

3.6 Property Files

The properties files, sometimes referred to as aggregate function files, facilitate analysis by allowing visualization according to certain attributes of nodes and edges. These files can be used where aggregation of data is to be performed based on certain unique values. This ultimately enhances the power of visualization. The properties files are located in 'Sci2/sampledata/scientometrics/properties'

All properties files follow the same pattern:

```
{(node|edge)}.new_attribute = table_column_name.{{target|source}}.function
```

The first part of file specifies whether action will be performed on a node or an edge.

The next part, *new_attribute*, will be a name selected by the user, which indicates the name of the attribute.

table_column_name is the name of the attribute we are going to operate on to create a new value for the final node, this can be the name of any of the node attributes in the network

function determines how we aggregate, or combine, the values for the new node

Aggregate functions:

- **arithmeticmean** - finds the average of an independent node attribute
- **geometricmean** - finds the average of a dependent node attribute
- **count** - counts the instances of appearance of a node attribute
- **sum** - the sum of each node's attribute values. Example use: When you have two author nodes who are really the same author and you want to combine the number of citations they have accumulated under both names.
- **max** - the maximum value of each node's attribute values. Example use: When you have two author nodes who are really the same author, which have two differing author ages, you might want to assume that the younger age was based on an old record, and keep the older age
- **min** - the minimum value of each node's attribute values.
- **mode** - reports the most common value for an attribute

Below is an example of a property file:

```
node.numberOfWorks = numberOfWorks.sum
node.timesCited = timesCited.sum
edge.numberOfCoAuthoredWorks = numberOfCoAuthoredWorks.sum
edge.weight = weight.ignore
```

The idea here is that, for each set of nodes that are each others' duplicate, we combine each of the old attributes' values for those nodes using some mathematical function, in order to create the new attribute values of the final merged node. If an attribute is not mentioned, the algorithm will pick one of the attribute values from the duplicate nodes to use in the resulting node. This is acceptable if the two values will always be the same, but not if they could differ. Read the [Detect Duplicate Node](#) page for a more detailed explanation of this algorithm.

Below is a list of some of the various property files available for Sci2, grouped by the type of files they can be used to operate on. Along with a brief description of each property file, there are instructions on how to use the file. Note: this list is not complete. Be sure to check back as more property files are documented.

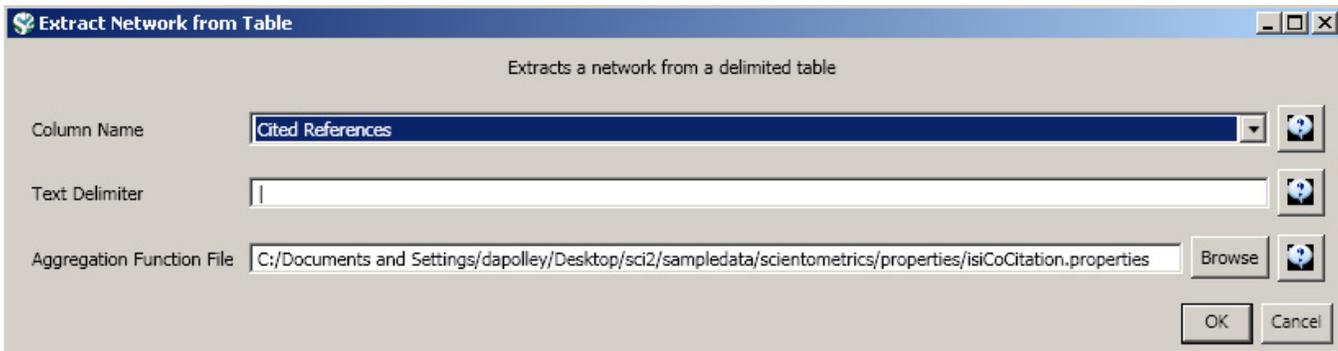
ISI Files

isiCoCitation.properties

This property file can be used with a co-occurrence extraction of cited references. The isiCoCitation.properties file will add the properties: "times cited," and "times appearing." It is useful for visualizing citation networks.

Select any ISI file to load in Sci2. Load the file in ISI flat format.

Choose 'Data Preparation > [Extract Co-Occurrence Network](#)' and apply the following parameters, making sure to select the aggregation function file indicated below:

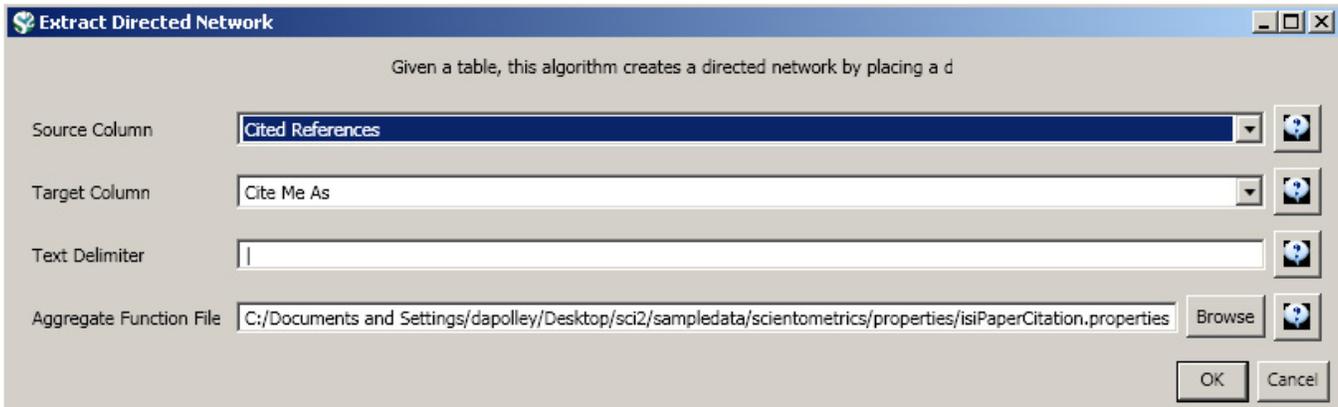


When the resulting network is visualized in GUESS, two properties will be appended to the nodes and will be accessible from the Graph Modifier pane. The isiCoCitation.properties file it will add "timesappearing" and "timescited" properties that will allow users to edit nodes according to these properties, allowing for more advanced visualizations.

isiPaperCitation.properties

This property file can be used in ISI paper citation data extractions. The file will add the properties: "global citation count", "in original data set", and "location citation count".

Load ISI file in Sci2. To use this file, users will want to extract a directed network, '*Data Preparation > Extract Directed Network*', and use the following parameters:



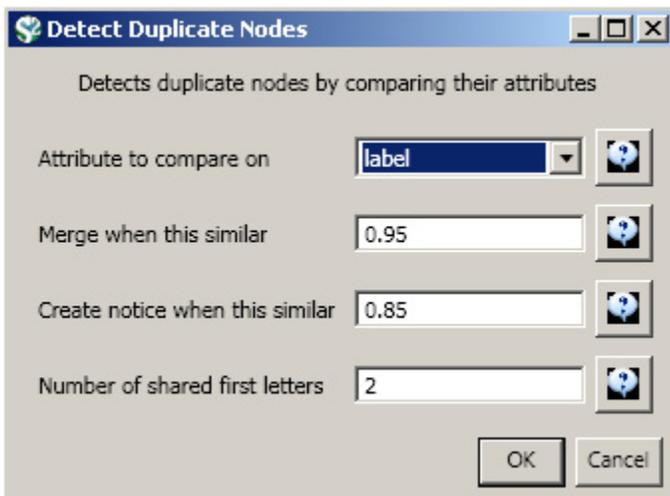
The resulting network will have the properties "globalcitationcount", "inoriginaldataset", and "localcitationcount" to the nodes in the network, allowing the nodes to be edited by these properties.

mergelsiAuthors.properties

This property file allows for the merging of duplicate nodes in an author table and updates the corresponding network.

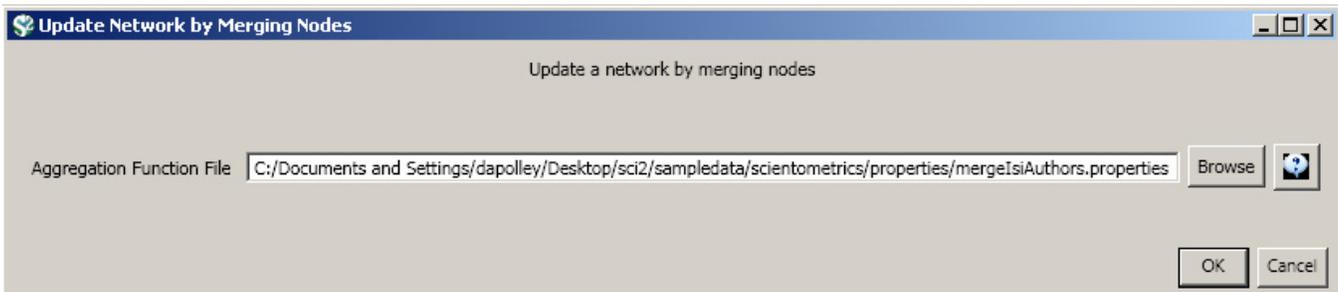
Load any ISI file in Sci2. This property file will be used when you extract a co-author network. Run '*Data Preparation > Extract Co-Author Network*'.

Next, you will want to detect any duplicate nodes in the network. Run, '*Data Preparation > Detect Duplicate Nodes*' using the following parameters:



Running this algorithm will result in a merge table and two text files. Specifically, the first log file provides information regarding which nodes will be merged, while the second log file lists nodes which are similar but will not be merged. The automatically generated merge table can be further modified as needed.

To merge identified duplicate nodes, select both the "Network with directed edges from Cited References to Cite Me As" network and "Merge Table: based on label" by holding down the 'Ctrl' key. Run '*Data Preparation > Update Network by Merging Nodes*'. This will produce an updated network as well as a report describing which nodes were merged. To complete this workflow, an aggregation function file must also be selected from the pop-up window:



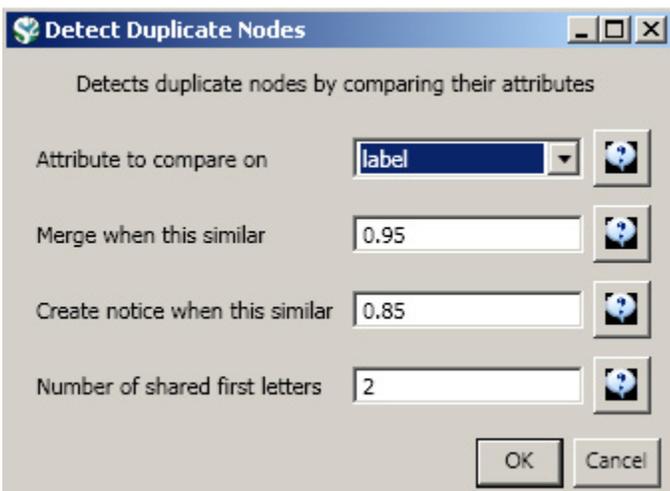
The resulting ISI co-author network will have duplicate author nodes merged.

mergelsiPaperCitation.properties

The property file allows for the merging of duplicate nodes in a paper citation table and updates the corresponding network.

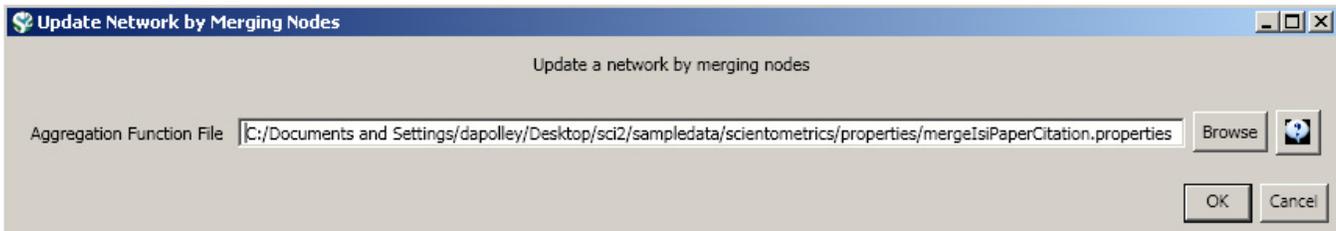
Load any ISI file in Sci2. To use this file, users will want to extract a directed network, '*Data Preparation > Extract Directed Network*', and use the following parameters:

Next, you will want to detect any duplicate nodes in the data set. Run, '*Data Preparation > Detect Duplicate Nodes*' using the following parameters:



Running this algorithm will result in a merge table and two text files. Specifically, the first log file provides information regarding which nodes will be merged, while the second log file lists nodes which are similar but will not be merged. The automatically generated merge table can be further modified as needed.

To merge identified duplicate nodes, select both the "Network with directed edges from Cited References to Cite Me As" network and "Merge Table: based on label" by holding down the 'Ctrl' key. Run '*Data Preparation > Update Network by Merging Nodes*': This will produce an updated network as well as a report describing which nodes were merged. To complete this workflow, an aggregation function file must also be selected from the pop-up window:



The resulting ISI paper citation network will have the properties "globalcitationcount", "inoriginaldataset", and "localcitationcount" to the nodes in the network and the duplicate nodes will have been merged.

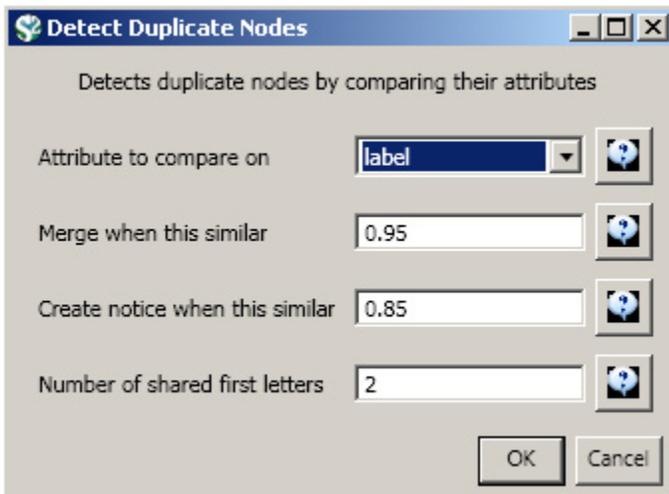
Endnote Files

mergeEndnoteAuthors.properties

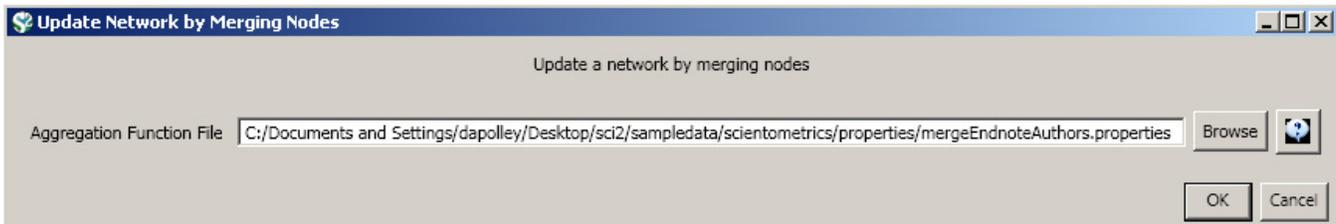
This property file allows for the merging of duplicate nodes in the author table and updates the corresponding network

Load any EndNote file in Sci2. Extract a co-author network by running '*Data Preparation > Extract Co-Author Network*'.

Run, '*Data Preparation > Detect Duplicate Nodes*' with the following parameters:



To merge identified duplicate nodes, select both the "Extracted Co-Authorship Network" network and "Merge Table: based on label" by holding down the 'Ctrl' key. Run '*Data Preparation > Update Network by Merging Nodes*': This will produce an updated network as well as a report describing which nodes were merged. To complete this workflow, an aggregation function file must also be selected from the pop-up window:



The resulting network will have duplicate author nodes merged.

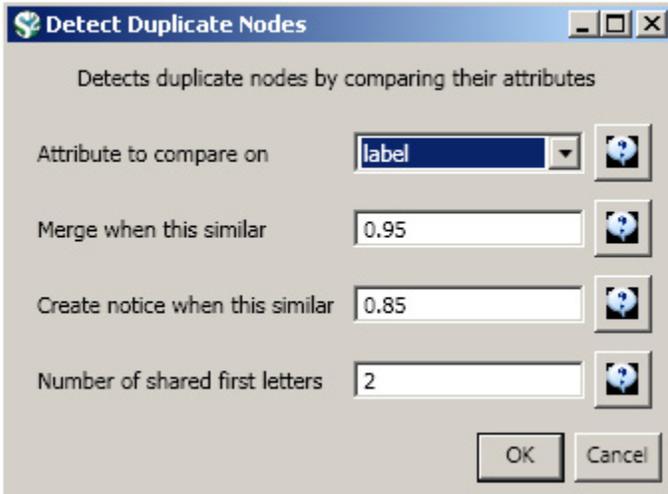
Scopus Files

mergeScopusAuthors.properties

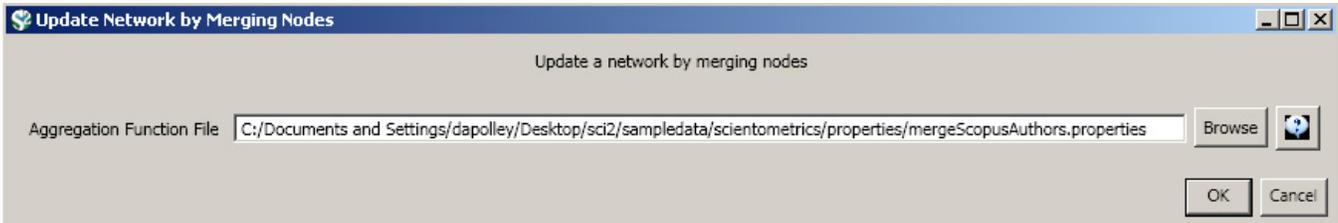
This property file allows for the merging of authors and updates the corresponding network.

Load any Scopus in Sci2. Extract a co-author network by following '*Data Preparation > Extract Co-Author Network*'

To detect any duplicate authors, select '*Data Preparation > Detect Duplicate Nodes*' with the following parameters:



To merge identified duplicate nodes, select both the "Extracted Co-Authorship Network" and "Merge Table: based on label" by holding down the 'Ctrl' key. Run '*Data Preparation > Update Network by Merging Nodes*'. This will produce an updated network as well as a report describing which nodes were merged. To complete this workflow, an aggregation function file must also be selected from the pop-up window:



The resulting network will have duplicate author nodes merged.