TOPOLOGICAL FEATURES AND PARAMETERS OF BIOCHEMICAL NETWORK STRUCTURE

Tatjana Rubina, Egils Stalidzans
Latvia University of Agriculture, Faculty of Information Technology,
Department of Computer Systems, Latvia, Jelgava, Liela 2
Tatjana.Rubina@llu.lv, Egils.Stalidzans@llu.lv

ABSTRACT

In the cell, tissue, organ and organism under different conditions operate metabolic, protein interaction, gene or genetic regulatory and signaling networks that describe biochemical reactions, biochemical or biophysical process. The biochemical networks can present the relationship between genes and gene products, proteins, metabolites and etc. The exploration of these networks is the key task in systems biology in the last few years and helps to understand some cellular process, functions or properties of biological system better.

We provide researchers and coterie information on network structure analysis, existing and used topological features with the following goals: 1) to accumulate the existing knowledge about the network structure analysis; 2) to provide a list of topological features and characterize the most of them, mentioning its biological sense; 3) to provide the information of the existing software tools for the structure analysis.

In this research paper, we focus on the structure of the networks, the topological parameters such as a degree distribution, a path length, a clustering coefficient, motifs such as feedback loops, a feedforward, self-loops and the other features.

1 INTRODUCTION

The goal of systems biology is to understand living organisms at the systems level, combining quantitative information on individual components in order to understand the emergent behaviors that result [14]. According to the systems theory, a system is defined as the set of objects with relations between objects and these properties. The living organisms can be considered as biological systems. In most cases, the biological systems are very complex. System biologists build a system-level understanding of how the biological world works and solve problems by understanding systems and then applying that knowledge to control them [20].

During the last several years, the biological research produces increasing volumes of data describing genome sequence of biological organisms, cellular components, their interactions, and states of biochemical networks for model organisms [12]. This available information about biological entities and their interactions enable us to consider different organisms as biological systems which control the genetic information [22].

To understand biology at the system level, we must examine the structure and dynamics of cellular and organismal function (Kitano) [1]. According to the systems theory, a structure is defined as the set of elements and relations (links) between them that define functions of system and distribution of the purposes [7]. Analyzing the structure, usually the researchers are interested in the properties of the system, which remain invariable for a long period of time or functioning on the system. Living organisms are subjected to the process of genetic mutations dependant on environmental conditions (external perturbations), it can be necessary to detect invariable and varyable system properties. It can be done with the purpose to compare several systems of one type, to detect influence of the certain and separate factors on the system.

In this research paper, we focused on the structure of biochemical networks and their exploration.

2 EXPLORATION OF BIOCHEMICAL NETWORKS

The goals of systems biology are to understand the mechanisms of how biochemical networks generate particular cellular functions in response to environmental stresses or genetic changes [18]. Structural or topological analysis of cellular interaction networks contributes to a deeper understanding of network-wide interdependencies, causal relationship, and basic functional capabilities. Structural analysis, towards a functional analysis of the structure is not based on quantitative and dynamic properties and can thus only provide qualitative answers [13].

It is important and necessary to explore, for example, protein interaction or gene regulatory networks, while they play a key role in disease understanding and accurate drug targets [11] finding. Structural or topological (parameter-free) and dynamic models are used to explore the networks. It is necessary to notice that the basis of each model is the structure. In this paper, we focused on a structural model.

In structural models networks are represented in directed or undirected graph form [13, 28, 29, 31] and mixed graphs [28], that consist of nodes and edges. Nodes represent genes, gene products, proteins, chemical compounds or small molecules. Graph links (edges in undirected graph or arcs in directed graph) represent functional relationships [8] i.e., various types of interactions or associations between the pair of nodes, e.g. metabolic reactions or events, gene modification, protein/protein-nucleotide interactions, protein modification, regulatory relationships such as transcriptional and translational regulation, signaling pathways etc. [25, 29]. Connections can be directed or undirected dependent on graph type; they can have physical meaning, denote general associations; they can represent shared characteristics [28] between components. Nodes depend on the number of connections and the type of graph can be hubs (highly connected nodes) [30] and metanodes [28].

The choice of the network representation is often dictated by a research issue. Directed networks are suitable when the interactions between two components have a well-defined direction, for example, the direction of metabolic flow from substrates to products, or the information flow from transcription factors to the genes that they regulate. For example, Zhao et al in their work [27] represent Homo sapiens metabolic network by a directed graph, which nodes correspond to metabolites and the arcs – to reactions between these metabolites, in which irreversible reactions are represented as directed arcs while reversible ones as bi-directed arcs (see Fig.3.). Transcription factor binding networks [28] can be represented too as directed networks.

Undirected networks, such as protein interaction networks, represent a mutual relationship: if protein A binds to protein B, then protein B binds to protein A. This type of representation also often applies to predictions made by high-throughput proteomic or genomic analysis, or indirect links based on shared genes or protein components between pathways and complexes [28].

3 FEATURES AND PROPERTIES OF NETWORK STRUCTURE

Networks have "emergent" properties that are distinct from those of their individual components. Emergent properties are non-linear, aggregated and combinatory effects generated by the interaction of the components of the network. For example, properties such as topology, information flow and the stable state of a network can only be detected at the network level, not by examining the individual components such as genes or proteins. The structural and dynamic features of genetic networks ultimately contribute to biological functions, robustness and evolvability of the networks [9].

Examining scientific literature, publications and analyzing software tools, we found many network structure properties and features, which can be categorized in five groups: network metrics, network motifs, topological parameters, topological features, and structure qualitative parameters (see Fig.1.). These properties and features are not the only one, but mostly used in studies and researches.

3.1 Topological parameters

In this paragraph we will describe such topological parameters as degree, degree distribution, paths, path length, path distribution, network diameter. The following definitions will be used graph node, vertex – a network component, graph edge or arc – link or connection.

Degree of a network component. The degree (or connectivity [3, 21]) of an undirected network element or node, k_i, is the number of edges (links) that it has with other elements that is incident with i:

$$k_i = \sum_j^n k_{ij} \qquad (1)$$

A degree is also a feature that distinguishes hubs (highly connected nodes) from leaves or orphans (weakly or non-connected nodes) in the network [31]. In protein interaction and genetic interaction networks, for example, the degree of a hub (highly connected component) is often its importance and essentiality for cell function [28], process or whole system.

For example, Han et al. [9] and Partil and Nakamura [19] defined hubs as nodes with degrees of more than 6 ($k \geq 6$).

Hase et al [11] in their study of Structure of Protein Interaction networks (PIN) and their implications on Drug design, has analyzed budding yeast and Human PINs to identify these topological features. Depending on degree value, there are 3 types of node degrees for each PIN: low-degree, middle-degree and high-degree nodes. **In the budding yeast PIN:**

- ✓ **Low-degree nodes** are nodes with degree of less than 6 ($k<6$),
- ✓ **Middle-degree nodes** are nodes with degrees from 6 to 38 ($6 \leq k \leq 38$),
- ✓ **High-degree nodes** are nodes with degrees of more than 38 ($k>38$).

In the Human PIN they define:

- ✓ **Low-degree nodes** are nodes with degree of less than 6 ($k<6$).
- ✓ **Middle-degree nodes** are nodes with degrees from 6 to 30 ($6 \leq k \leq 30$).
- ✓ **High-degree nodes** are nodes with degrees of more than 30 ($k>30$).

The degree of node can be used to examine the network structure. By examining the network structure, researcher should examine these elements, i.e., nodes by using degree of node. If the network has isolated [35] elements (which degree is 0), it means that in construction, description or creation of network structure have been committed errors.

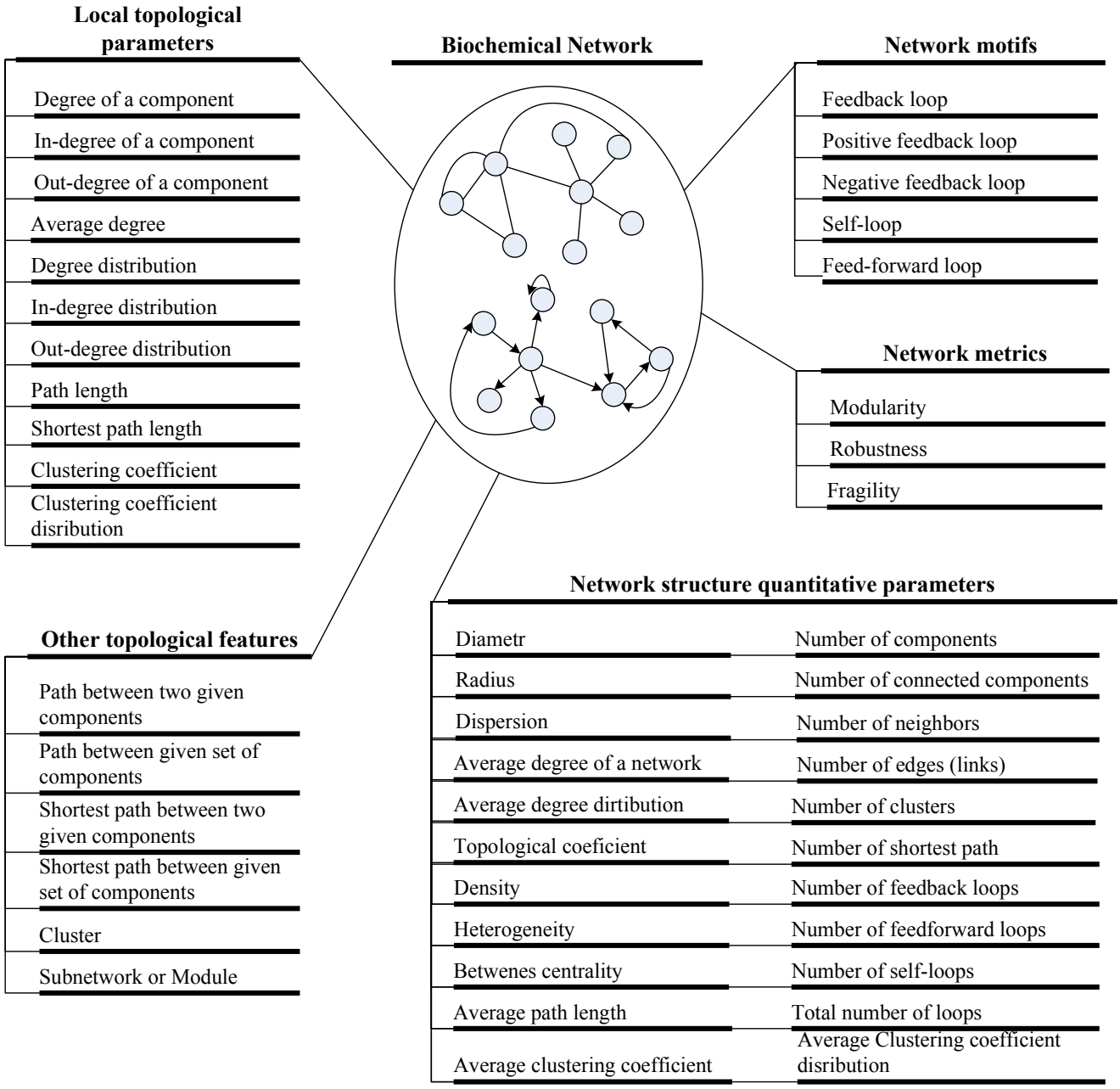


Fig. 1. Biochemical network measures and properties

In-degree and Out-degree distribution. For directed networks, such as transcription factor binding networks, the degree is separated into 'in'-degree and 'out'-degree, depending on the directions of interaction between two given elements [28].

Incoming degree (In-degree). The incoming degree is a number of links that point to the network component *I* [21]*:*

$$k_{+i} = \sum_{j}^{n} k_{ij} \qquad (2)$$

Outgoing degree (Out-degree). The outgoing degree is a number of links that start with network component [21].

$$k_{i+} = \sum_{j}^{n} k_{ij} \qquad (3)$$

Average degree. For an undirected network, the average degree in an undirected network [3, 5]:

$$K = < k > = \frac{2L}{N} \qquad (4)$$

, where N – the total number of components

L – the total number of links

Average degree of a network. For an undirected network, the average degree of the whole network is the average of the degrees of all individual network components [5]:

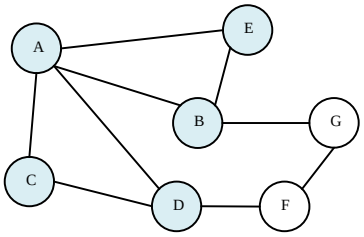
$$K = \frac{\sum_{i \in V} k_i}{N} \quad (5)$$

, where k_i - the degree of node i

N - the total number of nodes

Degree distribution. Degree distribution d_k is the number of nodes with degree k ($k=1,2,\ldots n$) [5, 21]. In other words, degree distribution is the number of neighbors of network component i.

For directed networks the degree distribution is separated into **in-degree** and **out-degree distribution**.



$k_A=4$

$k_B=3$

Fig. 2. Undirected network

Let K be the degree of a network element. Then a statistical model for the degree distribution is represented by:

$$P(K = k) = f(k) = \frac{N_k}{N} \quad (6)$$

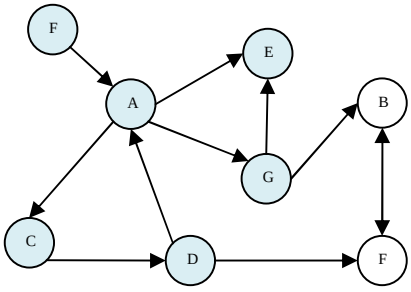
,where $f(k)$ is a probability distribution

N_k – a number of nodes with degree $k=1,2\ldots n$

N – the total number of nodes

The distribution of degrees $f(k)$ in undirected network, gives the probability that a selected component has degree k [3]. In the case of directed networks one needs to consider two distributions, $P(k_{in})$ and $P(k_{out})$ [4].

The distribution of degrees among network components is useful for characterizing the topology and scale of a network, and often has meaningful biological interpretation [28]. In general, degree distribution allows to separate different network types or classes.



$k_{Ain}=2$

$k_{Aout}=3$

Fig. 3. Directed network

The degree distribution of many types of real-life networks, such as metabolic, scientific collaboration networks [21, 26]:

$$P(K = k) \sim k^{-\gamma} \quad (7)$$

- this function is called a power law

,where γ - a constant or the degree exponent

The smaller the value of γ, the more important the role of the hubs is in the network. Whereas for $\gamma>3$ the hubs are not relevant, for $2>\gamma>3$ there is a hierarchy of hubs, with the most connected hub being in contact with a small fraction of all nodes and for $\gamma= 2$ a hub-and-spoke network emerges, with the largest hub being in contact with a large fraction of all nodes [4, 3].

A **power-law degree distribution** indicates that a few hubs hold together numerous small components or nodes [3]. A network with this degree distribution is called scale-free, indicating that there is a high diversity of node degrees and no typical nodes in the network that could be used to characterize the remaining nodes [28]. In a scale-free networks most nodes have only one or two functional links, whereas a small number of nodes, the hubs, have many links [8]. Nearly all biological networks, including regulatory, interactome and metabolic networks are scale-free [3, 4].

3.2 Topological parameters and features

In studying the function of pathways, the property of interest is often how a given gene or protein is related to (or responds to) an up- or downstream signal. Given a large data set of interactions, it may be useful in some contexts to find the most direct path between two genes, proteins, complexes or pathways; for example, the overall lengths of such pathways may be related to the immediacy or breadth of signal response [28].

The Path [35] form n_0 to n_k, according to the graph theory, is the sequence of nodes.

There can be different types of paths: chain (have all different edges), simple chain (have all different nodes), closed chain or cycle (starts and ends with the same node).

The path length l_{ij} is the number of edges in path from node i to j.

The path between two given nodes (see Fig.4.). In case of signaling networks, the computation of all paths between pair of species helps to recognize all the different ways in which a signal can propagate between two nodes, e.g. all the different ways by which a certain transcription factor (or any other species from the output) can be activated or inhibited by signals riving the input layer.

In metabolic networks, researchers are particularly interested in the reactions (edges), because they correspond to enzymes that are subject to regulatory processes and can be knocked-out in experiments. In metabolic pathway analysis, a statistical or combinatorial analysis of the participation and co-occurrences of

reactions in elementary modes proved to be useful for obtaining system-wide properties, such as the detection of essential reactions/enzymes or correlated reaction sets (enzyme subsets) [13].

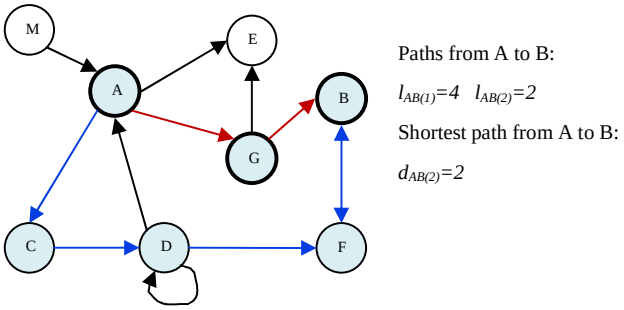


Fig. 4. Paths between two given nodes

The shortest path between two nodes is a path with the minimum number of edges that is necessary to traverse from one node to the other.

In weighted directed graph the shortest path [32] is the path between two nodes (source and target nodes, start and end nodes) with minimal sum of weights of the edges making the path.

The suboptimal path [32**]** is constructed by removing all edges in all shortest paths one by one and one at a time and finding the shortest path.

Average Path length of a network is the average of the shortest path lengths between every node pair [5].

$$L = \frac{2\sum_{\substack{i,j\in V \\ i<j}} d_{ij}}{N(N-1)} \qquad (8)$$

, where d_{ij} – the shortest path length between nodes i and j.

Average path length reflects how closely nodes are connected within the network and offers a measure of a network's overall navigability [5]. The average shortest path also indicates the well-known 'small-world' property of many real-life networks [28, 5].

Shortest path distribution. *SP(l)* is defined as the proportion of shortest paths with a specified length l in a network and reflects the diversity of the graph distances between two nodes in the network. In most real-life networks, there is a relatively short path between any two nodes, and the length is in the order of logarithm of the network size. This property is known as "small world" [31].

Diameter of the network is the (longest) distance between two most remote nodes [5]:

$$D = \max_{i,j\in V} d_{ij} \qquad (9)$$

, where d_{ij} – the shortest path length between nodes i and j.

In a network, different nodes have different levels of connectivity. It is necessary to evaluate which node is the most important. Network centrality is a local quantitative measure for assessing the position of a node relative to the other nodes, and can be used to estimate its importance or role in a global network organization. Different information sources reveal several centrality measures such as degree centrality, closeness centrality, and betweenness centrality [5]. Betweenness centrality better predicts the essentiality of a node then degree centrality [8]

The **closeness** of a node is defined as the inverse of the average distance from all other nodes. The **betweenness** is one of the standard measures of node centrality, originally introduced to quantify the importance of an individual in a network [4]. The betweenness *bi* of a node *i*, is defined as:

$$bi = \sum_{\substack{j,k\in N \\ j\neq k}} \frac{n_{jk}(i)}{n_{jk}} \qquad (10)$$

, where n_{jk} - the number of shortest paths connecting nodes j and k,

$n_{jk}(i)$ - the number of shortest paths connecting nodes j and k and passing through i.

The concept of betweenness can be extended also to the edges. The **edge betweenness** is defined as the number of shortest paths between pairs of nodes that run through that edge [4].

3.3 Network motifs

Identifying topological features in networks is an important part of understanding the relationship between structure and function of network motifs, e.g. feedback loops, feedforward loops (see Fig.5.).

Structural model characterize and provide information of the connectivity (topology) of the interactions involved in a biological process or system.

However, some insights into the dynamic properties can nevertheless often be obtained, because fundamental properties of the dynamic behavior are often governed by the network structure [13]. Feedback and biological regulation are two sides of the same coin, reflecting the need of the living cell to deal with changing environtments, to generate cell to cell heterogeneity and to optimize cellular metabolism to a given external condition [15]. By using feedback loops a network component promotes its own accumulation or activation. For example, a protein can activate its own transcription factor or inhibit its own proteolysis. It may activate its own activator, or somehow remove its own inhibitor. [1]

According to the graph theory, the **feedback loop** is a directed simple cycle or circuit that is defined as a sequence of arcs that starts and ends at the same node (vertex) i and visit no node twice.

Feedback loops may be positive or negative, depending upon the parity of the number of negative interactions in the loop, and they play key role in controlling the

dynamics [24] of a wide range of biological systems. Negative feedback loops tend to act within biological systems to maintain homeostasis [10] (are essential for homeostatic mechanisms (i.e. for adjusting and maintaining levels of system variables) [13] and are necessary for the existence of periodic behavior [24]. Systems involving negative feedback loops tend to settle to a steady state, which may be stable or unstable [10]. The behavior of a negative feedback loop depends partly upon its length. A one element loop generates a single, stable, steady state; a two element loop produces a single steady state which is approached to and departed from in a periodic way; while a loop with three or more elements can generate damped or stable oscillations depending upon parameter values [10]. Networks with larger number of independent negative feedback loops tend to have longer limit cycles and thus may exhibit more random or chaotic behavior [24].

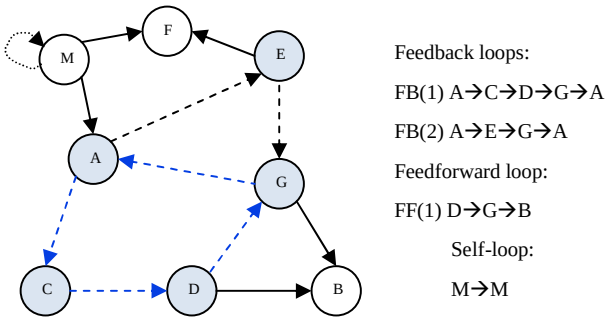


Fig. 5. Feedback, feedforward end self- loops

Positive feedback loops are responsible and even required for multiple steady state behaivior in dynamical systems. In biological systems, multistationarity plays a central role in differentiation processes and for epigenetic and switch-like behavior [13]. Multistationarity is the existence of a number of different stable state [10].

Self-Loop. Self-loop is an arc connecting a species with itself.

3.4 Subnetworks and clusters

It is possible to construct two networks with identical topological measures including degree distribution and clustering coefficient but different hierarchical structure. Therefore, local topological features such as modularity, motif, and network clustering are likely to be key concepts in understanding cellular mechanisms and biological functions in biomolecular networks [5].

The functions of biomolecular networks are closely related to their topologies and facilitated by characteristic topological patterns. Components of cellular networks including genes, proteins, and other molecules usually act in collaboration to carry out specific biological processes and biochemical activities, by forming relatively isolated functional units called modules. From the topological perspective, a module can be understood as a subnetwork that is densely connected within itself but sparsely connected with the rest of the network. For example, in cellular networks, a module refers to a group of physically or functionally connected biomolecules that work together to achieve some desired cellular function [5]. But, in the network of gene expression data, interaction subnetworks is a connected sets of interactions, whose genes show particularly high levels of differential expression. The interactions contained in each subnetwork provide hypotheses for the regulatory and signaling interactions in control of the observed expression changes [5, 23].

Cluster. Depending on the type of network, clusters may mean different things. For instance, clusters in a protein-protein interaction network have been shown to be protein complexes and parts of pathways. Clusters in a protein similarity network represent protein families [5, 23]. In gene regulatory networks, the cluster is a group of genes that are active at the same time are likely to be involved in the same regulatory process [16].

Cluster analysis is applied, for example, to microarrays data that have been collected over a variety of conditions or a series of time points. Its goal is to determine the original groupings of the genes. This technique assumes that genes that are active at the same time are likely to be involved in the same regulatory process. It assumes that genes are grouped and within the group the genes would produce the same expression profile [16].

Clustering coefficient. Clustering coefficients is the ratio of the number of existing connection or links between a node's neighbors to the maximum number of possible links between them [5, 28, 3, 33]:

$$C_i = \frac{2n_i}{k_i(k_i - 1)} \qquad (11)$$

, where n_i is the number of links between k_i neighbors.

While node *A* have 5 neighbors (*C,D, B, E, H*), the number of links between node *A* neighbors $n_A=2$ node A degree is $k_A=5$, the node *A* clustering coefficient is:

$$C_A = \frac{2 \cdot 2}{5 \cdot (5-1)} = \frac{1}{5}$$

Clustering coefficient measures how well the neighbors of a node i are locally interconnected [33] and the tendency of a network to have highly connected clusters. In large-scale mass spectrometric networks in yeast, this property can be used to identify groups of proteins involved in the assembly of the ribosome [28].

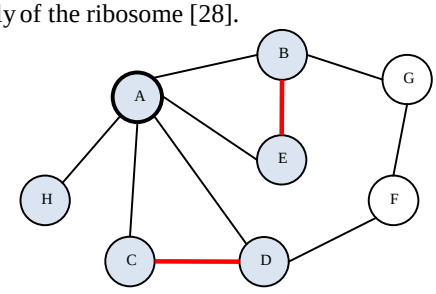


Fig. 6.Clustering coefficient of node A

Average clustering coefficient is the average of all individual clustering coefficients of the nodes in the

network. It is the clustering coefficient of a network that characterizes the overall tendency of nodes [3, 4] to form clusters or groups:

$$C = \frac{\sum_{i \in V} C_i}{N} = \frac{\sum_{i \in V} \frac{2n_i}{k_i(k_i - 1)}}{N} \quad (12)$$

, where n_i – the number of edges existing between the k_i nodes that are connected to node i,

k_i – the degree of node i.

Watts and Strogatz [34] have shown in their work that the clustering coefficient of many real systems networks is orders of magnitude larger than the one expected for a random network and, therefore, they are far from being random [33].

In particular, clustering coefficient, degree, the minimum path distance between pair of nodes, have attracted the attention of the physics community. Watts and Strogats have shown that real networks are characterized by a small average minimum path distance and a large clustering coefficient that together are named as *small world effect*. The topology of real networks is characterized by degree correlation and clustering hierarchy [33].

Clustering-Coefficient distribution. Cluster coefficient distribution *C(k)* is defined by averaging the clustering coefficients of nodes that have the same degree *k* [3]. It characterizes the diversity of cohesiveness of local neighborhoods and the overall tendency of nodes to form clusters or groups [5]. For many real networks, the exponential degree of the log-linear fit

$$C(k) = k^{-\gamma} \quad (13)$$

holds, which implies a network's hierarchical character or, in other words, is an indication of a network's hierarchical character. This function *C(k)* is a measure of the network's structure [3] and it captures a network's generic features and therefore can be used to classify various networks.

The **modularity metric** is defined as the gap between the fraction of arcs within clusters and the expected fraction of arcs if the arcs are wired with no structural bias [27]:

$$M = \sum_{i=1}^{r} \left[e_{ii} - \left(\sum_{j} e_{ij} \right)^2 \right] \quad (14)$$

, where r – the number of clusters,

e_{ij} – the fraction of arcs that leads between nodes (vertices) of cluster i and j.

The maximum modularity metric corresponds to the partition that comprise as many as within-module links and as few as possible inter-module links.

3.5 Software tools for structure analysis

Software tools NetworkAnalyzer and Visant analyze network structure.

Cytoscape is a complex network visualization and analysis tool supporting a standards (SBML, BIOPAX) and customizable network display styles [25]. Cytoscape specializes in the representation of interaction networks. Automatic layout algorithms help to organize massive amounts of interaction data relating to a set of molecules [5].

NetworkAnalyzer is the versatile Cytoscape plug-in that computes a comprehensive list of simple and complex topology parameters (single values and distributions) for directed and undirected networks using efficient graph algorithms. Simple parameters are the number of nodes, edges, self-loops, and connected components, the average number of neighbors, the network diameter, radius, density, centralization, heterogeneity, degree of component and clustering coefficient, the number of shortest paths, and the characteristic path length. Complex parameters are distributions of node degrees, neighborhood connectivity's, neighbors connectivity distribution, average clustering coefficients, average degree distribution, topological coefficients, shortest path lengths, shortest path distribution, average shortest path length and shared neighbors of two nodes. Network-Analyzer displays the distributions as histograms or scatters plots and allow export them as chart images in the formats JPG/PNG/SVG or as tables in plain text files [2].

VisANT is a web-based software framework (Java application) for visualizing and analyzing many types of networks of biological interactions and associations, as well as an especially useful tool for integrating information from a wide variety of sources. VisANT explicitly allows creation of mixed networks involving different types [28], i.e. networks containing both directed and undirected links [25]. Networks can also be analyzed for topological characteristics to identify larger global properties, such as degree (in-degree, out-degree) of a component, degree (in-degree, out-degree) distribution, average degree distribution, path length, shortest path length, clustering coefficient, clustering coefficient distribution calculations [25,29], clustering coefficient distribution and average clustering coefficient distribution, connections finding between a given set of genes/proteins, i.e., shortest path between two given components or given set of components, feedback loops (cycles), feedforward loops and self-loops.

VisANT provides:

(1) visualization, mining, analysis and modeling of the biological networks, which extend the application of GO [30].

(2) supporting exploratory pathway analysis using metagraphs, which includes multi-scale visualization of multiple pathways, editing and annotating pathways using a KEGG compatible visual notation and visualization of expression data in the context of pathways [29].

(3) the statistical and analytical tools needed for extracting topological properties of the user-defined networks.

4 CONCLUSIONS

The structural or topological analysis of biochemical networks contributes to a deeper understanding of network-wide interdependencies, causal relationships, and basic functional capabilities. The structural analysis is not based on quantitative and dynamic properties and can thus only provide qualitative answers.

Networks have “emergent” properties that are distinct from those of their individual components. Emergent properties are non-linear, aggregated and combinatory effects generated by the interaction of the components of the network.

The network structure properties and features can be categorized in the following groups: network metrics, network motifs, topological parameters(a Cytoscape plug-in) and VisANT are the most powerful tools for biochemical network structure analysis.

5 REFERENCES

[1] Alberghina L., Westerhoff H. (Eds.) (2005) Systems Biology. Definitions and Perspectives.Springer Verlag Berlin Weidelberg, 2005

[2] Assenov Y., Ramirez F., Schelhorn S.-H., Lengauer T., Albrecht M. (2008) Computing topological parameters of biological networks. Bioinformatics Applications Note, Vol.24 no. 2 2008, pp.: 282-284

[3] Barabasi A.L., Oltvai Z.N. (2004) Network biology: understanding the cell's functional organization. Nat Rev Genet 5: 101–113

[4] Boccaletti, S., Latora, V., Moreno, Y.,Chavez, M., Hwang, D.-U.(2006) Complex networks: Structure and dynamics. Physics Reports 424: 175-308.

[5] –Chen L., Wang R.-S., Zhang X.-S. (2009) Biomolecular networks: Methods and Applications in System Biology. John Wikey & Sons, Inc., Hoboken, New Jersey, 2009 ISBN 978-0-470-24373-2

[6] Cline M.S., Smoot M., Cerami E., Kuchinsky A., Landys N., Workman C., Christmas R., Avila-Campilo I., Creech M., Gross B., Hanspers K., Isserlin R., Kelley R., Killcoyne S., Lotia S., Maere S., Morris J., Ono K., Pavlovic V., Pico A.R., Vailaya A., Wang PL., Adler A., Conklin B.R., Hood L, Kuiper M., Sander C., Schmulevich I., Schwikowski B., J Warner G., Ideker T., D Bader G. (2007) Integration of biological networks and gene expression data using Cytoscape. Nature Protocols 2, 2007, Page(s): 2366 – 2382, doi:10.1038/nprot.2007.324

[7] Grundspeņķis J. (2006) Sistēmu teorija un vadība. ESF projekta ietvaros izveidots metodiskais materiāls.

[8] Han J.-D. J. (2008) Understanding biological functions through molecular networks. Cell Research (2008) 18:224-237

[9] Han J.D., Bertin N., Hao T., Goldberg D.S., Berriz G.F., Zhang L.V., Dupuy D., Walhout A. J., Cusick M.E., Roth F.P., and Vidal M. 2004. Evidence for dynamically organized modularity in the yeast protein–protein interaction network. Nature 430**:** 88–93.

[10] Hallinan J.S., Jackway P.T. Network Motifs, Feedback Loops and The Dynamic of Genetic Regulatory Networks.

[11] Hase T., Tanaka H., Suzuki Y., Nakagawa S., Kitano H. (2009) Structure of Protein Interaction Networks and Their Implications on Drug Design. PLoS Computational Biology, Volume 5, Issue 10, October 2009

[12] Herrgard M., Lee B.S., Portnoy V., Palsson B.O. (2006) Integrated analyses of regulatory and metabolic networks reveals

[13] Klamt S., Saez-Rodriguez J., A Lindquist J., Simeoni L., D Gilles E. (2006) A methodology for the structural and functional analysis of signaling and regulatory networks. BMC Bioinformatics2006, 7:56 doi:10.1186/1471-2105-7-56

[14] Kholodenko B., Bruggeman F., Sauro H. (2005) Mechanistic and modular approaches to modeling and inference of cellular regulatory networks, Systems Biology: Definitions and Perspectives. Springer, 2005

[15] Krishna S., Semsey S., Sneppen K. (2007) Combinatorics of feedback in cellular uptake and metabolism of small molecules. PNAS, Vol. 104, No.52, December 26, 2007, 20815-20819

[16] - Myers C. J. (2010) Engineering Genetics Circuits. Chapman & HALL/CRC Mathematical and computational Biology Series. 2010. Taylor and Francis Group, LLC. Page(s): 278.

[17] Newman, M.E.J. (2003) The structure and function of complex networks. SIAM Review 45:167-256. Press. August 25, 2007. [Available at: http://bioinformatics.oxfordjournals.org/cgi/reprint/btm401v1.pdf]

[18] Nishio Y., Usuda Y., Matsui K., Kurata H. (2008) Computer-aided rational design of the phosphotransferase system for enhanced glucose uptake in Escherichia coli. Molecular Systems Biology 4: Art.Nr. 160

[19] Patil A., Nakamura H. (2006) Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. FEBS Lett 580: 2041–2045

[20] Paxson R., Zannella K.. System biology: Studying the world’s most complex dynamic systems// Journal TheMathWorksNews&Notes - June (2007), pp.4-7

[21] Robins G, Pattison P., Koskinen J. (2008) Technical report. Network degree distribution. University of Melbourne, 2008

[22] Rubina T., Brusbardis V. (2009) Applications of biochemical networks discovering control mechanisms in systems biology. Proceedings of the Annuals Students International Scientific Conference “YOUTH IN SCIENCE AND PROFESSIONAL PRACTICE”, 2009
novel regulatory mechanisms in Saccharomyces cerevisiae. Genome research, 2006

[23] - Shannon P., Markiel A., Ozier O., Baliga NS., Wang JT., Ramage D., Amin N., Schwikowski B., Ideker T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research 2003 Nov; 13(11), Page(s): 2498-504

[24] Sontag E., Vilz-Cuba A., Laubenbacher R., Jarrah A.S. (2008) The Effect of Negative Feedbach Loops on the Dynamics of Boolean Networks. Biophysical Journal, Volume 95, July 2008, 528-526

[25] Suderman M., Hallett M. (2007) Tools for visually exploring biological networks. 2006. Published by Oxford University

[26] Zhang J., Shakhnovich E.-I. (2008) Sensitivity-dependent model of protein-protein interaction networks. Phys. Biol. 5 036011 (6pp.)

[27] Zhao J., Ding G.-H., Tai L., Yu H., Yu Z.-H., Luo J.-H., Cao Z.-W., Li Y.-X. (2007) Modular co-evolution of metabolic networks. BMC Bioinformatics 2007, 8:311

[28] - Zhenjun Hu, Joe Mellor, Jie Wu, Takuji Yamada, Dustin Holloway, Charles DeLisi. (2005) VisANT: data-integrating visual framework for biological networks and modules. Nucleic Acids Research 2005, Volume 33(Web Server Issue):W352-W357

[29] - Zhenjun Hu, David M. Ng., Takuji Yamada, Chunnuan Chen, Shuichi Kawashima, Joe Mellor, Bolan Linghu, Minoru Kanehisa, Joshua M. Stuart, Charles DeLisi. (2007) VisANT 3.0: new modules for pathway visualization, editing, prediction and construction. Nucleic Acids Research, 2007.

[30] - Zhenjun Hu, Jui-Hung Hung, Yan Wang, Yi-Chien Chang, Chia-Ling Huang, Matt Huyck, Charles DeLisi. (2009) VisANT 3.5multi-scale network visualization, analysis and inference based on the gene onthology. Nucleic Acids Research, 2009, Vol. 37, Web Server issue W115-W121

[31] –Zinovyev A., Viara E., Calzone L., Emmanuel Barillot. (2008) BiNoM: a Cytoscape plugin for manipulating and analyzing biological networks. Bioninformatics applications note. Vol. 24 no. 6 2008, Page(s): 876-877

[32] Zinovyev A., Calzone L. Binom Manual Version 1.0. InstitutCurie.

[33] Vazquez A. (2003) Growing network with local rules: preferential attachment, clustering hierarchy, and degree correlations. Phys Rev E Stat Nonlin Soft Matter Phys 67: 056104

[34] Watts D.J., Strogatz S.H. (1998) Collective dynamics of ‘small-world’ networks. Nature, Vol. 393, No. 6684. (04 June 1998), pp. 440-442.

[35] Wilson R.J. (1972) Introduction to Graph Theory. Oliver and Boyd. Edinburgh, 1972