

User manual for the network failure node detection plugin for Cytoscape

Michael Schubert

21st February 2011

1 Introduction

When working with networks, it is often important to know how resilient the given network is against an intended or random blocking or removal of nodes. The software usually employed when analysing network data is Cytoscape. However, its base system often lacks the tools to perform a specific usage task.

This is why this plugin to identify and analyse failure nodes in a network was written. The failure nodes therein might be an essential power plant in the electric grid or in the biological sense they represent potential drug targets because they are able to disrupt pathways. For more information on the latter topic, see the use case scenario.

2 Installation

In order to install the plugin, first locate or download the plugin file named „NetworkFailurePlugin.jar“. Know where you save it so you can copy it in the Cytoscape plugin directory later on. For this guide, it is assumed that Cytoscape is already installed and working.

2.1 Choose the directory to install the plugin in

2.1.1 Windows

The standard installation path of Cytoscape is „C:\Program Files\Cytoscape“. If you have chosen another directory during installation it is assumed that you find it again yourself. In this directory, there is a subdirectory called „plugins“. Copy the plugin file „NetworkFailurePlugin.jar“ in the „plugins“ directory.

2.1.2 Linux

There are two ways of installing, for your user only and for all users of the machine you are on. For the latter method you will need administrative privileges which you are unlikely to have if the machine is not yours. In that case, choose the method to install the plugin for your user only. The files shared

between users are in the directory „/usr/share/java/cytoscape“, the plugins again in the „plugins“ subdirectory. The Cytoscape settings directory for your user only is located at „~/cytoscape/2.8/“ (where the tilde sign represents your home directory) or whatever version you are using. Again, there is a subdirectory called „plugins“. Depending on your user privileges and if you want to install the plugin for all users or only yourself, choose option one or two, respectively and copy the plugin jar file in the corresponding plugins directory.

3 Calling the plugin from within Cytoscape

When installed successfully, the plugin appears in the „Plugins“ menu of Cytoscape under the name „Identify failure nodes“. If you click on it, the following window (cf. figure 1) appears.

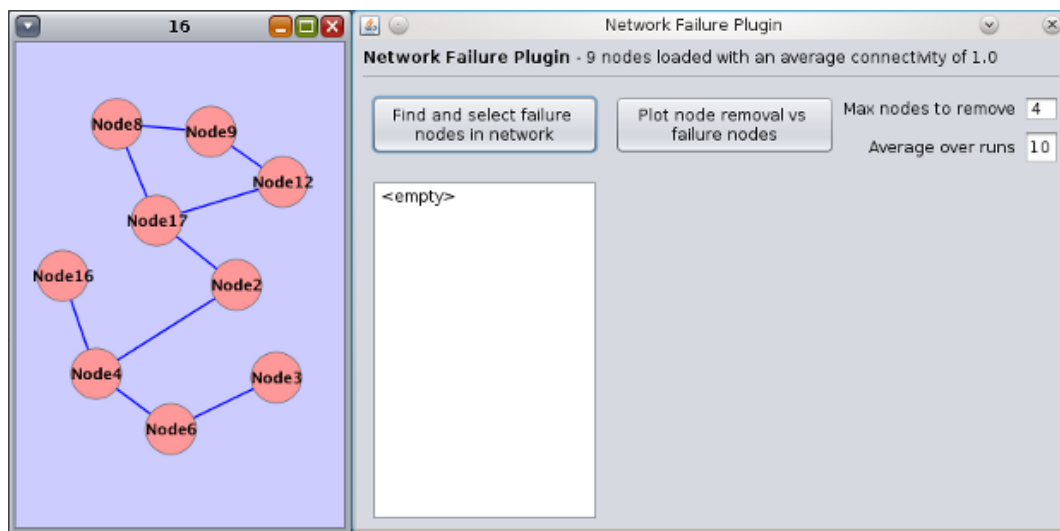


Figure 1: Screenshot of the plugin user interface when selected from the plugins menu. A sample network is also shown.

There are two main functions that can be used: „Identify failure nodes in network“ and „Plot node removal vs failure nodes“.

3.1 Identify failure nodes in network

When you click this button, the currently loaded network is searched for failure nodes that, when removed, split the network in two. The nodes are then selected in the Cytoscape network view and their names are then listed beneath the button instead of the „<empty>“ mark. You can select an item in this list, but also more than one by either holding down the Shift or the Ctrl Key.

In these cases, all nodes between your two clicks and the two nodes you clicked on are selected, respectively. In order to select all nodes in the list, you can click at any item and press the keys Ctrl+A. You can copy the selected nodes to your clipboard using the Ctrl+C key combination.

The result of this function is shown in figure 2.

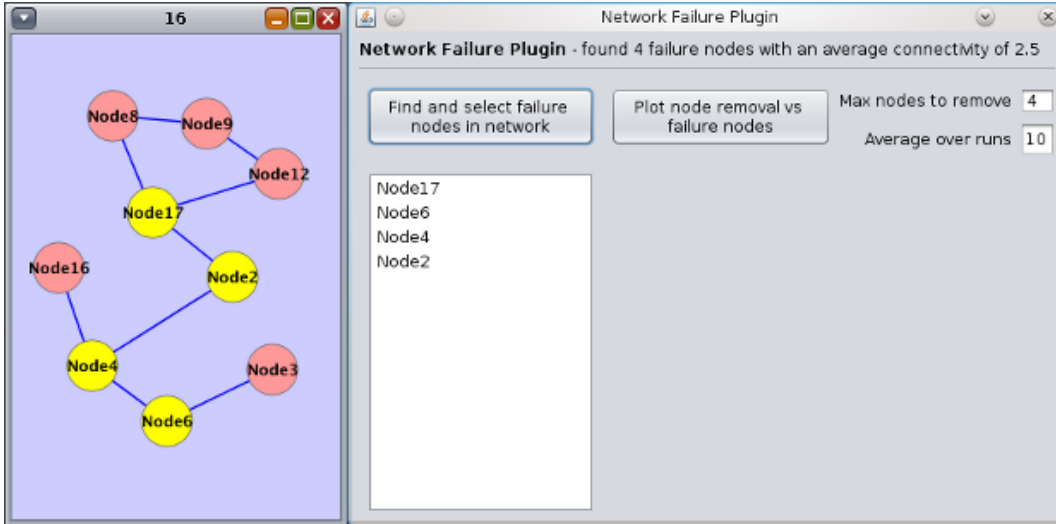


Figure 2: Result of identification of failure nodes in the sample network. One can see that the list is now filled with node identifiers that were determined to be failure nodes and the nodes being selected in the network.

3.2 Plot node removal vs failure nodes

This function is used to determine how robust the current network is against random failure of nodes. The user can choose a number of nodes to randomly delete from the network and the number of failure nodes is calculated for each number of removed nodes up to the one the user specified.

There are two options for this operation. With the first, you can choose the maximum number of nodes to remove in total. This means that at first there are no nodes removed, then one, then two, and so on until it reaches your desired limit. The nodes to be removed are selected randomly. You may not choose over 10 nodes or over half of the nodes of your network. With the second option, you can choose how many times each number of removed nodes is repeated and averaged. Depending on your time and the size of the network, choosing a higher number of repeats is likely to offer a more reproducible result.

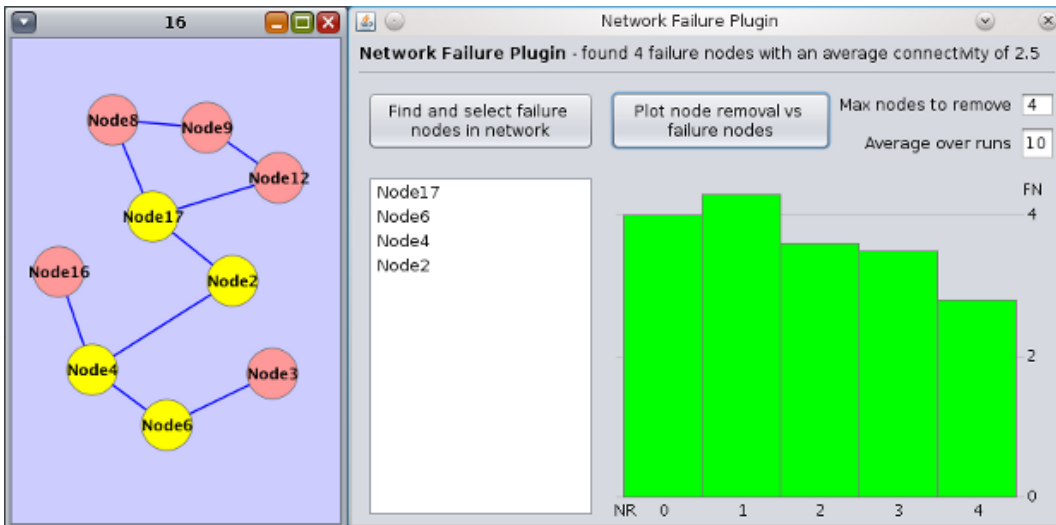


Figure 3: Result of the node removal vs number of failure nodes plot. One can now additionally see the number of removed nodes on the X-axis and the run average on the Y-axis on the right-side plot.

The result of this function is shown in figure 3.

4 Usage example on a protein interaction network

Here, the plugin was used to analyse the interaction network of the tumour suppressor protein p53. Therefore, the following steps were taken:

1. In the menu, choose File>Import>Network from Web services (which uses the PSICQUIC plugin that is installed by default)
2. As the data source, choose the „PSICQUIC Universal Web Service Client“
3. In the query field, write „p53“ without the quotes
4. Select „GET_BY_QUERY“ in the Search Property tab (cf. left part of figure 4)
5. Click the search button
6. Leave all databases selected, click on „Merge“
7. When asked which type of merge to perform, choose „Union“ (this is the default)

When following these steps, a network of direct as well as low-degree p53 interaction partners are loaded into your Cytoscape network, which then comprises about 2000 nodes. In the next step, you will want to remove all but the major interaction network. To do this, click on Layout>y-Files>Organic, select all stray subnetworks and remove them with the Delete key. You should be left with a network as seen in figure 4 (right side) with about 1600 nodes remaining.

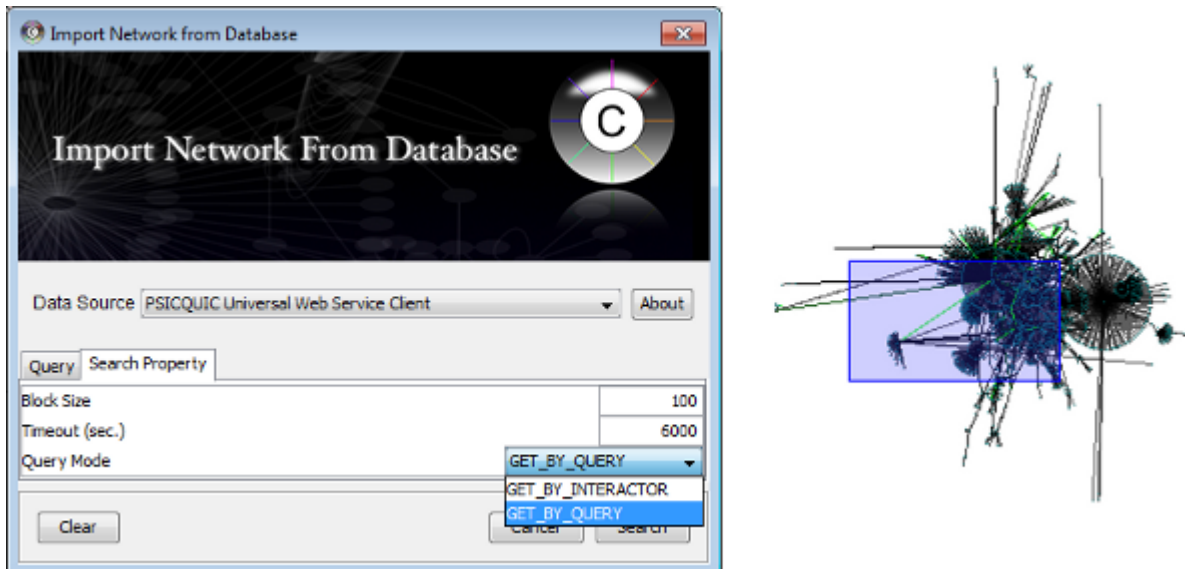


Figure 4: Suggested menu selection for p53 interaction partner search (left) and organic view of the resulting main network (right)

The next step will be to start the analysis using the plugin. A description on how to do that is provided in the previous section. When you receive an error message that the network is already split you have

missed to remove a stray network. Otherwise, the plugin reports to have found 1606 (this number may slightly vary because of database updates) nodes with an average connectivity of 2.22 edges per node.

When trying to the failure nodes in this network, the plugin reports 87 with an average connectivity of 44.63 (cf. figure 5). One can immediately see that the failure nodes in this type of network are also the highly connected hubs. This, however, is only biologically significant when there was little bias in the dataset (which can not be assumed since database query is likely to favour a couple of closely related proteins).

When plotting node removal vs failure nodes, one can see that in a network this size, the removal of up to 10 non-failure nodes has next to no effect on their counterpart's number. In this step, only one run was performed each since the size of the network impedes calculation speed. It still took about 30 minutes to compute, which leads to the conclusion that random node removal in large networks is not feasible.

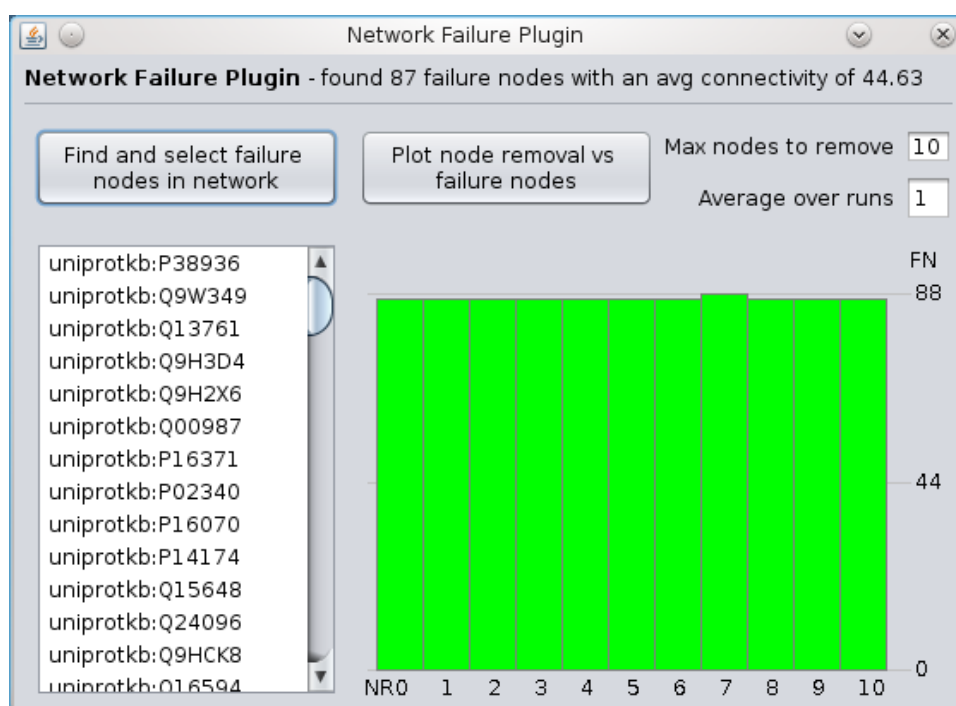


Figure 5: Analysis result of the p53 interaction network

The full list of UniProt IDs found is the following: P38936, Q9W349, Q13761, Q9H3D4, Q9H2X6, Q00987, P16371, P02340, P16070, P14174, Q15648, Q24096, Q9HCK8, Q16594, Q9Y4K3, Q8N726, P21675, O15151, Q9PST7, Q64364, Q92793, Q24040, Q8CHK4, Q6PCD5, Q64451, P51532, Q96RU2, P04637, Q96S44, P22814, Q9Y6I7, P04637-1, Q08509, Q9Y5J5, Q14694, P07193, Q9UER7, Q9UNL4, Q8BKX1, Q96EB6, P54132, P0CG48, Q9Y2X8, Q9NXV6, Q13535, Q9VBX5, Q9Y371, Q9QZR5, O70445, Q5VTR2, P10415, Q8T0S6, Q92993, Q9BXH1, P23508, P10809, P04233, Q12824, P61289, Q93009, Q923E4, P23497, P45481, Q9UM54, Q96PM5, P29590, P38398, Q13315, Q9UPN9, Q99ML1, P10361, P43246, O35618, Q8IMZ4, Q9VCR6, O75832, P49841, Q01196, Q71UM5, Q9N6D8, Q13362, Q9UQB8, Q9NZQ3, O60934, P23804, O94776, and Q15910.