresponding link is less than the value of the thresholding coefficient, the relationship between these genes is broken. In this way a topology of gene network is formed.

4.3 Technique of Gene Regulatory Network Reconstruction Based on ARACNE Inference Algorithm

Fig. 2 presents the structure block-diagram of the algorithm to estimate the thresholding coefficient optimal value which corresponds to optimal gene network topology in terms of used topological parameters.
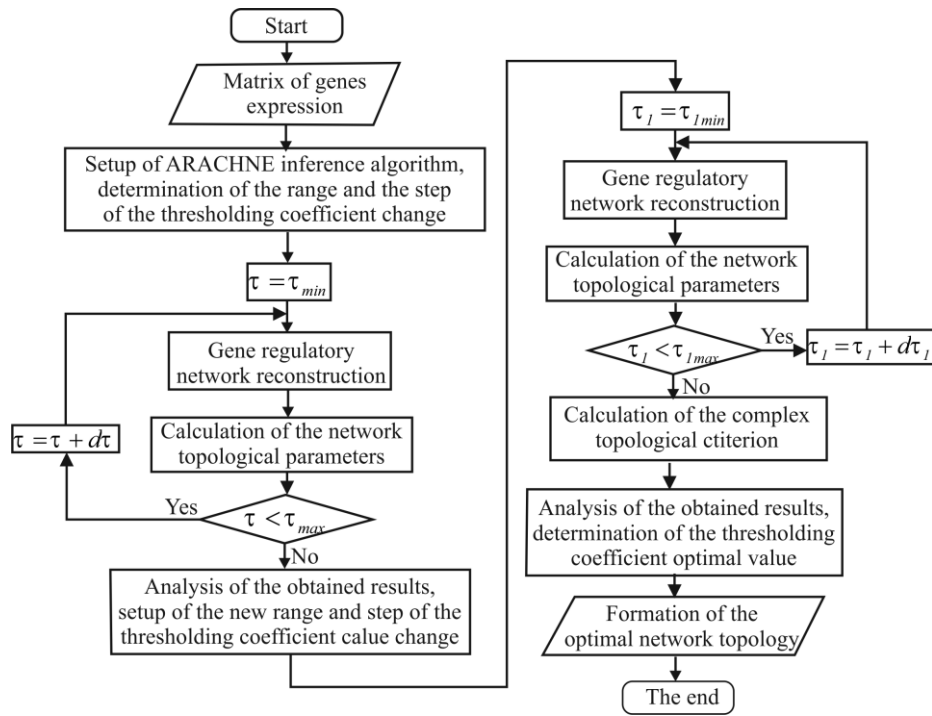

Fig. 2. A structure block-diagram of algorithm to form the gene regulatory network optimal topology

Implementation of the algorithm assumes the next steps:

1. Problem statement and the experimental dataset formation. The data is formed as a matrix, where rows and columns are the investigated genes and samples respectively.

2. Data preprocessing. Implementation of this step involves data normalizing and non-informative genes reducing. The number if the investigated genes significantly decreases at this step.

3. Setup the range and the step of the thresholding parameter value variation.

4. Reconstruction of gene regulatory network within the range of the thresholding coefficient change from minimum to maximum value. Calculation of the topological parameters (1)–(7) for each thresholding coefficient value.

5. Analysis of the obtained results. Setup a new significantly less range and step of the thresholding parameter value variation.

6. Reconstruction of gene regulatory network within the new range of the thresholding parameter variation. Calculation of the general topological parameters using the formulas (8)–(10).

7. Result analysis. Determination of the optimal value of the thresholding parameter. This value corresponds to the maximum of the general topological parameter in the case of maximum quantity of genes in network.

8. Gene regulatory network reconstruction using optimal thresholding coefficient value.

5 Experiment, Results and Discussions

Simulation of the gene network reconstruction based on the ARACNE inference algorithm was performed based on the *CytoScape* software [24] using gene expression profiles of *moe430a* dataset from *ArrayExpress* database [25,26]. The data contains the gene expression values of *mus* musculus organism obtained as a result of DNA-microarray experiment implementation. Data preprocessing was performed in accordance with the technique which is described in detail in [17]. The 147 the most informative genes were selected at this step for the following gene network reconstruction. Initially, the value of the thresholding coefficient was changed within the range from 0.1 to 0.9 with step 0.1. The charts of the topological parameters versus the thresholding coefficient value are presented in Fig. 3. The network clustering coefficient was equal to zero, that indicates an absence of the relationships between neighbours of genes in network.
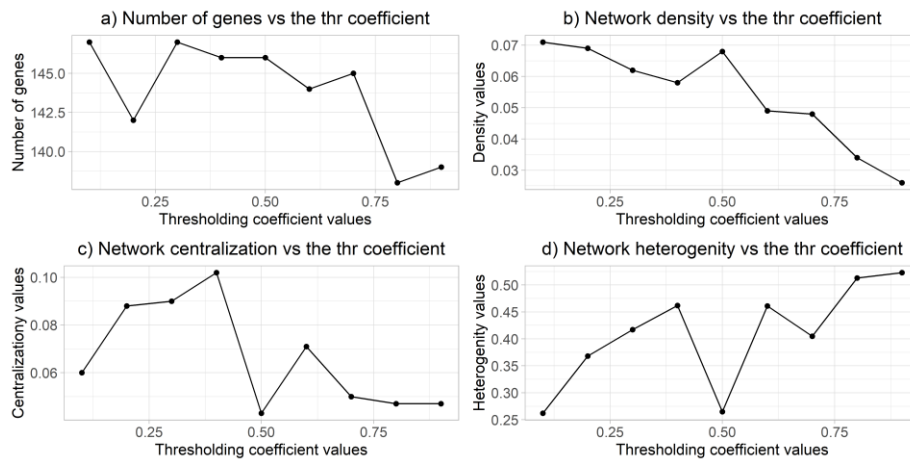

Fig. 3. Charts of network topological parameters versus the thresholding coefficient values in the case of range of the thresholding coefficient variation from 0.1 to 0.9

The analysis of the obtained charts allows us to conclude that the optimal value of the thresholding coefficient is in the range from 0.3 to 0.5, since the values of the centralization and heterogeneity coefficients in this range achieve the local maxima and the value of the density of the network nodes achieves the local minimum. The number of genes in this case varies from 147 to 146, what is quite acceptably. Fig. 4 shows the same charts for the case of range of the thresholding coefficient from 0.3 to 0.5. Step of this coefficient change was 0.02.
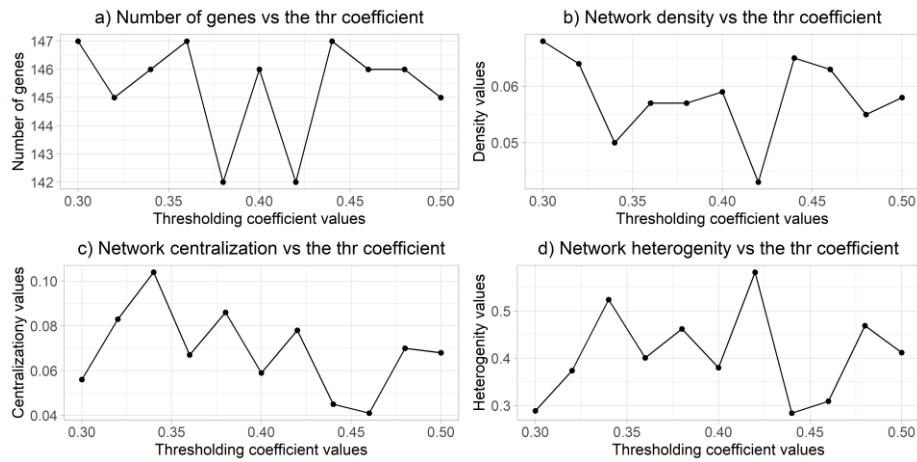

Fig. 4. Charts of network topological parameters versus the thresholding coefficient values in the case of range of the thresholding coefficient variation from 0.3 to 0.5

As it can be seen from the Fig. 4, the analysis of the obtained charts does not allow us to determine the optimal value of the thresholding coefficient, since the coefficients of centralization, heterogeneity and density have three local extrema, which disagree between each other. In accordance with hereinbefore presented algorithm it is necessary to calculate the general topology parameter for each thresholding parameter value. The plot of this parameter versus the thresholding coefficient is showed in Fig. 5.
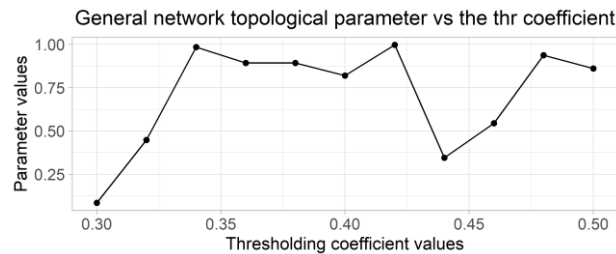

Fig. 5. Plot of general topological parameter versus the thresholding parameter

The obtained results analysis allows concluding that the maximum value of the topological parameter (10) achieves in the cases of 0.33 and 0.42 values of the thresholding coefficients. However, it should be noted that in the second case, the gene network contains five genes less, so the value of the thresholding coefficient of 0.34 is more acceptable in terms of the number of genes in the network. Fig 6 shows the charts of the topological parameters distribution for reconstructed gene regulatory network.
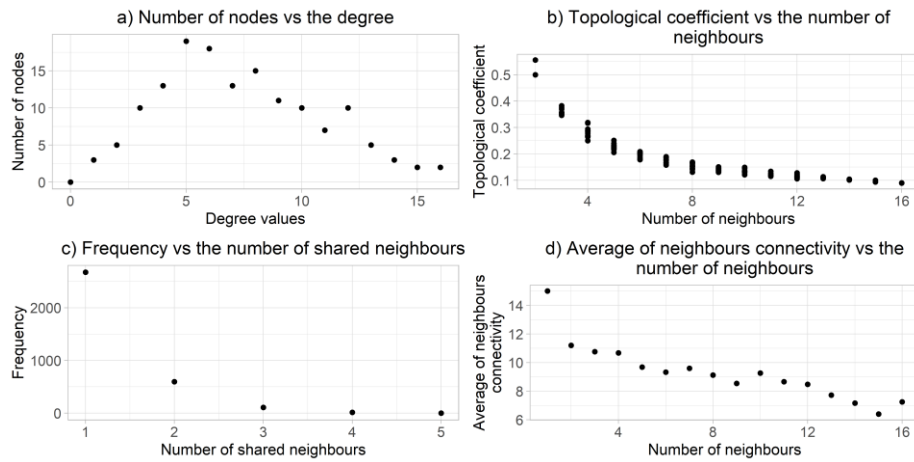

Fig. 6. Charts of topological parameters distribution: a) distribution of degree of the network nodes; b) distribution of the topological coefficient; c) distribution of number of shared neighbours; d) distribution of average of the neighbours connectivity

The obtained results indicate about high level of structuredness of the reconstructed gene network, since the number of genes with high degree is not so much, the values of the topological coefficient, number of shared neighbours and average of neighbours connectivity are decreased monotonically during the number of neighbours increase. Moreover, the comparison of the obtained charts with appropriate charts for other reconstructed valid gene networks [24] shows high level of their similarity. This fact also allows concluding about the correctnesss of the proposed technique. Of course, the effectiveness of the reconstructed gene regulatory network can be estimated by the further simulation process. This stage is a further perspective of the authors' research.

6 Conclusions

In the paper, we have presented the technique of gene regulatory networks reconstruction using ARACNE inference algorithm. The technique is presented as structure block-diagram of algorithm of the procedure of data processing in order to determine the optimal parameter of the ARACNE inference algorithm operation. This

parameter corresponds to maximum value of general topological parameter which includes as the components the network topological parameters. The simulation process of the gene regulatory network reconstruction has been performed based on the *CytoScape* software using gene expression profiles of *moe430a* dataset from *ArrayExpress* database. This dataset contains the gene expression profiles of *mus musculus* organism obtained as a result of DNA-microarray experiment implementation. The initial data contained 147 genes and 20 samples. The result of the simulation has shown that optimal network topology is achieved in the case 0.34 thresholding coefficient value. The charts of the topological parameters distribution for reconstructed gene regulatory network have been created as the simulation results. Their analysis has allowed us to conclude about high level of structuredness of the reconstructed gene network, since the number of genes with high degree is not so much, the values of the topological coefficient, number of shared neighbours and average of neighbours connectivity are decreased monotonically during the number of neighbours increase. Moreover, the comparison of the obtained charts with appropriate charts for other reconstructed valid gene networks also shows high level of their similarity.

The further perspective of the authors' research is the simulation of the reconstructed gene regulatory networks using modern techniques of intelligent data analysis.

References

1. D'haeseleer, P., Wen, X., Fuhrman, S., Somogyi, R.: Linear modeling of mRNA expression levels during CNS development and injury. Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing, pp. 41-52 (1999)
2. Liang, S., Fuhrman, S., Somogyi, R.: Reveal, a general reverse engineering algorithm for inference of genetic network architectures. Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing, pp. 18-29 (1998)
3. Friedman, N., Linial, M., Nachman, I., Pe'er, D.: Using Bayesian networks to analyze expression data. Journal of Computational Biology, 7(3-4), pp. 601-620 (2000) doi: 10.1089/106652700750050961
4. Chen, T., He, H.L., Church, G.M.: Modeling gene expression with differential equations. Pacific Symposium on Biocomputing. Pacific Symp. on Biocomputing, pp. 29-40 (1999)
5. Wong, K.-C., Li, Y., Zhang, Z.: Unsupervised learning in genome informatics. Unsupervised Learn. Algorithms, pp. 405-448 (2016) doi: 10.1007/978-3-319-24211-8_15
6. Emmert-Streib, F., Dehmer, M., Haibe-Kains, B.: Gene regulatory networks and their applications: Understanding biological and medical problems in terms of networks. Frontiers in Cell and Developmental Biology, 2(AUG), art. no. 38 (2014) doi: 10.3389/fcell.2014.00038
7. Wang, K., Zhang, L., Liu, X.: A review of gene and isoform expression analysis across multiple experimental platforms. Chinese Journal of Biomedical Engineering, 36(2), pp. 211-218 (2017) doi: 10.3969/j.issn.0258-8021.2017.02.012
8. Yan, B., Guan, D., Wang, C., et.al.: An integrative method to decode regulatory logics in gene transcription. Nature Communications, 8 (1), art. no. 1044 (2017) doi: 10.1038/s41467-017-01193-0

9. Alkallas, R., Fish, L., Goodarzi, H., Najafabadi, H.S.: Inference of RNA decay rate from transcriptional profiling highlights the regulatory programs of Alzheimer's disease. Nature Communications, 8 (1), art. no. 909 (2017) doi: 10.1038/s41467-017-00867-z

10. Nair, A., Chetty, M., Wangikar, P.P.: Improving gene regulatory network inference using network topology information. Molecular BioSystems, 11 (9), pp. 2449-2463 (2015) doi: 10.1039/c5mb00122f

11. Ma, X., Zhou, G., Shang, J., Wang, J., Peng, J., Han, J.: Detection of complexes in biological networks through diversified dense subgraph mining. Journal of Computational Biology, 24 (9), pp. 923-941. (2017) doi: 10.1089/cmb.2017.0037

12. Ma, C.-Y., Chen, Y.-P.P., Berger, B., Liao, C.-S.: Identification of protein complexes by integrating multiple alignment of protein interaction networks. Bioinformatics, 33 (11), pp. 1681-1688 (2017) doi: 10.1093/bioinformatics/btx043

13. Van Someren, E.P., Wessels, L.F.A., Backer, E., Reinders, M.J.T.: Genetic network modeling. Pharmacogenomics, 3 (4), pp. 507-525 (2002) doi: 10.1517/14622416.3.4.507

14. Gardner, T.S., Faith, J.J.: Reverse-engineering transcription control networks. Physics of Life Reviews, 2 (1), pp. 65-88 (2005) doi: 10.1016/j.plrev.2005.01.001

15. Schlitt, T., Brazma, A.: Current approaches to gene regulatory network modelling. BMC Bioinformatics, 8 (SUPPL. 6), art. no. S9 (2007) doi: 10.1186/1471-2105-8-S6-S9

16. Emmert-Streib, F., Dehmer, M., Haibe-Kains, B.: Gene regulatory networks and their applications: Understanding biological and medical problems in terms of networks. Frontiers in Cell and Developmental Biology, 2 (AUG), art. no. 38 (2014) doi: 10.3389/fcell.2014.00038

17. Babichev, S., Lytvynenko, V., Skvor, J., Korobchynskyi, M., Voronenko, M.: Information Technology of Gene Expression Profiles Processing for Purpose of Gene Regulatory Networks Reconstruction. Proceedings of the 2018 IEEE 2nd International Conference on Data Stream Mining and Processing, DSMP 2018, art. no. 8478452, pp. 336-341 (2018) doi: 10.1109/DSMP.2018.8478452

18. Babichev, S.: An evaluation of the information technology of gene expression profiles processing stability for different levels of noise components. Data, 3(4), art. no. 48 (2018) doi: 10.3390/data3040048

19. Tkachenko, R., Izonin, I., Kryvinska, N., Chopyak, V., Lotoshynska, N., Danylyuk, D.: Piecewise-linear approach for medical insurance costs prediction using SGTM neural-like structure. CEUR Workshop Proceedings, 2255, pp. 170-179 (2018)

20. Lytvyn, V., Vysotska, V., Dosyn, D., Lozynska, O., Oborska, O.: Methods of Building Intelligent Decision Support Systems Based on Adaptive Ontology. Proceedings of the 2018 IEEE 2nd International Conference on Data Stream Mining and Processing, DSMP 2018, art. no. 8478500, pp. 145-150 (2018) doi: 10.1109/DSMP.2018.8478500

21. Assenov, Y., Ramírez, F., Schelhorn, S.-E.Sven-Eric, Lengauer, T., Albrecht, M.: Computing topological parameters of biological networks. Bioinformatics, 24 (2), pp. 282–284 (2008) doi: 10.1093/bioinformatics/btm554

22. Harrington, J.: The desirability function. Ind. Qual. Control, 21(10), pp. 494–498 (1965)

23. Margolin, A.A., Nemenman, I., et.al.: ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC Bioinformatics, 7 (SUPPL.1), art. no. S7 (2006) doi: 10.1186/1471-2105-7-S1-S7

24. CytoScape Homepage, http://apps.cytoscape.org.

25. ArrayExpress Homepage, https://www.ebi.ac.uk/arrayexpress/experiments/browse.html.

26. El. Res.: http://www.affymetrix.com/support/technical/byproduct.affx?product=moe430