User Manual : CompNet

1. Loading networks

Networks to be compared can be loaded on CompNet in three user-friendly ways (Figure1). (1) Users can load multiple interaction networks in the form of simple text files containing list of edges in a network. In case the input text files contain multiple fields, users can specify the field-delimiters manually and choose the appropriate columns for the source and target edges. The tool gives a preview of the selection and allows renaming of the corresponding networks for user convenience. (2) Users can also upload a background network in the form of an edge-list, and subsequently provide a list of nodes to be overlaid on the background network. Such overlaying of different lists of nodes will result in creation of different networks, which can then be compared/analyzed. (3) CompNet also supports uploading a list of paths instead of edge-lists to create multiple networks.

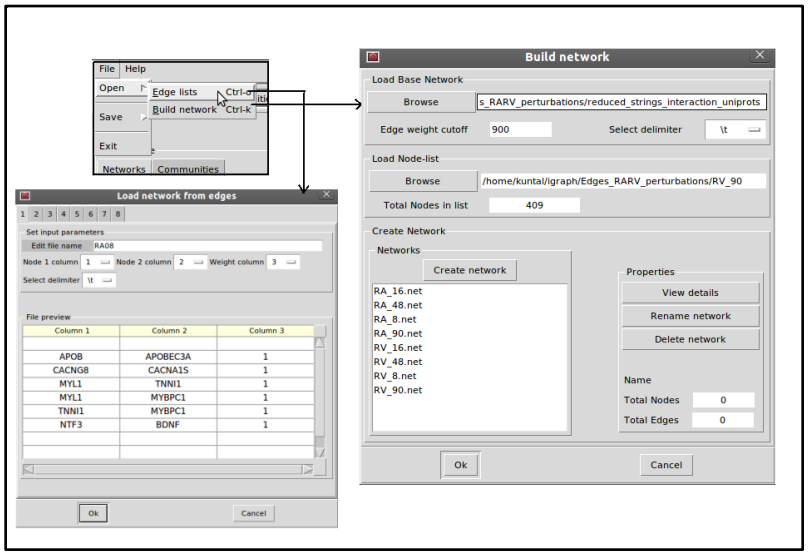

Figure 1 : Loading the networks in CompNet

2. Visualizing the network on the canvas

Once the networks are loaded, CompNet computes the union of all the networks and displays the same on the canvas (Figure 2A). For the convenience of users, a small window displaying a miniature canvas is loaded alongside to assist in easy navigation and to allow convenient selection as well as display of different parts of the union network. Users can view a desired portion of the network by mouse selection. The zoom button(s) and scroll bar(s) may be used to get an enlarged view of any such selected area of the canvas. The nodes in the union network are represented as 'pie-nodes' which depict how a node is shared across the different networks (Figure 2B). Hence, with a first glance at the network, one can easily identify the nodes which are unique to a network and the ones that are shared across multiple networks. CompNet additionally detects 'communities' in the union network (Figure 2A), colors them distinctly, and lists the communities in the 'Community' tab in a descending order of their sizes. A particular set of nodes of interest can also be selected using mouse action, to create a custom sub-network. Each of the detected communities, or any other chosen sub network(s) can be selectively visualized and analyzed using CompNet. Please refer to the GUI manual at the end of this document for a detailed guide.

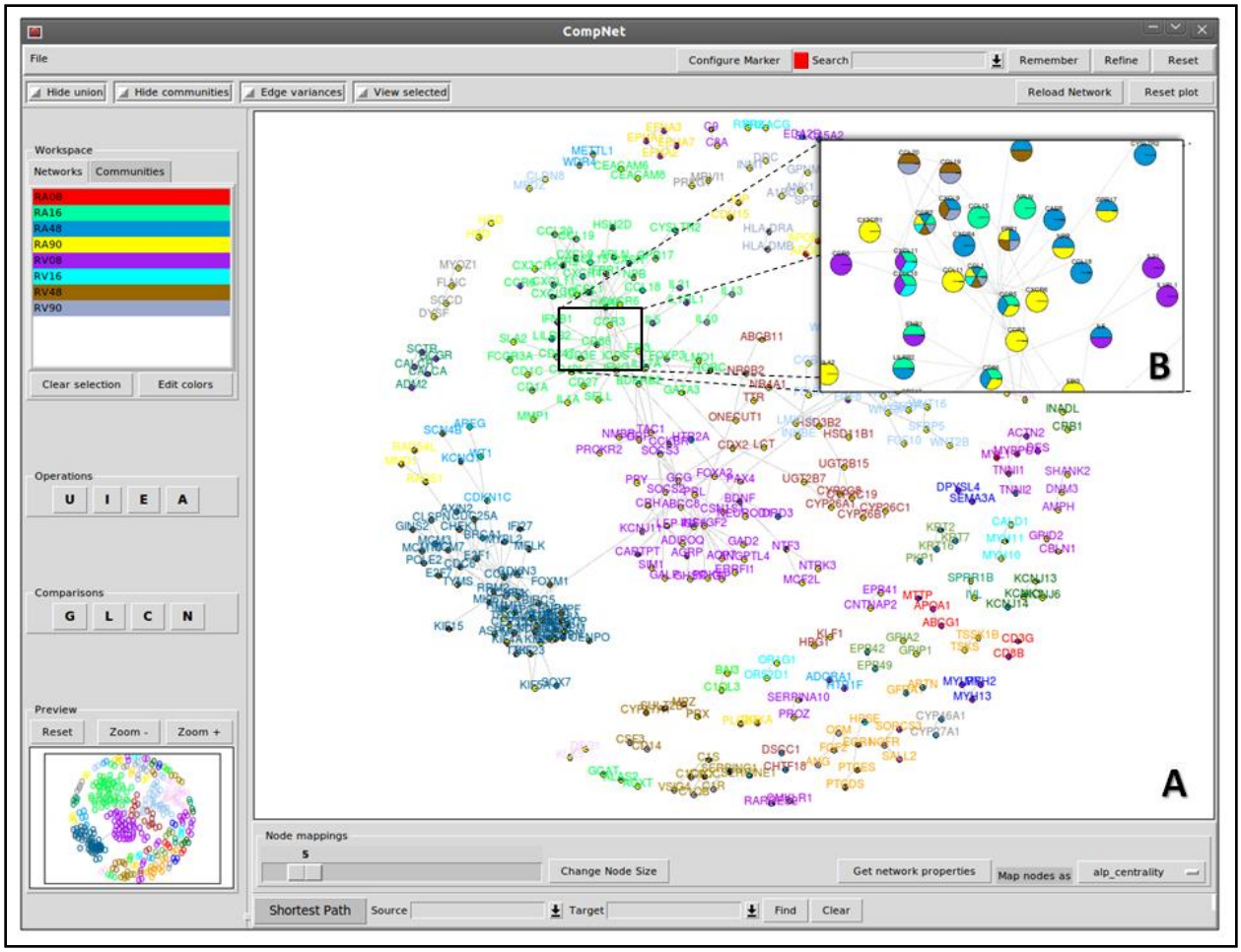

Figure 2: (A) CompNet canvas displaying the union of eight loaded networks. The names of nodes belonging to different communities (modules) in this union network, are marked with different colors. (B) The 'pie-nodes' representation enables a user to identify the presence/absence of individual nodes across the compared networks. A user can zoom into the network or increase the node sizes for utilizing this feature.

3. Interacting with the layout

The primary objective of the CompNet tool is to compare multiple interaction networks. Accordingly, CompNet facilitates interactive visualization of the union, intersection and complement regions for any selected set of networks. The union network can be turned off (invisible) while viewing an intersection or complement region, by using the feature 'Hide union network' (check-box). An efficient search feature is incorporated in the tool to search and select nodes in the network using regular expressions. Particular nodes can be searched and highlighted in the canvas using user configurable markers. When one or more nodes are selected by either mouse operation or using the search feature, all the edges connecting these nodes are highlighted simultaneously, thereby allowing users to visualize the connections between them. Once a particular set of nodes (forming a sub-network) is selected, the 'Analyze selected' functionality (button '**A**' under 'Operations') can be used to study the distribution of edges across different networks as a pie-chart-matrix (Figure 3). This particular feature is especially helpful while studying the distribution of edges in a community across the loaded networks.

4. Comparison

CompNet provides a series of functionalities to compare networks on the basis of similarity of edges, general network properties, community composition and node neighborhoods.

4.1. Comparing node and edge similarities: CompNet enables users to visualize the similarities/differences between the networks in terms of constituent nodes as well as edges. Nodes which are unique to a network can easily be identified by looking at its color. When present across multiple networks, a node, represented as a 'pie', will have differently colored pie-slices. The colors in these pies indicate the corresponding source networks. The similarities and dissimilarities between edges can be found out in the following way. The 'Union' ('**U**') and 'Intersection' ('**I**') buttons can be used to visualize the edges corresponding to the union and intersection of the selected networks, respectively. Similarly, the 'Exclusive' ('**E**') option helps to find out the edges exclusively present in the selected network(s) (complement region). The edges can also be colored according to the variance of their weights across the networks. Once a particular set of nodes have been selected (either by mouse action or by using the search feature), the check-boxes 'View selected' and 'Edge variances' can be used to apply the coloring. A red-blue color gradient is used, wherein an edge whose presence/weights varying significantly across the loaded networks is colored red, and an edge with low variance is colored blue. Clicking on the 'Analyze selected' additionally displays the selected edges in a triangular 'edge-pie' matrix plot (Figure 3), wherein the divisions of the pie represent how a particular edge is present across the loaded networks.

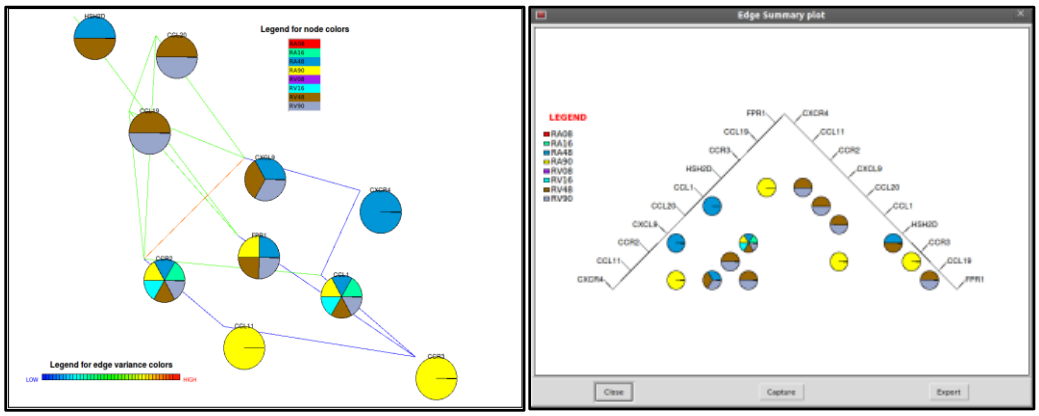

Figure 3: Visualizing pie-nodes, edge variance and edge presence (edge-pie chart)

4.2 Comparing network properties: A selection of networks can be compared by looking at the distribution of global graph properties like total nodes, total edges, density, clustering coefficient and diameter, using the feature 'Global properties' (Figure 4). A more detailed comparison can be made on the basis of node specific properties like degree, centrality, betweenness, closeness, eccentricity and coreness. This operation can be performed either on the currently displayed network on the canvas, by using the feature 'Get network properties', or individually on any of the loaded networks. Options of filtering for selected nodes (based on user-defined cut-offs) are also provided in the GUI. CompNet additionally allows mapping specific graph properties on the canvas by appropriately altering the node sizes (proportionally to the selected metric), through using the 'map property' feature. The property to be mapped can be selected from a drop-down menu.

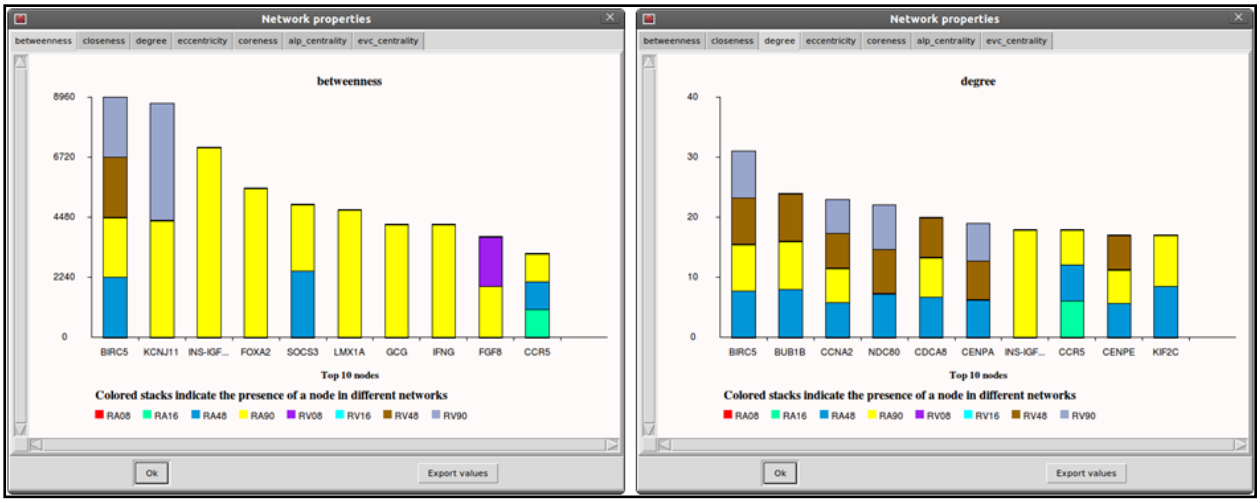

Figure 4: Cumulative bar plot for graph properties on the displayed network.The coloured stacks in the bar-plot represent the presence of a gene/node in the individual networks qualitatively (colors shown in Legend). The heights of the bars indicate the value of the selected graph property (e.g., betweenness, degree, etc.).

4.3 Comparing community composition: CompNet can be used to compare networks based on the detected communities. The communities are sorted in descending order of their sizes (number of constituent nodes) under the 'Community' tab. Users can use the feature 'Compare communities' (button **'C'**) for any selected set of communities and visualize the contribution of individual networks towards formation of the selected community(s) (Figure 5). The selected community assignments can also be exported as text file. An additional option allows finding and visualizing the communities which are enriched in a particular network by using the feature 'Enrichment' (which appears after selecting 'Compare communities'). Enrichment of a community in a network is calculated as a ratio of its local enrichment in the network (ratio of total nodes in the community in the selected network to total nodes in the network) to its global enrichment (ratio of total nodes in the community in the union network to total nodes in the union network).

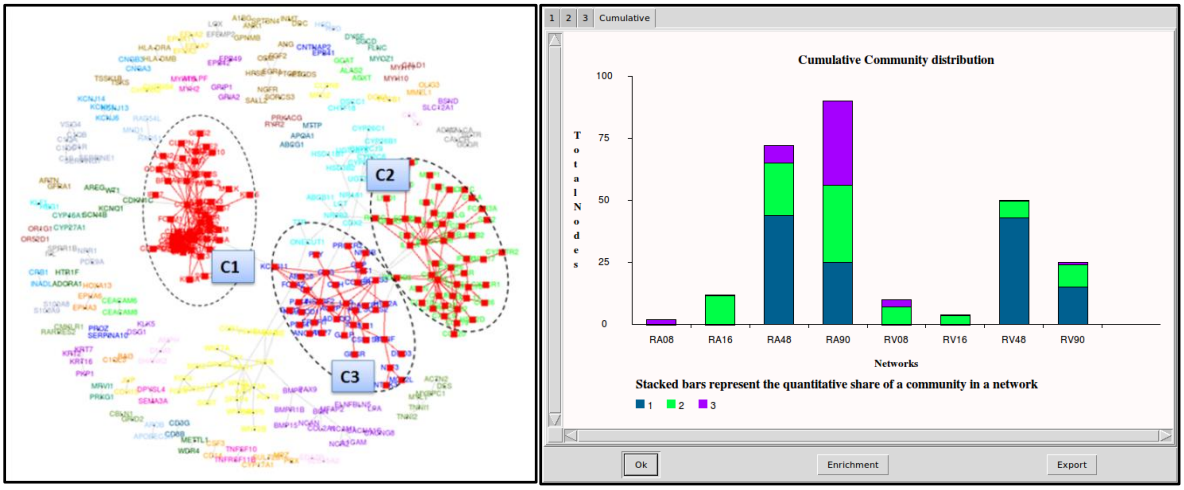

Figure 5: Community Analysis using CompNet. Stacks in the bar-plot represent the quantitative share of the communities (as shown in the Legend for each figure) in a network (displayed as a bar).

4.4 Node neighborhood: CompNet incorporates a method, based on neighborhood similarities of the constituent nodes, for comparison of multiple interaction networks. Two nodes (from two compared networks)

are deemed to be more similar if the lists of their immediate neighbors overlap. An overall similarity score (called CNSI or CompNet neighbor similarity index), cumulated for all constituent nodes, is finally used to designate the similarity between two compared networks. CNSI is calculated by cumulating the Jaccard similarity (in terms of first neighbors) between corresponding nodes from two compared networks, as -

$$CNSI = \sum_{i=1}^{N} \frac{f_{n_i}^{A} \cap f_{n_i}^{B}}{f_{n_i}^{A} \cup f_{n_i}^{B}}$$

Where n_i is the 'i'th node in the union of compared networks A and B (consisting of a total of N nodes), and f_{ni}^{A} and f_{ni}^{B} are the first neighbors of n_i in the networks A and B respectively.

The results of this comparison is presented by CompNet in form of a bubble chart as well as in form of a dendrogram (Figure 6), depicting the different networks grouped according to their similarities. The node by node similarity (as first neighbors) and individual score can be exported as text file for further analysis.

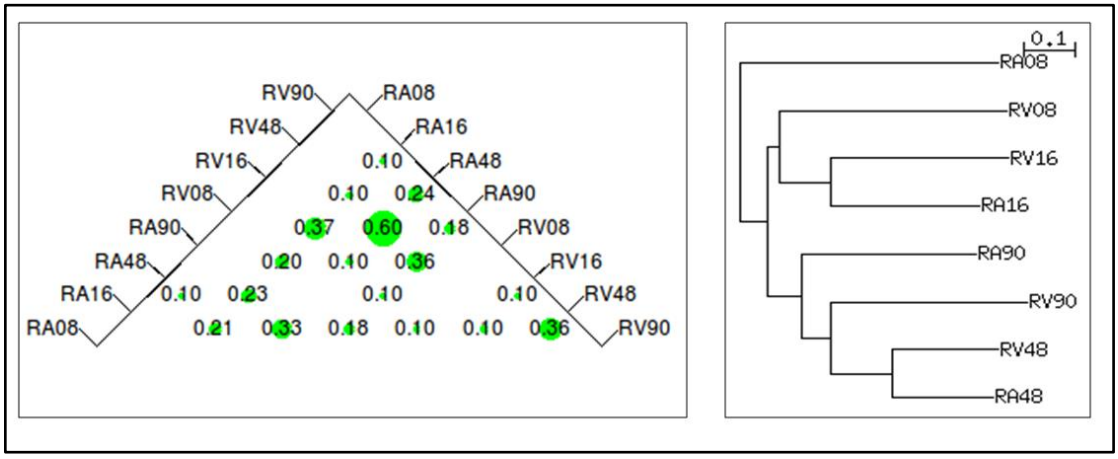

Figure 6: CompNet's neighbour similarity index based similarity matrix and the corresponding tree generated for the loaded networks

4.5 Visualizing and comparing shortest paths: Another important aspect while analyzing biological networks is to look for shortest paths between a set of nodes of interest. For a similar set of networks, like those representing time series data or protein interactions from healthy vs. diseased tissues/cells, the changes in shortest paths might provide valuable insights in understanding a biological mechanism. CompNet allows users to identify such shortest paths from multiple networks, and visually trace/compare these paths. Users need to enter the names of the source and target nodes and select the required networks (for which the shortest paths need to be traced). Once this is done, a checkerboard plots (Figure 7), showing all possible shortest paths between the desired nodes across the selected networks, are displayed. Shortest paths between multiple sets of sources and targets can also be computed with CompNet by providing two separate files containing lists of sources and targets.

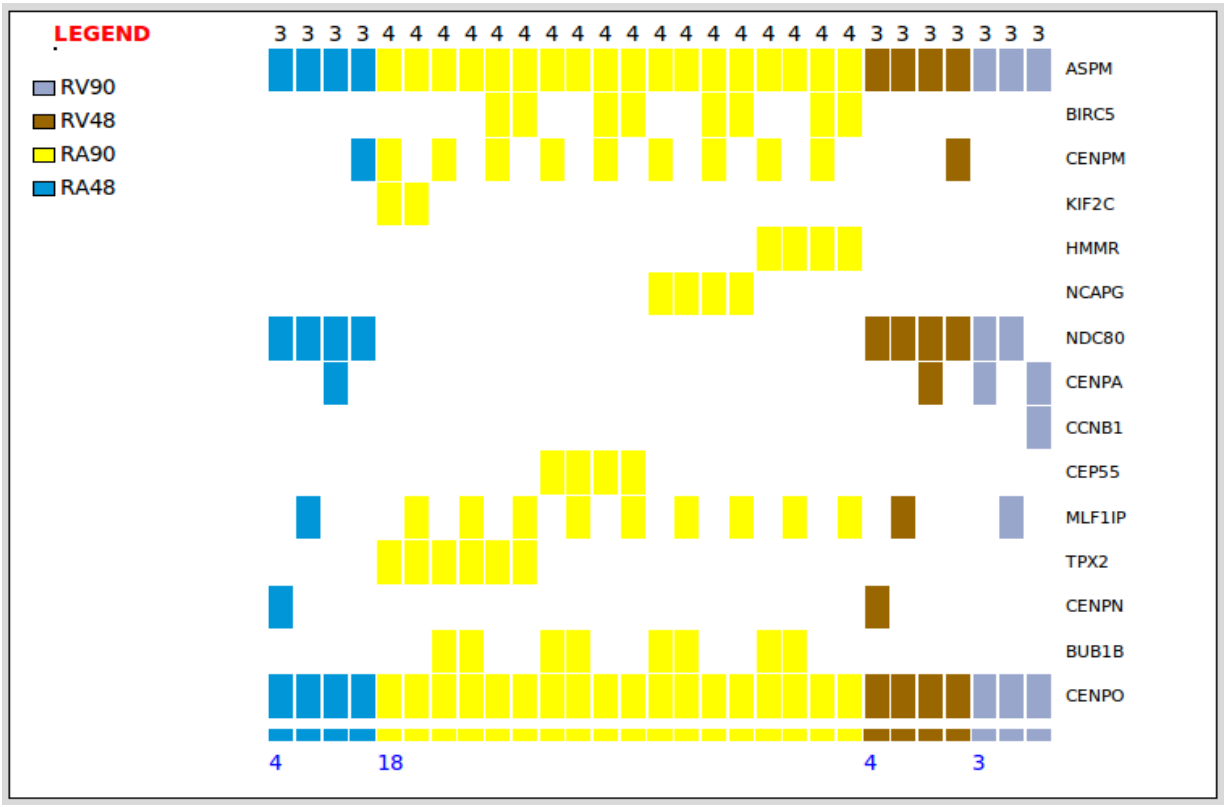

Figure 7: Shortest path analysis using CompNet. The numbers at the top represent the path lengths while those at the bottom represent the total number of shortest paths found in each network (colours corresponding to the legend).

Detailed descriptions of the CompNet user interface

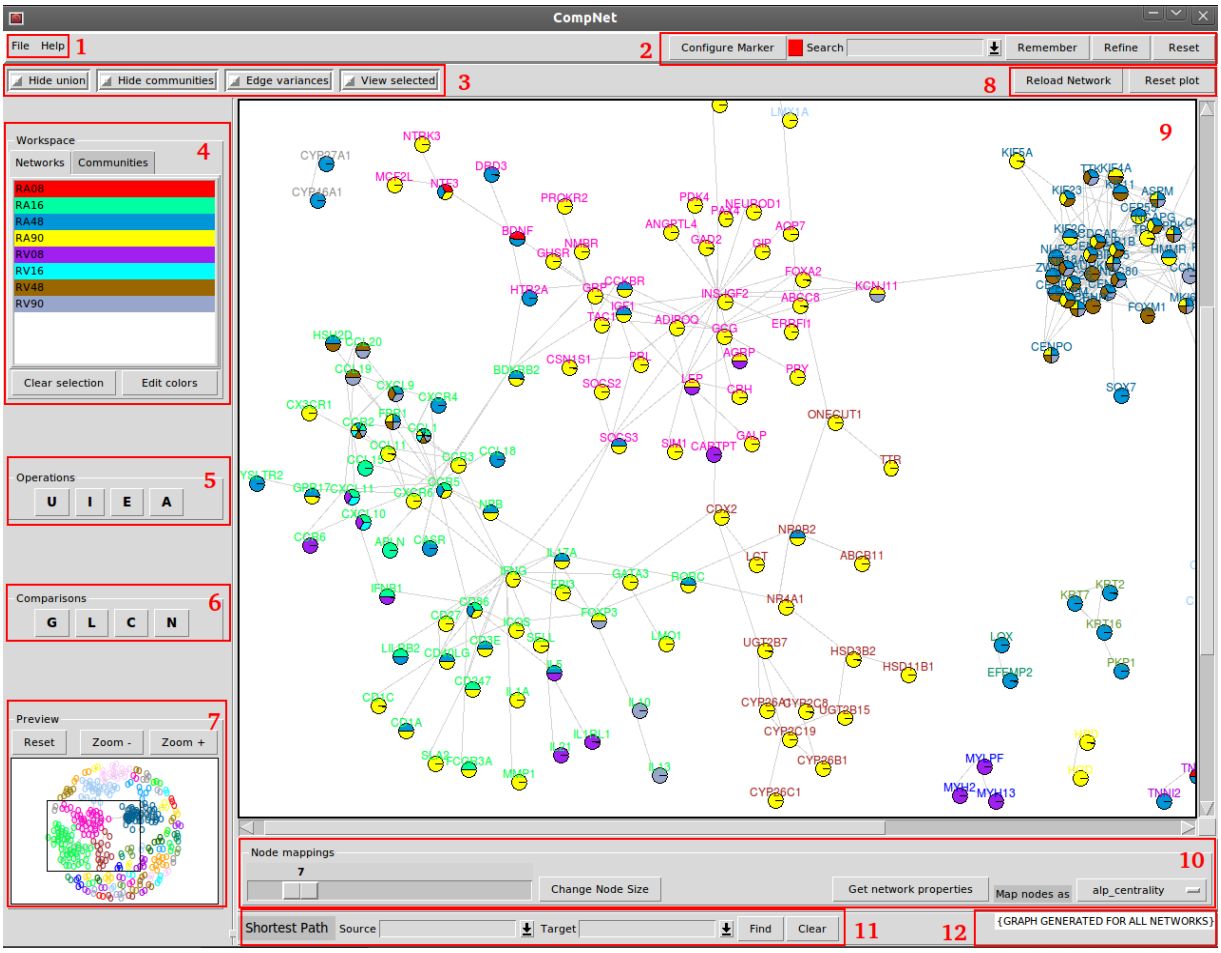

Figure 8: The graphical user interface (GUI) of CompNet

The GUI components required to access different features of CompNet have been highlighted (and numbered) in Figure 8. A detailed description of each of these components is provided below (following the number keys).

1. The File menu is used to load interaction networks as well as save outputs like selected nodes and selected edges as text files.

2. This component can be used to search and highlight one/more node using regular expressions. Configurable markers can be set using the 'Configure Marker' button to highlight a search. A search can further be refined using the 'Refine' button or remembered using 'Remember' button. The 'Reset' button can be used to reset the highlight markers.

3. This includes various options to modify the network displayed in the canvas. The function of the individual check buttons are described below:

'Hide union': After selecting a particular set of edges in the network using key operation 5, this checkbox if selected will turn of the remaining unselected background network.

'Hide communities': This check-box is used to show/hide the displayed communities.

'Edge variances': After selecting a set of desired nodes by mouse selection (left mouse-click and drag to select area) or from the Communities tab (Double left clicking on a community in key operation 4), this checkbox if selected will color edges as per the variance of their weights across the loaded networks. The color is red for higher values and blue for lower (analogous to the heated and cooled color of a flame).

'View selected': This checkbox can be used to view a set of selected nodes. This is often useful after selecting 'Edge variances' to view and analyze the edge distribution.

4. Once a set of networks are loaded, they appear in the 'Network' tab. The detected communities in the union network are listed (in descending order of the constituent node abundance) in the 'Communities' tab.

5. This includes various operations for selecting and analyzing Union (**U**), Intersecting (**I**) and Exclusive (**E**) edges for a selected set of networks in the 'Networks' tab. The Analyze (**A**) option can be used to analyze a selected set of edges when selected by mouse selection (left mouse-click and drag to select area) or from the Communities tab (Double left clicking on a community in key operation 4).

6. This includes various options to compare global properties (**G**) and local properties (**L**) of a selected set of networks from the 'Netwoks' tab as well as to compare a selected set of communities (**C**) from the 'Communities' tab. The (**N**) feature can be used to cluster a selected set of networks based on some indices.

7. This component is a preview canvas for the actual canvas (key operation 9) which can be used to select and view a particular area of the actual canvas, reset the view and perform zooming operations on the actual canvas.

8. The 'Reload network' can be used to destroy the current view and generate the union network afresh. The 'Reset Plot' option can be used to quickly reset colours and markers in the plot.

9. The display canvas showing the nodes and edges. The nodes in the union network are represented as 'pie-nodes' which depicts how a node is shared across the different networks.

10. This component is used to modify node properties like changing the node sizes (using the slider) or to map a particular network property to the node size (using the dropdown). The 'get network properties' button needs to be first clicked to calculate the properties of the currently displayed network which also displays the result as a stacked bar plot. Once the mapping is done for a network (the one currently displayed), a property can be mapped to the nodes in the displayed network.

11. A 'Source' and 'Target' node can be selected from the network (or manually entered) and the shortest paths between them can be compared across a selected set of networks. Selecting a single node in the network by mouse selection sets it as the 'source' and a subsequent similar single node selection automatically selects it as the 'target' (and appears in the corresponding text boxes).

12. This component shows all relevant notifications of the tool.