Proceedings of the International Workshop on Intelligent Information Systems
IIS2011, September 13-14, 2011, Chisinau, Republic of Moldova

Feedback Loop Centrality to Decide Gene Essentiality in Cellular Signaling Networks

V. Rogojin, S. Hautaniemi

Abstract

We introduce the new concept of graph-node centrality measure based on feedback loops (elementary cycles) in cellular signaling networks in order to identify key genes/proteins that are essential for the malignant cellular behavior. Our hypothesis is that the essential nodes participate in many feedback loops (elementary cycles). We expect our feedback loop centrality scores may help to direct experiments to genes and complexes that most probably play a crucial role in drug resistance and invasion. Feedback loop centrality involves counting/enumerating all elementary cycles the node is participating in. Moreover, we suggest an optimization for cycle counting procedure on cellular networks.

Keywords: Node centrality measure, elementary cycle, feedback loop, signaling network, digraph, scale-free network, duplication-divergence network, cancer, malignant cell, robustness

1 Introduction

Malignant cells exercise a robust behavior that is deleterious for the organism [1]. The general goal of our research is to reveal the key mechanisms responsible for the robustness of malignant cellular behavior. In order to solve this task, we apply discrete mathematics to study the structural and the dynamical properties of cellular signaling networks associated to cancers. For a review on using graph theory to study cellular networks we refer for instance to [2].

On the level of cellular networks, the robustness is in a great part due to feedback control loops [6]. Feedback loops play a decisive role in maintaining cellular functions in the face of internal or external uncertainties. Thus, we hypothesize that by destroying major part of feedback loops in a cellular signalling network we may compromise the robustness of its characteristic behavior. This idea enables us to predict the essential nodes in a cellular network by considering number of feedback loops they participate in.

In this paper, we propose a new centrality score to measure the "importance" of a network node basing on the number of elementary cycles it is participating in and also an idea of a cycle enumeration method optimized specifically for cellular networks. This work is presented in details in [5].

2 Compressing cellular networks

Formally, a cellular signaling network may be represented as a directed graph $G = (V, E)$ with nodes standing for the network's component and directed edges standing for the interactions [2]. For more background on graph theory and the notations we refer to [8].

A number of statistical studies on topological features of cellular networks [2] have revealed that great part of networks found in nature possess a number of groups of nodes that are connected to the same neighbors, see for instance [3].

We can compress a cellular network represented as graph a G to a network represented by G' such that there is a bijection between G and G' in the following manner: each set of nodes S that share same neighbors we substitute by a vertex p_S and connect it to all the neighbors of members of S.

As a consequence, the total number of paths, and hence, of elementary cycles in G' can be much lower than in G. Figure 1 demonstrates the idea.

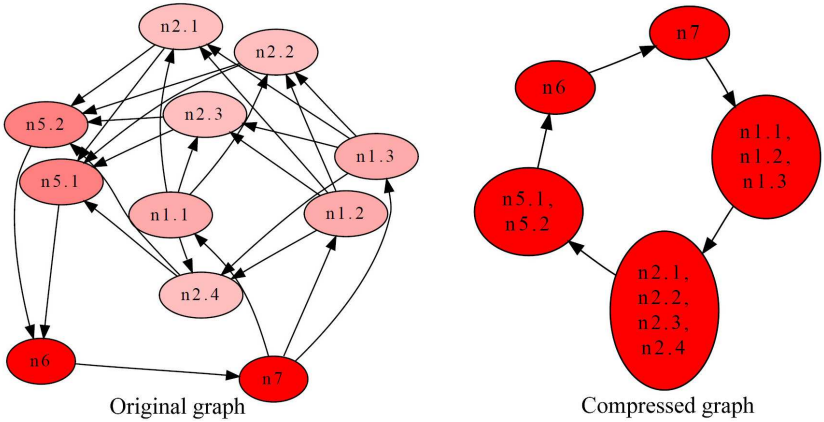

Figure 1. The original graph has 24 FBLs while the compressed graph has only one FBL. Groups of nodes $\{n1.x\}, \{n2.x\}$ and $\{n5.x\}$ share same neighbors and can be compressed.

3 Feedback loop centrality measure

The intuition behind our feedback loop centrality measure is as follows: an important node participates in a great amount of feedback loops.

Computing FBL scores involves counting all FBLs in a network. We have implemented Szwarcfiter and Lauer's cycle enumeration algorithm [7] in Java as an Anduril component [4] for directed graphs (digraphs) with the time complexity $O(n + m(c + 1))$ and space complexity $O(n + m)$, where n is the number of nodes in a digraph, m is the number of its edges and c is the total number of elementary cycles.

The problem of cycle counting/enumeration in a digraph is NP-complete. We propose here an idea how to optimize cycle counting on cellular networks: instead of counting cycles on a graph G representing a cellular network, we can count cycles on compressed graph G', then derive information about cycles in the original graph G basing on cycles from its compressed version. This optimization, however, remains to be developed.

References

[1] D.Hanahan, and R.A.Weinberg, *Hallmarks of Cancer: The Next Generation.* Cell, **144**, Elsevier, (2011), pp. 646–674.

[2] O.Mason, and M.Verwoerd, *Graph theory and networks in Biology.* IET Systems Biology, **1**(2), (2007), pp. 89–119, doi:10.1049/iet-syb:20060038.

[3] M.Middendorf, et al., *Discriminative topological features reveal biological network mechanisms.* BMC Bioinformatics, **5**:181, doi:10.1186/1471-2105-5-181.

[4] K.Ovaska, M.Laakso, et al., *Large-scale data integration framework provides a comprehensive view on glioblastoma multiforme.* Genome Medicine, (2010), 2:65, doi:10.1186/gm186.

[5] V.Rogojin, and S.Hautaniemi, *On feedback Loop centrality.* (To be submitted), 2011.

[6] J.Stelling, U.Sauer, Z.Szallasi, F.J.Doyle, and J.Doyle, *Robustness of Cellular Functions.* Cell, **118**, Cell Press, (2004), pp. 675–685.

[7] J.L.Szwartsfiter, and P.E.Lauer, *A search strategy for the elementary cycles of a directed graph.* BIT, (16), 1976, pp. 192–204.

[8] D. B.West, *Introduction to Graph Theory*, Prentice Hall, Upper Saddle River, NJ (1996).

Vladimir Rogojin, Sampsa Hautaniemi,

Vladimir Rogojin, Sampsa Hautaniemi
Computational Systems Biology Laboratory, Genome-scale Biology Research Program, University of Helsinki
Biomedicum Helsinki, B524a P.O.Box 63 (Haartmaninkatu 8) 00014 University of Helsinki
Phone: +358 919 125 407, +358 9 191 25419
E–mail: vladimir.rogojin@helsinki.fi, sampsa.hautaniemi@helsinki.fi