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

Metabolic networks [2] are used to model metabolism in various organisms. We usually have two types of vertices in a metabolic network: reactions and chemicals produced and consumed by the reactions. The most correct way to represent this kind of structure is with a bipartite network, where edges connect chemicals to reactions. Furthermore the edges are directed indicating whether the chemical was produced or consumed.

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?

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[1].

We have used a whole-cell metabolic network of the Chinese hamster ovary (CHO) cell that was taken from the BiGG database[3,1]. 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.

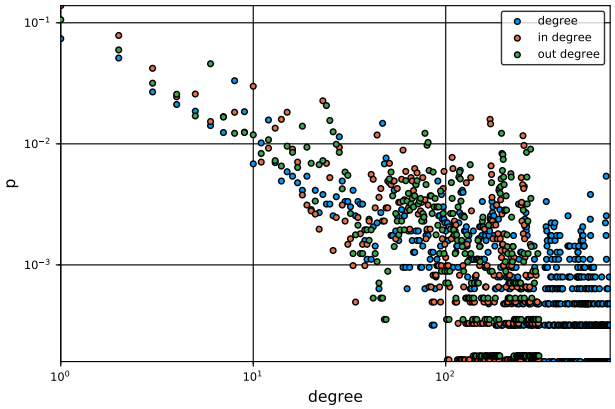

Fig. 1. The in-degree, out-degree and degree distributions of the network.

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 $\gamma = 2.0$.

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. 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.
2. 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.
3. 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.