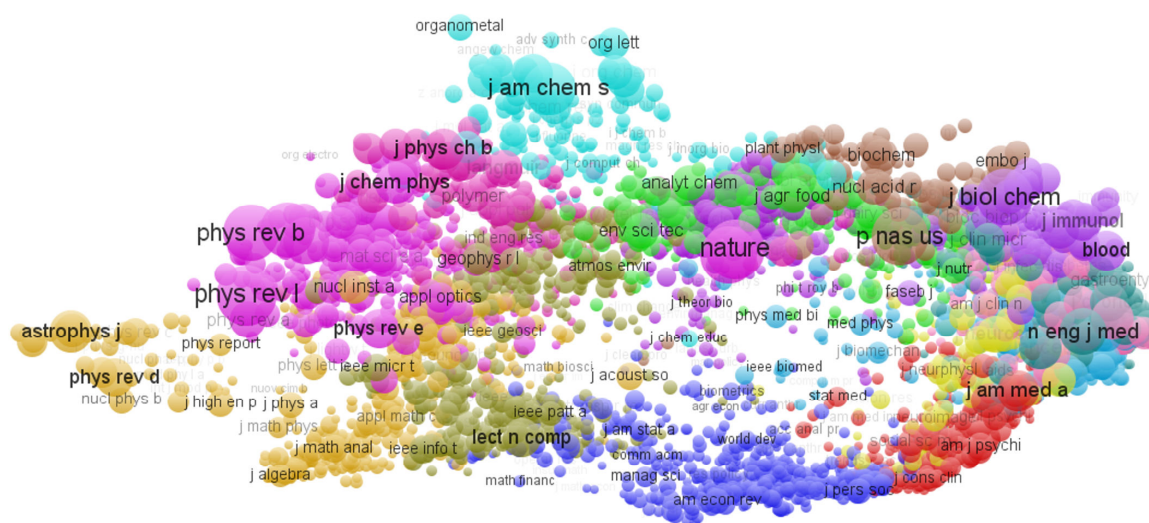


VOSviewer Manual

Nees Jan van Eck and Ludo Waltman

28 May 2012



Manual for VOSviewer version 1.5.0



Universiteit Leiden



Table of contents

1	Introduction	3
2	Functionality	4
2.1	Main panel	5
2.2	Information panel	9
2.3	Overview panel	9
2.4	Action panel	10
2.4.1	Action tab.....	10
2.4.2	Items tab	12
2.4.3	Map tab.....	13
2.4.4	Options tab.....	15
3	File types	18
3.1	Map file.....	18
3.2	Network file.....	19
3.3	Thesaurus file	20
3.4	Cluster colors file and score colors file	21
4	Advanced topics	23
4.1	Using command line parameters	23
4.2	Making a map available on the internet.....	24
4.3	Increasing the availability of memory	25
	References	26

1 Introduction

VOSviewer is a computer program that can be used for the following purposes:

- VOSviewer can be used to create maps based on network data. Maps are created using the VOS mapping technique and the VOS clustering technique.¹
- VOSviewer can be used to view and explore maps. It can show a map in various different ways, each emphasizing a different aspect of the map. It offers functionality such as zooming, scrolling, and searching, which facilitates the detailed examination of a map. The viewing capabilities of VOSviewer are especially useful when working with maps containing at least a moderately large number of items (e.g., at least 100 or 200 items).

VOSviewer is primarily intended to be used for analyzing bibliometric networks. The program can for example be used to create maps of publications, authors, or journals based on a citation, co-citation, or bibliographic coupling network or to create maps of keywords based on a co-occurrence network. However, the use of VOSviewer is not restricted to bibliometric networks. In fact, VOSviewer can be used to create maps based on any type of network.

VOSviewer has been written in the Java programming language, which means that it runs on most hardware and operating system platforms. VOSviewer can be obtained from the web site www.vosviewer.com. The program can be used freely for any purpose.

This manual is concerned with version 1.5.0 of VOSviewer. We explain the functionality of VOSviewer in Chapter 2, and we discuss the file types used by VOSviewer in Chapter 3. Some advanced topics are covered in Chapter 4. For additional information about VOSviewer, we refer to a paper that we have written (Van Eck & Waltman, 2010). In this paper, a general introduction to VOSviewer is provided. Also, the technical implementation of specific parts of the program is discussed in considerable detail.

¹ Together, these two techniques provide a unified framework for mapping and clustering. For more information about the techniques, we refer to Van Eck, Waltman, Dekker, and Van den Berg (2010) and Waltman, Van Eck, and Noyons (2010).

To get some hands-on experience with VOSviewer, we encourage the reader to use the map file `journal_map.txt`, which is distributed together with VOSviewer. This file can be used to reproduce the figures in this chapter. The file contains a map of 232 journals in the fields of economics, management, and operations research (for more details, see Van Eck & Waltman, 2010). To open the map, press the **Open** button on the **Action** tab in the action panel, select the map file `journal_map.txt`, and press the **OK** button.

2.1 Main panel

As can be seen in Figure 1, the main panel of VOSviewer is used to show a selected area of the currently active map. The zoom and scroll functionality of VOSviewer can be used to determine which area of the currently active map is shown. Depending on the available map data, the main panel consists of three or four tabs. Each tab provides a different view on the currently active map. We now discuss each of the tabs:

- **Label View.** In the label view, items are indicated by their label and, by default, also by a circle. For each item, the font size of the item's label and the size of the item's circle depend on the weight of the item.² The circle of an item is displayed in the color of the item. If items have been assigned to clusters, the color of an item is determined by the cluster to which the item belongs. If scores have been given to items (using the *score* column in a map file; see Section 3.1), the color of an item is determined by the score of the item, where by default colors range from blue (score of 0) to green (score of 1) to red (score of 2).³ Alternatively, the color of an item may have been specified directly (using the *red*, *green*, and *blue* columns in a map file; see Section 3.1). By default, to avoid overlapping labels, only a subset of all labels is visible. Also, by default, no lines between items are displayed. However, this can be changed on the **Options** tab in the action panel. The label view is particularly useful for

² The weight of an item is determined by the *weight* or *normalized weight* column in a map file (see Section 3.1). When a new map is created without providing a map file with a *weight* or *normalized weight* column, the weight of an item is set equal to the total strength of all links of the item. When an existing map is opened without providing a map file with a *weight* or *normalized weight* column, all items in the map are given the same weight.

³ If scores have been given to items, a color bar is shown in the lower right corner of the main panel. This color bar indicates which colors correspond with which scores. By clicking on the color bar, it is possible to import an alternative color scheme from a score colors file (see Section 3.4).

detailed examinations of a map. An example of the label view is shown in Figure 2.

- **Density View.** In the density view, items are indicated by their label in a similar way as in the label view. Each point in a map has a color that depends on the density of items at that point. By default, this color is somewhere in between red and blue. The larger the number of items in the neighborhood of a point and the higher the weights of the neighboring items, the closer the color of the point is to red. Conversely, the smaller the number of items in the neighborhood of a point and the lower the weights of the neighboring items, the closer the color of the point is to blue. The density view is particularly useful to get a quick overview of the important areas of a map. We refer to Van Eck and Waltman (2010) for a detailed discussion of the technical implementation of the density view. An example of the density view is shown in Figure 3.
- **Cluster Density View.** The cluster density view is available only if items have been assigned to clusters. The cluster density view is similar to the ordinary density view except that the density of items is displayed separately for each cluster of items. In the cluster density view, the color of a point in a map is close to the color of a certain cluster if there are a large number of items belonging to that cluster in the neighborhood of the point. Like in the ordinary density view, items with high weights count more heavily than items with low weights. The cluster density view is particularly useful to get a quick overview, for each cluster separately, of the important areas of a map. We refer to Van Eck and Waltman (2010) for a detailed description of the technical implementation of the cluster density view. An example of the cluster density view is shown in Figure 4.
- **Scatter View.** In the scatter view, items are indicated by a small circle. Like in the label view, the circle of an item is displayed in the color of the item. No labels are displayed in the scatter view. By default, lines between items are not displayed either, but this can be changed on the **Options** tab in the action panel. The scatter view can be useful to get an overview of the general structure of a map. The way in which a map is shown in the scatter view is similar to the way in which a map is shown in the overview panel discussed below. An example of the scatter view is shown in Figure 5.

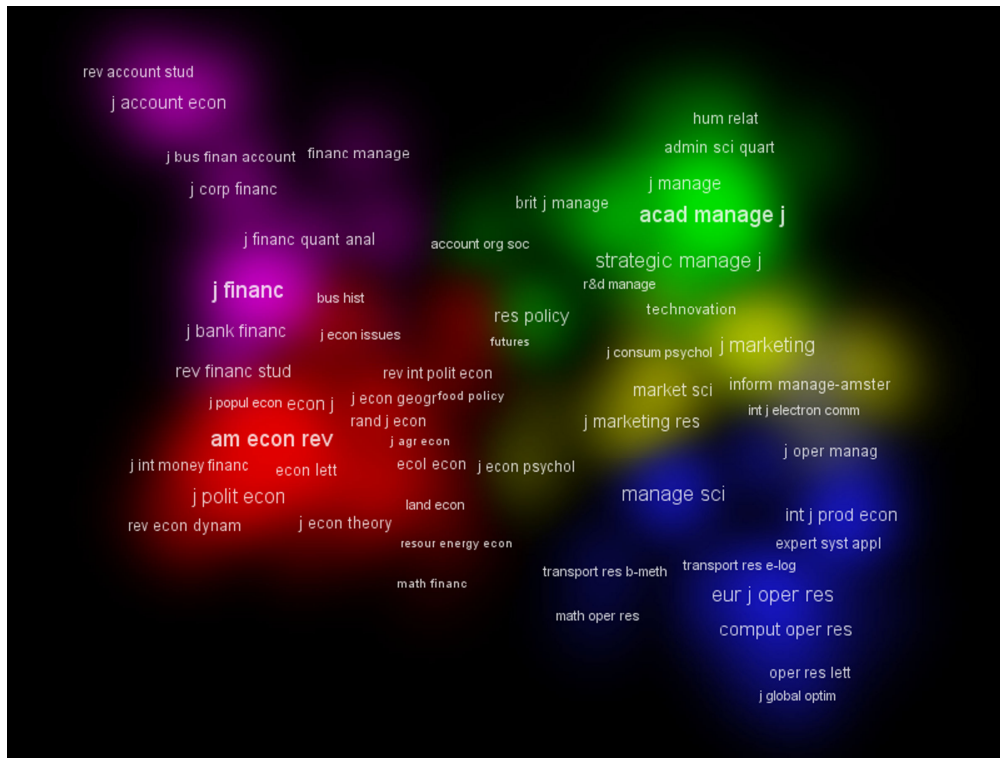


Figure 4. The cluster density view.

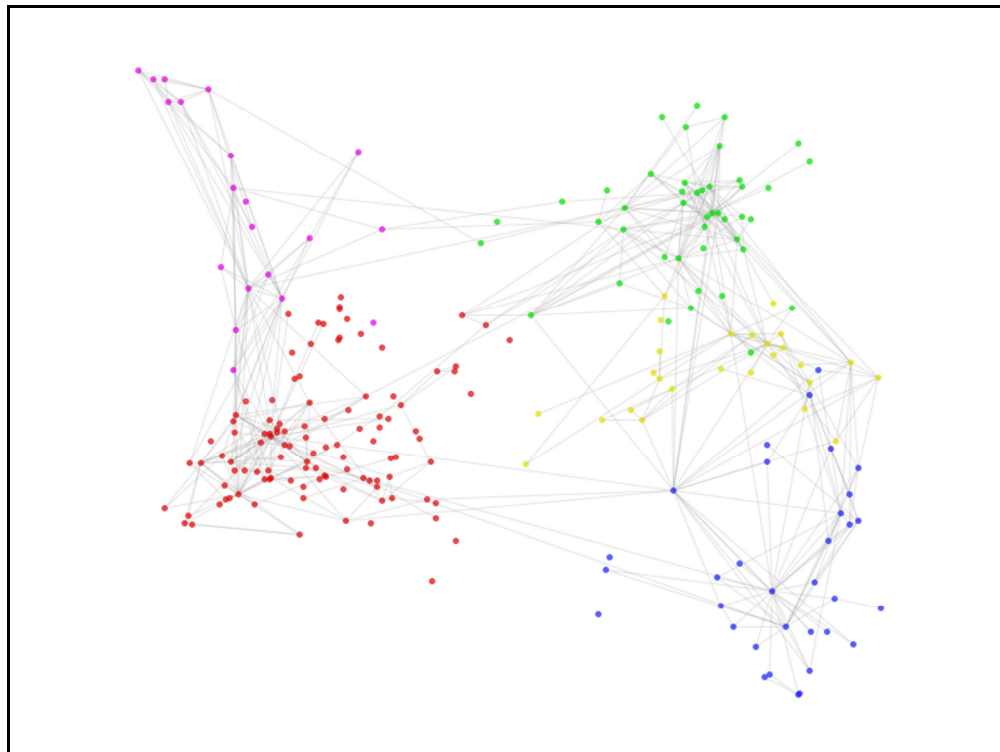


Figure 5. The scatter view.

To facilitate the detailed examination of a map, VOSviewer offers zoom and scroll functionality. In the main panel, zooming and scrolling can be done in the following three ways:

- *Using the mouse.* To zoom in, move the mouse upwards while keeping the right mouse button pressed. Conversely, to zoom out, move the mouse downwards while keeping the right mouse button pressed. As an alternative, the mouse wheel can be used to zoom in and out. To scroll through a map, move the mouse while keeping the left mouse button pressed.
- *Using the navigation buttons in the upper right corner of the main panel (see Figure 1).* Use the plus and minus buttons to zoom in and out. Use the arrow buttons to scroll through a map.
- *Using the keyboard.* Use the plus and minus keys to zoom in and out. Use the arrow keys to scroll through a map.

2.2 Information panel

In the information panel of VOSviewer, information about an item in the currently active map can be shown. For example, when one moves the mouse cursor over the label of an item in the main panel, information about the item will be shown in the information panel. If item descriptions have been provided (using the *description* column in a map file; see Section 3.1), the description of the item will be shown. If no item descriptions have been provided, the label of the item will be shown along with the number of the cluster to which the item belongs (assuming that items have been assigned to clusters).

Information about a link between two items can also be shown in the information panel. This is possible only if lines are displayed in the main panel. When one moves the mouse cursor over the line between two items, information about the link between the items will be shown in the information panel. The labels of the items and the strength of the link will be shown.

The height of the information panel can be adjusted. This can be done by dragging the area between the information panel and the main panel using the mouse.

2.3 Overview panel

In the overview panel of VOSviewer, an overview of the currently active map is shown. Each item in the map is indicated by a small colored dot. A rectangular frame is displayed in the overview panel to indicate the area of the currently active

map that is shown in the main panel. By left-clicking in the overview panel, the area of the currently active map that is shown in the main panel can be changed.

2.4 Action panel

The action panel of VOSviewer can be used to undertake all kinds of actions. The panel consists of four tabs: The **Action** tab, the **Items** tab, the **Map** tab, and the **Options** tab. We now discuss each of these tabs.

2.4.1 Action tab

The **Action** tab can be used to perform a number of basic actions. The following buttons are available on the **Action** tab:

- **Map.**
 - **Create.** Use this button to create a new map. The button brings up the **Create Map** wizard. There are two ways in which a new map can be created using this wizard:
 - **Create a map based on a network.** This option requires the adjacency matrix of a network. The adjacency matrix indicates which pairs of items in the network are linked to each other, and for each pair of linked items it indicates the strength of their link. The matrix can be read from a network file. In addition to a network file, one may also provide a map file. The map file may for example contain labels and descriptions of items. We refer to Sections 3.1 and 3.2 for a detailed discussion of map files and network files. Instead of a map file and a network file, it is also possible to use a Pajek file.⁴ A third possibility is to use a Web of Science file.⁵ The use of Web of Science files allows one to easily create maps of scientific publications, scientific journals, researchers, or research organizations based on either co-citation relations (i.e., multiple items being cited by the same publication) or bibliographic coupling relations (i.e., multiple items citing the same publication).

⁴ Pajek is a well-known computer program for social network analysis. It is available at <http://pajek.imfm.si/doku.php>.

⁵ Web of Science is a major bibliographic database of scientific publications produced by Thomson Reuters. It contains bibliographic data on publications from more than 10,000 journals in the sciences, social sciences, arts, and humanities. Web of Science is available at <http://webofknowledge.com>. A subscription is required. We note that VOSviewer supports only tab-delimited Web of Science files.

- **Create a map based on a text corpus.**⁶ This option requires a text corpus. The corpus is stored in a corpus file. A corpus file is a text file that contains on each line a piece of text. This piece of text can be an entire document or some part of a document, for example a section, a paragraph, or a sentence. All texts are assumed to be in English. Using natural language processing techniques, VOSviewer extracts terms from the corpus file, where a term is defined as a sequence of nouns and adjectives (ending with a noun). Based on the extracted terms, VOSviewer creates a term map. This is a map in which terms are located in such a way that the distance between two terms provides an indication of the number of co-occurrences of the terms. In general, the smaller the distance between two terms, the larger the number of co-occurrences of the terms. Two terms are said to co-occur if they both occur on the same line of the corpus file. It is also possible to create a term map based on a Web of Science file instead of a corpus file. In that case, terms are extracted from the titles and abstracts of scientific publications.
- **Open.** Use this button to open an existing map. The button brings up the **Open Map** dialog box. Map data can be read from a map file. The map file must contain coordinates of items. The file may also contain, for example, labels, descriptions, weights, and cluster numbers of items. In addition to a map file, one may also provide a network file. The network file contains the adjacency matrix of a network. Based on this matrix, lines can be displayed between the items in a map. Instead of a map file and a network file, it is also possible to use a Pajek file. For a detailed discussion of map files and network files, we refer to Sections 3.1 and 3.2.
- **Save.** This split button offers three options:
 - **Save map.** This option is the default choice. Choose this option to save the currently active map in a map file. Map files are discussed in Section 3.1. Instead of a map file, it is also possible to use a Pajek file.
 - **Save network.** This option can be chosen only if the adjacency matrix of a network is available. Choose this option to save the adjacency matrix in a network file. Network files are discussed in Section 3.2. Instead of a network file, it is also possible to use a Pajek file.

⁶ We refer to Van Eck and Waltman (2011) for more information about this option.

- **Save normalized network.** This option can be chosen only if the adjacency matrix of a network is available. Choose this option to save the normalized adjacency matrix in a network file. Normalization of an adjacency matrix is discussed in Subsection 2.4.3. Network files are discussed in Section 3.2. Instead of a network file, it is also possible to use a Pajek file.
- **Print.** Use this button to print a screenshot of the main panel.
- **Screenshot.** This split button offers two options:
 - **Save to file.** This option is the default choice. Choose this option to save a screenshot of the main panel. The screenshot resembles the main panel as closely as possible. However, the navigation buttons in the upper right corner of the main panel are not shown in the screenshot. If the **Optimized screenshots** check box on the **Options** tab is checked, the number of labels visible in the screenshot is optimized. This means that some labels not visible in the main panel may be visible in the screenshot. Screenshots can be saved in a number of graphic file formats. For most applications, we recommend the PNG format. Some formats, such as EPS, PDF, and SVG, use vector graphics when saving a screenshot. This has the advantage that the screenshot can be resized without loss of quality. We note that Figures 2 to 5 were obtained using the screenshot functionality of VOSviewer.
 - **Copy to clipboard.** Choose this option to copy a screenshot of the main panel to the clipboard. The screenshot can for example be pasted into a Word document or a PowerPoint presentation.
- **Info.**
 - **VOSviewer Manual.** Use this button to open the VOSviewer manual.
 - **About VOSviewer.** Use this button to view some general information about VOSviewer. The button brings up the **About VOSviewer** dialog box. This dialog box shows the version number of VOSviewer, a copyright notice, a license text, the address of the VOSviewer web site, and a list of software libraries used by VOSviewer.

2.4.2 Items tab

The **Items** tab provides a list of items in the currently active map. By default, a list of all items in the map is provided. However, a filter can be used to restrict the list to a subset of the items in the map. To do so, enter a filter string in the **Filter** text

box and press the **Filter** button. This yields a list of all items whose label contains the filter string.

The **Group items by cluster** check box determines the way in which items are listed. If the check box is unchecked, items are simply listed alphabetically. If the check box is checked, items are first grouped by cluster and then listed alphabetically within each cluster.

The **Show** button shows the currently selected item on the **Items** tab in the main panel. Instead of using the **Show** button, it is also possible to double-click on an item.

2.4.3 Map tab

The **Map** tab can be used to construct or transform a map. Maps are constructed using the VOS mapping technique and the VOS clustering technique.⁷ The **Map** tab can also be used to change the parameters of these techniques. The following parameters and controls are available on the **Map** tab:

- **Normalization parameters.**
 - **Method.** This parameter determines the method that is used to normalize the adjacency matrix of a network. There are three options:
 - **No normalization.** Using this option, no normalization is performed. In general, we discourage this option.
 - **Normalization method 1.** This is the default normalization method. This method uses the so-called association strength measure (Van Eck & Waltman, 2009).
 - **Normalization method 2.** This is an alternative normalization method. The difference between this method and the default method is rather technical, and we will therefore not discuss it here.
 - **Ignore self links.** This parameter determines whether self links in a network are ignored. In other words, the parameter determines whether the main diagonal of the adjacency matrix of a network is ignored. If self links are ignored, the elements on the main diagonal of the adjacency matrix are set equal to zero. In general, we strongly recommend to ignore self links.
- **Mapping parameters.**

⁷ We refer to Van Eck et al. (2010) and Waltman et al. (2010) for more information about these techniques.

- **Random starts.** This parameter determines the number of times the optimization algorithm of the VOS mapping technique is run. Each time the optimization algorithm is run, a somewhat different mapping solution will be obtained. The best mapping solution obtained in all runs of the optimization algorithm will be used as the final mapping solution. The larger the value of the **Random starts** parameter, the more accurate the final mapping solution that will be obtained.
- **Convergence.** This parameter determines the convergence criterion used by the optimization algorithm of the VOS mapping technique. The parameter must have a value between 0 and 0.01. The closer the value of the parameter is to zero, the more accurate the mapping solution that will be obtained. In general, the default value of the parameter works well and does not need to be changed.
- **Max. iterations.** This parameter determines the maximum number of iterations performed by the optimization algorithm of the VOS mapping technique. The larger the value of the parameter, the more accurate the mapping solution that will be obtained. In general, the default value of the parameter works well and does not need to be changed.
- **Clustering parameters.**
 - **Resolution.** This is a very important parameter that determines the level of detail provided by the VOS clustering technique. The parameter must have a non-negative value. The larger the value of the parameter, the larger the number of clusters that will be obtained. Our recommendation is to try out different values for the parameter and to use the value that yields the level of detail considered most satisfactory for the particular application at hand.
 - **Random starts.** This parameter determines the number of times the optimization algorithm of the VOS clustering technique is run.⁸ Each time the optimization algorithm is run, a somewhat different clustering solution will be obtained. The best clustering solution obtained in all runs of the optimization algorithm will be used as the final clustering solution. The larger the value of the **Random starts** parameter, the more accurate the final clustering solution that will be obtained.

⁸ The optimization algorithm of the VOS clustering technique is partly inspired by the work of Rotta and Noack (2011). In the terminology of Rotta and Noack, the algorithm uses a combination of multilevel coarsening and multilevel refinement.

- **Run.** The **Run** button is available only after a new map has been created (using the **Create** button on the **Action** tab in the action panel). The button can be used to run the VOS mapping and clustering techniques. The option buttons determine whether both techniques are run or only one of them. The **Random seed** parameter is used to randomly generate initial solutions for the optimization algorithms of the VOS mapping and clustering techniques. A given random seed always yields the same initial solutions. The only exception is random seed 0. This random seed yields different initial solutions each time the VOS mapping and clustering techniques are run.
- **Rotate/flip.**
 - **Rotate.** Use this button to rotate the currently active map. The **Degrees to rotate** parameter determines the number of degrees by which the map will be rotated.
 - **Flip Horizontally.** Use this button to flip the currently active map in horizontal direction.
 - **Flip Vertically.** Use this button to flip the currently active map in vertical direction.

2.4.4 Options tab

The **Options** tab can be used to change various settings of VOSviewer. The following settings can be changed:

- **Labels.**
 - **Size effect.** The higher the weight of an item, the larger the size of the font that is used to display the item's label in the main panel. The **Size effect** slider determines the strength of this effect.
 - **Max. length.** This text box determines the maximum length of a label displayed in the main panel. If the length of a label exceeds the maximum length, the last part of the label is not displayed.
 - **No overlap.** This check box determines whether labels are allowed to overlap each other in the main panel. If overlapping labels are not allowed, it depends on items' weights which labels are displayed and which are not. A label of an item with a higher weight has priority over a label of an item with a lower weight.
 - **Blurred background.** This check box determines whether a so-called blurred background is used in the main panel. Using a blurred background,

labels that are not displayed in the ordinary way (due to overlap with other labels) are displayed in a blurred way. The **Blurred background** check box has an effect only if the **No overlap** check box is checked.

- **Gradual appearance.** This check box determines whether in the main panel labels appear gradually or instantaneously when one zooms in. The **Gradual appearance** check box has an effect only if the **No overlap** check box is checked.
- **Optimized screenshots.** This check box determines whether the number of labels visible in a screenshot of the main panel is optimized. Optimization of the number of labels visible in a screenshot means that some labels not visible in the main panel itself may be visible in a screenshot of the main panel. The **Optimized screenshots** check box has an effect only if the **No overlap** check box is checked.
- **Lines.**
 - **Number of lines.** If the adjacency matrix of a network is available, lines can be displayed in the label view and the scatter view of the main panel. Lines can be displayed between any pair of items with a non-zero link strength. The **Number of lines** text box determines the maximum number of lines to be displayed. If the number of links with a non-zero strength exceeds the maximum number of lines to be displayed, lines are displayed only for the strongest links.
 - **Use normalization.** This check box determines whether lines are displayed based on normalized or unnormalized link strengths. In general, the use of unnormalized link strengths yields more satisfactory results.
- **Visualization.**
 - **Label view.** Use this drop down list to choose the visualization used in the label view. There are three options. If the **Show labels only** option is chosen, items are indicated only by their label. If the **Show labels and circles** option is chosen, items are indicated both by their label and by a circle. If the **Show labels and frames** option is chosen, items are indicated by their label displayed within a frame.
 - **Density view kernel width.** This slider determines the value of the kernel width parameter of the density view and the cluster density view.

- **Density view transformation.** This slider determines the value of the color transformation parameter of the density view and the cluster density view.⁹
- **Density view color scheme.** Use this drop down list to choose the color scheme used in the density view.
- **White background in cluster density view.** This check box determines whether a black or a white background color is used in the cluster density view.
- **Choose Cluster Colors.** This button is not available if scores have been given to items (using the *score* column in a map file; see Section 3.1) or if the colors of items have been specified directly (using the *red*, *green*, and *blue* columns in a map file; see Section 3.1). The button brings up the **Choose Cluster Colors** dialog box. This dialog box can be used to choose the colors of clusters (and, consequently, the colors of the items belonging to the clusters). Using the **Choose Cluster Colors** dialog box, cluster colors can also be imported from or exported to a cluster colors file (see Section 3.4). If items have not been assigned to clusters, the dialog box can be used to choose the color of all items.

⁹ We refer to Van Eck and Waltman (2010) for more information about the kernel width parameter and the color transformation parameter.

3 File types

The two primary file types used by VOSviewer are the map file and the network file. Map files and network files are simple text files that can easily be viewed and edited using a text editor (e.g., Notepad) or a spreadsheet program (e.g., Excel). Map files and network files consist of multiple columns. Hence, each line in a map file or a network file contains multiple pieces of information. Different pieces of information on the same line are separated from each other by a comma, a semicolon, or a tab. If a piece of information (e.g., the label of an item) itself contains a comma or a semicolon, the whole piece of information needs to be enclosed within double quotes.

We discuss map files and network files in more detail in Sections 3.1 and 3.2. Three additional file types, the thesaurus file, the cluster colors file, and the score colors file, are discussed in Sections 3.3 and 3.4.3. We note that all example files mentioned in this chapter are distributed together with VOSviewer.

3.1 Map file

A map file is a text file that contains information about the items in a map. Each line in a map file corresponds with an item. The only exception is the first line. This is a header line that indicates what is contained in each of the columns of a map file. Below, we list the columns that can be used in a map file. For each column, we provide the column header and we specify what the column contains.

<i>id</i>	The ID of an item. Items need to have an ID only if a map file is used in combination with a network file.
<i>label</i>	The label of an item.
<i>sublabel</i>	The sublabel of an item. In the main panel, the sublabel of an item is displayed below the item's ordinary label. Sublabels are displayed in a smaller font.
<i>description</i>	The description of an item. The description of an item is used to provide information about the item in the information panel. HTML formatting can be used in this column.
<i>url</i>	The URL of an item. This column can be used to associate a web page with an item. When one clicks on the label of an item in the main panel, the web page associated with the item will be opened in a web browser.
<i>x</i>	The horizontal coordinate of an item.
<i>y</i>	The vertical coordinate of an item.
<i>weight</i>	The weight of an item. Only non-negative numbers are allowed in this column. The higher the weight of an item, the more prominent the item is presented in the

	main panel.
<i>normalized weight</i>	The normalized weight of an item. Only non-negative numbers are allowed in this column. The higher the normalized weight of an item, the more prominent the item is presented in the main panel. The default presentation of an item is obtained by setting the item's normalized weight to 1.
<i>score</i>	The score of an item. The score of an item determines the color of the item. By default, colors range from blue (score of 0) to green (score of 1) to red (score of 2).
<i>cluster</i>	The number of the cluster to which an item belongs. Only integers between 1 and 1000 are allowed in this column.
<i>red</i>	The red component of the color of an item. Only integers between 0 and 255 are allowed in this column.
<i>green</i>	The green component of the color of an item. Only integers between 0 and 255 are allowed in this column.
<i>blue</i>	The blue component of the color of an item. Only integers between 0 and 255 are allowed in this column.

In a map file, one always uses only a subset of the above columns. The order in which the columns are used is not important.

There are a number of restrictions on the columns that are used in a map file:

- There must be an *id* column or a *label* column. (If there is no *label* column, the ID of an item is used as the item's label.)
- If there is a *sublabel* column, there must be a *label* column as well.
- The *x* and *y* columns must be used together.
- The *red*, *green*, and *blue* columns must be used together.
- The *weight* column and the *normalized weight* column cannot be used together.
- The *score* column and the *cluster* column cannot be used together.
- The *score* column and the *red*, *green*, and *blue* columns cannot be used together.
- The *cluster* column and the *red*, *green*, and *blue* columns cannot be used together.

For an example of a map file, see the file `journal_map.txt`.

3.2 Network file

A network file is a text file that contains the adjacency matrix of a network. The adjacency matrix of a network is a square matrix that indicates for each pair of

items in the network the strength of the link between the items. The strength of a link is given by a non-negative number. If there is no link between two items, the strength of the link between the items equals zero. VOSviewer works with symmetric adjacency matrices. If a network file contains an asymmetric adjacency matrix, VOSviewer will make the matrix symmetric by averaging corresponding elements on both sides of the main diagonal.

A network file has either a full format or a sparse format:

- *Full format.* The entire adjacency matrix, including elements that are equal to zero, is stored in the network file. The file consists of n lines and $n + 1$ columns, where n denotes the number of items in the network. The element in the i th row and the j th column of the adjacency matrix is stored on the i th line and in the $(j + 1)$ th column of the network file. The first column of the network file contains IDs of items. This column indicates for each row and column of the adjacency matrix the ID of the corresponding item. For an example of a network file in full format, see the file `journal_network_full.txt`.
- *Sparse format.* Only the non-zero elements of the adjacency matrix are stored in the network file. The file consists of two or three columns. The first two columns contain IDs of items. The third column contains the non-zero elements of the adjacency matrix. This column indicates the strength of the link between the items referred to in the first two columns. If there is no third column, the strength of the link between the items referred to in the first two columns always equals one. In the case of a symmetric adjacency matrix, it is sufficient to store only half of the matrix (either the upper triangular part or the lower triangular part) in the network file. For an example of a network file in sparse format, see the file `journal_network_sparse.txt`.

A network file is usually used in combination with a map file. For each item ID in the network file, there must then be a corresponding item ID in the map file.

3.3 Thesaurus file

A thesaurus file is a text file that can be used in the following situations:

- When creating a map based on a Web of Science file (see Subsection 2.4.1), a thesaurus file can be used to merge different variants of a source title, an author name, an organization name, or a cited reference. This may for example be useful when the name of a researcher is written in different ways in different publications (e.g., with first initial only and with all initials). A thesaurus file can

then be used to indicate that different names in fact refer to the same researcher.

- When creating a map based on a corpus file (see Subsection 2.4.1), a thesaurus file can be used to merge synonyms into a single term. This may be useful not only for merging different terms referring to the same concept, but also for merging different spellings of the same term (e.g., *behavior* and *behaviour*). It may also be useful for merging an abbreviation of a term with the term itself.

Each line in a thesaurus file contains a label and indicates an alternative label that replaces the original label. The only exception is the first line. This is a header line that indicates what is contained in each of the columns of a thesaurus file. A thesaurus file must have two columns: A *label* column and a *replace by* column. The *label* column contains a label. Depending on the situation (see above), a label may represent a source title, an author name, an organization name, a cited reference, or a term. The *replace by* column contains an alternative label that replaces the original label. The *replace by* column may also be empty. In that case, the original label is not replaced by an alternative one, but instead the original label is simply ignored. When creating a map based on a corpus file, this allows one to use a thesaurus file as a kind of stop word list. A thesaurus file may for example indicate that certain uninteresting terms (e.g., *copyright* and *Elsevier* in the case of a corpus of abstracts of scientific publications) should be ignored. For examples of thesaurus files, see the files `thesaurus_authors.txt` and `thesaurus_terms.txt`.

3.4 Cluster colors file and score colors file

A cluster colors file is a text file that contains the colors of clusters. Each line in a cluster colors file corresponds with a cluster. The only exception is the first line. This is a header line that indicates what is contained in each of the columns of a cluster colors file. A cluster colors file must have four columns: A *cluster* column, a *red* column, a *green* column, and a *blue* column. The *cluster* column contains cluster numbers. Only integers between 0 and 1000 are allowed in this column. Cluster number 0 is a special case. This cluster number does not refer to an actual cluster. Instead, cluster number 0 can be employed to specify the color that is used if items have not been assigned to clusters and all have the same color. The *red*, *green*, and *blue* columns contain the red, green, and blue components of the colors of clusters. Only integers between 0 and 255 are allowed in these columns. For an example of a cluster colors file, see the file `cluster_colors.txt`.

A score colors file is similar to a cluster colors file except that the colors it contains correspond with scores rather than with clusters. This means that instead of a *cluster* column a score colors file has a *score* column. If scores have been given to items (using the *score* column in a map file; see Section 3.1), these scores are matched with the scores in the *score* column of the score colors file and in this way the colors of items are determined. Exact matching of scores is usually not possible, and in that case the colors in the score colors file are interpolated. For an example of a score colors file, see the file `score_colors.txt`.

4 Advanced topics

In this chapter, some advanced topics are addressed. We first consider the use of command line parameters (Section 4.1). We then discuss how a map can be made available on the internet (Section 4.2) and how the amount of memory available to VOSviewer can be increased (Section 4.3).

4.1 Using command line parameters

VOSviewer supports a number of command line parameters. These parameters can for example be used to automatically open a map when VOSviewer is started or to override some of the default settings of VOSviewer. The command line parameters supported by VOSviewer are listed below.

cluster_colors	Use this parameter to specify a cluster colors file (see Section 3.4). Cluster colors will be imported from this file when VOSviewer is started. Cluster colors are used only if items have been assigned to clusters.
encoding	Use this parameter to specify the character encoding that is used by VOSviewer to read and write text files. For a list of the available encodings, see http://docs.oracle.com/javase/6/docs/technotes/guides/intl/encoding.doc.html . If this parameter is not used, VOSviewer will attempt to automatically determine the correct encoding when reading a text file (which in some cases may result in the use of an incorrect encoding), while it will use the default encoding when writing a text file.
file_location	Use this parameter to specify the folder that is used by VOSviewer as the default file location.
label_size_effect	Use this parameter to specify the initial value of the Size effect slider on the Options tab in the action panel (see Subsection 2.4.4).
map	Use this parameter to specify a map file (see Section 3.1). The map in this file will be opened when VOSviewer is started.
network	Use this parameter to specify a network file (see Section 3.2). An adjacency matrix will be read from this file when VOSviewer is started. The adjacency matrix can be used to display lines between the items in a map or to create a new map.
n_lines	Use this parameter to specify the initial value of the Number of lines text box on the Options tab in the action panel (see Subsection 2.4.4).
score_colors	Use this parameter to specify a score colors file (see Section 3.4). Score colors will be imported from this file when VOSviewer is started. Score colors are used only if scores have been given to items (using the <i>score</i> column in a map file; see Section 3.1).
view	Use this parameter to specify which tab in the main panel will be selected when VOSviewer is started (1 for the Label View tab, 2 for the Density View tab, 3 for the Cluster Density View tab, and 4 for the Scatter View tab).

zoom_level Use this parameter to specify the initial zoom level in the main panel. The zoom level must have a value of at least 1. The higher the zoom level, the more the main panel will be zoomed in on the center of a map.

To use the above command line parameters, VOSviewer needs to be run from the command line. When using the Windows executable of VOSviewer,¹⁰ this can for example be done as follows:

```
VOSviewer -map journal_map.txt -view 2 -zoom_level 2.5
```

In this way, VOSviewer will be started and the map in the map file `journal_map.txt` will be opened. Also, the **Density View** tab in the main panel will be selected, and the main panel will be zoomed in on the center of the map. When instead of the Windows executable of VOSviewer the VOSviewer JAR file is used, VOSviewer can for example be run in the following way:

```
java -jar VOSviewer.jar -map journal_map.txt -view 2 -zoom_level 2.5
```

4.2 Making a map available on the internet

Suppose one wants to make a map available on the internet. One way to do this is simply by making a map file and a network file (or only a map file) available on the internet. To open the map, one then needs to take two steps. In the first step, the map file and the network file are downloaded from the internet. In the second step, VOSviewer is started and the downloaded map file and network file are used to open the map.

The above two-step approach is not very convenient, and therefore an alternative approach is available as well. In this alternative approach, a map can be opened in VOSviewer directly from a web page. The alternative approach works as follows. A map file and a network file (or only a map file) need to be made available on the internet. Suppose these files are located at

```
http://www.example.com/map.txt
```

and

```
http://www.example.com/network.txt
```

The map can then be made available on a web page by creating a hyperlink that points to

¹⁰ The Windows executable of VOSviewer is available at www.vosviewer.com/download/. The VOSviewer JAR file can be obtained here as well.


```
http://www.vosviewer.com/vosviewer.php?map=http://www.example.com/map.txt&network=http://www.example.com/network.txt
```

Using this hyperlink, the map can be opened directly in VOSviewer.

We note that all command line parameters discussed in Section 4.1 can also be used in a hyperlink. For example, to open a map, to go directly to the density view, and to zoom in on the center of the map, the following hyperlink can be used:

```
http://www.vosviewer.com/vosviewer.php?map=http://www.example.com/map.txt&network=http://www.example.com/network.txt&view=2&zoom_level=2.5
```

4.3 Increasing the availability of memory

When using VOSviewer to work with large maps (e.g., maps containing several thousands of items), the memory requirements can be quite substantial. VOSviewer will produce an out of memory error if there is not enough memory available. The availability of memory can be increased by running the VOSviewer JAR file from the command line and specifying the amount of memory that needs to be available. For example, if 1000 MB of memory needs to be available, the VOSviewer JAR file can be run as follows:

```
java -Xmx1000m -jar VOSviewer.jar
```

When working with large maps, VOSviewer may also produce a stack overflow error. To prevent this error from occurring, the stack size needs to be increased. This can be done by running the VOSviewer JAR file from the command line in the following way:

```
java -Xss1000k -jar VOSviewer.jar
```

In this case, the stack size is set to 1000 KB, but other values are possible as well.

References

- Rotta, R., & Noack, A. (2011). Multilevel local search algorithms for modularity clustering. *Journal of Experimental Algorithmics*, 16(2), article 2.3.
- Van Eck, N.J., & Waltman, L. (2009). How to normalize cooccurrence data? An analysis of some well-known similarity measures. *Journal of the American Society for Information Science and Technology*, 60(8), 1635–1651.
- Van Eck, N.J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538.
- Van Eck, N.J., & Waltman, L. (2011). Text mining and visualization using VOSviewer. *ISSI Newsletter*, 7(3), 50–54.
- Van Eck, N.J., Waltman, L., Dekker, R., & Van den Berg, J. (2010). A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. *Journal of the American Society for Information Science and Technology*, 61(12), 2405–2416.
- Waltman, L., Van Eck, N.J., & Noyons, E.C.M. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, 4(4), 629–635.