HiVis: ~~A~~ a portable, scalable tool for hierarchical visualization and analysis of biological networks

Zhiwen Qiang [1], Hao Chen [2], Xiaopeng Zhu [3], Shikui Tu [1*]

[1]Department of Computer Science and Engineering, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, DongChuan Road, Shanghai, 200240, ~~Shanghai,~~ China

[2]Bioinformatics Program, Boston University, Boston, 02215, ~~Boston,~~ USA

[3]School of Computer Science, Carnegie Mellon University, Pittsburgh, 15213, ~~Pittsburgh,~~ USA

Abstract

It is very important to analyze the network structure of interacting genes, proteins, RNAs, etc. in large-scale biological networks which represent complex biological systems. Many tools have been made for this purpose. However, they are either unable to display the hierarchically structured view of the networks or not easy to build in cross-platforms. Here, we present a navigation tool called HiVis for biological network visualization. HiVis provides a hierarchical view of the networks through a zoom-in or zoom-out function powered by k-means and fast approximate spectral clustering algorithms. It is a cross-platform, portable, fast desktop application to large-scale networks.

~~Introduction~~Background

To the Editor: motivated by HiMap ~~Lei Shi~~ (Shi et al. ~~(~~ **2009**), we present HiVis, an open-source software that provides hierarchical interactive visualization and analysis for biological networks. It contains three main features:

- Not only is it able to visualize networks which ~~contains~~ contain tens of thousands interactions in a hierarchically structured way, but also it allows the analysis of the local connections centered on a specific element.
- It provides a suite of interactive methods for the user to analyze structural detail of the network.
- HiVis is a portable cross-platform application, so that users ~~don't~~ do not need to go through the tedious installation process and bothered by installing a lot of dependent packages.

Methods

We use a top-down approach to cluster the biological networks into a multilevel node-link graph. We demonstrate the main features in ~~Figure~~ Fig. **1** through a ~~dataset~~ data set of STRING interaction data ~~Protein-Protein Interaction Networks (0000~~(https://string-db.org/cgi/input.pl). It is well known that many biological networks contain clique and hub genes which ~~is~~ are highly clustered and ~~contains~~ contain self-similar structure. ~~So~~ Therefore, clustering them together into one node can reveal their hierarchical information and make the visualization readable.

40535_2018_50_Fig1_print.png

Fig. 1 Illustration of features in HiVis, a software that provides hierarchical visualization and analysis of biological networks

Currently, we provide two cluster algorithms, *k*-means clustering, and fast approximate spectral ~~clusteringYan~~ clustering (Yan et al. (**2009**), which is faster than normal spectral clustering method and is more suitable for network clustering than *k*-means algorithm. The idea is to use k-means as a preprocessor ~~,~~ and then to perform the normal ~~spectural~~ spectral clustering method. Here, we also implemented a heuristic approach to determine the number of clusters by checking the slope of neighboring points of sorted eigenvalues.

Implementation

It is difficult to display the biological network meaningfully in one scope, HiVis ~~solve~~ solves it by cluster them into a hierarchical tree and only reveal the nodes in the current view and the view in the next hierarchy. To be more specific, here, we present a example illustrated in ~~Figure~~ Fig. **1**, ~~1a~~ **1**a is the overview of HiVis, the genes are clustered using *k*-means, and each node represents a group of them. It can switch the algorithm to modified ~~spectrual~~ spectral clustering ~~,~~ and the result is shown in ~~Figure~~ Fig. **1**d. HiVis offers smooth animation methods to further analyze the computed clusters. When selecting one cluster, it will show a full scope of its inner structure~~,~~ ; as shown in ~~Figure~~ Fig. **1**b, users can also return to the previous stage, or they can further analyze the current cluster with the same procedure. Apart from that ~~,~~ HiVis will show each genes within the cluster, as is shown in the left part of ~~Figure~~ Fig. **1**b, when choosing one gene, HiVis will display the gene and its most related genes in a network, as is shown in ~~Figure~~ Fig. **1**c. Thus, it will help to analyze the network locally.

Conclusion

In this paper, we present HiVis, a software built on electron for hierarchical visualization and analysis of biological networks. We implemented the *k*-means clustering, fast approximate spectral clustering algorithm, and force-direct layout algorithm to let HiVis capable of handling large-scale biological data. It is worth noting that this kind of displaying technique is very general and can be applied to many other field like social network, security network, etc.

40535_2018_50_OnlinePDF.pdf

~~Author~~Authors' ~~s~~ contributions

The authors discussed the problem and the solutions were proposed all together. All authors participated in drafting and revising the final manuscript. All authors read and approved the final manuscript.

manuscript.

References

Lei Shi, N., Cao, S.X., Liu, W.H., Qian, L., Tan, G.D., Wang, J.M., Sun, C.Y., Lin: Himap: Adaptive

References

Protein–protein interaction networks. https://string-db.org/cgi/input.pl

Shi L, Cao N, Liu SX, Qian WH, Tan L, Wang GD, Sun JM, Lin CY (2009) Himap: adaptive visualization of large-scale online social networks. In: IEEE Pacific ~~Visualization Symposium~~visualization symposium, PacificVis ~~2009 - Proceedings, 41–48 (2009)~~

Yan D, Huang L, Jordan MI (2009) Fast approximate spectral clustering. In: Proceedings of the 15th ACM SIGKDD international conference on ~~Knowledge~~ knowledge discovery and data ~~mining - KDD '09, 907 (2009)~~mining—KDD '09. p 907