EPJ manuscript No.
(will be inserted by the editor)

Metabolic subsystems and network science

Rok Novosel[1] and Matija Čufar[2]

Fakulteta za računalništvo in informatiko

the date of receipt and acceptance should be inserted later

Abstract. Subsystems are parts of a metabolism that perform different important tasks in a cell. In this article, we will explore these subsystems from a network science point of view. We will attempt to find ways of detecting subsystems in a metabolic network and compare their structure. We will present our results on the Chinese hamster ovary cell, a mammalian cell that is commonly used in biomedical research and in biotechnology.

PACS. XX.XX.XX No PACS code given

1 Introduction

Since the beginning the turn of the century, life sciences have been evolving rapidly. Advances in data acquisition, storage and analysis technology have allowed scientist to gather immense amounts of data and build complex models from it[4]. These complicated models have brought people of various backgrounds, such as physics, mathemathics and computer science into the field of biology.

One of such fileds, itself a very recent development, is network science, which is often used to analyze different kinds of networks that appear in the various subfileds of modern biology, including ecology[8], systems biology[1] and neuroscience[9].

Metabolic networks[5] are used to model the metabolisms of various organisms. They are usually represented with a bipartite graph composed of two types of vertices: reactions and chemicals produced and consumed by the reactions. The where edges in such a network connect chemicals to reactions. Furthermore, the edges are directed indicating whether the chemical was produced or consumed. A third kind of vertices can be added to represent enzymes that catalyze the reaction, but do not directly partake in it. Other commonly used representations are simplified reprpresentations, where one of the types of vertices is omitted[7].

In this article, we will analyze metabolism subsystems and how they can be analyzed, compared and detectied using approaches from network science.

2 Methods

Some initial ideas for methods we could use:

- Community detection: how do different algorithms detect subsystems?
- Motifs: what motifs are the most common and which motifs appear in different subsystems? We will look for feedback and feed-forward loops.
- Maybe some other techniques to compare the subsystems with?
- Try the idea from [3].

3 Resutls

3.1 Network global structure overview

In this article, we will analyse a metabolic network of the Chinese hamster ovary (CHO) cell. The CHO cell is frequently used in biological and medical research and in the production of biopharmaceuticals[2].

We have used a whole-cell metabolic network of the Chinese hamster ovary (CHO) cell that was taken from the BiGG database[6, 2]. The original network contains 4,456 metabolites that take part in 6,663 reactions. The reactions and metabolites are annotated with additional metadata, such as name, BiGG ID, subsystem etc.

We have simplified the network to a simple directed graph, where reactions are represented with nodes. If one reaction produces a metabolite that is used by another reaction, they are connected by an arc. This network has 6,663 nodes and 656,609 arcs.

The network has a very large connected component of 6,036 nodes, while the other components are very small, as they are composed of at most 4 nodes. The largest connected component contains a strongly connected component of 5,307 nodes, while the other nodes are isolated. These probably represent sources and sinks of the metabolism.

The network has scale-free structure. Its in-degree, out-degree and degree distributions are plotted in figure 1. The estimated scale factors for the network are $\gamma_{in} = 2.5$, $\gamma_{out} = 2.3$ and $\gamma = 2.0$.

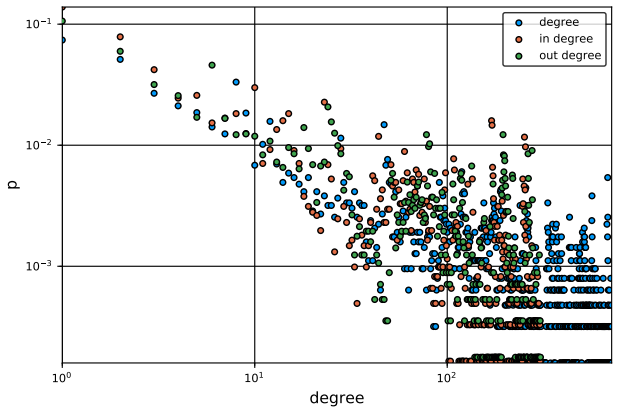

Fig. 1. The in-degree, out-degree and degree distributions of the network.

4 Authors contributions

All the authors were involved in the preparation of the manuscript. All the authors have read and approved the final manuscript.

References

1. Albert-Laszlo Barabasi and Zoltan N Oltvai. Network biology: understanding the cell's functional organization. *Nature reviews genetics*, 5(2):101–113, 2004.
2. Hooman Hefzi, Kok Siong Ang, Michael Hanscho, Aarash Bordbar, David Ruckerbauer, Meiyappan Lakshmanan, Camila A Orellana, Deniz Baycin-Hizal, Yingxiang Huang, Daniel Ley, et al. A consensus genome-scale reconstruction of chinese hamster ovary cell metabolism. *Cell Systems*, 3(5):434–443, 2016.
3. Petter Holme, Mikael Huss, and Hawoong Jeong. Subnetwork hierarchies of biochemical pathways. *Bioinformatics*, 19(4):532–538, 2003.
4. B. Ingalls. *Mathematical Modelling in Systems Biology: An Introduction*. Applied Mathematics, University of Waterloo, 2012.
5. Hawoong Jeong, Bálint Tombor, Réka Albert, Zoltan N Oltvai, and A-L Barabási. The large-scale organization of metabolic networks. *Nature*, 407(6804):651–654, 2000.
6. Zachary A. King, Justin Lu, Andreas Drger, Philip Miller, Stephen Federowicz, Joshua A. Lerman, Ali Ebrahim, Bernhard O. Palsson, and Nathan E. Lewis. Bigg models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Research*, 44(D1):D515, 2016.
7. M. Newman. *Networks: An Introduction*. OUP Oxford, 2010.
8. Stephen R Proulx, Daniel EL Promislow, and Patrick C Phillips. Network thinking in ecology and evolution. *Trends in Ecology & Evolution*, 20(6):345–353, 2005.
9. Olaf Sporns. Contributions and challenges for network models in cognitive neuroscience. *Nature neuroscience*, 17(5):652–660, 2014.