ModuLand plug-in for Cytoscape: determination of hierarchical layers of overlapping network modules and community centrality

Máté Szalay-Bekő1,†, Robin Palotai1,†, Balázs Szappanos2, István A. Kovács1,3,4, Balázs Papp2 and Péter Csermely1,*

1Department of Medical Chemistry, Semmelweis University, Budapest 1444, 2Institute of Biochemistry, Biological Research Centre, Hungarian Academy of Sciences, Szeged 6726, 3Department of Physics, Loránd Eötvös University and 4Institute for Solid State Physics and Optics, Wigner Research Centre, Hungarian Academy of Sciences, Budapest 1525, Hungary

ABSTRACT
Summary: The ModuLand plug-in provides Cytoscape users an algorithm for determining extensively overlapping network modules. Moreover, it identifies several hierarchical layers of modules, where meta-nodes of the higher hierarchical layer represent modules of the lower layer. The tool assigns module cores, which predict the function of the whole module, and determines key nodes bridging two or multiple modules. The plug-in has a detailed JAVA-based graphical interface with various colouring options. The ModuLand tool can run on Windows, Linux or Mac OS. We demonstrate its use on protein structure and metabolic networks.


Contact: csermely.peter@med.semmelweis-univ.hu

Supplementary information: Supplementary data are available at Bioinformatics online.

Received on November 21, 2011; revised on June 13, 2012; accepted on June 14, 2012

1 INTRODUCTION

Nodes of biological networks often belong to multiple network communities. Recently, a number of methods were published to determine tightly or extensively overlapping network modules (Adamscek et al. 2004; Abe et al. 2010; Borgnat et al. 2010; Kovács et al. 2011; Mithalik and Csermely 2011; Path et al. 2009). Our ModuLand framework (Kovács et al. 2010) introduced community landscapes. The x–y plane of a community landscape is a conventional 2D visualization of the network, whereas the z-axis represents community centrality. Community centrality of a given edge (or node) was defined as the sum of local influence zones of all network edges (or nodes) including the given edge (or node; Supplementary Fig. S1). Thus community centrality represents an integrated measure of the whole network’s influence to one of its edges or nodes. Hills of the community landscape correspond to network modules (Supplementary Fig. S1) yielding extensive overlaps. This concept led to the development of the ModuLand family of network modularization methods (Kovács et al. 2011).

2 SOFTWARE OVERVIEW

The widely used Cytoscape program (Shannon et al. 2003) has several very useful clustering plug-ins available (Bader and Houge 2003; Morey et al. 2011; Rhiissonraksu and Gonsalvez 2011; Rivera et al., 2011; Su et al. 2011). However, these methods do not focus on extensive modular overlaps, and do not build a modular hierarchy, where meta-nodes of the higher level represent modules of the lower level. Moreover, existing plug-ins do not provide measures identifying the centre of the module, as well as key nodes bridging two or multiple modules (see Supplementary Table S9 and Supplementary Discussion). Here, we introduce the Cytoscape plug-in of the most widely applicable version of the ModuLand method family (Kovács et al. 2011). We demonstrate its ability to determine biologically relevant, extensively overlapping network modules, hierarchical layers of modules, module cores and key inter-modular nodes using protein structure and metabolic networks.
Fig. 1. The hierarchical modular structure determined by the ModuLand Cytoscape plug-in. The left side shows the protein structure network of *E.coli* Met-tRNA synthase and its three hierarchical levels as determined by the plug-in. Each meta-node of a higher hierarchical level represents a module of the level right below. All networks are coloured according to the five modules identified on hierarchical level 1. This colouring option can be performed by the two clicks of the plug-in main dialog box shown on the right.

The lower number of modules at higher hierarchical levels may be visualized either using the meta-nodes of the higher hierarchical level itself, or projecting this higher level modular structure to the nodes of the original network. On any level of module hierarchy, nodes or meta-nodes can be visualized assigning them the colour of the module they mostly belong to. This shows a non-overlapping assignment of nodes to modules. Node colours can also visualize several node (or meta-node) measures including weighted degree, betweenness centrality, community centrality, overlap and bridgeness (see Supplementary Fig. 2).

The plug-in has an option to merge highly similar module pairs containing roughly the same nodes or meta-nodes with the same intensity. For merging of modules the plug-in offers a correlation histogram, and allows the user to select an appropriate correlation threshold. The runtime complexity of the plug-in version remained \( \sim O(n^3) \), as defined earlier [Kovács et al., 2010]. To enhance the performance of the plug-in for calculating the higher hierarchical layers further, we introduced a user-selected optimization. This results in the disappearance of meta-edges with very small weights at the higher hierarchical levels. These low intensity meta-edges are derived from the minor overlaps of distant modules of the lower level. This optimization allowed a speedup in running time by a factor of 7 for larger networks [Supplementary Table S10].

The plug-in is capable to generate overview reports for each hierarchical level. These reports list the number of the nodes (meta-nodes), edges (meta-edges) and modules, the effective number of modules (see Kovács et al., 2010) and the size of each module. The overview also contains the list of the 10 nodes of each module having maximal module assignment value to the corresponding module at one level below in the hierarchy.

Node colours can also visualize several node (or meta-node) measures including weighted degree, betweenness centrality, community centrality, overlap and bridgeness (see Supplementary Fig. 2). The plug-in has an option to merge highly similar module pairs containing roughly the same nodes or meta-nodes with the same intensity. For merging of modules the plug-in offers a correlation histogram, and allows the user to select an appropriate correlation threshold. The runtime complexity of the plug-in version remained \( \sim O(n^3) \), as defined earlier [Kovács et al., 2010]. To enhance the performance of the plug-in for calculating the higher hierarchical layers further, we introduced a user-selected optimization. This results in the disappearance of meta-edges with very small weights at the higher hierarchical levels. These low intensity meta-edges are derived from the minor overlaps of distant modules of the lower level. This optimization allowed a speedup in running time by a factor of 7 for larger networks [Supplementary Table S10].

The plug-in is capable to generate overview reports for each hierarchical level. These reports list the number of the nodes (meta-nodes), edges (meta-edges) and modules, the effective number of modules (see Kovács et al., 2010) and the size of each module. The overview also contains the list of the 10 nodes of each module having maximal module assignment value to the corresponding module at one level below in the hierarchy.

Node colours can also visualize several node (or meta-node) measures including weighted degree, betweenness centrality, community centrality, overlap and bridgeness (see Supplementary Fig. 2). The plug-in has an option to merge highly similar module pairs containing roughly the same nodes or meta-nodes with the same intensity. For merging of modules the plug-in offers a correlation histogram, and allows the user to select an appropriate correlation threshold. The runtime complexity of the plug-in version remained \( \sim O(n^3) \), as defined earlier [Kovács et al., 2010]. To enhance the performance of the plug-in for calculating the higher hierarchical layers further, we introduced a user-selected optimization. This results in the disappearance of meta-edges with very small weights at the higher hierarchical levels. These low intensity meta-edges are derived from the minor overlaps of distant modules of the lower level. This optimization allowed a speedup in running time by a factor of 7 for larger networks [Supplementary Table S10].

The plug-in is capable to generate overview reports for each hierarchical level. These reports list the number of the nodes (meta-nodes), edges (meta-edges) and modules, the effective number of modules (see Kovács et al., 2010) and the size of each module. The overview also contains the list of the 10 nodes of each module having maximal module assignment value to the corresponding module at one level below in the hierarchy.

Node colours can also visualize several node (or meta-node) measures including weighted degree, betweenness centrality, community centrality, overlap and bridgeness (see Supplementary Fig. 2). The plug-in has an option to merge highly similar module pairs containing roughly the same nodes or meta-nodes with the same intensity. For merging of modules the plug-in offers a correlation histogram, and allows the user to select an appropriate correlation threshold. The runtime complexity of the plug-in version remained \( \sim O(n^3) \), as defined earlier [Kovács et al., 2010]. To enhance the performance of the plug-in for calculating the higher hierarchical layers further, we introduced a user-selected optimization. This results in the disappearance of meta-edges with very small weights at the higher hierarchical levels. These low intensity meta-edges are derived from the minor overlaps of distant modules of the lower level. This optimization allowed a speedup in running time by a factor of 7 for larger networks [Supplementary Table S10].

The plug-in is capable to generate overview reports for each hierarchical level. These reports list the number of the nodes (meta-nodes), edges (meta-edges) and modules, the effective number of modules (see Kovács et al., 2010) and the size of each module. The overview also contains the list of the 10 nodes of each module having maximal module assignment value to the corresponding module at one level below in the hierarchy.

Node colours can also visualize several node (or meta-node) measures including weighted degree, betweenness centrality, community centrality, overlap and bridgeness (see Supplementary Fig. 2). The plug-in has an option to merge highly similar module pairs containing roughly the same nodes or meta-nodes with the same intensity. For merging of modules the plug-in offers a correlation histogram, and allows the user to select an appropriate correlation threshold. The runtime complexity of the plug-in version remained \( \sim O(n^3) \), as defined earlier [Kovács et al., 2010]. To enhance the performance of the plug-in for calculating the higher hierarchical layers further, we introduced a user-selected optimization. This results in the disappearance of meta-edges with very small weights at the higher hierarchical levels. These low intensity meta-edges are derived from the minor overlaps of distant modules of the lower level. This optimization allowed a speedup in running time by a factor of 7 for larger networks [Supplementary Table S10].

The plug-in is capable to generate overview reports for each hierarchical level. These reports list the number of the nodes (meta-nodes), edges (meta-edges) and modules, the effective number of modules (see Kovács et al., 2010) and the size of each module. The overview also contains the list of the 10 nodes of each module having maximal module assignment value to the corresponding module at one level below in the hierarchy.

Node colours can also visualize several node (or meta-node) measures including weighted degree, betweenness centrality, community centrality, overlap and bridgeness (see Supplementary Fig. 2). The plug-in has an option to merge highly similar module pairs containing roughly the same nodes or meta-nodes with the same intensity. For merging of modules the plug-in offers a correlation histogram, and allows the user to select an appropriate correlation threshold. The runtime complexity of the plug-in version remained \( \sim O(n^3) \), as defined earlier [Kovács et al., 2010]. To enhance the performance of the plug-in for calculating the higher hierarchical layers further, we introduced a user-selected optimization. This results in the disappearance of meta-edges with very small weights at the higher hierarchical levels. These low intensity meta-edges are derived from the minor overlaps of distant modules of the lower level. This optimization allowed a speedup in running time by a factor of 7 for larger networks [Supplementary Table S10].

The plug-in is capable to generate overview reports for each hierarchical level. These reports list the number of the nodes (meta-nodes), edges (meta-edges) and modules, the effective number of modules (see Kovács et al., 2010) and the size of each module. The overview also contains the list of the 10 nodes of each module having maximal module assignment value to the corresponding module at one level below in the hierarchy.

Node colours can also visualize several node (or meta-node) measures including weighted degree, betweenness centrality, community centrality, overlap and bridgeness (see Supplementary Fig. 2). The plug-in has an option to merge highly similar module pairs containing roughly the same nodes or meta-nodes with the same intensity. For merging of modules the plug-in offers a correlation histogram, and allows the user to select an appropriate correlation threshold. The runtime complexity of the plug-in version remained \( \sim O(n^3) \), as defined earlier [Kovács et al., 2010]. To enhance the performance of the plug-in for calculating the higher hierarchical layers further, we introduced a user-selected optimization. This results in the disappearance of meta-edges with very small weights at the higher hierarchical levels. These low intensity meta-edges are derived from the minor overlaps of distant modules of the lower level. This optimization allowed a speedup in running time by a factor of 7 for larger networks [Supplementary Table S10].

3 RESULTS AND CONCLUSION
ModuLand-derived communities of various yeast protein–protein interaction networks gave a functionally meaningful description of the yeast interactome [Kovács et al., 2010]. Function of module core proteins proved to be good indicator of the function of the whole
module Mihalik and Csermely 2011. Here, we demonstrate the use of the ModuLand Cytoscape plug-in on the protein structure network of Escherichia coli Met-tRNA synthase, since an elegant study Ghosh and Vishveshwara 2007 showed the existence of four alternative communication paths in this enzyme. The five major sub-domains of Met-tRNA synthase were well reflected by the five modules obtained at the second hierarchical level of the protein structure network (Fig. B Supplementary Table S3). Key amino acids of the most frequently used communication path (Ghosh and Vishveshwara, 2006) either belonged to the module cores or were positioned (like module cores and bridges). As shown by several studies, the five major functions of extensively overlapping modules, and determines key network positions (like module cores and bridges). As shown by several case studies, modules identified by the plug-in correspond to biologically meaningful groups, module cores help the identification of biological functions and inter-modular nodes have a key role in a variety of biological networks.

**ACKNOWLEDGEMENTS**

The authors thank the Editor and the anonymous Reviewers for their comments and suggestions, Amit Ghosh (Lawrence Berkeley National Laboratory, Berkeley, CA, USA) and Saraswathi Vishveshwara (Indian Institute of Science, Bangalore, India) for the 3D coordinates of E.coli Met-tRNA synthase, and members of the LINK Group (www.linkgroup.bat) for their discussions and help.

**Funding:** EU [FP6-518230, TÁMOP-4.2.2/B-101/2010.0013 and PD-75261]; International Human Frontiers Science Program Organization (BP); Momentum Program of the Hungarian Academy of Sciences (BP); and residence at the Rockefeller Foundation Bellagio Center (PC). Conflict of interest: none declared.

**REFERENCES**


Bellagio Center (PC).