# Introduction

Signaling pathways are sets of directed interactions between biological entities, initiated by input signals and resulting in realization of target processes. Assessment of pathway activities is crucial for determination of pathways which play role in development of various phenotypes (diseases, developmental stages, etc.)

The Pathway Signal Flow Calculator (PSFC) is a Cytoscape app for calculation of pathway signal flow based on gene expression data and pathway topology. The app provides a flexible interface for setting flow computation rules and visualization of signal propagation on the network.

Main use case
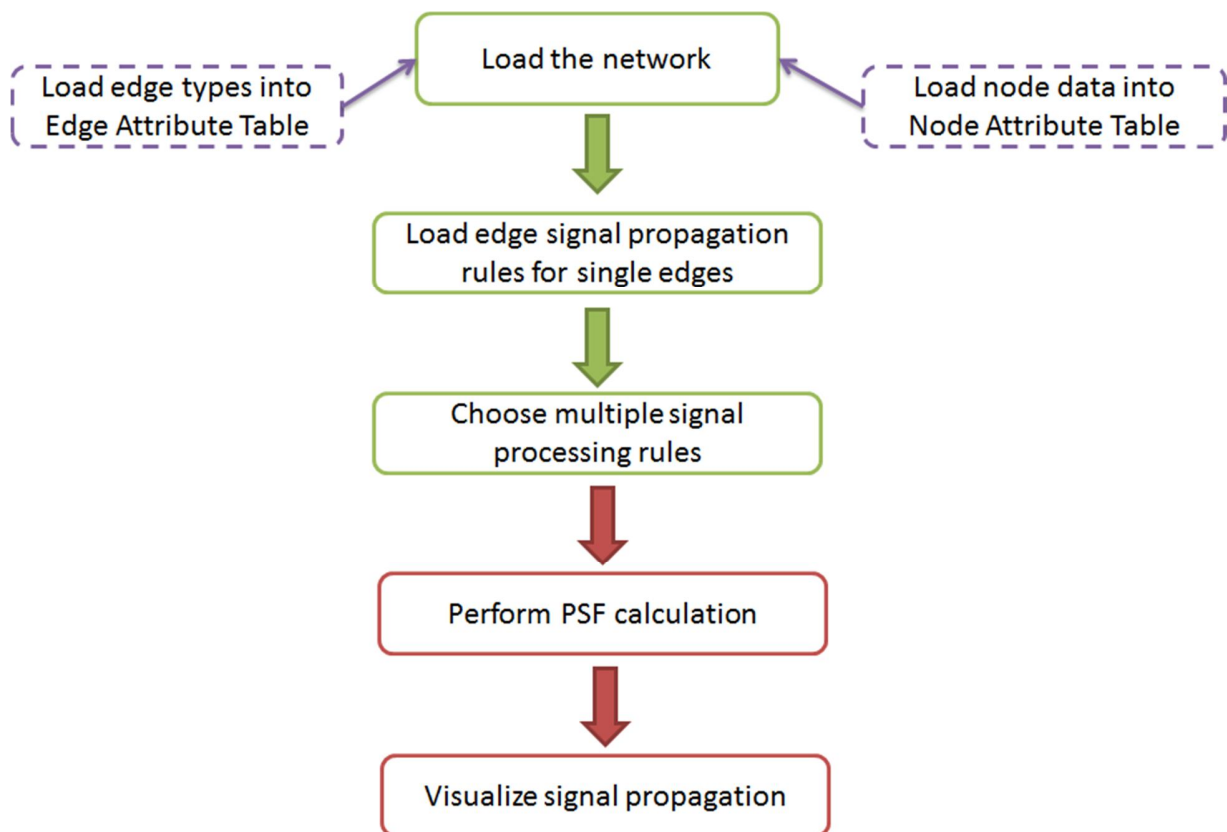
Application example
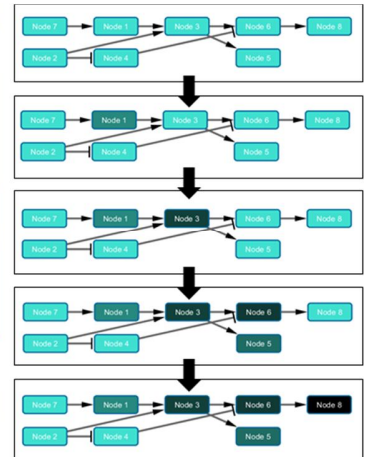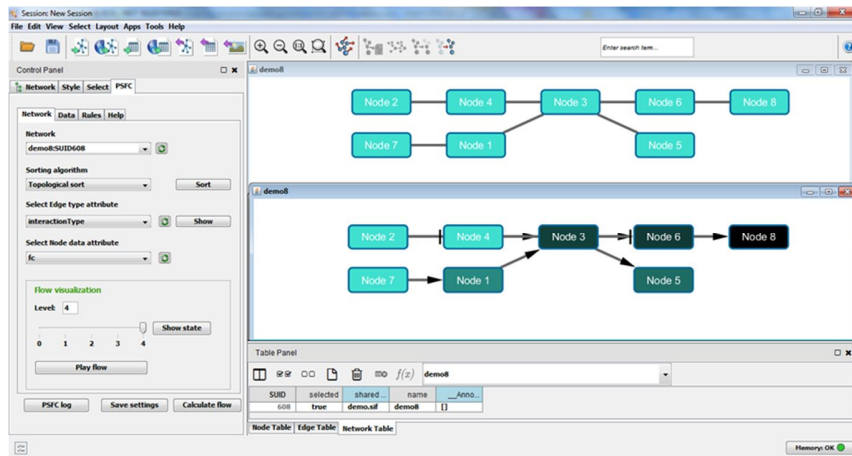
Getting started

Installing the app from web

In Cytoscape, go to Apps -> "App Manager" , choose PSFC from the list and click on the install button.

*Installing the plugin from file*

If you have already downloaded the PSFC jar file , go to Apps -> "App Manager", click on "Install from File" button at the bottom of the dialogue and point to the directory of the .jar file. In case of successful installation, the PSFC panel should appear among east panel group in Cytoscape window.

[Move example here? and outline it as a step-by-step tutorial]
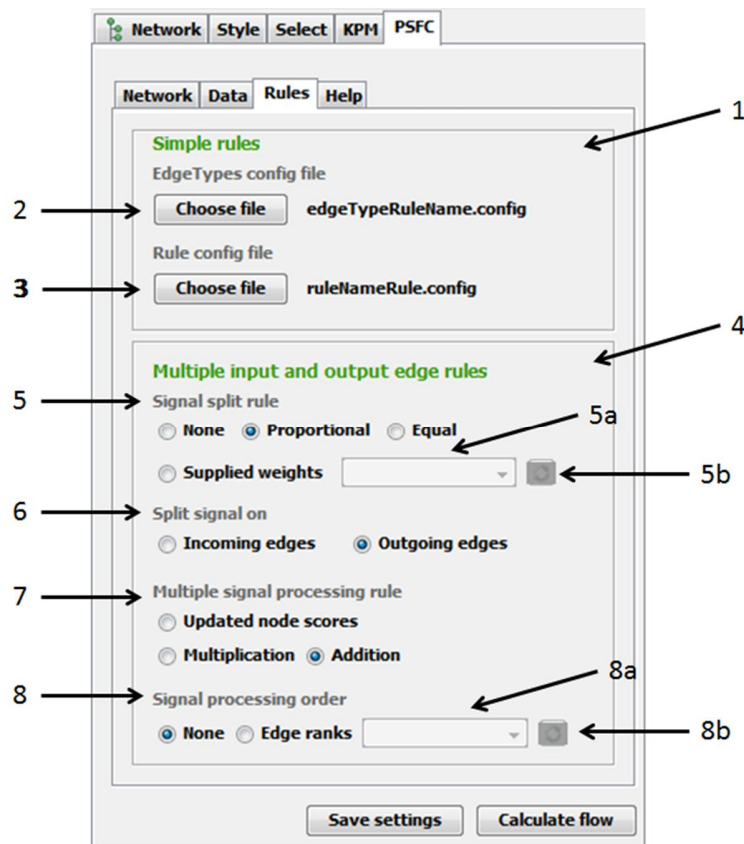
GUI components

Network tab
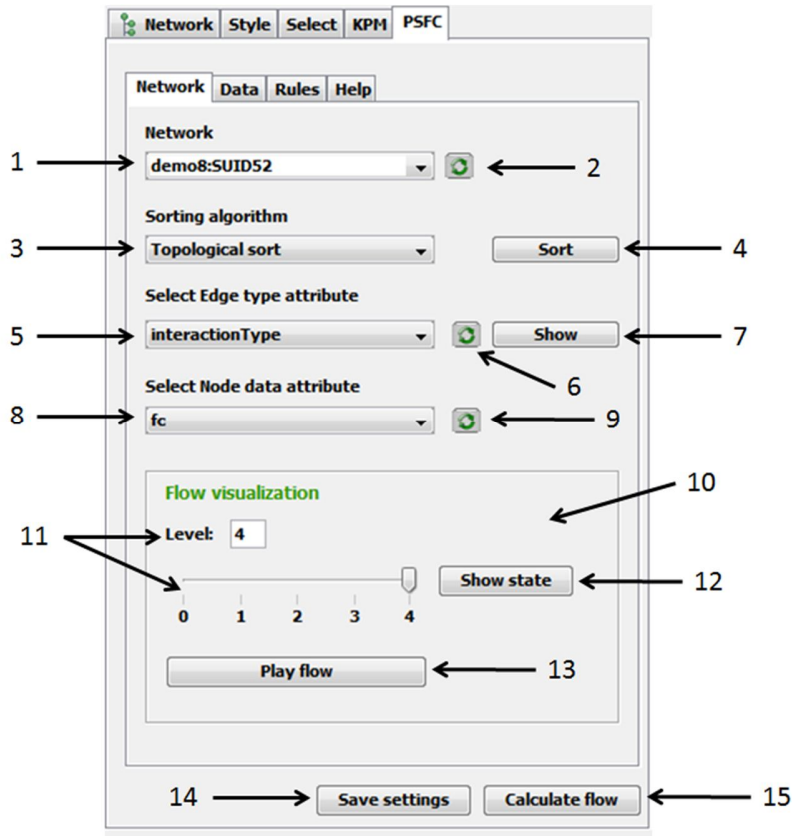
Contains the following components:

1. *Network* combo box - for selection of the network under interest. The networks are represented in the form *networkName*:SUID*networkSUID*.
2. Refresh button for *Network* combo box - updates the list of currently loaded networks. Should be used after a new network has been loaded or removed.

3. *Sorting algorithm* combo box - for selection of the sorting algorithm to be applied on the network. Currently only topological sorting algorithm is supported (see *Network Sorting* section).
4. *Sort* button - for applying selected sorting algorithm onto selected network and updating the network layout respectively.

5. *Select Edge type attribute* combo box - for selecting the Edge type attribute from selected network's Default Edge Table. The Edge type attribute is used for setting edge type specific rules (see *Flow rule configurations* section).
6. Refresh button for *Select Edge type attribute* combo box - updates the list of columns of the selected network's Default Edge Table. Should be used when a column is added or removed from the table.

7. *Show* button for *Select Edge type attribute* combo box - shows the list of unique values in the selected Edge type column. The list is just for checking the edge type attribute selection, as well as for copying the unique values for edge type-specific rules file generation (see *Flow rule configurations* section).

8. *Select Node data attribute* combo box - for selecting the Node data attribute from selected network's Default Node Table. The Node data attribute is used for setting the initial values of nodes for flow calculation. Node data values can be either row of logarithmic expression values, fold change or log-fold change of expression values, etc.

9. Refresh button for *Select Node data attribute* combo box - updates the list of columns of the selected network's Default Node Table. Should be used when a column is added or removed from the table.

10. *Flow visualization* panel - for performing visual mapping of node signals after SPF calculation.

11. *Level* textbox and slider - for setting the PSF state at the specified level of nodes in the sorted network.

12. *Show state* button - performs color gradient visual mapping based on the node signals at the chosen level.

13. *Play flow* button - performs color gradient visual mapping based on the node signals for all the levels in order.

14. *Save settings* button - Saves the options selected from *Network, Data* and *Rules* tab.

15. *Calculate flow* button - Performs PSF calculation on selected network.

Rules tab

The *Rules* tab contains the following components:
1. *Simple rules* panel - for setting edge type specific rules onto single edges.
2. *Choose EdgeTypes config file* button - sets the configuration file containing edge type : rule name mapping.
3. *Choose Rule config file* button - sets the configuration file containing rule name : rule specification mapping.

4. *Multiple input and output edge rules* panel - for setting rules to process multiple signals incoming to- or outgoing from nodes.
5. *Signal split rule* - specifies whether to split multiple signals coming to a node or outgoing from a node by dividing to the number of signals (*Equal*), proportional to source or target node signals (*Proportional*), based on supplied weights (*Supplied Weights*) or not apply any splitting (*None*).
   - 5a. The CyColumn in Default Edge table of the selected network, which contains edge weight attribute values. Should be of Floating Point type. Is active when *Supplied weights* option is selected.
   - 5b. Refresh button for supplied weights column -updates the list of columns of the selected network's Default Edge Table. Should be used when a column is added or removed from the table.

6. *Split signal on* - specifies whether signals incoming to single node should be processed as a single set, or those outgoing from a node.
7. *Multiple signal processing rule* - specifies how multiple signals received by a single node should be processed - be added (*Addition*) or multiplied (*Multiplication*) and replace the previous node signal, or the signals should be processed one after another, each of them replacing the previous node signal (*Updated node scores*).
8. *Signal processing order* - specifies whether there is a specific order multiple signals should be processed in. This option is meaningful if the *Multiple signal processing rule* is set to *Updated node scores.*
    8a. The CyColumn in Default Edge table of the selected network, which contains edge rank attribute values. Should be of Floating Point type. Is active when *Edge ranks* option is selected.
    8b. Refresh button for edge ranks column - updates the list of columns of the selected network's Default Edge Table. Should be used when a column is added or removed from the table.

Network sorting

For computation of pathway signal flow propagation, the network should initially be sorted based on distances of intermediate nodes from the input nodes. Where input nodes, i.e. nodes that don't have incoming edges, should be assigned level 0. In order to see network sorting results before pathway flow calculation, the user may press the *Sort* button (4) in *Network tab* panel. PSFC will sort the selected network and apply the level based layout onto network view. *psfc.level* column will appear in the default node table of the selected network, where the level of each node will be indicated.

Currently, only topological sorting algorithm is provided as an option. For directed acyclic graphs (DAGs) topological sorting algorithm results in such an ordering of the graph, where for every edge *uv* for nodes *u* and *v, u* comes before *v* in the ordering. Biological pathways, however, mostly contain cycles that represent positive or negative feedback loops. In such cases the *Topological sort* option will firstly remove backward edges by applying a greedy algorithm described in [Ispolatov 2007], perform topological sort on the remaining DAG and put back the edges. Note, that since the greedy algorithm may find different backward edges if applied many times, the sorting algorithm may produce different results for the same network topology.

Flow rule configurations

Rules for signal flow propagation for single edges are defined with two files.
● *EdgeTypes config* file (set with no. 2 button in *Rules* tab) should be a tab delimited file, where the first column indicates the edge types, and the second column indicates rule names corresponding to each edge type. All the edge types in the selected network should be present in the config file in order for PSF calculation to be performed.
This is an example of EdgeTypes config file:
```
activation act
```

```
inhibition inh
phosphorilation ph
binding    +
```

- *Rule config* file (set with no. 3 button in *Rules* tab) should be a tab delimited file, where the first column indicates the rule names, and the second column indicates the rules corresponding to each rule name. All rule names corresponding to all the edge types in the selected network should be present in this file. Rules are defined as mathematical equations with variables named "source" and "target" for source and target nodes respectively. An example of *Rule config* file is as follows:

```
act   source*0.5*target
inh   1/source * target
ph    source*exp(1/target)
+     source + target
-     source - target
const source*2
```

The parsable functions include a number of mathematical operations, including *sin*, *cos, log* (natural logarithm), and constants, such as *Math.PI, Math.E.*

With rules, defined as such, the algorithm will update the signal at each edge, by taking its source and target node values and applying respective mathematical operations.

Pathway flow calculation

Pathway flow calculation is the process of signal propagation from nodes at the first level to the nodes at the last level of the network. At each level assigns weights to the edges, according to the options set with *Signal split rule* and *Split signal on* (button group no. 5 and 6 in the *Rules* tab). For each edge, the signal at the edge is computed based on the edge type-specific rule and the source and target node values, and multiplied by its weight. If a node has several incoming edges, the edge signals are processed and accumulated in the target node, according to *Multiple signal processing rule* and *Signal processing order* (button group no. 7 and 8 in the *Rules* tab).

To perform pathway flow calculation on the selected network, after setting the respective options simply press the *Calculate flow* button. Nodes and edge signals will be updated and the signals of all the nodes and edges will be visible in newly created or updated columns in default node and edge tables of the selected network in Cytoscape. The column names will have the suffix "psfc.signal_" followed by the level value. Signal values are also kept in the score backup xls file, kept in the apps directory, usually located at C:\Users\User\CytoscapeConfiguration\app-data\PSFC\ (the address is indicated in TaskMonitor). The log of PSF calculation steps are debugged in PSFC log file, which is located in the apps directory, usually located at C:\Users\User\CytoscapeConfiguration\app-data\PSFC\PSFC.log, and can be opened by pressing the *PSFC log* button in the bottom of the panel.

Flow visualization

The *Flow visualization* panel (no. 10 in the *Network* tab) allows the user to see the updated node signals by applying color based visual mapping of the values in the "psfc.signal_" columns. The *Show state* button will show the node signals at the state when the flow propagation has reached the nodes of the level indicated on the levels' slider or level textbox. The *Play flow* button will sequentially visualize node signals at all the levels.

Development

License
MIT
Apache2
GNU GPL