

VOSviewer: A Computer Program for Bibliometric Mapping

Nees Jan van Eck and Ludo Waltman

ERIM REPORT SERIES <i>RESEARCH IN MANAGEMENT</i>	
ERIM Report Series reference number	ERS-2009-005-LIS
Publication	February 2009
Number of pages	19
Persistent paper URL	http://hdl.handle.net/1765/14841
Email address corresponding author	nvaneck@ese.eur.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus Universiteit Rotterdam P.O.Box 1738 3000 DR Rotterdam, The Netherlands Phone: + 31 10 408 1182 Fax: + 31 10 408 9640 Email: info@erim.eur.nl Internet: www.erim.eur.nl

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website:
www.erim.eur.nl

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES *RESEARCH IN MANAGEMENT*

ABSTRACT AND KEYWORDS	
Abstract	We present VOSviewer, a computer program that we have developed for constructing and viewing bibliometric maps. VOSviewer combines the VOS mapping technique and an advanced viewer into a single easy-to-use computer program that is freely available to the bibliometric research community. Our aim in this paper is to provide an overview of the functionality of VOSviewer and to elaborate on the technical implementation of specific parts of the program.
Free Keywords	bibliometric mapping, visualization, VOSviewer, VOS, density view
Availability	The ERIM Report Series is distributed through the following platforms: Academic Repository at Erasmus University (DEAR), DEAR ERIM Series Portal Social Science Research Network (SSRN), SSRN ERIM Series Webpage Research Papers in Economics (REPEC), REPEC ERIM Series Webpage
Classifications	The electronic versions of the papers in the ERIM report Series contain bibliographic metadata by the following classification systems: Library of Congress Classification, (LCC) LCC Webpage Journal of Economic Literature, (JEL), JEL Webpage ACM Computing Classification System CCS Webpage Inspec Classification scheme (ICS), ICS Webpage

VOSviewer: A computer program for bibliometric mapping

Nees Jan van Eck^{*†}

Ludo Waltman^{*}

^{*}Econometric Institute, Erasmus School of Economics

Erasmus University Rotterdam

P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

E-mail: {nvaneck,lwaltman}@ese.eur.nl

[†]Centre for Science and Technology Studies, Leiden University

P.O. Box 905, 2300 AX Leiden, The Netherlands

Abstract

We present VOSviewer, a computer program that we have developed for constructing and viewing bibliometric maps. VOSviewer combines the VOS mapping technique and an advanced viewer into a single easy-to-use computer program that is freely available to the bibliometric research community. Our aim in this paper is to provide an overview of the functionality of VOSviewer and to elaborate on the technical implementation of specific parts of the program.

Keywords

Bibliometric mapping, visualization, VOSviewer, VOS, density view.

1 Introduction

In this paper, we present VOSviewer, a computer program that we have developed for constructing and viewing bibliometric maps. VOSviewer is freely available to the bibliometric research community (see www.neesjanvaneck.nl/vosviewer/). The program can for example

be used to construct maps of authors or journals based on co-citation data or to construct maps of keywords based on co-occurrence data. However, VOSviewer is not only a tool for constructing bibliometric maps, but it also is an easy-to-use viewer that allows bibliometric maps to be examined in full detail. VOSviewer 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). Such maps are difficult to show on paper (e.g., Van Eck, Waltman, Noyons, and Buter, 2008b), and a simple static representation on a computer screen usually also does not provide satisfactory results. To the best of our knowledge, no other computer programs are available with the same viewing capabilities as VOSviewer.

The two main components of VOSviewer are the mapping technique used for constructing maps and the viewer used for examining maps. VOSviewer uses the VOS mapping technique, where VOS stands for *visualization of similarities*. The VOS mapping technique was introduced by Van Eck and Waltman (2007a), who used it in a number of papers (Van Eck and Waltman, 2007b; Van Eck, Waltman, van den Berg, and Kaymak, 2006; Van Eck et al., 2008b). A computer program that implements the VOS mapping technique is freely available (see www.neesjanvaneck.nl/vos/). The viewer component of VOSviewer is based on the viewer software used by Van Eck and Waltman (2007b) and Van Eck et al. (2006, 2008b). We have enhanced and extended this viewer software in numerous ways. VOSviewer is the result of integrating the VOS mapping technique and the improved viewer software. We note that, even though the VOS mapping technique is an important part of VOSviewer, it is perfectly possible to use VOSviewer to view maps that have been constructed using other mapping techniques. VOSviewer can for example be used to view maps that have been constructed using a multidimensional scaling program such as the PROXSCAL program in SPSS. Another more technical thing to note is that VOSviewer has been written entirely in the Java programming language. Because of this, VOSviewer can be used on almost any hardware and operating system platform. It can also be started directly from a web page on the internet.

In the remainder of this paper, we first discuss for what type of bibliometric maps VOSviewer is intended to be used and we then provide an overview of the functionality of VOSviewer. We also discuss the VOS mapping technique, and we elaborate on the idea of a density view. We

note that we do not want VOSviewer to be a ‘black box’ computer program. Especially for scientific purposes, we believe it to be important that specifications of the main algorithms used by a program such as VOSviewer are publicly available. For that reason, one of our aims in this paper is to describe the technical implementation of specific parts of VOSviewer in considerable detail.

2 Types of bibliometric maps

In this section, we discuss for what type of bibliometric maps VOSviewer is intended to be used. In general, two types of maps can be distinguished that are commonly used in bibliometric research. We refer to these types of maps as distance-based maps and graph-based maps. Distance-based maps are maps in which the distance between two items indicates the strength of the relation between the items. A smaller distance generally indicates a stronger relation. In most cases, items tend to be distributed quite unevenly in distance-based maps. On the one hand this makes it easy to identify clusters of related items, but on the other hand this sometimes makes it difficult to label all the items in a map without having labels that overlap each other. Graph-based maps are maps in which the distance between two items need not indicate the strength of the relation between the items. Instead, lines are drawn between items to indicate relations. Items are usually distributed in a fairly uniform way in graph-based maps. This has the advantage that it is relatively easy to avoid the problem of overlapping labels. In our opinion, a disadvantage of graph-based maps compared with distance-based maps is that it typically is more difficult to see the strength of the relation between two items. Clusters of related items are also more difficult to detect.

In Table 1, we list some mapping techniques that are used in bibliometric research to construct distance-based and graph-based maps (some additional techniques are discussed by Börner, Chen, and Boyack, 2003). For constructing distance-based maps, multidimensional scaling (e.g., Borg and Groenen, 2005) is by far the most popular technique in the field of bibliometrics. An alternative to multidimensional scaling is the VOS mapping technique (Van Eck and Waltman, 2007a,b; Van Eck et al., 2006; available at www.nees.janvaneck.nl/vos/). In certain cases, this technique produces much better structured maps than multidimensional scaling (Van Eck, Waltman, Dekker, and van den Berg, 2008a). A third technique for

Table 1: Some mapping techniques for constructing distance-based and graph-based maps.

Distance-based maps	Graph-based maps
Multidimensional scaling	Pajek
VOS	Pathfinder networks
VxOrd	

constructing distance-based maps is VxOrd (Davidson, Wylie, and Boyack, 2001; Klavans and Boyack, 2006; available at www.cs.sandia.gov/~smartin/software.html, where it is called DrL). This technique is especially intended for constructing maps that contain a very large number of items (more than 700,000 items in Klavans and Boyack, 2006). A disadvantage of VxOrd is that a complete specification of how the technique works is not available. For constructing graph-based maps, researchers in the field of bibliometrics (e.g., Leydesdorff and Rafols, 2009) frequently use the computer program Pajek (De Nooy, Mrvar, and Batagelj, 2005). This program relies on a mapping technique developed by Kamada and Kawai (1989). Another technique that can be used to construct graph-based maps is the technique of pathfinder networks. The use of this technique for bibliometric mapping was proposed by White (2003). For various examples of graph-based bibliometric maps, we refer to Chen (2004).

As discussed above, distance-based and graph-based maps both have advantages and disadvantages. In general, however, we believe that distance-based maps are easier to interpret than graph-based ones. For that reason, we have chosen to work with distance-based maps in VOSviewer. VOSviewer can be used to view any distance-based map, regardless of the mapping technique with which the map has been constructed. One can use VOSviewer to view multidimensional scaling maps produced using statistical packages such as SAS, SPSS, and R, but one can also use VOSviewer to view maps constructed using less common techniques such as VxOrd. Because the VOS mapping technique shows a very good performance (Van Eck et al., 2008a), this technique has been fully integrated into VOSviewer. This means that VOSviewer can be used not only to view VOS maps but also to construct them. Hence, no separate computer program is needed for constructing VOS maps.

3 Functionality of VOSviewer

In this section, we provide an overview of the functionality of VOSviewer. To demonstrate some of the features of VOSviewer, we use a data set that consists of co-citation frequencies of journals belonging to at least one of the following five closely related subject categories of Thomson Reuters: *Business*, *Business-Finance*, *Economics*, *Management*, and *Operations Research & Management Science*. The co-citation frequencies of journals were determined based on citations in articles published between 2005 and 2007 to articles published in 2005. A journal was included in the data set only if it had at least 25 co-citations. There were 232 journals that satisfied this condition. We note that the data set is a subset of a data set studied by Van Eck and Waltman (2009). We also note that, based on a clustering technique, the journals in the data set were divided into five clusters.

The viewing capabilities of VOSviewer partly depend on the use of colors. Unfortunately, since this paper is printed in black and white, some of the functionality of VOSviewer cannot be shown in a satisfactory way. To get a more complete impression of the functionality of VOSviewer, the interested reader is therefore encouraged to have a look at the program itself. The program is available at www.neesjanvaneck.nl/vosviewer/. The data set that we use in this section to demonstrate some of the features of VOSviewer is available there as well.

A screenshot of the main window of VOSviewer is shown in Figure 1. As can be seen in the figure, the main window consists of the following three panels:

Main panel In this panel, a selected area of the currently active map is shown. VOSviewer has zoom and scroll functionality that can be used to determine which area of the currently active map is shown in the main panel.

Overview panel In this panel, an overview of the currently active map is shown. A rectangular frame is displayed in the overview panel to indicate which area of the currently active map is shown in the main panel.

Action panel This panel can be used to undertake various actions, such as opening a map, saving a screenshot, searching for an item, running the VOS mapping technique, and changing settings.

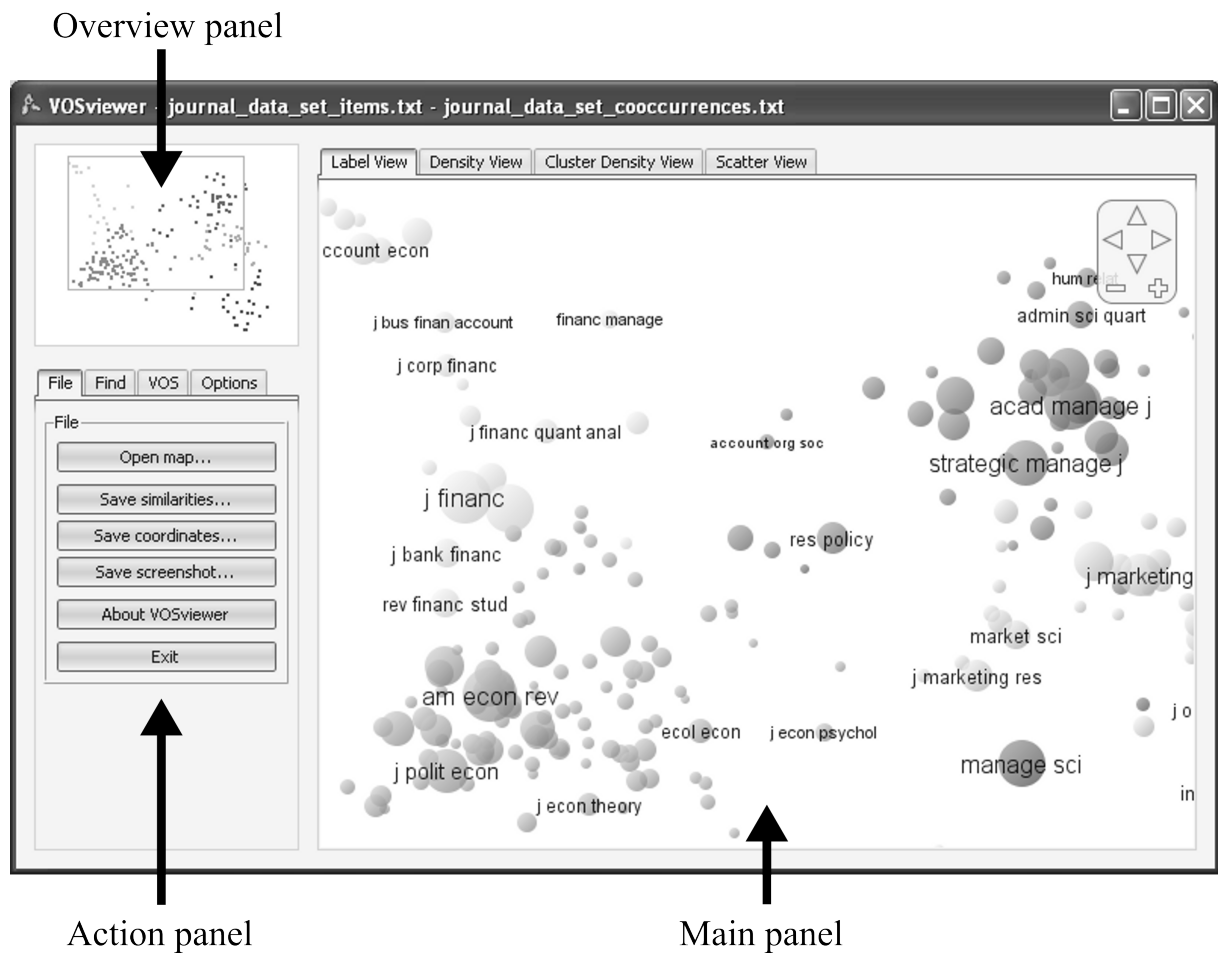


Figure 1: Screenshot of the main window of VOSviewer.

Depending on the available data, the main panel of VOSviewer consists of three or four tabs. Each tab provides a different view on the currently active map. The views provided by the tabs are referred to as the label view, the density view, the cluster density view, and the scatter view. The views all emphasize a different aspect of the currently active map. We now discuss each of the views:

Label view In this 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 area of the item's circle depend on the weight of the item. (The weight of an item usually equals the total number of occurrences or co-occurrences of the item.) If colors have been assigned to items, each item's circle is displayed in the color of the item. By default, to avoid overlapping labels, only a subset of all labels is visible. The label view is particularly useful for a detailed examination of a map.

An example of the label view is shown in Figure 2. The map shown in the figure was constructed based on the journal co-citation data set discussed at the beginning of this section. Gray scales indicate the cluster to which a journal was assigned by the clustering technique that we used. It can be seen that the map reveals a clear clustering of journals and that there is strong agreement between this clustering and the clustering obtained using our clustering technique. The clusters are easy to interpret and correspond with the following five research fields: accounting/finance, economics, management, marketing, and operations research. Although this is not directly visible in Figure 2, we note that there is a large overlap in the map between the *Business* and *Management* subject categories of Thomson Reuters. This indicates an important difference between the clustering that we found and the clustering provided by the subject categories of Thomson Reuters.

Density view In this 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 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 items, the closer the color of the point is to blue. The density view is particularly useful to get an overview of the important areas of

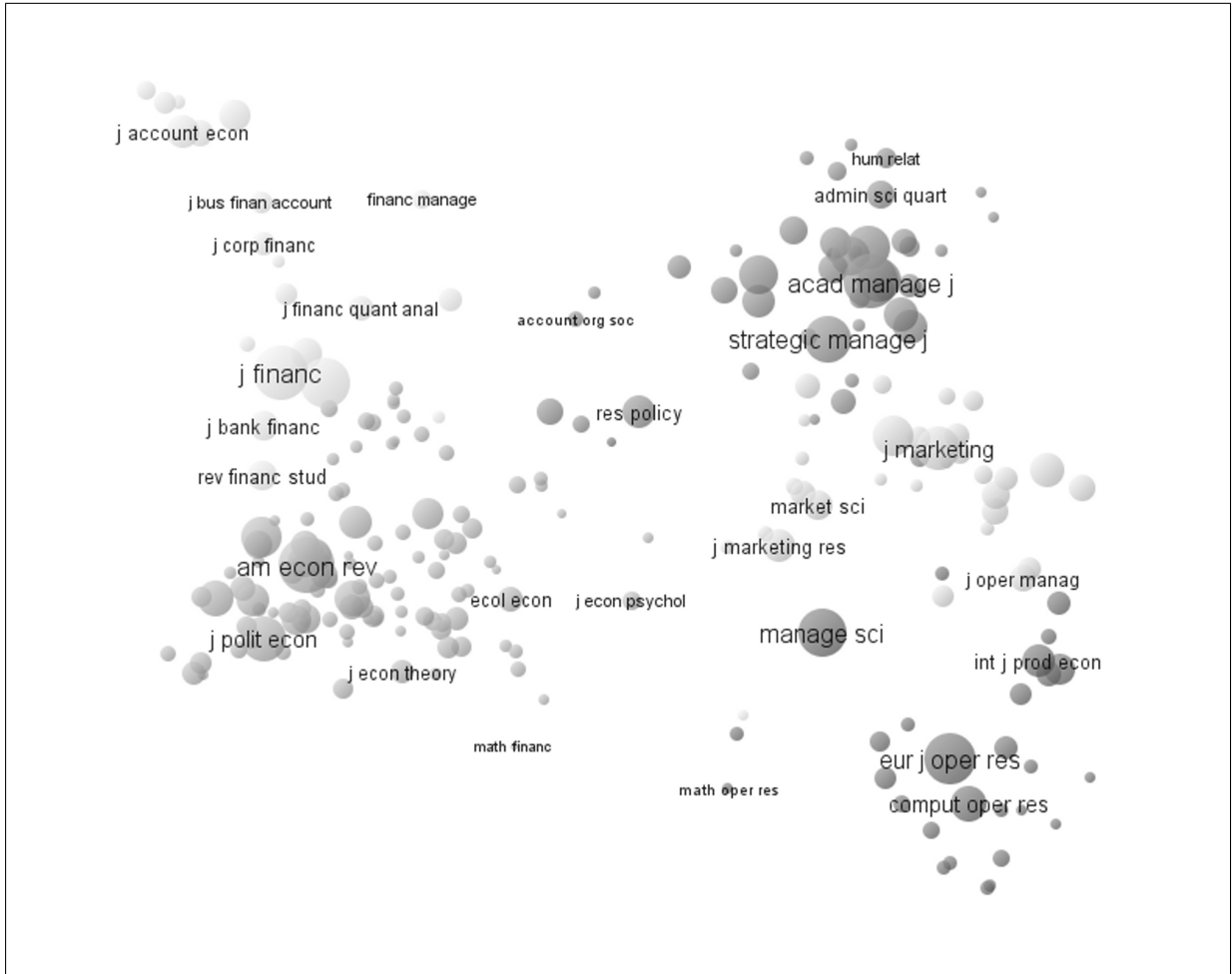


Figure 2: Screenshot of the label view.

a map. A detailed description of the technical implementation of the density view will be provided later on in this paper.

An example of the density view is shown in Figure 3. The map shown in the figure is the same as the one shown in Figure 2. Using gray scales only, the density view is probably of quite limited value. However, when shown in full color, the density view immediately reveals the important areas of a map. Especially the economics and management areas turn out to be important. These areas are very dense, which indicates that overall the journals in these areas receive a lot of citations.

Cluster density view This 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

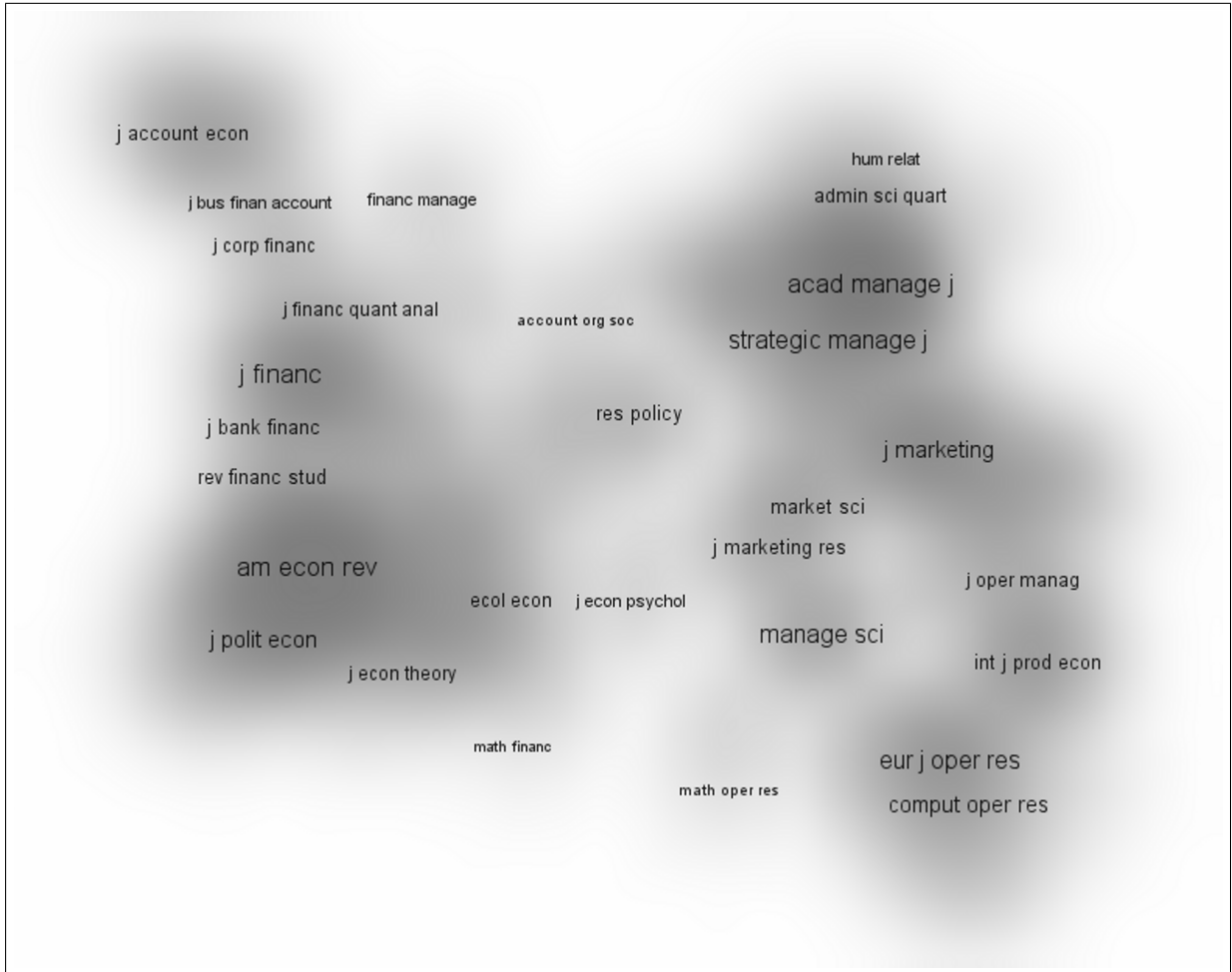


Figure 3: Screenshot of the density view.

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 an overview, for each cluster separately, of the important areas of a map. A detailed description of the technical implementation of the cluster density view will be provided later on in this paper.

Unfortunately, since colors are essential to the cluster density view, we cannot show an example of the cluster density view in this paper.

Scatter view In this view, items are indicated by a small circle. If colors have been assigned to items, each item's circle is displayed in the color of the item. No labels are displayed. The scatter view is particularly 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 above.

In addition to the four views discussed above, there are a number of other features of VOSviewer to which we would like to draw special attention:

- VOSviewer offers extensive zoom and scroll functionality. One can zoom in on a specific area of a map at any desired level of detail. This greatly simplifies the detailed examination of a map. The zoom and scroll functionality of VOSviewer is indispensable when working with maps containing large numbers of items.

An example of the use of the zoom and scroll functionality of VOSviewer is shown in Figure 1. The map that is displayed in this figure is the same as the one that is shown in Figure 2. However, we have zoomed in on a specific area of the map, namely the area in which the economics, finance, and management journals are located. Most of the operations research journals are therefore not visible.

- VOSviewer offers functionality to search for an item in a map. Searching works in a straightforward way. First, one enters a search string. VOSviewer will then provide a list of all items whose label contains the search string. Next, one selects one of the items in the list. VOSviewer will then zoom in on the selected item. The search functionality of VOSviewer is especially useful when working with maps containing large numbers of items.
- When showing a map, VOSviewer uses a special algorithm to determine which labels can be displayed and which labels cannot be displayed without having labels that overlap each other. The further one zooms in on a specific area of a map, the more labels become visible. Labels of items with high weights have priority over labels of items with low weights.
- VOSviewer can handle maps containing large numbers of items. It can construct maps that contain several thousands of items, and it can show maps that contain more than 10,000 items. Due to VOSviewer's zoom, scroll, and search functionality and its advanced labeling algorithm, even such large maps can be examined in full detail.
- Data can easily be transferred between VOSviewer and other computer programs. VOSviewer works with plain text files that have very simple formats. Spreadsheet software such as Excel and mathematical and statistical packages such as MATLAB, SAS, SPSS, and R

can easily handle the file formats used by VOSviewer. A detailed description of the file formats used by VOSviewer can be found in the manual of VOSviewer (available at www.neesjanvaneck.nl/vosviewer/).

- VOSviewer provides special support for taking screenshots of the contents of the main panel. Screenshots can be saved directly in a number of popular graphic file formats. Screenshots of the label view and the scatter view can even be saved using vector graphics. Saving a screenshot using vector graphics has the advantage that the screenshot can be resized without any loss of quality. We note that Figures 2 and 3 were obtained using the screenshot functionality of VOSviewer.

We have now provided an overview of the functionality of VOSviewer. In the remainder of this paper, we focus our attention on the technical implementation of specific parts of the program.

4 Construction of a map

In this section, we discuss how VOSviewer constructs a map based on a co-occurrence matrix. The construction of a map is a process that consists of three steps. In the first step, a similarity matrix is calculated based on the co-occurrence matrix. In the second step, a map is constructed by applying the VOS mapping technique to the similarity matrix. And finally, in the third step, the map is translated, rotated, and perhaps also reflected. We now discuss each of these steps in more detail.

4.1 Step 1: Similarity matrix

The VOS mapping technique requires a similarity matrix as input. A similarity matrix can be obtained from a co-occurrence matrix by normalizing the latter matrix, that is, by correcting the matrix for differences in the total number of occurrences or co-occurrences of items (Waltman and van Eck, 2007). The most popular similarity measures for normalizing co-occurrence data are the cosine and the Jaccard index. VOSviewer, however, does not use one of these similarity measures. Instead, it uses a similarity measure known as the association strength (Van Eck and Waltman, 2007b; Van Eck et al., 2006). This similarity measure is sometimes also referred

to as the proximity index (e.g., Peters and van Raan, 1993; Rip and Courtial, 1984) or as the probabilistic affinity index (e.g., Zitt, Bassecoulard, and Okubo, 2000). Using the association strength, the similarity s_{ij} between two items i and j is calculated as

$$s_{ij} = \frac{c_{ij}}{w_i w_j}, \quad (1)$$

where c_{ij} denotes the number of co-occurrences of items i and j and where w_i and w_j denote either the total number of occurrences or the total number of co-occurrences of, respectively, item i and item j . It can be shown that the similarity between items i and j calculated using (1) is proportional to the ratio between on the one hand the observed number of co-occurrences of items i and j and on the other hand the expected number of co-occurrences of items i and j under the assumption that occurrences of items i and j are statistically independent. We refer to Van Eck and Waltman (2009) for an extensive discussion of the advantages of the association strength over other similarity measures, such as the cosine and the Jaccard index.

4.2 Step 2: VOS mapping technique

We now discuss how the VOS mapping technique constructs a map based on the similarity matrix obtained in step 1. A more elaborate discussion of the VOS mapping technique, including an analysis of the relation between the VOS mapping technique and multidimensional scaling, is provided by Van Eck and Waltman (2007a).

Let n denote the number of items. The aim of the VOS mapping technique is to construct a two-dimensional map in which the items $1, \dots, n$ are located in such a way that the distance between any pair of items i and j reflects their similarity s_{ij} as accurately as possible. Items that have a high similarity should be located close to each other, while items that have a low similarity should be located far from each other. The idea of the VOS mapping technique is to minimize a weighted sum of the squared Euclidean distances between all pairs of items. The higher the similarity between two items, the higher the weight of their squared distance in the summation. To avoid trivial maps in which all items have the same location, the constraint is imposed that the average distance between two items must be equal to 1. In mathematical notation, the objective function to be minimized is given by

$$E(\mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{i < j} s_{ij} \|\mathbf{x}_i - \mathbf{x}_j\|^2, \quad (2)$$

where the vector $\mathbf{x}_i = (x_{i1}, x_{i2})$ denotes the location of item i in a two-dimensional map and where $\|\cdot\|$ denotes the Euclidean norm. Minimization of the objective function is performed subject to the constraint

$$\frac{2}{n(n-1)} \sum_{i < j} \|\mathbf{x}_i - \mathbf{x}_j\| = 1. \quad (3)$$

We note that the distances $\|\mathbf{x}_i - \mathbf{x}_j\|$ in the constraint are not squared. The constrained optimization problem of minimizing (2) subject to (3) is solved numerically in two steps. The constrained optimization problem is first converted into an unconstrained optimization problem. The latter problem is then solved using a so-called majorization algorithm. The majorization algorithm used by VOSviewer is a variant of the SMACOF algorithm described in the multi-dimensional scaling literature (e.g., Borg and Groenen, 2005). To increase the likelihood of finding a globally optimal solution, the majorization algorithm can be run multiple times, each time using a different randomly generated initial solution.

4.3 Step 3: Translation, rotation, and reflection

The optimization problem discussed in step 2 does not have a unique globally optimal solution. This is because, if a solution is globally optimal, any translation, rotation, or reflection of the solution must also be globally optimal (for a discussion of this issue in the multidimensional scaling context, see Borg and Groenen, 2005). It is of course important that VOSviewer produces consistent results. The same co-occurrence matrix should therefore always yield the same map (apart from differences caused by local optima). To accomplish this, it is necessary to transform the solution obtained for the optimization problem discussed in step 2. VOSviewer applies the following three transformations to the solution:

Translation The solution is translated in such a way that it becomes centered at the origin.

Rotation The solution is rotated in such a way that the variance on the horizontal dimension is maximized. This transformation is known as principal component analysis.

Reflection Let i and j denote the items with, respectively, the lowest and the highest coordinate on the horizontal dimension, and let k and l denote the items with, respectively, the lowest and the highest coordinate on the vertical dimension. If $i > j$, the solution is reflected in the vertical axis. If $k > l$, the solution is reflected in the horizontal axis.

These three transformations are sufficient to ensure that VOSviewer produces consistent results.

5 Density view

In this section, we elaborate on the idea of a density view. We distinguish between the ordinary density view of a map and the cluster density view. We note that in VOSviewer the cluster density view is available only if items have been assigned to clusters.

5.1 Ordinary density view

We first discuss the ordinary density view (see also Van Eck and Waltman, 2007b). For related ideas, we refer to Eilers and Goeman (2004) and Van Liere and de Leeuw (2003).

In the ordinary density view, the color of a point in a map is determined based on the item density of the point. Let \bar{d} denote the average distance between two items, that is,

$$\bar{d} = \frac{2}{n(n-1)} \sum_{i < j} \|\mathbf{x}_i - \mathbf{x}_j\|. \quad (4)$$

The item density $D(\mathbf{x})$ of a point $\mathbf{x} = (x_1, x_2)$ is then defined as

$$D(\mathbf{x}) = \sum_{i=1}^n w_i K\left(\frac{\|\mathbf{x} - \mathbf{x}_i\|}{\bar{d}h}\right), \quad (5)$$

where $K : [0, \infty) \rightarrow [0, \infty)$ denotes a kernel function, $h > 0$ denotes a parameter called the kernel width, and w_i denotes the weight of item i , that is, the total number of occurrences or co-occurrences of item i . The kernel function K must be non-increasing. VOSviewer uses a Gaussian kernel function given by

$$K(t) = \exp(-t^2). \quad (6)$$

It follows from (5) that the item density of a point in a map depends both on the number of neighboring items and on the weights of these items. The larger the number of neighboring items and the smaller the distances between these items and the point of interest, the higher the item density. Also, the higher the weights of the neighboring items, the higher the item density. We note that the calculation of item densities using (5) is similar to the estimation of a probability density function using the technique of kernel density estimation (e.g., Scott, 1992).

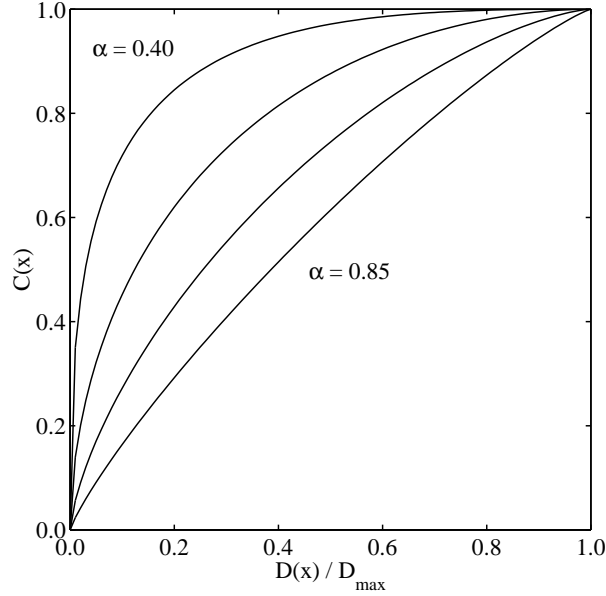


Figure 4: Relation between the color strength $C(\mathbf{x})$ and the normalized item density $D(\mathbf{x})/D_{max}$ for $\alpha = 0.40$, $\alpha = 0.55$, $\alpha = 0.70$, and $\alpha = 0.85$.

Based on the item density of a point in a map, the point's color strength is calculated. The color strength $C(\mathbf{x})$ of a point \mathbf{x} is defined as

$$C(\mathbf{x}) = 1 - \left(1 - \left(\frac{D(\mathbf{x})}{D_{max}} \right)^\alpha \right)^{1/\alpha}, \quad (7)$$

where $\alpha > 0$ denotes a parameter called the color transformation and D_{max} denotes the maximum item density, that is,

$$D_{max} = \max_{\mathbf{x}} D(\mathbf{x}). \quad (8)$$

As shown in Figure 4, the color strength $C(\mathbf{x})$ of a point \mathbf{x} can be seen as a simple monotone transformation of the normalized item density $D(\mathbf{x})/D_{max}$ of the point. The transformation serves to enhance the visual appeal of the density view. Color strengths calculated using (7) are translated into colors using a color scheme. VOSviewer offers two color schemes. By default, a red-green-blue color scheme is used. In this color scheme, red corresponds with the maximum item density and blue corresponds with the minimum item density. The other color scheme offered by VOSviewer consists of gray scales. Darker gray scales correspond with higher item densities in this color scheme. The color scheme consisting of gray scales is used in the density view shown in Figure 3.

The default values used by VOSviewer for the kernel width parameter h and the color trans-

formation parameter α are, respectively, 0.125 and 0.550. These values generally seem to work fine. However, if necessary, one can change the values of both parameters. We further note that the above-described calculation of the color of a point in a map is performed only for a limited number of points. These points are located on a grid. The colors of points that do not lie on this grid are obtained through interpolation.

5.2 Cluster density view

We now discuss the cluster density view. In this view, the item density of a point in a map is calculated separately for each cluster. The item density of a point \mathbf{x} for a cluster p , denoted by $D_p(\mathbf{x})$, is defined as

$$D_p(\mathbf{x}) = \sum_{i=1}^n I_p(i) w_i K\left(\frac{\|\mathbf{x} - \mathbf{x}_i\|}{dh}\right), \quad (9)$$

where $I_p(i)$ denotes an indicator function that equals 1 if item i belongs to cluster p and that equals 0 otherwise. Like in the ordinary density view, the Gaussian kernel function given by (6) is used in the cluster density view.

After calculating item densities, the color of a point in a map is determined in two steps. Each cluster is associated with a color. In the first step, the colors of the clusters are mixed together. This is done by calculating a weighted average of the colors, where the weight of a color equals the item density for the corresponding cluster, as given by (9). In the second step, the color obtained in the first step is mixed with the background color of the cluster density view. This background color can be either black or white in VOSviewer. Mixing is again done by calculating a weighted average. The weights depend on the color strength $C(\mathbf{x})$ given by (7). The weight of the color obtained in the first step equals $C(\mathbf{x})$, while the weight of the background color equals $1 - C(\mathbf{x})$. In this way, the lower the total item density of a point, the closer the color of the point is to the background color.

6 Conclusion

In this paper, we have presented VOSviewer, a computer program for constructing and viewing bibliometric maps. VOSviewer combines a powerful mapping technique and an advanced viewer into a single easy-to-use computer program. As far as we know, no other computer programs are available with the same viewing capabilities as VOSviewer. We expect to rely

heavily on VOSviewer in future research on bibliometric mapping. By making VOSviewer freely available to the bibliometric research community, we hope that others will benefit from it as well. We welcome any feedback on VOSviewer, and we are open to suggestions for further improvements.

References

- I. Borg and P.J.F. Groenen. *Modern multidimensional scaling*. Springer, 2nd edition, 2005.
- K. Börner, C. Chen, and K.W. Boyack. Visualizing knowledge domains. *Annual Review of Information Science and Technology*, 37(1):179–255, 2003.
- C. Chen. *Information visualization*. Springer, second edition, 2004.
- G.S. Davidson, B.N. Wylie, and K.W. Boyack. Cluster stability and the use of noise in interpretation of clustering. In *Proceedings of the IEEE Symposium on Information Visualization 2001*, pages 23–30, 2001.
- W. de Nooy, A. Mrvar, and V. Batagelj. *Exploratory social network analysis with Pajek*. Cambridge University Press, 2005.
- P.H.C. Eilers and J.J. Goeman. Enhancing scatterplots with smoothed densities. *Bioinformatics*, 20(5):623–628, 2004.
- T. Kamada and S. Kawai. An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31(1):7–15, 1989.
- R. Klavans and K.W. Boyack. Quantitative evaluation of large maps of science. *Scientometrics*, 68(3):475–499, 2006.
- L. Leydesdorff and I. Rafols. A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, 60(2):348–362, 2009.
- H.P.F. Peters and A.F.J. van Raan. Co-word-based science maps of chemical engineering. Part I: Representations by direct multidimensional scaling. *Research Policy*, 22(1):23–45, 1993.

- A. Rip and J.-P. Courtial. Co-word maps of biotechnology: An example of cognitive scientometrics. *Scientometrics*, 6(6):381–400, 1984.
- D.W. Scott. *Multivariate density estimation*. John Wiley & Sons, 1992.
- N.J. van Eck and L. Waltman. VOS: A new method for visualizing similarities between objects. In H.-J. Lenz and R. Decker, editors, *Advances in Data Analysis: Proceedings of the 30th Annual Conference of the German Classification Society*, Studies in Classification, Data Analysis, and Knowledge Organization, pages 299–306. Springer, 2007a.
- N.J. van Eck and L. Waltman. Bibliometric mapping of the computational intelligence field. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 15(5):625–645, 2007b.
- N.J. van Eck and L. Waltman. How to normalize co-occurrence data? An analysis of some well-known similarity measures. Technical Report ERS-2009-001-LIS, Erasmus University Rotterdam, Erasmus Research Institute of Management, 2009.
- N.J. van Eck, L. Waltman, J. van den Berg, and U. Kaymak. Visualizing the computational intelligence field. *IEEE Computational Intelligence Magazine*, 1(4):6–10, 2006.
- N.J. van Eck, L. Waltman, R. Dekker, and J. van den Berg. An experimental comparison of bibliometric mapping techniques, 2008a. Paper presented at the 10th International Conference on Science and Technology Indicators, Vienna.
- N.J. van Eck, L. Waltman, E.C.M. Noyons, and R.K. Buter. Automatic term identification for bibliometric mapping. Technical Report ERS-2008-081-LIS, Erasmus University Rotterdam, Erasmus Research Institute of Management, 2008b.
- R. van Liere and W. de Leeuw. GraphSplatting: Visualizing graphs as continuous fields. *IEEE Transactions on Visualization and Computer Graphics*, 9(2):206–212, 2003.
- L. Waltman and N.J. van Eck. Some comments on the question whether co-occurrence data should be normalized. *Journal of the American Society for Information Science and Technology*, 58(11):1701–1703, 2007.

H.D. White. Pathfinder networks and author cocitation analysis: A remapping of paradigmatic information scientists. *Journal of the American Society for Information Science and Technology*, 54(5):423–434, 2003.

M. Zitt, E. Bassecoulard, and Y. Okubo. Shadows of the past in international cooperation: Collaboration profiles of the top five producers of science. *Scientometrics*, 47(3):627–657, 2000.

Publications in the Report Series Research* in Management

ERIM Research Program: “Business Processes, Logistics and Information Systems”

2009

How to Normalize Co-Occurrence Data? An Analysis of Some Well-Known Similarity Measures

Nees Jan van Eck and Ludo Waltman

ERS-2009-001-LIS

<http://hdl.handle.net/1765/14528>

Spare Parts Logistics and Installed Base Information

Muhammad N. Jalil, Rob A. Zuidwijk, Moritz Fleischmann, and Jo A.E.E. van Nunen

ERS-2009-002-LIS

<http://hdl.handle.net/1765/14529>

Open Location Management in Automated Warehousing Systems

Yugang YU and René B.M. de Koster

ERS-2009-004-LIS

<http://hdl.handle.net/1765/14615>

VOSviewer: A Computer Program for Bibliometric Mapping

Nees Jan van Eck and Ludo Waltman

ERS-2009-005-LIS

<http://hdl.handle.net/1765/14841>

* A complete overview of the ERIM Report Series Research in Management:

<https://ep.eur.nl/handle/1765/1>

ERIM Research Programs:

LIS Business Processes, Logistics and Information Systems

ORG Organizing for Performance

MKT Marketing

F&A Finance and Accounting

STR Strategy and Entrepreneurship