

Construction of Sub-Clusters in the Cluster of Graph Representing the Protein Network

Tsitsiashvili Gurami*

Institute for Applied Mathematics FEB RAS, Vladivostok, Russia

Abstract

Previously an algorithm is constructed to replace each cluster (a class of cyclically equivalent vertices) in a directed graph (representing a protein network) with an acyclic sub-graph preserving all input and output vertices. This algorithm is continuing by an introduction of minimal number of edges between the output vertices and the input vertices in order to introduce feedbacks, stabilizing the functioning of the network.

Keywords: Cluster; Acyclic digraph; Sub-graph; Protein network

Background

A problem of a visual display in protein networks is discussing in previous study [1,2]. Consider a directed graph (digraph) G with a set of vertices U and a set of edges (oriented) V . The set U may be clustered with respect to cyclic equivalence by the algorithm of previous study [3]. The vertices u_1, u_2 of the digraph are calling cyclically equivalent if there is a cycle containing these two vertices in the graph G . On the set of vertices of the digraph, we determine the ratio of partial order $u_1 \geq u_2$, if there is a path from the vertex u_1 to the vertex u_2 .

In a study [4] the algorithm of replacement in the digraph G of each cluster $g \in G$ on an acyclic sub-graph \bar{g} preserving all its input $p(g)$ and output $QW(g)$ vertices in the digraph is constructed. Under the entrance vertices of the cluster, g we mean the vertices, which include edges from the outside of the cluster, under the output vertices we mean the vertices, which have edges going to the outside of the cluster. After replacing clusters g by appropriate acyclic digraphs \bar{g} , all edges linking clusters are preserved.

The algorithm for converting a cluster g into an acyclic digraph \bar{g} consists of the following steps. First, we construct a directed acyclic sub-graph \bar{g} of the digraph g containing a non-empty subset $U(\bar{g}) \subseteq U(g)$ of the set of input vertices and the set $V(g)$ of all output vertices. For this purpose, the cluster g implements a recurrent wave front algorithm from the set of all input vertices $U(g)$ until the front passes through all output vertices $V(g)$.

Then we construct an acyclic sub-graph \bar{g} , consisting of paths starting at some vertex of the set $U(\bar{g}) = U(g) \setminus U(\bar{g})$ and ending at some vertex of the set $P(\bar{g})$. For this purpose, in the cluster g , we implement a recurrent algorithm of the wave backward front from the set of all vertices $P(\bar{g})$ until the front passes through all input vertices $U(\bar{g})$.

The algorithm ends with the union $\bar{g} = \bar{g} \cup \bar{g}$, forming an acyclic graph satisfying the required inclusion $U(g) \cup V(g) \subseteq P(\bar{g})$.

Results

The idea of new edges introduction into digraph describing protein network is earlier discussing in previous study [5]. It follows from the algorithm of the previous section that in the acyclic sub-graph \bar{g} those vertices of the set $V(g)$ are maximal, of which there are no edges in the graph \bar{g} . Due to the equality all, these vertices are maximal in the graph \bar{g} . Thus, all maximal vertices of the graph \bar{g} are output in cluster g .

Similarly, we obtain that in the graph \bar{g} all minimal vertices

belong to the set $U(\bar{g})$, if they are not connecting by paths from the set $U(\bar{g})$ to some vertex of the set $U(\bar{g}) \subseteq P(\bar{g})$. In turn; in the graph, \bar{g} the vertices of the set $U(\bar{g})$, that are not including paths from the set, $U(\bar{g})$, are also minimal. Thus, all minimal vertices of the graph \bar{g} are inputs to the cluster g .

It also following from the construction of the graph \bar{g} , any output vertex that is not maximal lies on one of paths, starting at some minimal vertex and ending at some maximal vertex of the graph \bar{g} .

Similarly, any input vertex that is not minimal lies on one of the paths, starting at some minimal vertex and ending at some maximal vertex of the graph \bar{g} .

We now pose the question, how to make the introduction of a minimum number of new edges such that each input vertex belongs some cycle containing some output vertex, and each output vertex belongs some cycle containing some input vertex. The answer to this question is giving by the following construction.

We include the edges coming from the set $Max(\bar{g})$ of all maximal vertices to the set $Min(\bar{g})$ of all minimal vertices of the graph \bar{g} . For any pair of vertices $u_1 \in Min(\bar{g})$, $u_2 \in Max(\bar{g})$, connecting in some way, we introduce a new edge $u_2 \rightarrow u_1$. The resulting digraph \bar{g} decomposes into a disconnected clusters. Moreover, if the vertex u_1 is the input in the graph \bar{g} , the vertex u_2 is the output, and these vertices are connecting in some way, then they are containing some cycle of the digraph \bar{g} .

Note that if there is a path from the vertex $u_1 \in Min(\bar{g})$, to the vertex $u_2 \in Max(\bar{g})$ in the graph \bar{g} , then the undirected edge (u_1, u_2) can be including between these vertices. In addition, the resulting undirected graph of these edges splits into connectivity components whose vertices are including in the graph \bar{g} .

***Corresponding author:** Gurami T, Institute for Applied Mathematics FEB RAS, Vladivostok, Russia, Tel: 49228734875; E-mail: guram@iam.dvo.ru

Received August 19, 2018; **Accepted** August 27, 2018; **Published** August 31, 2018

Citation: Gurami T (2018) Construction of Sub-Clusters in the Cluster of Graph Representing the Protein Network. J Biom Biostat 9: 405. doi: [10.4172/2155-6180.1000405](https://doi.org/10.4172/2155-6180.1000405)

Copyright: © 2018 Gurami T. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

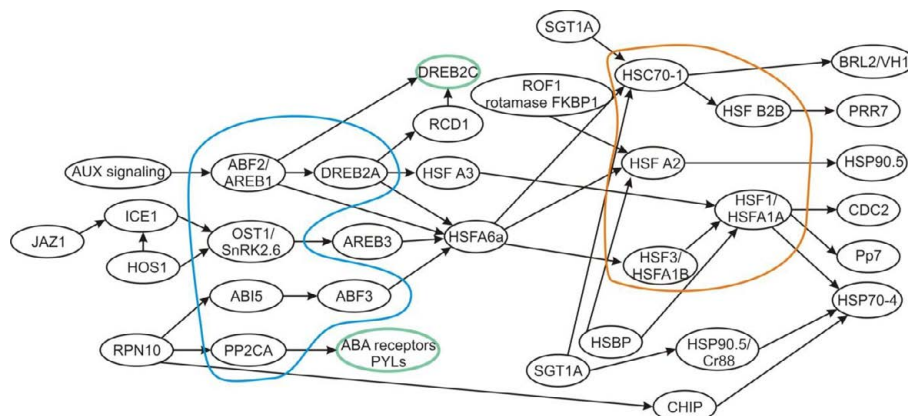


Figure 1: Replacement of clusters by shortest ways between input and output nodes in clusters.

Real Data Applications

In a study [4] the protein network using for research of plants stability to drought and extreme temperatures is clustering and replacing by acyclic digraph (Figure 1). Main cluster of this network, containing 14 proteins, is replacing by acyclic digraph, outlining by blue color. Using suggested algorithm it is possible add the following edges: DERB2A → ABF2/AREB1, ABF3 → ABI5 supplying necessary feedbacks.

Acknowledgement

Partially supported by Russian Fund of Basic Researches (Project 17-07-00177).

References

1. Pircon N, Fortemaion Ch, Jacobs S, Dremier J, Dumont E, et al. (2000) The visual display of regulatory information and networks. Trends in Cell Biology 10: 404-408.
2. Vishwakarma K, Upadhyay N, Kumar N, Yadav G, Singh J, et al. (2017) Absciscic Acid Signaling and Abiotic Stress Tolerance in Plants: A Review on Current Knowledge and Future Prospects. Front Plant Sci 20: 8-161.
3. Tsitsiashvili G (2013) Sequential algorithms of graph nodes factorization. Reliability: Theory and Applications 8: 30-33.
4. Tsitsiashvili G, Bulgakov VP, Losev AS (2018) Replacement of Directed Graph by Acyclic Directed Graph and Its Application in Biostatistic. Journal of Biometrics and Biostatistics 9: 390.
5. Tsitsiashvili G, Bulgakov VP, Losev AS (2018) Introduction of Edges in Directed Graph. Applied Mathematical Sciences 12: 903-910.