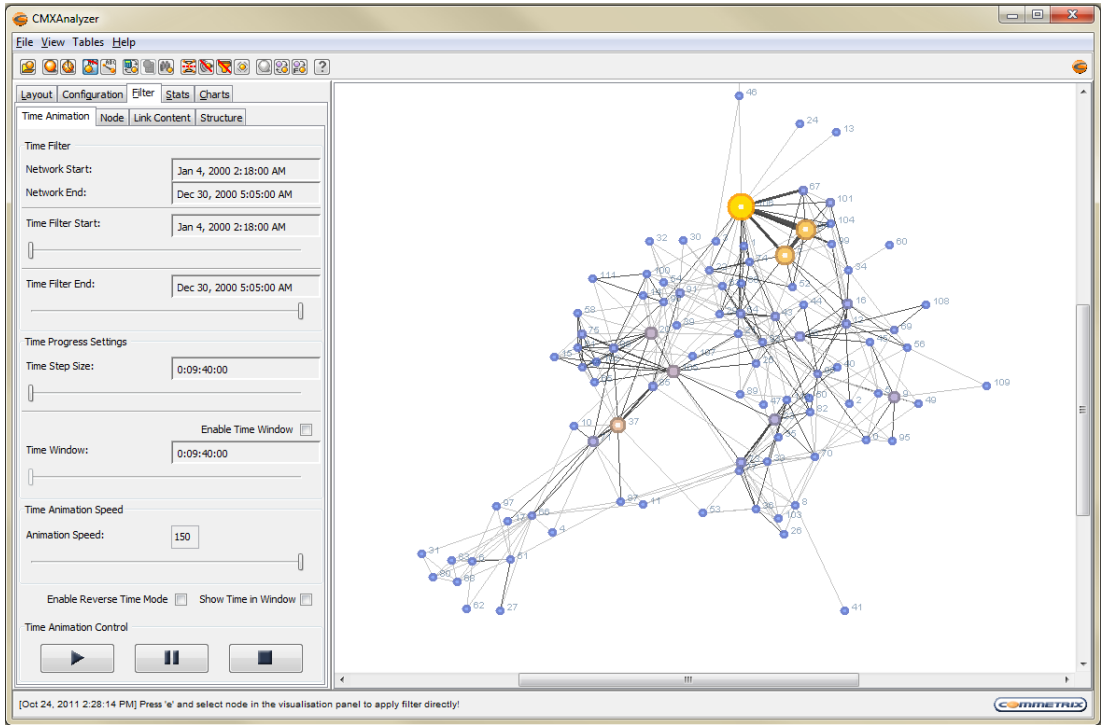


**USER MANUAL**  
**VERSION 2011-10-01**



**Commetrix Analyzer**  
**Exploration of Network Dynamics and Contents**

Product Website: [www.commetrix.net](http://www.commetrix.net)  
Vendor Contact : [info@trilexis.com](mailto:info@trilexis.com)



# Contents

1	Introduction .....	3
2	Installation and Activation .....	3
3	Starting Commetrix Analyzer.....	5
4	Main User Interface Elements .....	5
5	Opening a .cmx Network Datafile .....	6
6	Network Exploration.....	7
6.1	Layout Manipulation .....	7
6.2	Configuration of the network display .....	8
6.2.1	Labels .....	8
6.2.2	Size .....	9
6.2.3	Rings .....	10
6.2.4	Color.....	10
6.3	Filter and Search in Subnetworks .....	11
6.3.1	Structure Filters.....	11
6.3.2	Node Filter - Building Egonetworks.....	12
6.3.3	Keyword Search and topic related subnetworks .....	14
6.3.4	Time Filtering and Animation filtering time periods and animating transitions...	15
6.3.5	Animation.....	16
7	Network Analysis .....	17
7.1.1	Network Statistics .....	17
7.1.2	Node Statistics .....	19
7.2	Charting Metrics .....	21
7.3	Looking at original data tables .....	23
8	Changing Views.....	24
8.1	Different views 2D, sim 3D, 3D .....	24
8.2	Toggling Fading Mode .....	24
8.3	Toggling Label Mode.....	25
9	Dyad and Triad Analysis.....	25
10	Exporting Statistics to Tables (.csv,.xcl) .....	26
11	Contact .....	26

# 1 Introduction

**COMMETRIX** is a software framework for the analysis of social networks. Next to conventional static network analysis of cumulative snapshots of relationship networks, its event-based network model enables the examination of network evolution over time.

Commetrix offers an innovative **DATA MODEL**. Unlike in most other SNA tools, it does not store links directly as valued relationships but as individual relational events. This provides the opportunity of modelling networking processes with multiple types of relationships and comprehensive qualitative and quantitative node and link attributes in a single dataset. With this approach, users can encode aspects like topic descriptors (e.g. content coding, keywords), types of links, e.g. socialization, document exchange, affiliation, media, time stamps of links, links connecting multiple nodes, as well as arbitrary quantitative or qualitative variables classifying actors, e.g. affiliation, age, types, etc.

The **USER INTERFACE** is very easy to use. It is developed to visually support exploratory examination of a network dataset in order to identify and observe relevant substructures, periods, and processes of your network data. Commetrix computes time window measures and additionally provides very sophisticated functionality for displaying and animating the community evolution as an evolving graph to visually inspect the actors' activities. The animated graph, called *communigraph*, is one of the best existing visualizations of network change. Visual variables can be set by the user to represent node and link properties by label, node size, node colour (brightness, transparency), or a number of rings around the node.

Complex options for time, actor, relationship, and topic filtering help to **FOCUS RELEVANT STRUCTURES**, i.e. relevant actors, relationships, time periods, or even topics. For any time period and for any selection, typical social network measures can be computed and analyzed. Selections can be exported to tables for further analysis. This includes the export of network changes over time. All these features help to actually represent and visually trace change in a network and adds additional insight to the quantitative results.

## 2 Installation and Activation

The standard process involves

- Downloading the software from the website <http://www.commetrix.net> and
- Activating it afterwards in an (semi-automated) procedure based on a semi-automated e-mail interaction.

This way is recommended if you are installing Commetrix on your own PC for your personal use, e.g. to test the software with the demo dataset that comes with it. Note that there is a **FREE TRIAL** version.

If you cannot get this process to work, please simply request a license manually in a short mail to [licensing@trilexis.com](mailto:licensing@trilexis.com). We will then contact you via e-mail with more information. Via this e-mail, we can at all times inform you about pricing models or collaboration modes (if you are an academic researcher) and you can extend the license period.

Please note that you have to have an active internet connection, but only data directly necessary for the license activation request is being transmitted.

**Installation Process.** The homepage <http://www.commetrix.net> directs you to the download section. After entering basic data necessary for the licensing process, you are able to download the software system as a zipped file.

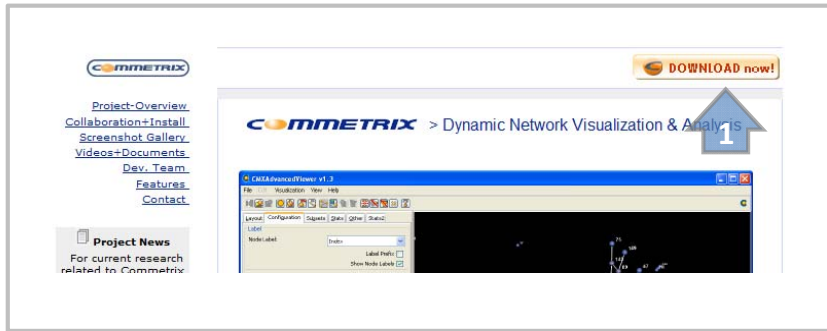


Figure 1: Internet Homepage of [www.commetrix.de](http://www.commetrix.de) with download button in the top right corner.

Click on that button to open the user registration page.

Title:	<input type="text" value="Mr"/>
First name:	<input type="text"/>
Name:	<input type="text"/>
Email:	<input type="text"/>
Username:	<input type="text"/>
Password:	<input type="password"/>
Validate Password:	<input type="password"/>
<input type="button" value="Submit"/> <input type="button" value="Reset"/>	

Figure 2: Web form for entering user data. Fill this out to create a user and submit to open the download page.

1) Create a folder at a location of your choice with a name you like. For this use your file browser (e.g. by pressing Windows key + e). Navigate to your intended location (e.g. drive C, folder program files). Click with the right mouse button in the right large panel of your window. Select “New”/”Folder” from the popup menu. Enter your own foldername, e.g. “Commetrix”.

2) Now, unzip the Commetrix zip file into this folder. This is usually done by first saving the attached zip file to your local computer and then opening it with a doubleclick. If you use windows, the unzip program should be installed and opens to show the files contained in the zip. Now you click on extract files. You are asked for a location. Select your created folder (e.g. “Commetrix” in the “Program Files” parent folder of your drive C). Click OK and the files are copied to that folder.

3) In your folder, you now see a file with the name “CMXAnalyzer.exe”. Click on this file to start Commetrix Analyzer. A login screen opens. Before you can login, an account needs to be created. Click on the button ‘Trial’ to trigger a license activation process (this also applies if you want to create a full license). A new window opens, where you need to enter your desired user name and a password. Please note that this password will not be shown again during the process, so please remember it well. An e-mail address is required to later receive your personal license activation file as an e-mail attachment. Please further read the terms and conditions and indicate your agreement in the little checkbox. Now you can press on the button “Request License” to continue with the process. The activation software now submits the entered information as an automated e-mail to Trilexis GmbH, the company offering Commetrix licenses.

4) If the semi-automated process should fail, a fallback process starts. You see a window, which displays a textbox with your user data and a license code. Please copy this text (e.g. by marking it with your mouse, and then right click then select copy from the menu or by marking it and pressing control key (ctrl) + c). Now you need to paste this text into an e-mail, which you need to send to [licensing@trilexis.com](mailto:licensing@trilexis.com).

5) Note that the activation in rare cases may take up to 24 hours as it includes a manual check. After your license activation request has been processed, you will receive an e-mail containing your personal license file "License\_Analyzer.dat" and some PDF documents with further information.

5) Copy the license file "License\_Analyzer.dat" into your Commetrix program folder (defined in step 1).

6) Double click on "CMXAnalyzer.exe" to start the Commetrix Analyzer. Enter your login data and password and then click on the button "Login" to start the Software System.

### 3 Starting Commetrix Analyzer

Make sure that your Commetrix program folder contains the license activation file "License\_Analyzer.dat" that you received via mail from Trilexis (previous step: installing and activating).

Once you have successfully installed Commetrix, you can start it by clicking on the executable file "CMXAnalyzer.exe" in the folder, where you installed Commetrix.

1) Login with the username and the password you have selected in the license activation process. Enter the password provided to you with the

2) The software starts with the welcome screen and you are ready to select datafiles in the Commetrix format .cmx and explore, measure, and visualize these networks. Note that a demo datafile "enron2000.cmx" is included in the install files. This is explained in detail in the next sections.

If a window opens, please read the message. It may inform you that your license expired. If you want to extend your license, you can at any time contact: "info@trilexis.com".

### 4 Main User Interface Elements

The following figure shows a screenshot of the application's interactive user interface.

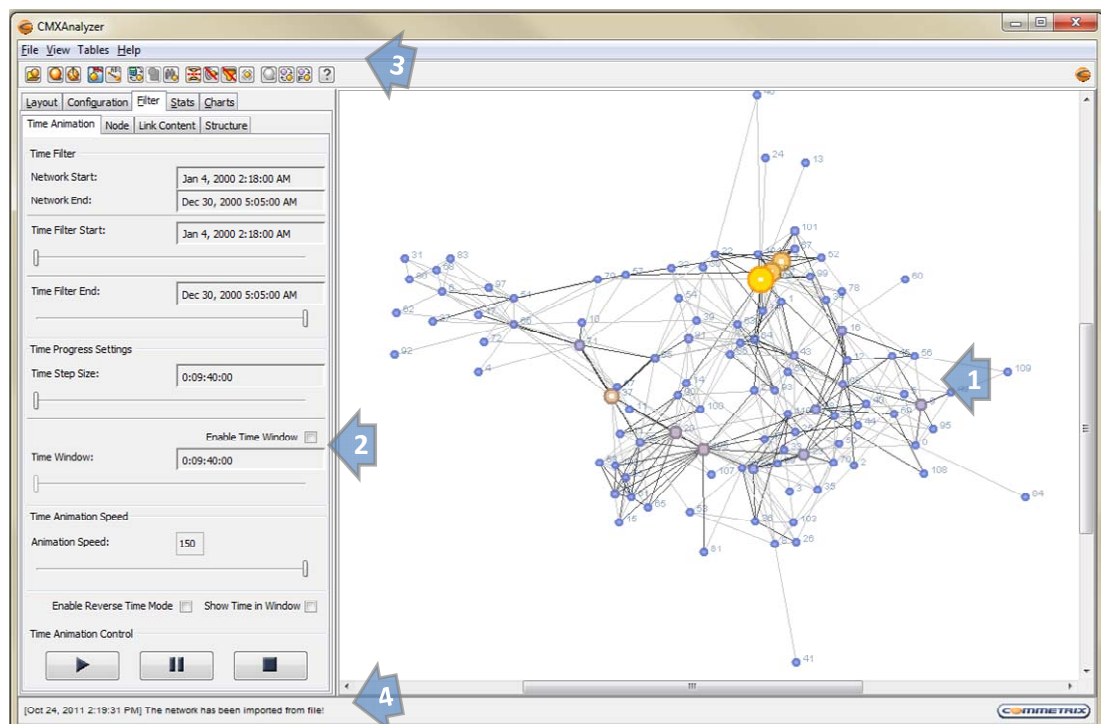


Figure 3: User Interface Commetrix Analyzer. (1) Main visualization window. (2) Control panels with different tabs and subtabs. (3) menu bar and shortcut icon bar. (4) status bar.

The main interface screen consists of various elements.

**Right.** The application consists of a main visualization window on the right hand side. Initially, here is a welcome screen, later you can observe your network in this main window. You can select nodes and deselect them by pressing in an empty area. When you have selected a node you can reposition it by pressing the right mouse button at your destination position.

**Left.** On the left hand side, you find the main control panel. It is divided into several tabs (e.g. layout, configuration, subsets, etc.) which can be selected with a mouse click, once you have loaded data into your Commetrix client. For example, you can click on the tab 'layout' to open an interface for configuring the layout of the network. Some tabs can have subtabs, e.g. the Tab "Subsets" has a subtab "Structure". Again, you can select those subtabs, e.g. 'structure' to open the interface for filtering subsets according to structural features.

**Top.** On the top you find a menu, which enables you to load network datafiles, to select the type of visualization, to view some related data sheets, or to access the help documents.



Figure 4: The icon bar. From left to right: (1) Selecting/Loading a new network. (2) Toggling 2D and 3D view. (3) Shortcut to the time animation control panel. (4) Toggling node labels on/off. (5) Toggling edge labels on/off. (6) Update network measures; requires computation time. (7) Export current status to tables; only if export is activated in the file menu. (8) Search network. (9) Toggling clustering on/off. (10) Freeze or unfreeze node movement. (11) Turn all filters off or on again. (12) Hide or show isolated nodes. (13) Switch grayscale/color display. (14) Update all variable scales, e.g. color shades of displayed nodes relative to the complete set of nodes in the datafile. (15) Update all scales, e.g. color shades of the nodes, relative to the selected and filtered subset of currently displayed nodes. (16) Access the help pages.

Under the menu tab you find a icon list with short cuts to the main functions of your client, including importing discourses, selecting discourses, changing view mode, storing your current discourse configuration, enabling or disabling labels, filters, time animation, freezing nodes, or accessing these help pages.

**Bottom.** On the bottom, you can find information about current processes and a progress bar in the status bar.

## 5 Opening a .cmx Network Datafile

To start the analysis, you need first to load a Commetrix network datafile (.cmx). For this you can either click the button on the icon list or you click in the top menu on File/Select Network. A dialog tab opens which provides you with a file explorer. If you have no own datafile, you can also use the demonstration network datafile coming with the download. It is located in your Commetrix program folder. See download instruction documentation to find this file.

Next to the demonstration file, you can of course select your own network files. One way to create such a file is to produced a .cmx datafile from raw data using Commetrix Producer. This tool is to be found in your Commetrix installation folder. Note that it is licensed separately. It enables you to connect to a database (in the case of Commetrix DB Producer) or to an Excel file (there is an excel coding template to aid coding network data in the right format). Commetrix Producer also offers a trial mode, in which production of network files is limited to 50 relational events (e.g. messages, relationships) between any number of nodes. Commetrix Producer has its own manual with more information. As an alternative to this production software Trilexis GmbH, the company licensing Commetrix, offers the "Production Service". If you want to learn more about this service, contact Trilexis using the e-mail address at the end of this document.

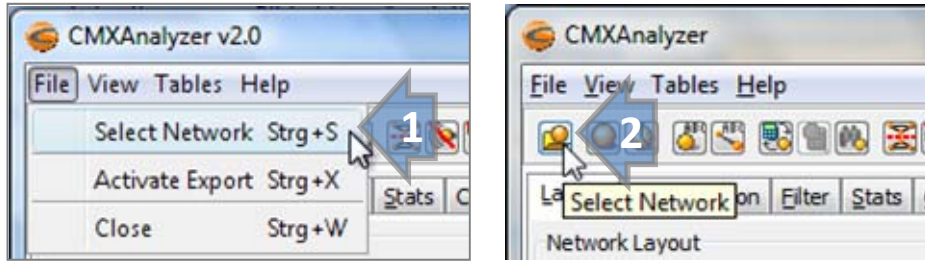


Figure 5: Selecting a network file (1) via menu or (2) via shortcut icon.

After you have selected a valid network datafile (.cmx), the network will be loaded. This may take a while according to the size of the discourse (even minutes sometimes).

After loading is finished you see your selected network with orange nodes connected via edges in the right hand visualization window. This network is now ready to be analyzed.

## 6 Network Exploration

Commetrix Analyzer enables you to tailor your observed network to your needs. You can precisely define what you want to see and how. This allows a unique way of exploring your network data. You can change the spatial layout, the use of colors, sizes, rings, and labels. You can utilize various filters to elicit interesting parts of the network for further analysis, e.g. only strong relationships, only reciprocal relationships, only active actors, only a limited set of actors, only certain keywords, only actors of a certain type, only a certain time window, and much more. Such a filtered part of a larger network is also sometimes called a subnetwork. You can highlight how your filtered set is embedded in the larger network. If time information is contained in your dataset, you can observe the evolution of your subnetwork set or of the complete network over time in a moving graph, just like in an interactive video.

After the exploration features are described the manual will continue with the analysis options.

### 6.1 Layout Manipulation

After successful selection of a discourse (click [here](#) for more information about this), you are ready to manipulate the layouting of the network. For this, click on the **Layout Tab** in the Control Window (on the left hand side). This includes the following steps:

**1. Zoom Level.** Define your zoom level by moving the zoom slider with your mouse and mouse button pressed until you have the desired magnification of your network data. Then release the mouse button. The number in the box corresponds with the selected zoom level. Simultaneously, your graph will change in size. Note, if you zoom in and take a screenshot (via a menu item), this picture will contain the whole (!) network at the selected zoom level, not just the part visible on the screen. In this way, you can create very large images of your network.

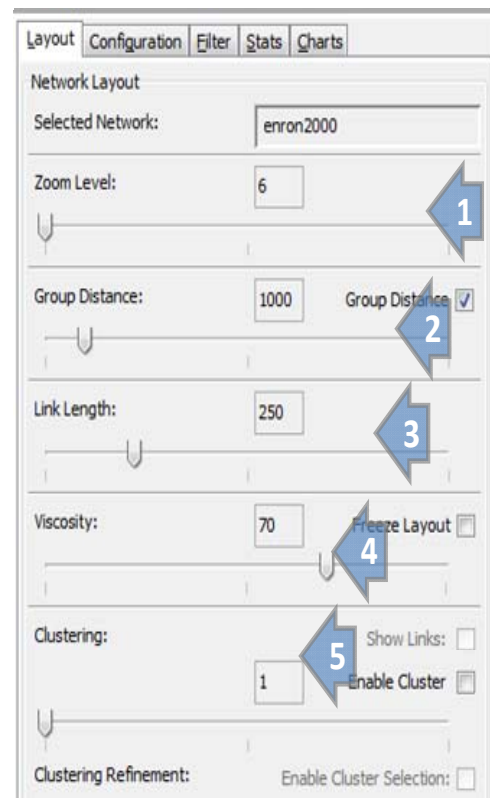


Figure 6: Layout tab. (1) Toggle zoom. (2) Toggle or deactivate group distances in the graph. (3) Change general link length to inflate or deflate the graph. (4) Change viscosity / inertia of nodes in the graph (5)

**2. Group Distance.** Moving the Group Distance slider changes the distance between nodes which belong to different groups (i.e. nodes which are not connected in the graph on the right hand side). This results in a layout, where different parts of the network are either close to each other or repulse each other, thus moving network fragments away from each other. In a fully connected network, this slider will have not much of an effect. The group distance feature can also be deactivated completely by unchecking the checkbox on the right hand side of the slider.

**3. Link Length.** Moving the Link Length slider affects the length of the links between nodes and thus changes the overall distance between nodes. The network will appear more or less compact. Increasing link length generally helps when labels are overlapping or structures are not transparent. Also check out the link length represents function in the Configuration tab, explained further below. You can experiment with combinations of adapting the zoom level (which also will yield transparent structures, but often parts of the network move outside the screen) and adapting the link length.

**4. Viscosity.** This slider affects the viscosity of the networks' movements. It is like configuring a jelly effect, as either your nodes move swiftly as in water or slowly like in jelly. In other words, this controls the inertia of nodes' movements. This sliders' setting is important to define how nodes should react to coordinate changes (movements). If you prefer to follow such changes slowly put the slider to the left, if you are interested in quick adaptations of the network to new settings put the slider to the right.

Activating the checkbox Freeze Layout (or alternatively clicking on the freeze icon in the icon list) results in a fixation of the current network layout. This can be helpful for massive graphs or complex calculations, as not every calculated change of node positions will result in rerendering the network visualization, thus freeing memory and computing power. A network is also often frozen, if you want to explore its temporal unfolding of relationships, without letting the nodes change their position.

**5. Clustering.** This function allows to check what would happen if hubs would be deleted from the network. By clicking on 'Enable Cluster', the algorithm deletes the top hubs from the network (most adjacent links). Nodes, which would be isolated remain with this eliminated node. The slider specifies how many of the top nodes should be treated in that way. This approach can be seen as a test of how the network would dissolve when hubs are attacked or eliminated. Further approaches to yield such clusters are possible via the subset tab. For example, you can delete weak relationships or nodes with few linking activity and observe how the network dissolves into activity-based substructures.

## 6.2 Configuration of the network display

After successfully selecting a network and after achieving a satisfactory layout you can visualize more information by setting node size and color to represent underlying variables. This can be done by clicking on the **Configuration Tab** in the Control Window (on the left hand side). Configuration includes the following elements.

### 6.2.1 Labels

**1. Node Label.** You can show labels next to the nodes. Clicking the selection box next to 'node label' provides you with a list of optional labels. These are dynamically derived from your dataset. They include index, group, database-ID, and keyword list. Then follow five individual attributes specifically defined for the dataset. Typically they will contain a unique name, an alias name, plus three further node variables (e.g. node's type, age, etc.). Finally, there is a large set of node related metrics. Such metrics are automatically recognized by the software as a possible node label. Examples are sent and received link events, or betweenness centrality. Labeling a node with such a metric can help to identify interesting nodes quickly.

Activating the “Show Node Labels” check box is necessary to add the labels to the visualized graph. Alternatively, this can also be done via clicking on the respective shortcut icon in the icon bar. By unchecking this box, all node labels are hidden in the visualized graph.

Activating this box will add the variable that the node label is currently representing, e.g. “name” or “index”. This can be useful to give the information of the labels’ type in exported graph images so that the reader knows, the node label shows e.g. centralities. For example, activating this box can change the label from ‘13’ to ‘Index 13’.

Link labels can be used similar to node labels by clicking the selection box next to ‘link labels’. The selection box provides you with a list of possible labels, including number of contacts, number of linkevents, first and last linkevent, numbers of linkevents initiated by node A or B of the relationship, keywords, or reciprocity (yes,no).

Just as node labels, link labels can be turned on or off using the checkbox next to the selection menu.

**2. Label Size.** A slider below the label selection menu allows to define the label size. If the slider is not active, there is no label turned on in the above node and link label selection menu. In this case, you first need to check either ‘show node labels’ checkbox and/or ‘show link label’ checkbox.

#### 6.2.2 Size

**3. Node Size represents variable.** You can set the node size to represent an underlying variable, like e.g. the activity or the number of linkevents. This results in large nodes representing large values of the underlying variable and smaller nodes representing lower values of the underlying variable. By default, this option is set to ‘constant’ size. If you do not want to show a variable by node size you can at every time select this option ‘constant’ in the selection box next to ‘node size represents’. This box provides you with a list of possible underlying variables, e.g. ‘Linkevents sent’, ‘Direct contacts’, ‘Betweenness Centrality’, etc. If you have defined node attributes, which lend themselves to visualization as node size (e.g. no nominal scales), then these attributes will also appear in the selection box.

**4. Node size maximum and minimum slider.** These two sliders let you define the difference between the largest node visible (upper slider) and the smallest node visible (bottom slider). Together this changes the degree of differences or the range of sizes of the underlying visualized variable. This allows you for example, to reduce unimportant nodes to the size of a dot in order to highlight important nodes.

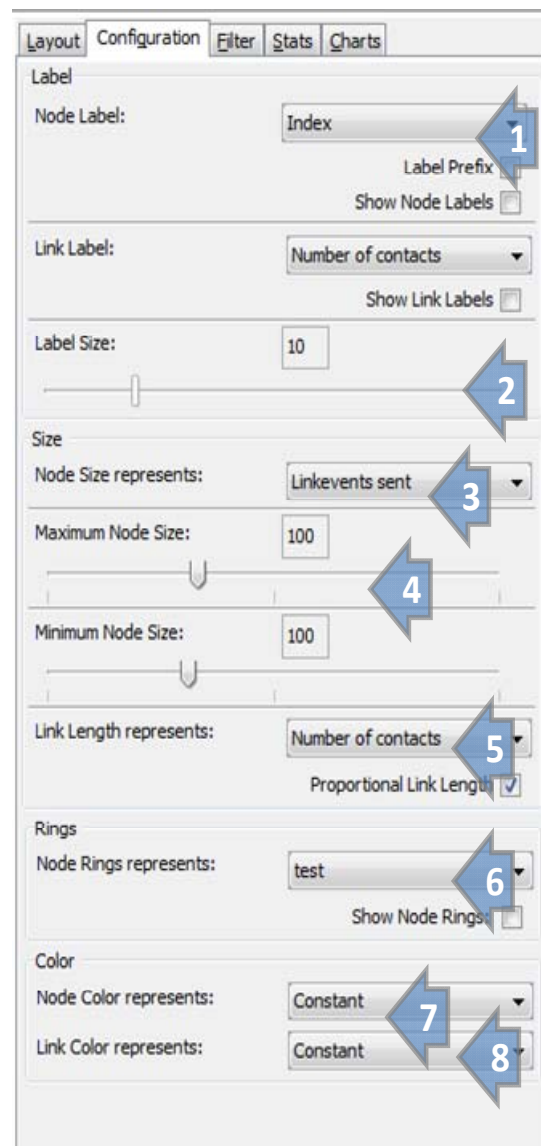


Figure 7: Configuration options. (1) Select variable to be used as a node and link label. (2) Toggle label size. (3) Toggle variable to be represented by node size. (4) Toggle parameters for larger versus small nodes. (5) Select the variable that is represented by link length. (6) Select

**5. Link Length represents variable.** You can set the link length to represent an underlying variable, like e.g. the number of link events that constitute the link. Clicking the selection box next to 'link length represents' provides you with a list of possible underlying variables, e.g. 'number of contacts', 'Number of linkevents', etc.

The link length Checkbox next to the slider is used to activate and deactivate the correspondence between link length and the selected underlying variable. For example, if the box is activated, this results in shorter links for tight relationships with high interaction rates versus longer links for weaker relationships. Default of this checkbox is checked as this will usually result in a more intuitive graph layout, representing dense clusters. Unselecting this check box will result in edge lengths which do not reflect the strength of the relationship and are only reflecting the embeddedness of the node in the network. This is usually unfolding the center of the network a bit.

### 6.2.3 Rings

**6. Rings.** In 2D mode you can select rings to visualize an underlying variable, like e.g. the activity of a node. The visualization proxy extends the options for displaying underlying variables in the graph (next to node size, color, and label). It generates rings around nodes in a number from 0 to 5. Clicking the selection box next to 'node rings represent' provides you with a list of possible underlying variables, e.g. 'Linkevents sent', 'Direct contacts', etc. This could for example be used to highlight large values of a node variable, e.g. a high rank of an actor.

The “Show Node rings” Check Box needs to be checked in order to add rings around the nodes. Unchecking the box will hide the rings again.

### 6.2.4 Color

**7. Node Color represents variable.** You can set the node color to represent an underlying variable, like e.g. the activity of a node. There are two color scales. One is a set of discernable colors used for nominal (categorical) variables, which have no ranking order. The second color scale is for metric variables: a gradient ranging from blue for low values to orange for high values (i.e. a temperature-like scale from cold to hot). For example, if you want to display the activity of nodes as a color, Commetrix automatically recognizes that this is a metric scale and applies the gradient to represent the different levels of node activity. This results in orange nodes with high activity and blueish nodes with lower sending activity. Clicking the selection box next to 'node color represents' provides you with a list of possible underlying variables, e.g. 'Linkevents sent', 'Direct contacts', 'Core Group Member', etc. Selecting 'Constant' will show all nodes in equal size again.

**8. Link color** can also be used to represent a link variable. Link color is a grey gradient with darker colors representing stronger links and lighter colors for weak links. The variable to be represented by link color can be set by using the selection box. For example, selecting the entry reciprocity in this box will result in light grey links showing unidirectional relationships and darker links showing reciprocal relationships.

The color scales used can be switched to a grey color set using the respective button (depicted as a large grey node) in the icon bar. This can be useful, if you need to produce a picture for a black and white print. By pressing this button again you can switch back to the original color set.

The Update Property Scale buttons (see figure of buttons, above) can be used ensure that the current view is also containing the updated values and colors of some more complex metrics (not all are computed after every user interaction). There are two modes, one is updating properties in relation to the complete dataset (all nodes, also those currently hidden in your view), the other is only considering the currently active set of nodes (ignoring the nodes that you have filtered out during your session). The latter mode will usually result in a fuller usage of the color gradient as the extreme values of the color gradient

are related to the extreme values of the currently active networks (and not to some node currently not visible).

## 6.3 Filter and Search in Subnetworks

By utilizing a complex set of filters you can elicit relevant parts of your network data. In Commetrix, these selected subsets of the overall network data are called subnetworks. You can do this by filtering out nodes and links that are not of interest for you. You can also search in the graph. The main window will always directly show the current result of your filtering. This enables an interactive and visual exploration process.

All filters are reached via the ‘filter’ tab. There are several subtabs for affecting the structure, selecting only certain nodes according to their properties (e.g. as ego-nodes), selecting a subset that matches certain keywords, or limiting the time-period shown as linkevents in the graph. All filters can be combined, leaving you with numerable options to elicit relevant aspects of your data. At each time your filters correspond with the graph, and new metrics are computed. Additionally, you can press Control+F to search the network data. Search is hence also filtering the network, e.g. to all link events and associated actors, which relate to a search term. In the next sections, filtering and searching is explained in more detail.

Filters and Search may change the visible network data. Not all metrics are computed on the fly in the background. Some require a manual update, because their computation may take some seconds or minutes. Whenever you see a ‘-‘ in your stats or if you want to ensure that more complex metrics are correctly updated (as they are only computed on request) you can press the **“Update Properties”** button in the icon bar (see the buttons and their functions in the figure above).

If you should want to turn off all filters at once (e.g. in order to compare the filtered set with the original dataset), you can click on the **“Filter off”** button in the icon bar. Often it is advisable to freeze the graph before you switch off all filters. When frozen, all nodes will remain at their current position and you can better compare the filtered graph with the unfiltered original. You can turn on all filters that you have set by again clicking on the **“Filter”** Button in the icon bar.

Another option for comparing your filtered sets with the complete dataset is to activate **“Fading mode”** via the **“View”** menu in the menu bar. This will result in a different mode of filtering out nodes. Instead of hiding them from your view, they are shown in a very light grey (i.e. faded out). You can now always see how your filtered network is embedded in the overall network. Using the **“Preferences”** in the **“Help”** menu in the menu bar and clicking on its **“Coloring”** tab enables you to change the visibility and transperence of these faded nodes.

### 6.3.1 Structure Filters

**1. Node’s Linkevent sent Filter.** If you move the slider of this filter to the right, a threshold/filter is increased for showing only nodes with at least the defined number of initiated link events in the network. For example if you set this slider to the value 10, only nodes that have triggered (sent) at least 10 linkevents are remaining in the graph. The rest is filtered out. This filter results in a reduced graph on the right hand side containing only the more active nodes.

The checkbox **“Weak Node Filter”** inverts the filter so that if you move the slider to the right now, nodes with the defined or a larger number of initiated linkevents are filtered out. If you mark this checkbox and set the filter to value 10, you filter out all actors with at least 10 linkevents. This results in a subgraph containing only the less active nodes.

**2. Link Strength Filter.** If you move the slider of this filter to the right, a threshold/filter is increased for showing only relationships with at least the defined strength. This strength relates to the exchanged linkevents between nodes. Your manipulation results in a reduced graph on the right hand side. The remaining structure only hides the weak relationships and reveals the core relationship network of your communication data.

The checkbox “Weak Link Filter” inverts the filter so that if you move the slider to the right now, links with the defined or a larger strength are filtered out. If you mark this checkbox and set the filter to value 10, you filter out all links with at least 10 linkevents.

This results in a subgraph containing only weak relationships.

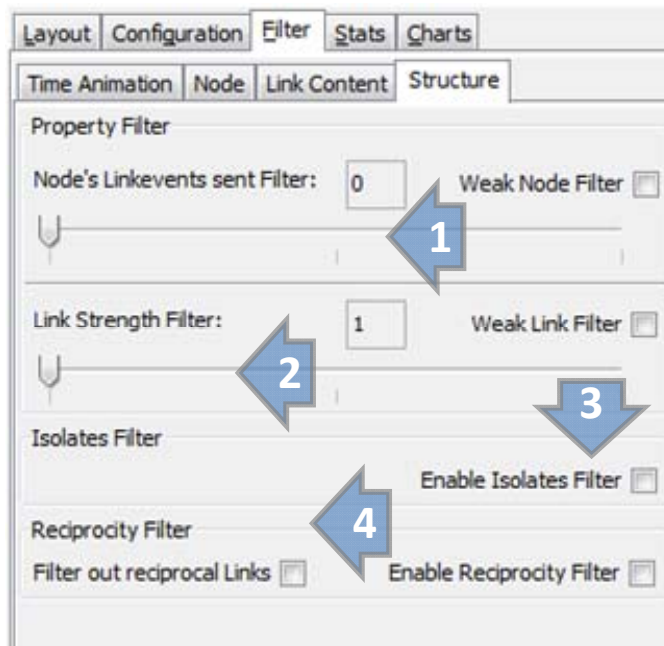


Figure 8: Controls for Structural Filtering.

**3. Enable Isolates.** Checking this checkbox will hide nodes from your network, who have no links to other nodes. Such authors are called isolates. Your manipulation will usually result in a reduced graph on the right hand side.

**3. Reciprocal Links.** The reciprocity filter consists of two checkboxes. The right checkbox turns this filter on (checked) and off (unchecked). The left checkbox switches between two effects of the activated right checkbox. Filtering out all reciprocal links (checked) or all non-reciprocal, unidirectional links (unchecked). Checking this checkbox will result in a graph that only shows unidirectional, non-reciprocal links.

Two examples help to demonstrate how the reciprocity filter works. If the right checkbox is checked without simultaneously checking the left checkbox, you will be able to filter out all non-reciprocal links. You can simply check this by setting the link label to reciprocity. This label shows “no” for non-reciprocal links and “yes” for reciprocal links. Now uncheck the enable button again and check the mode button “filter out reciprocal links”. If you then activate the “Enable Reciprocity Filter” button again, you will yield a subnetwork where the reciprocal links are filtered out and only non-reciprocal links are contained in the network.

### 6.3.2 Node Filter - Building Egonetworks

Nodefiltering or ego-filtering allows you to focus on one or more particular nodes in your network. Usually it is employed to visualize and measure the often star like network around one selected node, e.g. a certain person. An interesting features of an egonetnetwork is for example how many links the contacts of the selected nodes exhibit between each other (as a measure of their independence from the ego node).

In Commetrix you can filter out more than one node (even all but one if you like). You can thus use node filtering for creating interesting subsets of nodes.

You turn on node filtering options by first clicking on the “Enable Node Filter” checkbox.

**1. Depth around ego.** This slider affects the size of the egonet network. Putting the slider to the value 0 will only show the selected node(s) without their neighbors. Value 1 is the most common value and shows all direct relationships which of the selected node. Value 2 will also show paths with a length of two steps which start at the node (that is why it is called relationship depth 2). This means that not only the direct communication partners of the selected node are shown, but also the indirect at the length of two steps (the contacts of the node's contacts). The next values 3..6 will increase the allowed path length accordingly. Given the law of 6 degrees (every two persons in the world can be connected by a path length of only six), the whole network should be shown when the value 6 is selected.

**2. Selecting Node Property.** This selection box allows you to select a criterion which is used to list the available nodes in the large selection box. For example, selecting the entry "Name" will list available nodes by their name. If you have a property "Type", the box will list all available types for you to select. If more than one node is having the same value of the selected property, you select them all. This is a good way to select a group of nodes, e.g. all nodes of the type 'project manager'.

**Selecting the Nodes.** After you selected the sorting criterion the large list box presents a list of available nodes (or node attributes) in your dataset. The list enables you to select one or more nodes for which a reduced network should be generated (for one selected author this is called an egonet as it shows the environment of one person/ego). You can select a node by clicking on it. If more than one entry in the list shall be selected, you need to press control while clicking on various list entries. You can select all entries between one selected and a second selected entry by keeping the space bar pressed when you select the second entry. Marked entries are indicated by a colored bar.

**3. Apply Filter.** This button uses your defined selection of the node subnetwork and computes the according filtered graph in the visualization panel.

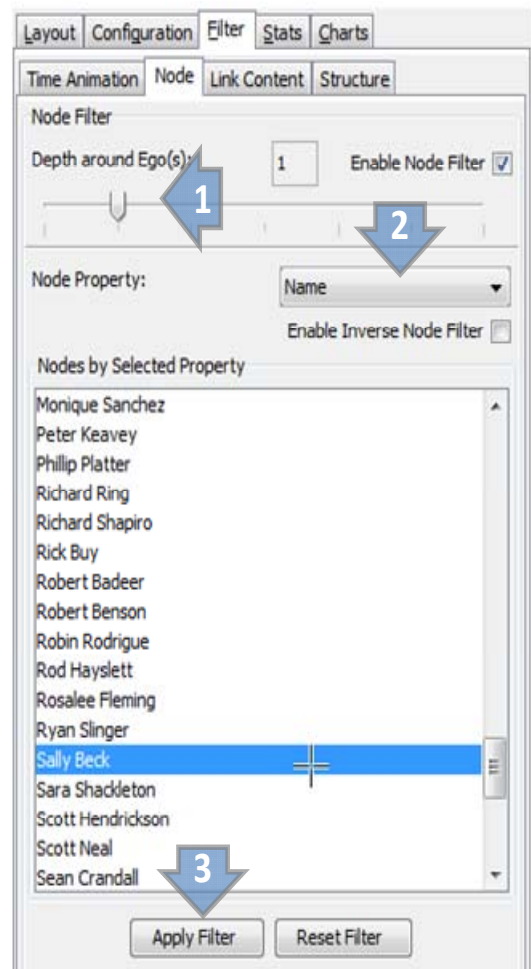


Figure 9: The ego view panel. (1) Define the path length around ego nodes (set to 1 if only direct contacts shall be shown). (2) Select the node property for the node list below, e.g. list nodes by name. (3) Confirm your selections and apply the filter or reset the filter again to deactivate it.

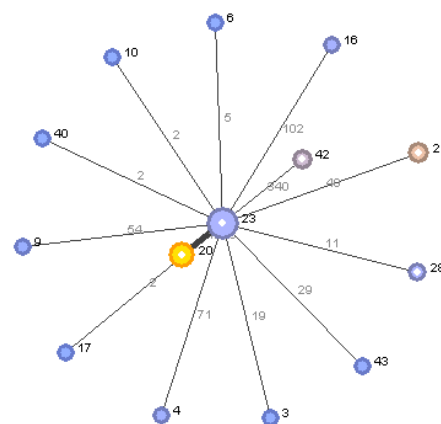


Figure 10: Example Ego Network around the central node.

### 6.3.3 Keyword Search and topic related subnetworks

Commetrix offers content filtering, e.g. searching for keywords. You thus can reduce your network to include only linkevents that relate to particular search terms. Examples are link references or discussion postings with a particular topic (e.g. ‘book’ or ‘laptop’). This can be done by clicking on the “Link Content” subtab of the “Filter” Tab. Content/keyword filtering includes the following elements.

**1. Content Keyword Filter.** The keyword filter has to be activated before the other functions of this tab can be utilized. Check the box „Enable Content Keyword Filter“ to activate.

**2. Select OR versus AND search.** Either all selected terms have to be in the link event or just one of the list.

**3. Select Search Term(s).** You can select one or more search terms from the terms identified in your dataset by using the LEFT keyword selection windows called ‘Primary INCLUDE Filter’. If more than one entry in the list shall be selected, you need to press control while clicking on various list entries. You can select all entries between one selected and a second selected entry (= a range) by keeping the space bar pressed when you select the second entry. Marked entries are indicated by a colored bar.

Clicking on a selected term when control key is pressed will deselect it again.

**3. Optionally select exclude keywords.** Next to selecting keywords that reduce your overall network to INCLUDE only the activity that relates to the defined keywords in the “Primary INCLUDE Filter”, you can optionally also specify one or more keywords in the “Secondary EXCLUDE Filter” list. Note, this list is considered in a second step to reduce from your INCLUDE list linkevents that contain certain keywords. For example, assume the OR Search is activated. You select the four keywords shown in the Figure. Now if you want to NOT consider a particular topic than mark it in the secondary EXCLUDE filter. This then reads: Give me a subset with all linkevents that contain keyword A or B or C or D (from the include filter) BUT NOT E (or F or G) from the EXCLUDE list.

**TIP: Create special selections.** Further note, that keyword filtering can be used as a powerful tool to create selections for your filtering. For example, if you want to reduce a network to include only link events of a certain type, e.g. “new” to indicate data that you added in a second step, you could manually code a specific signal-keyword to the respective set of link events. This must be done upfront in your dataset coding process. We recommend to use a “!” together with the term, e.g. “!new”. This will later show you, as the user, that this entry of the keyword list is actually a (artificial) signal keyword, that you

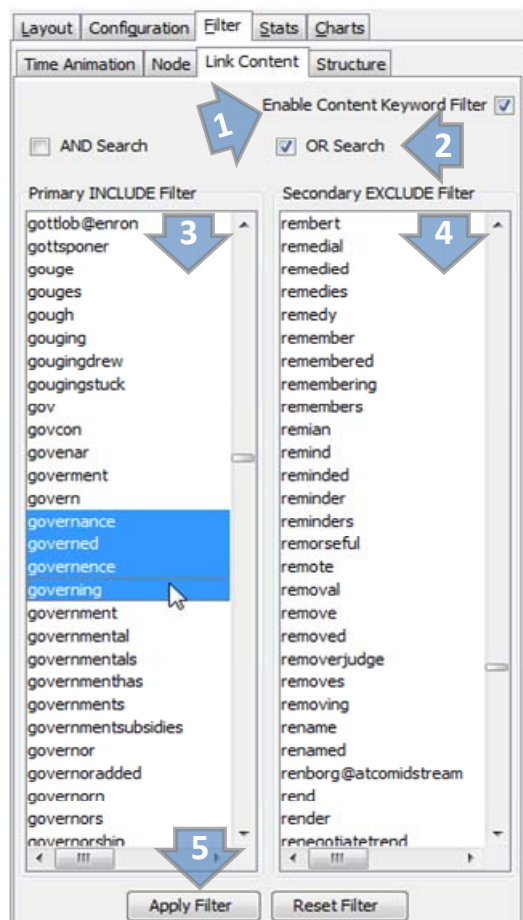


Figure 11: Keyword filtering. (1) Switch filter on or off. (2) Select AND versus OR search. (3) Select which keywords should be used for filtering a subset of related link events. (4) Optionally select keywords that exclude Linkevents, if they contain the keyword (this is used to refine the INCLUDE filter!) (5) Apply your current keyword filter to the graph or reset everything.

have assigned manually and not a normal keyword that appeared somewhere in the data (as in “...new colleague...”).

If you have coded some activity with such artificial signal keywords, you can select them easily with the content filter and include (or exclude) only nodes of type “!new” by selecting this code as a keyword.

**4. Apply Filter.** Pressing this button will use your current specifications of the keyword filter and filter the data with it. Alternatively you can press reset to eliminate all defined in or out keywords from the keyword filter and start over again.

#### 6.3.4 Time Filtering and Animation filtering time periods and animating transitions.

Time Filtering is an important function for dynamic network analysis. You select the corresponding control panel by clicking on the “Time Animation” Subtab of the “Subsets” Tab.

You can reduce your network to include only link events (e.g. messages), relating to a particular period of time. Commetrix can play through such periods (also called time windows) and by that animates and measures your network’s evolution. You can export all selected/filtered data to a table, which you can later edit with your preferred table calculation tool (e.g. Excel).

**1. Duration of your network data.** The start and end time (as defined by the first and the last linkevent) are shown.

**2. Time Filter Start Position.** This slider sets the start position of your animation. Very often, this slider remains in the left position. The corresponding absolute time is also displayed in the text box next to the slider. At all times, your visualized network shows all linkevents and adjacent actors after this point in time. If you move the slider to the right, you exclude all link events BEFORE that date from your analysis (and thus you create a subset of your graph). This can be useful, if you have a large network with a large time duration and you want to focus on the more recent events. In the final (cumulative) network they are more emphasized if you blend out old networking activity. For example, in a citation network you would yield a graph that ignores those old seminal papers who all cite and focus more on the recent additions to the academic discouse. The network will adapt to every movement of this slider and you can observe how the network transitions to blend out older linkevents and nodes and move more recent nodes to the center.

**3. Time Filter End Position.** This slider defines the end date of your observed network. At the same time this is corresponding with the current position of your animation (the “now”; i.e. the day of

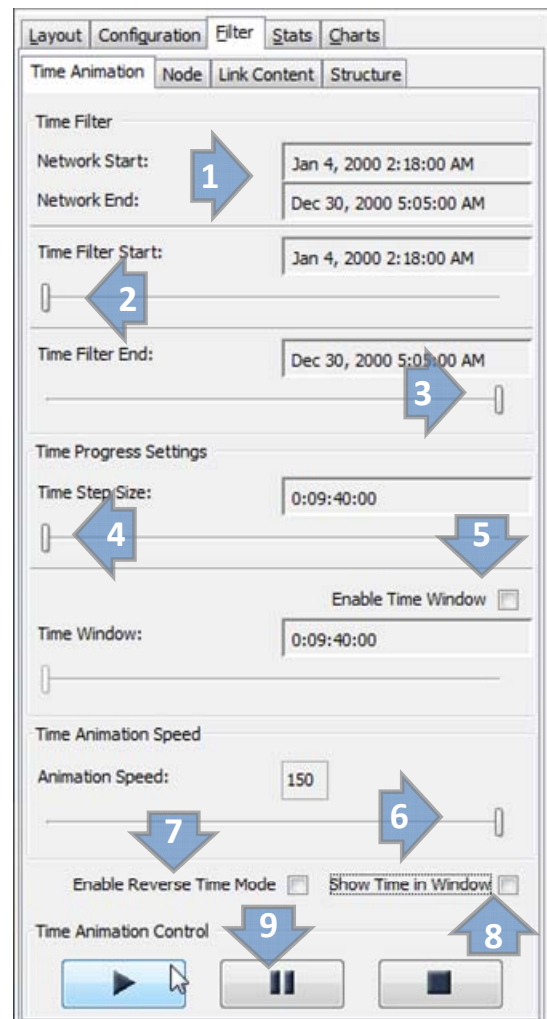


Figure 12: Time Filter Panel: (1) Duration of your network data. (2) Events before this date are filtered out. (3) Events after this date are filtered out. (4) Definition of step size through time period. (5) Activate sliding window mode and select sliding window size with slider. (6). Change animation speed. (7) Play animation in reverse order; i.e. backwards. (8) Show a little window with the current and the end time. You can move this window over the animation by dragging it with the mouse. It can be useful as a legend for screencast videos. (9) Time Animation controls, play, pause, stop.

observation). Moving this slider to the left will reduce your sample to not include messages exchanged AFTER the defined time. So for example, if you have a dataset of 100 days, you can create the network as of the 90<sup>th</sup> day by moving the slider 10 days to the left. The network will adapt to every movement of this slider and you can visually observe its shrinking. If you then slowly move the slider back to the right, you can visually trace, how the network developed through the last ten days.

**4. Time Step Size.** You can define the step size of the movement through time. The time shown is in “days:hours:minutes:seconds”. The animation is usually done by aggregating all events of the next time step and then updating the graph (in a transition) to integrate these events. If time step size increases, there are more events aggregated to one update step. That means, small time steps will show almost every event individually and large time steps will show more abrupt changes as more events “enter” the graph together. For example: setting this slider to 1 day will add 1 day of networking data after each step of the animation (i.e. in each new frame).

The feature can be used to increase animation speed as less “animation frames” are computed and thus the animation consists of fewer individual “pictures”. It is also useful to match time step to some underlying step size in your data. For example, if you have only annual observations in January, you will have an animation where much happens and then for a while nothing happens (e.g. from February to December). However, you can adapt time step size to a jump of 12 month, which should yield a smoother time animation. The months between your events are simply passed more quickly.

**5. Time Window.** This slider enables you to define a so-called time window. It can be thought of as a moving window with a certain range of visibility, which shows exactly your defined range of data when it moves along the whole sample period. For example, you are setting the window to a size of 10 days. Then starting your animation will show you the network from day 0 to day 10, from day 1 to day 11, from day 2 to day 12 and so on. This is effectively yielding an animation with a visible time window of 10 days. In practice, this enables you to watch only the changes (delta) of your networks evolvement (10 days of change, you could also set the time window to 1 day to observe daily change).

The checkbox “Enable time window” has to be marked to activate the time window function.

### 6.3.5 Animation

**6. Animation Speed.** This slider lets you manipulate the rendering speed of the animation of the graph's evolvement. Unlike the Time Step Size above it affects only the frames per second, i.e. the speed of the animation, and has no effects on the size of added data for each new frame. If the animation starts you can see how the time filter end slider moves as time is passing. If the animation is finished, it automatically starts over again. Please note, that in order to yield a better result, it is often advisable to let the animation run through a second time as some node positions will have moved closer to their optimum.

In the Help/Preferences Menu there is a tab “Time Animation” with further settings. For example, you can define the animation speed. The default setting is 150. You can simply enter another number and the slider will cover the range from 0 to that number. You can further increase the node movement rate in these preferences. It requires a bit of experimenting with the variables node movement, time step size, animation speed, and viscosity (in the layout tab) to optimize the smoothness of the animation. Please note that if no change effect is to be observed, it can help to move the animation speed slider a bit to update the settings.

**7. Reverse Time Mode.** Near the animation speed slider, there is a checkbox to switch on the “reverse time mode”. This is effectively playing the animation backwards and can be helpful if you want to study how the network reduces itself back to its beginning (like “dissolving” or “imploding”).

#### 8. Show a small information window.

It can be dragged across the screen and is useful to capture the network evolution as a movie and display in it information about the current and the end time.



Figure 13: Time Animation Controls, play, pause, and stop

**9. Time Animation Control.** Here you find the typical controls for animation and videos. You can start playing the animation. This will show the evolvement of the network. You can pause and resume the animation using the pause button. Finally, you can stop the animation and put the pointer to the beginning again.

## 7 Network Analysis

Next to visualizing the network, there are many metrics computed in the background. Most of them are social network analysis metrics, like degrees or centrality. You can consult the computed metrics by clicking on the “Stats” tab. The panel is separated into a box for metrics on the network level and one box for metrics on the node level. In order to study node metrics, you need to have a node selected. This is simply done by clicking on a node in the graph. The node is marked red and its metrics appear in the stats box for node metrics. If this box is empty, no node has been selected. If you want to unselect a node, click on another node or click on an empty area in the graph window.

Please note that Commetrix computes all quick measures on the fly (and keeps them current). However, there are many measures in network analysis, which require some computational effort and would, if updated after every user interaction, delay the whole usage experience. That is why there is an update button in the icon list. If you click on this button, everything will freeze for some time (depending on your network size, this can even take several minutes), and all metrics are updated.

Instead of the “Update Properties” button, you can also click on the “Update” button below the statistics boxes on the “Stats” tab.

All metrics, which Commetrix recognizes to be outdated due to a change of the dataset driven by the user, are marked with a “-“. After you clicked on the “Update Properties” button, these “-“ should be replaced with the latest computed values. If there are still some minuses left, this means that the value is infinite or empty in your current view of the dataset.

### 7.1.1 Network Statistics

**1. Network Name.** This item gives you the name assigned to the active network.

**2. First Link Event.** Here you can read the date of the first entry in your dataset: the first linkevent.

**3. Last Link Event.** Here you can read the date of the last entry in your dataset: the last linkevent.

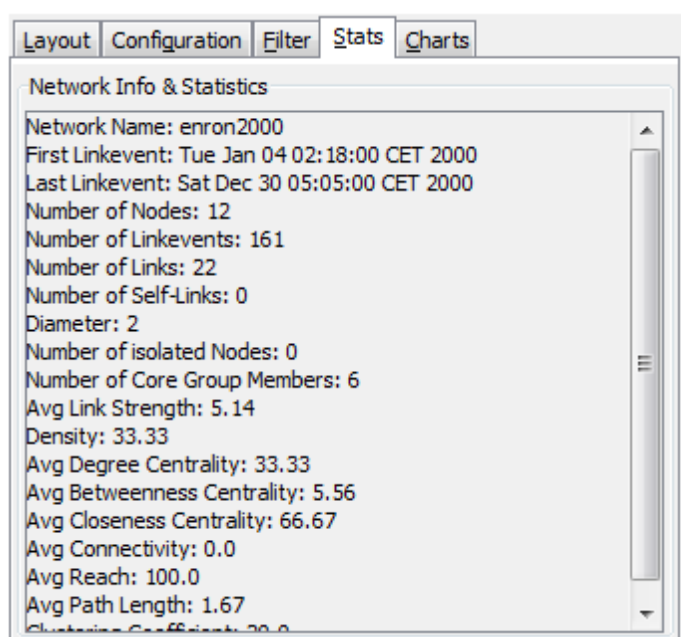


Figure 14: Network information and statistics.

**4. Number of Nodes, Linkevents, and Links.** You can quickly see the number of nodes in the currently visible range of your data set, the current number of linkevents in the network, and the number of relationships, i.e. links, between nodes. A relationship between two nodes emerges when the two nodes share a linkevent. For example, if an actor 1 is node 1 and actor 2 is node 2 and actor 1 sends a message to actor 2 (which is the linkevent), then a link of strength 1 results.

**5. Number of self-links.** This number indicates the relations which occur, because one node referenced to itself (e.g. by sending a linkevent to itself). Thus, the node created a circular relationship. These relationships have to be considered when metrics of the networks are interpreted. In the metrics they are kept separated.

**6. Diameter.** The diameter is indicating the size of the network. Of the many shortest paths between each two nodes of the network, the diameter is the longest (shortest) path, which can be found between any two nodes of the network. Such a diameter can indicate how many steps and with what probability an information or a change will move to a node on the opposite side of the network.

**7. Isolated Nodes.** Isolated nodes have no relationships to other nodes in the network. They can have triggered a linkevent without a receiving node (e.g. a self-link or a simple posting without a reference). The number of isolated nodes indicates the connectedness of the relationship network, the likelihood to receive answers on requests and the integration of new nodes.

**8. Number of Core Group Members.** To determine this core group, the most active nodes in a network are ranked. Then the top nodes are added to the core group until this group accounts for 80 percent of the overall network activity (i.e. link events initiated). If this number is set into relation with the total number of nodes, this results in the core group's share (in percent). It is defined as the size of the core group divided by the total number of available nodes in the network. This measure shows, if the core group is either very small compared to the rest, showing that there is a 'tall peak' in the networking activity with only a small but very active nucleus, or if the network is more evenly distributed in its activity. Core group members can also be visualized as node colors or sizes to indicate the most active nodes in a large network, the nucleus of communication.

**9. Average Link Strength.** This measure is calculated by counting and averaging the numbers of linkevents (e.g. messages from node 1 to node 2) for each relationship link between two nodes. High average relationship strength indicates that the nodes of the network have engaged in many linking events and have to some extent already developed a form of a tight connection.

**10. Density.** Density is defined as the actually occurring relationships in a network divided by the maximum possible relationships. It shows how tightly knit the network is, and how many direct relationships between nodes are actually formed (compared to only indirect relationships). Increased density indicates that nodes establish more direct connections to each other, thus increasing their own connectedness to others, which in turn makes a smooth flow of information more likely. It should be noted, that density values are often rather low.

**11. Average Degree Centrality, Betweenness, and Closeness.** Degree Centrality is a simple centrality measure, counting the relative share of contacts of a node in the network. If it is linked to 50% of all nodes then degree centrality is 50%. Betweenness Centrality is a measure of network control. It is computed as the number of shortest paths between pairs of nodes, which run through the observed node. In an e-mail network this could be the person who forwards important messages and thus is important for the information transfer between pairs of actors. This can be an important network position but is also critical for information transfer in a communication setting. Closeness centrality quantifies the distance of a node to all other nodes in the network measured with average shortest path length. This measure indicates how fast or efficient a node could "access" the network and how likely it is, that information or change reaches the node. For the network the nodes' average of these three measures is computed.

**12. Average Connectivity.** This measure computes the average node connectivity. Node connectivity is an indicator for the number of paths in the network contributed by a certain node. It's  $1 - \frac{\text{connectivity without node } i}{\text{connectivity with node } i - \text{direct contacts of } i}$  or short:  $[1 - (\text{connectivity without } i) / (\text{connectivity with } i - \text{direct contacts of } i)]$ . The resulting percentage score is high, if the node existence yields many paths which would not be existing without that node. The node is thus contributing to the connectivity of the graph. The average of that number can be interpreted as an average importance of a node.

**13. Average Reach.** Each node reaches a percentage of other nodes through its connections. In a large cluster, usually nodes reach all other nodes, if the network is completely connected and no islands exist. The average then computes how many nodes can be reached by a node

**14. Average Path Length.** This measure computes the length of all paths found in the network and then derives an average of that length. A low value of this measure indicates that all node can interact with each other via short routes. Typically it should have a range from 1 to 6 (six degrees of separation)

**15. Clustering Coefficient.** This measure is an indicator for the forming of direct relationships in the neighborhood of nodes and is a triadic measure. It looks for all closed triangles of three nodes that can be build around a focused node (the ego node and two other of its contacts). These triangles are related to all occurring combinations of three nodes.

### 7.1.2 Node Statistics

Author properties are only visible if an author has been selected. This is done by clicking left mouse button over a node in the visualization window on the right hand side. After successful selection, the node is marked red and a list is displayed in the Node Info & Statistics box on the "Stats" panel. If another node shall be selected, simply click on it. If you want to deselect all nodes, click in the empty space between nodes on the visualization panel.

The node info and statistics box displays the following information.

**1. Index.** Every node has an index number given by Commetrix when the datafile is being created.

**2. Group.** This number is showing the group of nodes, to which the current node is associated. If the network is a large completely connected component, there will be only one group (usually with number 0). If the network visualization displayed consists of several isolated groups, i.e. different unconnected subnetworks. The group number is documented with the export. It further is kept current, when you filter or cluster the network (i.e. for keywords, time durations, node properties, link properties). So if you disaggregate the network by filtering you will be able to detect groups of nodes in your exported tables (which have no visualization) simply by looking at the group number. Please note, that this group number is not always starting with 0 and continuing with 1,2,3,... rather, Commetrix associates a new index number to each group as they appear in your network exploration. You will at all times have different group numbers but not in a sequential form.

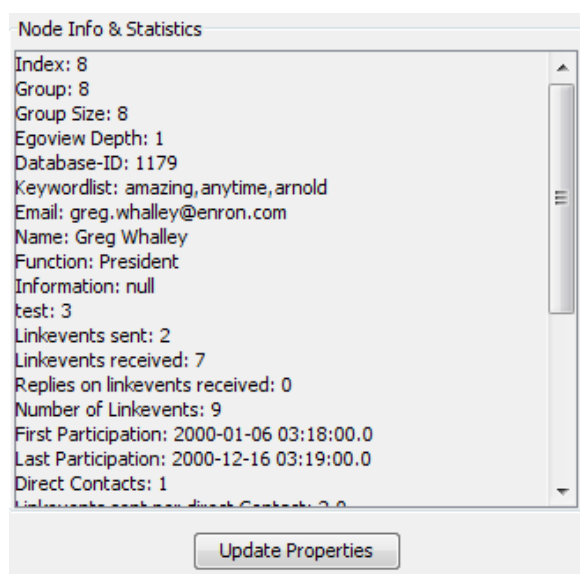


Figure 15: Node information and statistics.

For better visibility you can go to the “Configuration” tab and select “Group” as node color. This will color each group with a different signal color.

**3. Group Size.** This indicator computes the size of the group of the currently selected node. You can see to how many other nodes the selected nodes is currently connected by reducing this number by 1, as you need to exclude the selected node from the group to learn about the number of this node’s contacts in this group.

**4. Egoview Depth.** If you use the ego filter in the “Subset”/”Egoview” tab, you select one or sometimes more nodes as your focused nodes (called egonodes). If this filtering mode is active (a checked box), the value of the metric “Egoview Depth” will show, if you currently have selected an ego node (i.e. selected in the list on your “Egoview” subtab). In this case, the value will be 0 – indicating the distance of the node to an egonode. If you select a node and Egoview Depth computes “1”, you know that this selected node is a direct neighbor of your egonode(s). If it’s a “2”, then the selected node is connected to an egonode by a path length (“depth”) of two steps. Thus the number indicates the distance to (the closest of) your defined egonode(s).

**5. Database-ID.** Every node has an entry in the database with an according Database-ID. It is an internal number of the dataset and can be used to directly identify a node using its ID, or to anonymize it in the visualization.

**6. Keywordlist.** This measure shows the three most frequent keywords of a node (after filtering out some irrelevant terms during dataset production). It is a very rough representation of the contents of the messages associated with that node. This indicator remains empty if no node-related keywords are computed during the production process of your dataset (which is an optional step). If you have defined content descriptors as a very selected set of keywords for your linkevents, those which are most frequently associated with your node should appear here.

If you have changed your filtering (e.g. the time window, relationship strength, etc.), you will need to update the computation of these content descriptors using the update button in the icon bar.

**15. Defined Node Attributes.** The following entries (usually five) display the node attributes that have been assigned as details of your dataset during the dataset modeling and production process. Often, there will be two attributes for names and three for variables like node types or other properties. In the example shown in the figure, an e-mail address has been stored as a unique name identifier, followed by a real name (where for common names, one name can appear more than once, this could as well be a namecode or a nickname). Then follow the assigned dataset variables function (to define the node’s organizational function), information (obviously to store some other description item about a node, e.g. country of origin), and test (which stores some variable “3” for the user selected in the above figure).

**16. Linkevents sent.** This entry shows the number of linkevents initiated by the selected node. For example in a discourse, this would be messages sent.

**17. Linkevents received.** This entry shows the number of linkevents received by the selected node. For example in a discourse, this would be messages received.

**17. Number of linkevents.** This entry shows the number of linkevents associated with the selected node.

**18. First and Last Participation.** Lists the date of the first occurrence of the node in the currently active data and the last occurrence. Comparing this date with the beginning of the sampling period shows, when the selected author entered the discussion.

**19. Direct contacts.** This measure shows the number of other nodes, with which the selected node is in direct contact via linkevents.

**20. Linkevents sent per direct contact.** Shows the average linkevents sent by a node to its contacts. This is a proxy for the average relationship strength of the selected node. Some will have many but weak contacts, some have a few intense contacts. But note, that this measure shows only the average.

**21. Pulsetaker.** A pulsetaker is a node with few direct contacts, but those few direct contacts have again many contacts themselves. This results in a preferable network position: With low effort (interacting with only few other nodes) the pulsetaker can access many indirect contacts and thus might be aware of much information in the network. The pulsetaker value is the ratio of indirect relationships divided by the number of direct relationships. The higher the value the more the network position is a pulsetaker.

**22. Core Group Member.** This indicator is set 'yes' if the selected node is belonging to the core group of the network. To determine this core group, the nodes with most (linkevent sending) activity in an expert network are ranked. Then, in descending rank order, they are added to the core group until this group accounts for 80 percent of the network traffic. All nodes which are in this group are in the core group of active nodes, the others are in the less active non-core group.

**23. Closeness Centrality, Betweenness Centrality, and Degree Centrality.** Degree Centrality is a simple centrality measure, counting the relative share of contacts of the selected node in the network. For example, if the selected node is linked to 50% of all other nodes then its degree centrality is 50%. Betweenness Centrality is a measure of network control. It is computed as the number of shortest paths between pairs of nodes, which run through the selected node. In an e-mail network this could be the person who forwards important messages and thus is important for the information transfer between pairs of actors. This can be an important network position but is also critical for information transfer in a communication setting. Closeness centrality quantifies the distance of a node to all other nodes in the network measured with average shortest path length. This measure indicates how fast or efficient a node could “access” the network and how likely it is, that information or change reaches the node. For the network the nodes’ average of these three measures is computed.

**24. Reach.** Each node reaches a percentage of other nodes through its connections. In a large cluster, usually nodes reach all other nodes, if the network is completely connected and no islands exist. The reach then computes how many nodes of the network can be reached by a node (as a percentage score).

**25. Brokering Activity.** This measure indicates how many new paths (i.e. connections) between other nodes of the network have been created or shortened by the activities of the observed node. It is computed by eliminating the node and computing the increase (i.e. deterioration) in path lengths.

**26. Originating Brokering Activity.** This measure is a more restrictive version of brokering activity. Only new connections (paths) between other nodes that have been created through the activity of the selected nodes.

**27. Brokering Impact.** If Brokering activity counts the number of created and shortened paths as an impact measure of the selected node, Brokering Impact indicates the extend of shortened paths in the network as a percentage score. Large reductions in other nodes’ path length are recognized.

**28. Local Clustering Coefficient.** This measure is an indicator for the forming of direct relationships in the neighborhood of the selected node and is a triadic measure. It looks for all closed triangles of three nodes that exist around the selected node (the ego node and two other of its contacts). These triangles are related to all occurring combinations of three nodes (with the selected node and two of its neighbors).

## 7.2 Charting Metrics

Next to observing statistics in the “Stats” panel, Commetrix offers a simple and quick charting tool accessible via the “Charts” tab (more extensive charts can easily be created using the export functionality).

**1. Time Interval Slider.** The first element of this tab is the slider to define the number of time intervals for computing metrics. This slider defines the number of sections into which the overall time period is disaggregated and for which the selected measures shall be computed. “15” means, the user wants his overall network dataset period to be separated into 15 partial time periods (time intervals), for which measures are computed and charted.

**2. Measurement Panel.** The panel contains a reduced list of metrics: (1) ATBMS indicates the average time between linkevents as an indicator for the average time lag between activities. (2) Degree Distribution is a histogram showing how many nodes have which degree (in descending order of degree). This can indicate how centralized the network is. It is centralized if only very few nodes have very many contacts whereas a very large number has only a very small degree (the periphery). (3) Density is the existing relationships in the network versus the theoretically possible relationships. (4) Diameter is the longest shortest path between two nodes of the network. (5) Nodes simply is a node count. (6) Isolated Nodes shows how many isolated nodes existed in each measured time period. (7) Link Strength indicates the average number of linkevents per link. (8) Reciprocity shows the number of reciprocal links, i.e. links where there is a two way connection between the two nodes of the dyad.

**3) A,T,I Computation Type.** (A) “A” stands for Average. Marking this box will create an average of the respective measure. This average is shown in a speedometer visualization. (T) “T” stands for Trend Analysis. Checking this box yields a line graph with metric on the y-axis and time on the x-axis. The trend is computed as a floating average which takes the current and all past time periods into account. (I) “I” stands for interval. Checking this box yields a line graph with metric on the y-axis and time on the x-axis. The values are computed

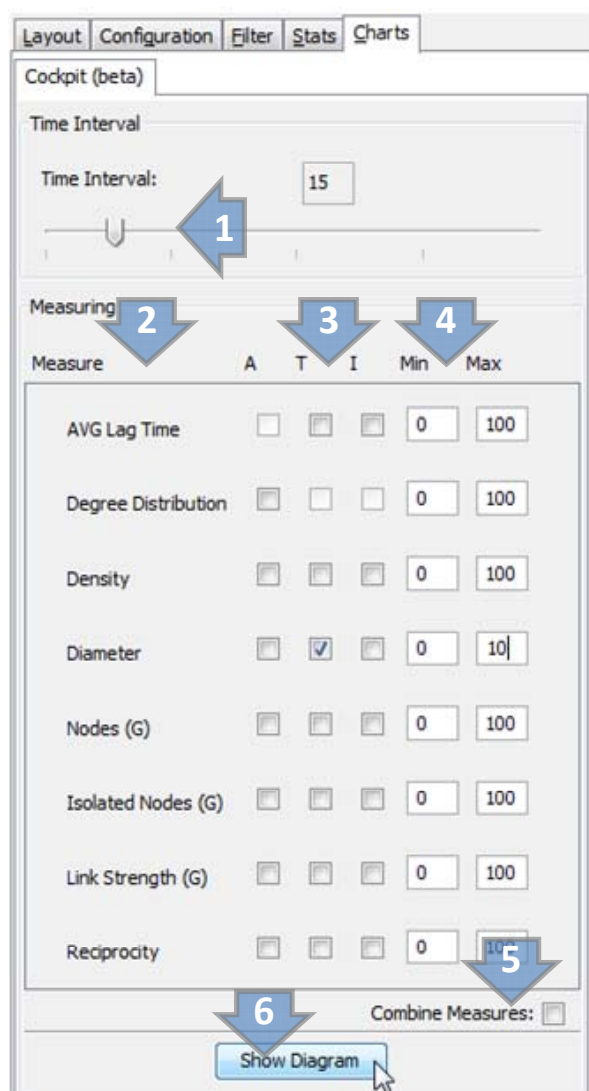


Figure 16: Chart Tab. (1) Defining the time interval granularity for the time chart. (2) Measures. (3) Chart Type (Average, Trend, Interval). (4) Min and Max of Y-Axis. (5) Combine Measures in one chart. (6) Compute and show the selected diagrams.

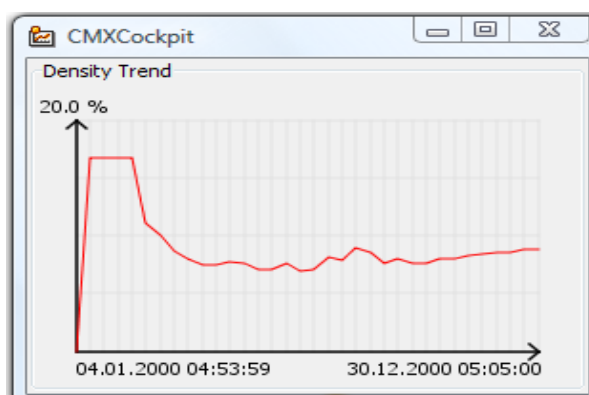


Figure 17: Example Chart showing the trend of the measure density over time.

separately for each time period segment. It will be more erratic over the periods but contains a concrete value corresponding to the network status in the respective time period.

**4) Min-Max.** Here you can define the range of the Y-axis. Per default the min is always set to 0 and the max is set to 100. Feel free to change these values by clicking in the box and entering another range min or max. This will directly affect the readability of the graph.

**5) Combine Measures.** If you check this box, the selected measurement results will be integrated into one graph. This can improve comparability between two measures

**6) Show Diagram.** This button is creating a pop-up window, which contains the diagrams you have selected. The window is closed again by pressing the X in the top right corner of the pop-up window.

Please note that the diagrams are not automatically updated. Close them and re-open your charts to create a current version after you have changed filters. The charts are only intended as a quick reference and may be extended in future versions. However, we recommend to use the exported data (e.g. as Excel tables) and the much more comprehensive charting facilities of spreadsheet tools like Excel.

## 7.3 Looking at original data tables

Via the Menu Item “Tables” you can access several data tables to complement your visual exploration. You can click on the column heading to sort the data in descending or ascending order. You further can use copy and paste to export this table into excel or other spreadsheet tools. You can close this window by clicking on the X in the top right corner. The tables correspond to the complete dataset and not to the currently selected subsection. This data is available through the export facility.

**1. Node Table.** This table shows all nodes and their stored attributes. These attributes correspond to the entries in the node label selection box (in the configuration tab) and with the node properties shown in the node info box (in the stats tab).

**2. Link Table.** This tables shows all links between each two nodes together with their number of linkevents (in affiliation networks also number of affiliated nodes), the first linkevent, the last linkevent, how much is sent by node A and by node B, and if this link is reciprocal. Some other values might also have been added according to the individual definition of the dataset.

**3. Linkevent Table.** This table shows all linkevents recorded in the dataset with their ID, user-defined variables, the timestamp, and contents (if contents have been stored).

**4. Keyword Table, Subject Keyword Table, and Content Keyword Table.** This table only contains entries if you have coded contents for your dataset. In this case, all keywords are listed together with their average frequency of occurrence per linkevent, the number of sent linkevents that contained the keyword, the number of send and received linkevents that contained the keyword, and the number of nodes that have used this keyword in their linkevents. Sorting the columns yields a toplist. The list can be separated into subjects and contents, if such a differentiation has been coded into the dataset.

**5. Keyword Matrix, Subject Keyword Matrix, and Content Keyword Matrix.** This matrix shows co-occurrence of keywords in the cells. Co-occurrence is when two keywords appeared in the same linkevent. It gives you an impression, which terms often appear near each other and can be used to specify broader terms.

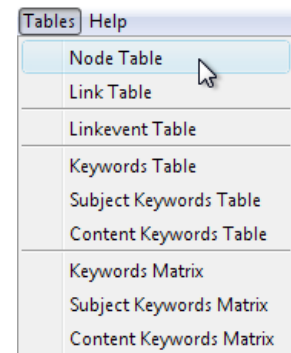


Figure 18: Available data tables.

## 8 Changing Views

The menu “views” gives you access to various different visualization modes. You can select between the default 2D view, a 3D view, and a simulated 3D view. You further can change into fading mode or you can select if certain special label tags are shown.

### 8.1 Different views 2D, sim 3D, 3D

**2D.** The default view activated after starting Commetrix is 2D. We recommend this as the view which will yield the most insight while consuming not much of computation power. All above explanations primarily relate to this view.

**3D.** When you switch to full 3D view, the computer is utilizing a special 3D graphics engine, which is not as comprehensive as our 2D view. However, it gives a very nice impression of an optimum layout as nodes have more room to move freely into a position, which ideally represents their true relationship to all other nodes. (A sphere around a node has more space than a circle for other nodes to move around it.) Most features of 2D mode have been adapted for the 3D engine. You can change the layout, configurations, the coloring, you can filter, etc. It is advisable to toggle node size and link length to create a good initial layout, as the 2D settings do not always apply well for the 3D world.

You now also have an “extra” tab, which allows you to export screenshots of your visualization or even movies. In the 2D mode, you would need special screenshot or screencast programs to do so. There are some available for free in the internet.

You further have special means to control the visualization. Pressing the left mouse button and dragging in any direction on the visualization panel allows rotating the view. Pressing the right mouse button and dragging up and down zooms the view in and out.

**Simulated 3D.** Simulated 3D is interesting, if you do not have enough computation power for the 3D view to run smoothly. The simulated 3D view simply lays out nodes in three dimensions but it does not use a special 3D graphics support to do this. The nodes are displayed on your normal 2D panel but three dimensions are used for positioning nodes. This results in some node sizes being bigger simply because the node is “closer” to you (in the z-axis). However, all remaining features are consistent with the 2D view. So it is easy to quickly switch between 2D and Simulated 3D. The difference between the Simulated 3D and the full 3D visualization is, that the latter computes light sources so that the network appears to consist of spheres not circles.

Note the “extra” tab in the control panel. It gives you three sliders, which you can use for rotating the view.

### 8.2 Toggling Fading Mode

This is a little but very useful option. Turning on fade mode will show all nodes which have been deactivated by your filtering (keywords, link strength, etc.) are shown in a semi-transparent light grey. This gives you an impression of the embeddedness of your filtered subset in the overall network. For example, you load a dataset, then you use the egoview filter to focus on a specific node and its direct contacts (ego depth = 1). Normally, you would now only see this egonet. If you switch on fading mode now, you show all excluded nodes (i.e. not belonging to the focused subset, the egonet) in light grey and semi-transparent. You can now see if this node and its contacts are located at the periphery or at the center of the overall network. Similarly, you can turn on fading mode, enter a keyword as a search term, and find out, where in the overall network the keyword exists.

## 8.3 Toggling Label Mode

**1. Activate ego label mode.** This option changes the way labels are presented in the graph. Switching it on and you will only see labels when you move the mouse over a certain node. This is useful for large networks, where labels would otherwise obstruct the view. In this mode you can observe interesting structures and then explore the node label by moving the mouse over interesting nodes.

Note, that there are more options in the “Preferences” menu, that can be opened via the menu entry “Help”/”Preferences”. In the preferences window go to the “Visualization” tab. There you can chose if inactive nodes shall also be labeled when you move over them with your mouse (i.e. the grey semi-transparent nodes, that are actually excluded after you filtered out a subnetwork, but which you can bring back as “ghost” nodes to the visualization by switching to “Fading Mode”). Further you can choose whether the direct links of the node touched with the mouse pointer shall also be labeled. This is very interesting for tracing direct connections of a selected node in a larger network. A third option will yield even more visibility for the selected nodes contacts. Turn on increased focused node size, and all labeled contacts with further be increased in their sizes. You need to confirm your changes in the preferences with the “Apply” button. If you want to leave the window, you can also apply and exit with the “OK” button.

**2. Activate Ego Legend Mode.** Switching this mode on will extend the labels to show all details for the nodes as a label, e.g. not only one entry, but a long list of all available data. This is obviously only applicable if you have the ego label mode on, otherwise the long labels will quickly obstruct your view.

## 9 Dyad and Triad Analysis

Commetrix Analyzer enables you to easily count dyads and triads in your dataset. The analysis is reached via Tables/Dyad and Triad Tables. After the computation (it may take some time to compute depending in the size of your dataset), you can switch between a triad and a dyad overview (see Figure). The tables also give hints how to interpret the different configurations.

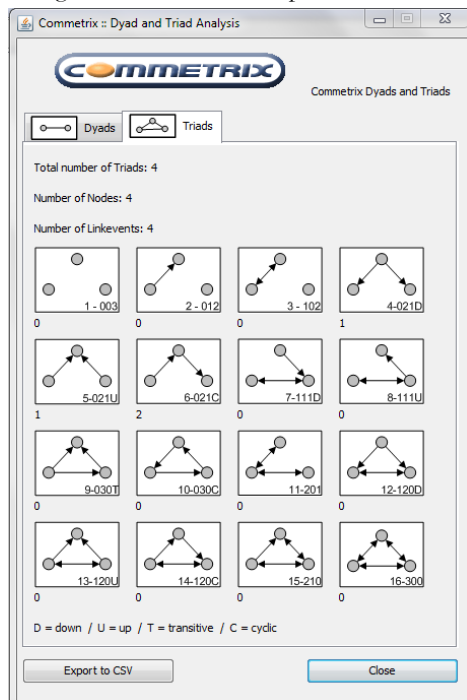


Figure 19: Example Triad Table for a simple network.

## 10 Exporting Statistics to Tables (.csv,.xcl)

**1. Activate Export.** Before you can export your current view into data tables, you need to activate this mode by clicking “Activate Export” in the “File” Menu.

Note, that Commetrix will ask for a destination folder for your export files. If you have not yet defined one, you need to go to the “Help” menu and select “Preferences”. In the preferences window, click on the “Export” tab and chose a destination folder by clicking on the button with the three dots (...). This will open a directory browser dialog, which you can use to point to your favorite location. Then click on “Open”. The directory location will be taken as the export destination. Click “OK” to apply and close the preferences tab again. Click now on “File” and “Activate Export” again until it is in active mode. This mode is signaled by the export icon in the icon bar. Normally it is turned off, now it is colored and active.

**2. Execute Export of current view.** At any time simply click on the export button in the icon bar to trigger an export. In the background your current network subset will be stored in two folders that are created in your destination. After the successful export, these files will be shown in a pop-up window. Click “OK” to proceed.

You can now switch to a spreadsheet tool like Excel and look at your files. These files are stored in a folder that has the name of your dataset. Within this folder there is now a “Nodes” and a “Links” folder. In both folders, you can see a “.csv” file, which can be opened with many spreadsheet tools, e.g. excel. If you have a large dataset, more than one file will be created. They are in sequential order and a unique time stamp code has been taken as a file name. Everytime you now press export another such file will be created and stored in these folders.

**3. The node export file.** The node export file contains the filter settings of Commetrix, as of when the file was created. Further the network properties as displayed in the “Stats” panel. Then all active nodes’ details and finally all computed metrics for each active node (in relation to the current filter settings, e.g. the selected time period). If some metrics are not computed try to press “Update Properties” button in Commetrix before you export into tables.

**3. The link export file.** The link export file complements the node export file (the filter settings are the same). It shows all active links (as a “nodeID-nodeID” relation code, e.g. 1194-1198 for the link between the nodes with ID 1194 and 1198. Further some link properties (also shown in the link label selection box in the “Configuration” tab) are shown in the table. You can now study every link and compute further aggregated measures or statistics.

## 11 Contact

If our Commetrix Framework Solution is of interest for you, we will be happy to arrange a license contract. Just contact us via: [info@trilexis.com](mailto:info@trilexis.com) (preferred) or via the following address.



Trilexis Information Management GmbH  
 Helmut-Kaeutner-Weg 5,  
 13591 Berlin  
 GERMANY  
 ID: DE 29 039 00437  
[info@trilexis.com](mailto:info@trilexis.com)