Available online at www.sciencedirect.com
ScienceDirect
Journal of Informetrics 1 (2007) 338–356
Journal of INFORMETRICS An International Journal
www.elsevier.com/locate/joi

Complete graphs and bibliographic coupling: A test of the applicability of bibliographic coupling for the identification of cognitive cores on the field level

Bo Jarneving

Swedish School of Library and Information Science, University College of Borås, Sweden

Received 15 June 2007; accepted 7 August 2007

Abstract

The method of bibliographic coupling in combination with the complete link cluster method was applied for mapping of the field of organic chemistry with the purpose of testing the applicability of a proposed mapping method on the field level. The method put forward aimed at the generation of cognitive cores of documents, so-called 'bibliographic cliques' in the network of bibliographically coupled research articles. The defining feature of these cliques is that they can be considered complete graphs where each bibliographic coupling link ties an unordered pair of documents. In this way, it was presumed that coherent groups of documents in the research front would be found and that these groups would be intellectually coherent as well. Statistical analysis and subject specialist evaluations confirmed these presumptions. The study also elaborates on the choice of observation period and the application of thresholds in relation to the size of document populations.

Keywords: Bibliographic coupling; Science mapping; Research fronts

1. Introduction

The sub-field of Information Science dealing with citation based mapping of research fields has widely been focused on the cocitation analytical method, which over time has been developed to an intelligence tool for science policy applications (Griffith, Small, Stonehill, & Dey, 1974; Small, 1973; Small & Griffith, 1974; Small & Sweeney, 1985). The "cocitation bibliometric modelling" has been claimed to provide with ". . .a detailed description of international research front" (Franklin & Johnston, 1988) and further development of the method, combining word analysis with cocitation clustering, has been accomplished (Braam, Moed, & van Raan, 1991a, 1991b). Less attention has been paid to another citation-based method, bibliographic coupling, which was introduces through a series of reports and articles in the early 1960s (Kessler, 1960, 1962, 1963a, 1963b, 1965). In these reports, with the focus on document retrieval, the applicability of bibliographic coupling as method for the coupling of similar documents was established. While the method of cocitation analysis prospered and was added with the author cocitation analytical method (White & Griffith, 1981), further important empirical findings did not show up until 1984 when Vladutz and Cook performed a successful large scale experiment testing the hypothesis that strong bibliographic coupling links imply strong subject relatedness between

E-mail address: bo.jarneving@hb.se.

1751-1577/$ – see front matter
doi:10.1016/j.joi.2007.08.001

articles. The question of cognitive resemblance between bibliographically coupled research papers was also elaborated in Peters, Braam, and van Raan (1995), where it was found that word profile similarity within document groups sharing a citation to a highly cited publication was significantly higher than between documents without such a relationship. One can conclude that the research on bibliographic coupling so far had not explicitly elaborated on the applicability of bibliographic coupling as a science mapping tool. This area of application for bibliographic coupling, however, was presented the same year in a proceeding by Glänzel and Czerwon, and subsequently in an extended version (1996), who claimed that bibliographic coupling could be used to identify "hot" research topics, meaning that documents connected by many and strong bibliographic coupling links can be applied for the mapping of research fronts. Their method was based on the identification of so-called "core documents", which were defined as documents coupled with at least 10 other documents with a set minimum of normalized coupling strength. The measure of normalized coupling strength applied had previously been suggested by Sen and Gan (1983) as the "coupling angle" (C.A.) between documents, analogous to the well known Salton cosine measure. Both empirical experience as well as theoretical consideration underlied Glänzel and Czerwon's definition, and they claimed that a too small number of coupling links may favour articles published in series and a too large number may exclude smaller research specialties. The authors recommended a threshold of normalized coupling strength of 0.25, which was deemed necessary for the filtering out of less significant links between documents. For their empirical experiment, Glänzel and Czerwon applied a whole volume of the *Science Citation Index* (CSI), comprising more than half a million research articles. Approximately 1% of these fulfilled the requirement of a core document. An important finding was that core documents tended to be cited above the average by subsequent research articles, which, from an additional angle, justified the label 'core document'. The authors concluded that a method of bibliographic coupling may well be applied for science mapping purposes. Most important, applying bibliographic coupling, a research article is available for analysis as soon as it is published, making the identification of emerging specialties more feasible in comparison with the cocitation method. In spite of its favourable features, the total research on bibliographic coupling as a method complementary to cocitation mapping has been meagre. The cause for this unobtrusive position of bibliographic coupling is not obvious and comparable and complementary results to the cocitation mapping approach have been reported (Jarneving, 2001; Persson, 1994; Sharabchiev, 1988).

This study had as its goal to elaborate further on this method's applicability in the context of science mapping. Considering the fact that only a small fraction of research articles would fulfil the requirements of a core document, the investigation of the applicability of bibliographic coupling as a method for science mapping should not be delimited to the analysis of core documents exclusively, and other models of document selection should be tried. On basis of theoretical considerations and previous findings, a method of bibliographic coupling is proposed and applied for the exclusive identification of cognitive cores in the research-front network, in accordance with requirements that stipulate both the shape of clusters as well as their internal strength of coherence.

1.1. Statement of purpose and research questions

The purpose of this study was to investigate the applicability of a compound method constructed for the purpose of the identification of coherent research themes within a defined science field. The two major components of the method put forward are (i) a method for the measuring of document–document similarity and (ii) a method for the partition of document populations into disjoint groups. With regard to (i), a normalized measure of bibliographic coupling strength was applied and with regard to (ii) a hierarchical agglomerative cluster method. The hypothesis was that this compound method would make it feasible to identify cognitive cores of documents, i.e. document groups such that the relationship between documents within groups is strong, and markedly stronger than the relationship to documents outside the group. The research questions stated were the following:

- Q1 How well does the suggested method perform with regard to the generation of internally coherent and externally isolated clusters?
- Q2 To what degree do generated clusters represent valid depictions of research themes?
- Q3 How do results from the application of the suggested method comply with a subject specialist's apprehension of the cognitive structure of the research field?

2. Methodology

The construction of the proposed method was done with a point of departure in previous research on bibliographic coupling, theory from multivariate statistical analysis and graph theory. It was hypothesized that the partition of a document population into statistically coherent groups on basis of bibliographic coupling would mirror coherent research themes. The notion of cluster coherence here is related to the original definitions of bibliographic coupling and criteria for bibliographically coupled document groups presented by Kessler (1962), where "a single item of reference shared by two documents is defined as a unit of coupling between them". Based on this unit, two graded criteria of coupling were defined:

Criterion A A number of articles constitute a related group G_A if each member of the group has at least one reference (one coupling unit) in common with a given test article, P_0. The coupling strength between P_0 and any member of G_A is measured by the number of coupling units between them. G_A^n is that portion of G_A that is linked to P_0 through n coupling units. (According to this criterion, there need not be any coupling between the members of G_A, only between them and P_0.)

Criterion B A number of articles constitute a related group G_B if each member of the group has at least one coupling unit with every other member of the group. The coupling strength of G_B is measured by the number of coupling units between its members. Criterion B differs from criterion A in that it forms a closed structure of interrelated articles, whereas criterion A forms an open structure of articles related to a test article.

Kessler studied the application of criterion A exclusively, but in this case only criterion B is of interest. Document groups generated in accordance with criterion B, or "bibliographic cliques" (Sen & Gan, 1983), are best described from a graph theoretical point of view. As the association between two articles through bibliographic coupling is symmetrical, we can apply the concept of undirected graphs. An undirected graph G, is constituted by a set V of vertices and a set E of edges such that each edge $e \in E$ is associated with an unordered pair of vertices. The existence of an unique edge e associated with the vertices v and w, implies the existence of an edge e associated with the vertices w and v and this is written as $e = (v, w)$ or $e = (w, v)$ (Johnsbaugh, 1997). In Fig. 1, G is constituted by the set $V = \{a, b, c, d\}$ of vertices and the set $E = \{e_1, e_2, \ldots, e_5\}$ of edges. Furthermore, a graph G' whose vertices and edges form subsets of the vertices and graph edges of a graph G, is a sub graph of G, and G is said to be a super graph of G'. A complete graph is a graph in which each pair of vertices is connected by an edge (Johnsbaugh, 1997). In Fig. 2, subsets of V and E constitute the sub graph G', which also is a complete graph.

Such graphs would always have a maximal degree of interconnectedness, i.e. a maximal *density* (D), where D is defined as

$$D = \frac{2(\#L(G))}{N(N-1)}, \tag{2.1}$$

where $\#L(G)$ is the number of edges connecting two vertices and N is the number of vertices (Otte & Rousseau, 2002).

The interval is [0,1] and the maximum value is reached when the value of $\#L(G)$ equals the value of $N(N-1)/2$. In this context this means that the maximal value is reached when all possible document pairs in a cluster are bibliographically coupled.

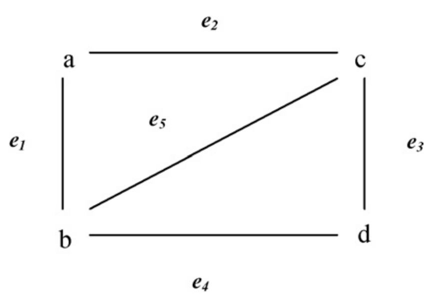

Fig. 1. The incomplete undirected graph G.

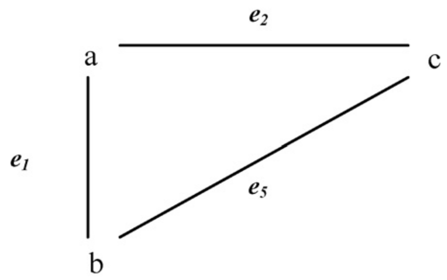

Fig. 2. The complete sub graph G' of the undirected graph G in Fig. 1.

Within the network of communicated formal research, one may imagine hypothetical large undirected and incomplete supergraphs where documents are vertices tied by edges of bibliographic coupling. One may also visualize subgraphs as parts of supergraphs, of which some constitute complete subgraphs. The shape of supergraphs and the number of complete graphs within these may be altered if one deletes edges on or below a certain threshold of coupling strength (cf. Fig. 3).

Moreover, one may assume that strong edges between a document A and document B and between A and C do not imply a strong edge between B and C. If so, one would prefer a method of partition that identifies complete graphs, and, if one can assume that strong relationships between A and B and A and C plausibly reflect a strong relationship between B and C, one that identifies incomplete graphs. This corresponds to two principles of agglomeration in hierarchical cluster analysis: (1) single link clustering and (2) complete link clustering. The latter is applicable for the generation of complete graphs and is in line with the first assumption of relationship, the former for the generation of incomplete graphs and the second assumption. In hierarchical agglomerative clustering the agglomeration of clusters starts with the most similar pair of objects and clusters are subsequently merged in order of similarity. How to define the similarity between clusters defines the cluster method. In single link the similarity between two clusters, X and Y, is measured as the largest similarity between x_i and y_i, where x_i is a member of X and y_i a member of Y. In complete link the opposite applies and the smallest similarity is considered for the agglomeration of clusters. Given a set minimum

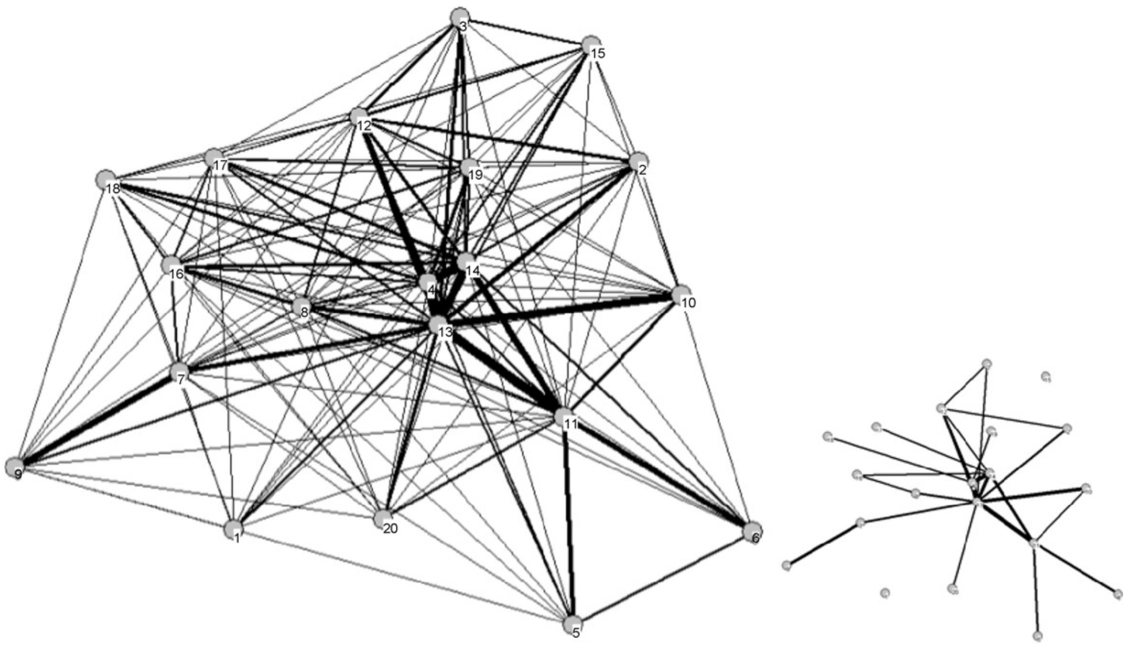

Fig. 3. Illustration of a super graph based on couplings between 20 objects. The width of edges and the distance between vertices illustrate the strength of association. Deleting edges <10 couplings the density of the super graph decreases and only a few complete graphs with size ≥3 vertices remain (see the miniature in the lower right corner).

similarity >0 between object-pairs, this means that all $n(n-1)/2$ object pairs in a cluster generated by the complete link method are bibliographically coupled, constituting a complete graph. In the case of the single link method, the less serious requirement to fulfil may imply prolongated clusters, with a low degree of interconnectedness. Hence, one may regard these methods each others opposites. As the objective in this study was to empirically test a cluster method for bibliographic coupling that is able to generate clusters that can be defined as complete graphs or cliques, the choice of cluster method was given.

The application of thresholds for the applied measure of similarity and link frequency for bibliographically coupled document pairs was inspired by findings and theoretical arguments presented in earlier research (Glänzel & Czerwon, 1995, 1996). The findings reported motivated the concept "core documents", the term referring to such documents' central positions in the formal communication network of the research front of science. This position is based on the many and strong associations with other documents with regard to a common focus on previous research.

Conclusively, the central motive of this research design, as opposed to one that aims at a more comprehensive overview of a field, was to exclusively monitor statistically coherent groups, testing the hypothesis that such coherence is mirrored by subject coherence. Metaphorically speaking, one may zoom in and out of a research field and the zooming operations can be geared by choice of cluster method and thresholds. The setting of thresholds is an important component of the applied method and is further elaborated under the following section.

3. Methods and data

In this section, the major aspects of the proposed method are presented. Specific methods applied for the evaluation of cluster solutions were found to be more comprehensible when presented in the running text of the reported experiment.

3.1. Measure of similarity and the cluster analysis

The similarity between two objects can be measured in two essentially different ways: as (i) the direct similarity between two objects or (ii) the pattern of relations with other third objects in the population studied. Where bibliometric studies are concerned: (i) has been applied in document cocitation cluster analysis (cf. Griffith et al., 1974; Small, 1973; Small & Griffith, 1974; Small & Sweeney, 1985) and also in research on bibliographic coupling (e.g. Persson, 1994; Sen & Gan, 1983; Vladutz & Cook, 1984).[1] In other instances: (ii) has been the prevalent approach as in author cocitation analysis, where the common method is to compare an ordered pair of vectors of author cocitations and calculate Pearson product-moment correlation between vectors (McCain, 1990; White & Griffith, 1981). A problem with (ii) is that bibliometric data frequently include relations of the type 'co-absence'. Co-absence refers to pairs of objects which might be seen as similar in the sense that both objects lack an association to other third objects in the population of study. Thus, for each empirical investigation there is a need to investigate whether co-absence contain useful information or not (cf. Everitt, Landau, & Leese, 2001, p. 36). In this study, there was the possibility of clustering documents on grounds of their lack of a common focus on the previous literature rather than on a shared focus. This, one may assume, could bring about the clustering of non-subject related documents from divergent specialties. When performing some preliminary tests where Euclidean distances and the Pearson product-moment correlation coefficients were applied, early cluster-fusions of documents on the basis of a small or zero distance or high positive correlation in combination with a low coupling strength or the complete absence of bibliographic couplings were frequent. Besides these findings, the objective of this study was to identify cognitive cores with the shape of complete graphs and this cannot be accomplished with a pattern-based similarity measure.

The original definition of bibliographic coupling strength may not be optimal. Besides the fact that a shared reference is not guarantee that both articles are referring to the same piece of information in the cited article (cf. Martyn, 1964), the significance of a single reference is unknown and from this follows that the sharing of a reference may not mirror equal impact of the cited reference on the citing articles. However, it seems reasonable to assume that the probability of a cognitive relationship between two documents should increase by the number of common references and that the significance of a bibliographic coupling unit associating two articles should be inversely related to the combined

[1] In this study, the number of shared cited authors was used as input values. Counts like this, where a higher value corresponds to a high similarity and vice versa, should be considered similarities.

lengths of the reference lists of both documents (cf. Vladutz & Cook, 1984). Therefore, a function that normalizes for the length of reference lists is needed. The choice in this study was the C.A. (Sen & Gan, 1983) which has been used in several studies in the past (e.g. Glänzel & Czerwon, 1995, 1996; Mubeen, 1995; Sharada & Sharma, 1993).

The C.A. was defined as

$$\text{C.A.} = \frac{(D_{\text{o}j} \cdot D_{\text{o}k})}{\sqrt{(D_{\text{o}j} \cdot D_{\text{o}j})(D_{\text{o}k} \cdot D_{\text{o}k})}} \tag{3.1}$$

(Sen & Gan, 1983), which is the cosine of the angle for the two vectors $D_{\text{o}j}$, $D_{\text{o}k}$ of document j, respectively, k in a binary $m \times n$ citation matrix. The range is [0,1] and the denominator has the normalizing function. For the application on an $n \times n$ non-binary matrix, we define the normalized coupling strength (NCS) for two documents, j and k, as

$$\text{NCS}_{j,k} = \frac{r_{jk}}{(n_j \cdot n_k)^{1/2}}, \tag{3.2}$$

where r_{jk} is the number of references common to both j and k, n_j the number of references in the reference list of article j, and n_k is the number of references in the reference list of article k.

The partition of the final set from the original document population by the complete link method required some special attention. Since agglomerative hierarchical methods reduce the initial N number of clusters by $N-1$ fusions to one final cluster, the search for an optimal number of clusters demands a decision when to stop. Different methods for finding the ‘best cut’ in the hierarchy may be applied, and there are more and less formal procedures. Generally, significant changes of fusion levels in hierarchies are assumed to indicate the optimal cut. In this case, common methods were rejected as the intention was to allow the set threshold of NCS to steer the generation of an optimal N. In practice this meant that all clusters emerging before the last level of agglomeration were accepted (Fig. 4).

Fig. 4. A section of the dendrogram representing the agglomeration of 268 clusters. All clusters below the last $N-1$ level of agglomeration were initially accepted. In this section, from top to bottom, clusters of the following sizes emerged: 5, 5, 4, 1, 1, 1, 7 and 2. Last level of fusion is indicated by the vertical line at the end of the dendrogram.

Table 1
The effect of thresholds of NCS on the number of articles and links for the document population studied

Intervals of NCS	No. of links	No. of articles
0.00–1.00	827,544	14,389
0.10–1.00	41,742	11,385
0.20–1.00	7,777	6,537
0.30–1.00	2,614	3,177
0.40–1.00	1,057	1,575
0.50–1.00	466	779
0.60–1.00	206	367
0.70–1.00	90	161
0.80–1.00	23	45
0.90–1.00	4	8

3.2. Thresholds and observation period

The issue of threshold settings is problematic as there exist no clear rules to apply and mostly thresholds are based on empirical experience or by rule of thumb. The principle though is to filter out random associations. Previous researches on bibliographic coupling have suggested quite severe thresholds of NCS. Sen and Gan (1983) suggested a cut off at 0.50 and Glänzel and Czerwon (1996) a smallest value of 0.25. Such thresholds, however, implies the analysis of larger fields of science, as smaller and/or younger fields may not have enough published articles to generate significant and enough many bibliographic coupling links. In this case, the field of organic chemistry is highly productive and more severe thresholds may be applied. However, the ratio between number of analyzed articles and number of articles in the selected document population should in each case be regarded. The immediate consequence of threshold setting is the reduction of bibliographic coupling links between articles. This means as thresholds rise, the density of graphs and the size of clusters decrease (cf. Fig. 1).[2] The filtering out of insignificant links may, however, be a desired effect as the deletion of more random associations between articles may lay clear significant structures otherwise not discernable. The reduction of the original population of articles and links between articles when different thresholds of NCS are applied can be seen in Table 1.

Table 1 contains important information with regard to choice of selection/filtering model. As can be expected, in the beginning, the pattern of reduction is very dissimilar. A first threshold setting removing links with a NCS < 0.10 brought about that the number of links falls of by a factor of 19.8 but the number of articles only by a factor of 1.3. Hence, at this threshold level, the noise reduction is strong but the reduction of articles moderate. At the next threshold level, the number of links falls of by a factor of 5.4 while the corresponding factor for number of articles is 1.7. It is first at the threshold level of 0.30 that differences in reduction levels out. Thus, increasing the threshold beyond 0.30, should be meaningless as this leads to a similar reduction of both articles and links. One can conclude, that the optimal threshold would be in the region of 0.10 if the goal is to filter out as much noise as possible and at the same time hinder a severe reduction of the original population. When raising the threshold near 'core-document' level, less than half to an approximate third of the population remains. Large reductions of original document population may be desired, on the other hand, if a more comprehensive mapping of a research area is wished for, this would not suffice. Hence, the objective of a study would gear the threshold setting. The important thing to know is what actually is being mapped. In this study, a threshold of 0.25 was set, thus, the method of clustering implied that a member in cluster had a coupling link of NCS of at least 0.25 to each other member of that cluster.

Bibliographic coupling techniques are somewhat sensitive to the length of the publication period, and one may hypothesize that the number of shared references would decrease as the period of observation is augmented. Put differently, one may assume that an increase of the distance in time between bibliographically coupled articles leads to a diminishing pool of shared references as there is a tendency to cite the more current articles. In Glänzel and Czerwon (1996), an observation period of one and a half to 2 years was suggested on these grounds. In order to substantiate these

[2] Therefore, all links of a generated cluster should be recovered before the final analysis. Though the internal minimal coherence is secured by setting a threshold of 0.25 NCS, the external links to other clusters (including links <0.25 NCS) must be identified for a complete analysis of the separation between clusters.

Table 2
The distribution of NCS over mean–distances between bibliographically coupled articles for the interval 0.05–0.40, NCS

Intervals of NCS	Mean–distance in publication years
0.30–0.40	0.50
0.25–0.29	1.00
0.20–0.24	1.29
0.15–0.19	1.34
0.10–0.14	1.67
0.05–0.09	2.13

The median NCS was 0.05.

assumptions, 20,616 articles from the *Journal of the American Chemical Society*, distributed over the publication years 1994–2003 were downloaded from the *SCI* CDROM for further analysis. This experiment aimed at testing if there is a tendency for stronger links to be distributed over shorter distances in time. This implied the adding of publication years to pairs of documents joined by NCS. The resulting file was sorted by NCS and the distribution cut at its midpoint. For the upper half of this distribution, the mean–distance in publication years between coupled articles at intervals of NCS was calculated (Table 2).[3]

The findings presented in Table 1 clearly indicate a strong temporal impact on the NCS between articles. Though some cut off point was no arrived at, one may conclude that a 2 years distance between source documents would be the maximum period of observation.

3.3. The original population of articles and the final sets

In this study, articles from the field of organic chemistry was applied as the test bed. Organic chemistry is a sub-discipline to chemistry and concerns the study of the structure, properties, composition, reactions, and synthesis of compounds that contain carbon. The structure of the carbon atom is unique among atoms and allows for a great array of compounds of importance. This gives rise to a large number of research foci and a high publication output.

The selection of data was based on a number of central journals. The identification of these journals was accomplished by a subject specialist. The final journal set comprised the following journals:

i. *Journal of the American Chemical Society*;
ii. *Tetrahedron Letters*;
iii. *Journal of Organic Chemistry*; and
iv. *Angewandte Chemie-International Edition*.

It is to be noted that *Journal of the American Chemical Society* covers a broader range than organic chemistry, but was deemed as central by the subject specialist. A final journal set was compiled by downloading all articles in these journals indexed in the 2002 and 2003 volumes of the SCI CD-ROM. This resulted in an original set of 14,389 articles containing 464,106 references. The average number of references in the articles was 32. In this case, the considerably large number of articles and references made it feasible to identify the more central documents in the network of bibliographically coupled articles. The setting of thresholds and noise reduction was accomplished by filtering out bibliographically coupled pairs with a NCS below 0.25 from the remaining articles. In the next step, only articles with at least five links to other articles were chosen, which rendered a total of 294 articles. The number of articles was also reduced somewhat as 26 articles were exclusively coupled to other articles than the selected 294. It is to be noted that this method of sampling does not per see imply the generation of a coherent set of strongly interrelated documents as a larger or smaller share of links may connect to documents extrinsic to the selected set of articles. In this case the selected set of 294 articles was further reduced to a final set of 268 articles due to such links, a loss of 8.8%. This method of sampling should be further developed, but was deemed sufficient for the purpose of this study.

[3] In the course of a publication (calendar) year, articles are published on different dates, hence, the maximal distance in time between two papers published during the same publication year is less but approximately a year, hence, the maximal error is less but approximately 1 year.

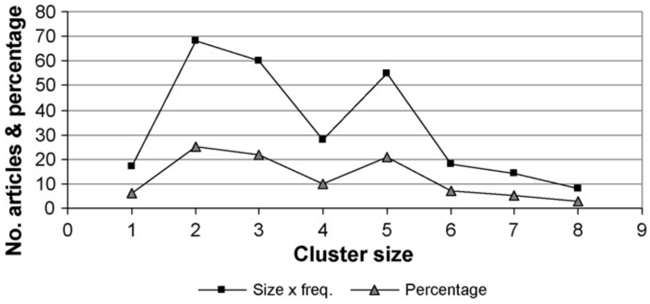

Fig. 5. The size–frequency distribution of clusters and the percentage of a certain cluster size.

4. The experiment: findings and discussion

A total of 268 articles from the original population of 14,389 articles were clustered, resulting in 95 clusters of which 17 were singleton clusters. As can be seen in Fig. 5, the most frequent cluster size is '2' making up for a quarter of the total number of articles. Following the size–frequency curve it is obvious that there is no consistent pattern in this distribution. For instance, more articles are contained within clusters with a size of five members than in clusters sized '4', and most articles belong to clusters with sizes 2, 3 and 5. As the goal of clustering is the arrival at some kind of meaningful summation of data in a smaller number of groups of objects, a confused pattern of numerous single objects and pairs would not contribute to such a goal. Hence, at this point of the analysis, it was decided that clusters with a size <3 should be excluded from the further analysis. Articles in these clusters made up for approximately 32% of the 268 selected articles.

4.1. How well does the proposed method perform with regard to the generation of internally coherent and externally isolated clusters?

The first research question, Q1, concerns the statistical properties of cluster quality. In this sense, we can define cluster quality with regard to (i) internal coherence and (ii) external isolation. Optimally, a cluster should be internally coherent and externally isolated, which would indicate a genuine grouping as opposed to artefactual constructs, which well may be the result of a clustering. As mentioned previously, the density D, was default 1.0 for a cluster (cf. Eq. (2.1)) as was a lowest mean internal coherence of 0.25. The aspect of cluster coherence may be measured as the mean of normalized coupling strength for each cluster. For this purpose the average normalized coupling strength for all clusters was computed as

$$\text{Avg NCS}(C) = \frac{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\text{NCS}(d_i d_j)}{\binom{n}{2}}, \tag{4.1}$$

where $\text{NCS}(d_i,d_j)$ is the coefficient of NCS between two articles, d_i,d_j and n is the number of articles in a cluster c.

This measure is the sum of all NCS-coefficients in a cluster divided by the $n(n-1)/2$ pairs. The other aspect of cluster quality, the isolation from other clusters, was next studied. For this purpose, the mean for links connecting articles from different clusters was computed as

$$\text{Avg NCS}(C, C') = \frac{\sum_{i=1}^{k}\sum_{j=1}^{m}\text{NCS}(d_i d_j)}{k \times m} \tag{4.2}$$

where $d_i \in C$, $d_j \in C'$ and $\text{NCS}(d_i,d_j)$ = the coefficient of NCS between two articles, d_i, d_j.

This measure is the sum of all NCS-coefficients connecting articles in two different clusters, C and C', divided by the $k \times m$ pairs generated by k articles in C and m articles in C'. It is to be noted that the computing of this measure requires the access to all original links (N = 827,544) as opposed to the NCS(C) which by application of method delimits the range of NCS for links (0.25-max).

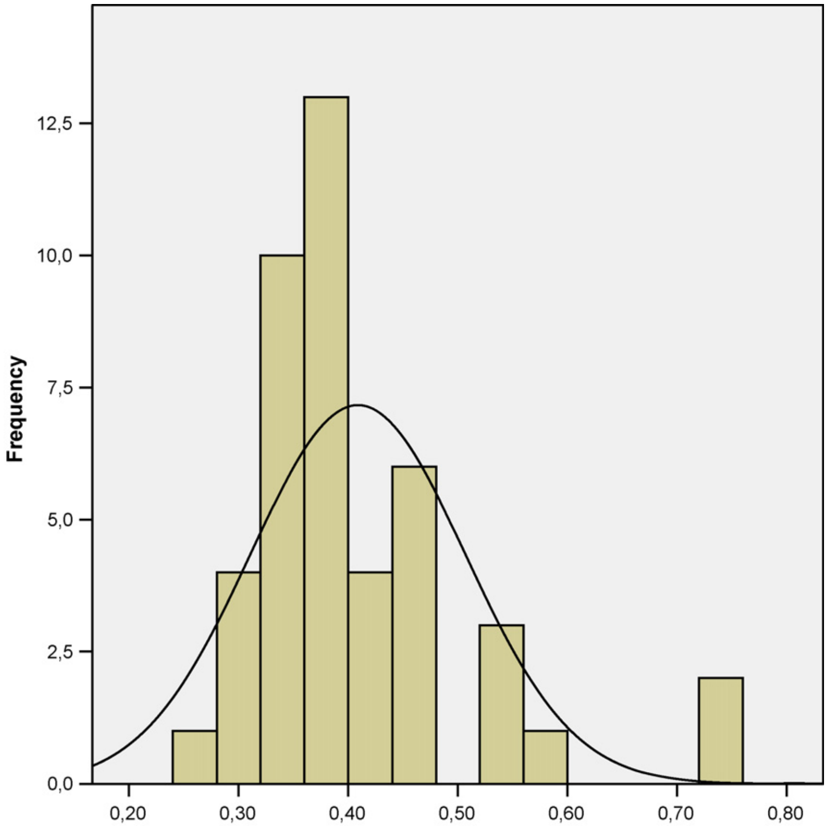

Fig. 6. The distribution of NCS(C).

Displaying the distributions of these measures in histograms, we can see that coefficients for NCS(C) are generally clustered near 0.40 (Fig. 6) while the distribution for NCS (C,C') is strongly clustered at the interval 0.00–0.02, with a smaller share of higher values at the upper tail of the distribution (Fig. 7). One may expect that some research themes become fragmented, as a few clusters have links to other clusters in the interval 0.20–0.40. In general, clusters were coherent and well separated. The descriptive statistics are presented in Table 3.

So far we can appreciate that clusters in general are well interconnected and that a minor share of links between clusters are strong. In order to further elaborate to what extent similar articles were grouped together and to what extent dissimilar articles were separated, the magnitudes of intra- and inter-cluster similarity was assessed by the Silhouette Index (cf. Everitt et al., 2001, p. 185). This index measures the standardized difference between $b(i)$ and $a(i)$ where $a(i)$ is the average dissimilarity of object i to all other objects in its own cluster and b(i) is the average dissimilarity of object i to all objects in the nearest cluster (Everitt et al., 2001, p. 104). This can be expressed as

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, \tag{4.3}$$

(Kaufman & Rousseeuw, 2005, p. 85).

This index, is taking on values in the interval [−1,1] and when it is near 1, an object i is closer to its own cluster than the cluster to which it has the smallest distance, and therefore considered well classified. When i is close to −1 it

Table 3
Descriptive statistics for the distributions of NCS(C) and NCS(C,C')

	N	Range	Minimum	Maximum	Mean	S.D.
AvgNCS(C)	44	0.470	0.267	0.737	0.409	0.098
AvgNCS(C,C')	117	0.346	0.000	0.346	0.057	0.095

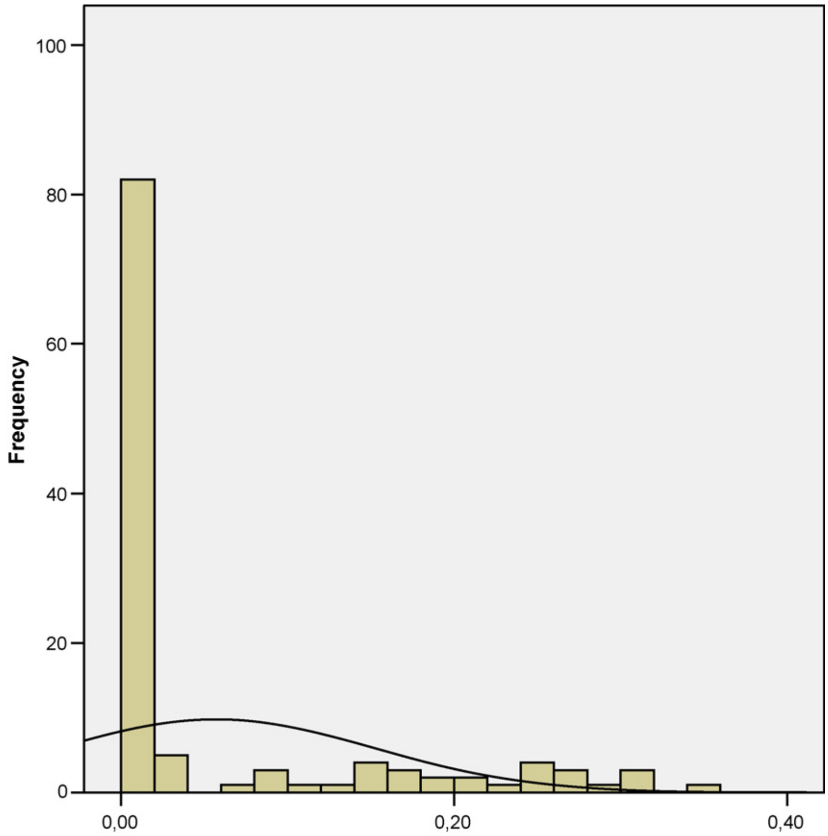

Fig. 7. The distribution NCS(C,C').

is considered as misclassified and should probably belong to another cluster. When the value is 0 or near 0, it is unclear to which cluster it should belong to. Silhouette data may be plotted as a vertical histogram, where the objects in a cluster are ranked in decreasing order by their Silhouette values. The bars in the plot are horizontal and their lengths correspond to the silhouette values of objects. Hence the width of a cluster reflects how genuine the cluster is and the height corresponds to the number of objects in a cluster. As can be seen in Fig. 8 from the profile of the Silhouette plot, the overall cluster structure is clearly homogenous, and no cluster member has a stronger association with a cluster member in another cluster, as would have been indicated by negative bars in the histogram. The overall Silhouette value for the cluster solution was 0.408.

We can conclude that the intentions behind the proposed method are well reflected in Figs. 6–8, and that a high statistical cluster quality is reached, which with consideration to the method is what should be expected.

Though we do not know on beforehand exactly the degree of coherence within clusters, the applied method would, in a general sense, not generate statistically in-coherent clusters. Hence, the more interesting part of Q1 concerns cluster isolation, which is related to the aforementioned plausible effect of fragmentation (cf. Fig. 7). As the isolation of a cluster seldom is total, the links between clusters may be used to illustrate and map the structure made up by the total clusters arrived at. In this case, 39 out of 44 clusters had links to other clusters. On basis of the NCC(C,C') between 39 clusters, an MDS map was produced adding complementary information to the statistical analysis (Fig. 9).

In Fig. 9, links connecting clusters, possibly indicating the augmentation or continuation of some research themes beyond the sole cluster level, is seen. When deleting links <0.25 NCS(C,C') a more sparse configuration is arrived at, with only the more significant links connecting clusters. Though the intra-/inter-cluster similarity ratio was high, the generation of graphs with sizes >1 on an inter-cluster level indicates links between clusters that presumably contain useful information. This aspect, if studied, requires careful expert examination and was for reasons of delimited resources not carried out in this study.

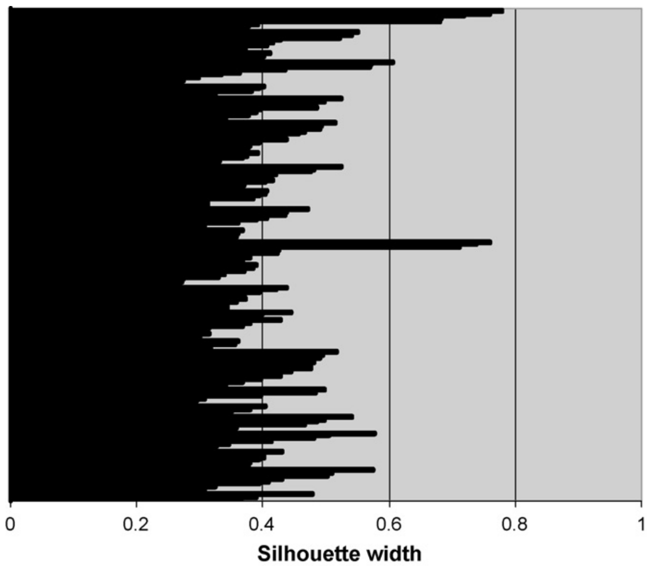

Fig. 8. Silhouette plot for 183 articles in 44 clusters, sorted ascending, cluster wise. Overall Silhouette width = 0.408.

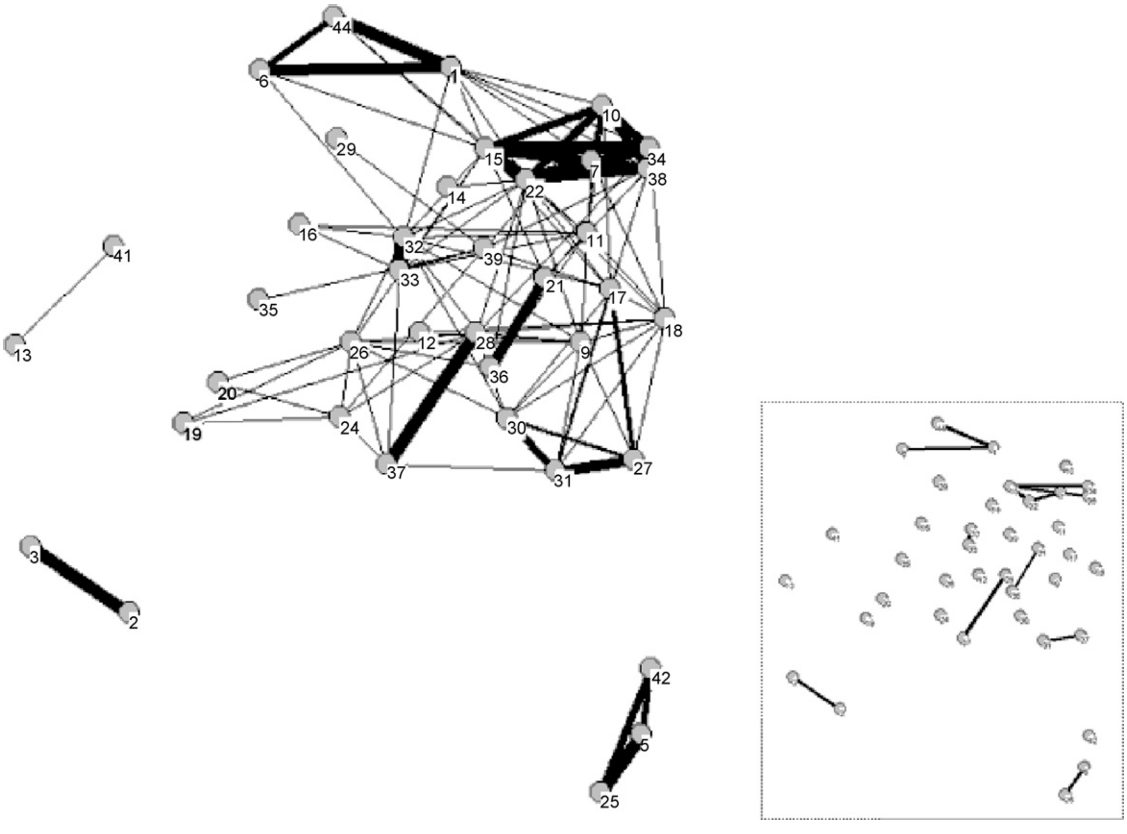

Fig. 9. MDS based on the NCC(C,C') between 39 clusters. The width of links correspond to the relative strength of association between clusters, and is somewhat misguiding as coefficients of NCC(C,C') were multiplied by 1000 in order to be able to graphically depict also weaker links. When links <0.25 NCS(C,C') are deleted, the configuration incorporated in the lower right corner of the map is arrived at. Kruskal's stress was 0.06.

4.2. To what degree do generated clusters represent valid depictions of research themes?

The ability of the applied method to identify related documents was assessed by letting a subject specialist judge the composition of clusters. The issue of cluster relevance was operationalized as the identification of common research foci for articles in clusters, assessed by subject specialists' examination of their subject content. Information concerning the subject matter of articles is contained in *Content Describing Elements* (CDE), where a CDE denotes an element in a bibliographic record which describes the content of an article in such a way that it is not easily mixed up with another (Noyons, 1999, p. 18). Such elements are document specific to a large extent, but some may be less specific for a particular article, such as author-names, journal-titles and cited references (Noyons, 1999). Titles, abstracts or publication specific keywords, supplied by the authors themselves, describe the subject content of the article and could be categorized as uncontrolled terms, whereas indexing terms externally supplied by professional indexers are controlled ones. Both can bee seen as representing problem domains or research themes (Tijssen, 1992, p. 73).

In this case, for each article in a cluster, first author names, publication years, journal titles, article titles, keywords[4] and abstracts were compiled and made available to the subject specialist. The subject specialist's examination of clusters' relevance was pursued by examination of each cluster, article by article, in order to detect inconsistencies as to subject content in clusters. Any article deviating from a common research theme of a cluster was regarded as 'misplaced' and marked. A cluster was regarded representing a research theme when more than 50% of articles shared a research focus and the remaining articles would be considered misclassified. When this condition was not fulfilled, all constituent articles would be counted as misplaced and the cluster regarded as irrelevant. Hence, should a tie occur, all constituent articles would be regarded as misplaced. In total, four articles in four different clusters were regarded as deviating from the cluster's research focus. This means that 98% of 183 articles in 44 clusters were successfully classified, which by any standards must be considered a high share.

When evaluating the partition generated by the complete link cluster method, the subject specialist concluded that "there exists several possible common denominators" for a single cluster solution, and these were often hard to decide. In fact, the proposed method supplied the expert with a new set of principles of division which were not easy to anticipate. According to the subject specialist, some of the classes brought about by the complete link clustering could clearly be of interest in relation to several research questions at issue. A number of clusters were based on the association between articles with a common focus on specific methods whereas others seemed to be based on the association between articles with a common focus on chemical compounds or classes of compounds.

4.3. How do results from the application of the suggested method comply with a subject specialist's apprehension of the cognitive structure of the own research field?

In order to compare the cluster structure generated by the complete link method with the subject specialist's apprehension of the field's intellectual structure, a manual classification of the 183 articles was performed, applying a card sorting method (Biglan, 1973; McCain, 1986; Miller, 1969) where bibliographic data from articles printed on cards were used. For each card representing a specific article, title, abstract, keywords, author name(s), journal title and publication date were printed and formatted to optimal convenience. In addition, complete bibliographic descriptions for all articles were available when needed. The subject specialist was instructed to sort these cards into categories (piles) on the basis of similarity of the subject matter between articles and to assign a proper label to each category. Cards were presented to the subject specialists without any order, and any number of piles, $1, \ldots, n$, where n is equal to the number of articles, was allowed and tentative clusters could be broken up and revised at any time.[5] The subject specialist were also asked to comment on her/his method of partition and perception of the analysed field's intellectual structure. It should be noted that the subject specialist's partitions were performed before the evaluation of clusters generated by the complete link cluster method, hence, her/his intellectual classification was not affected by impressions of the cluster structures generated by the complete link cluster method.

[4] Here, keywords were of two types: *author keywords* assigned by the author and *Keywords Plus* which are words or phrases that frequently appear in the titles of an article's references, but do not necessarily appear in the title of the article or in a list of author keywords.

[5] In order to avoid any systematic with regard to the order of presentation of cards to the subject specialist, a third party sorted and re-sorted the initial pile of cards.

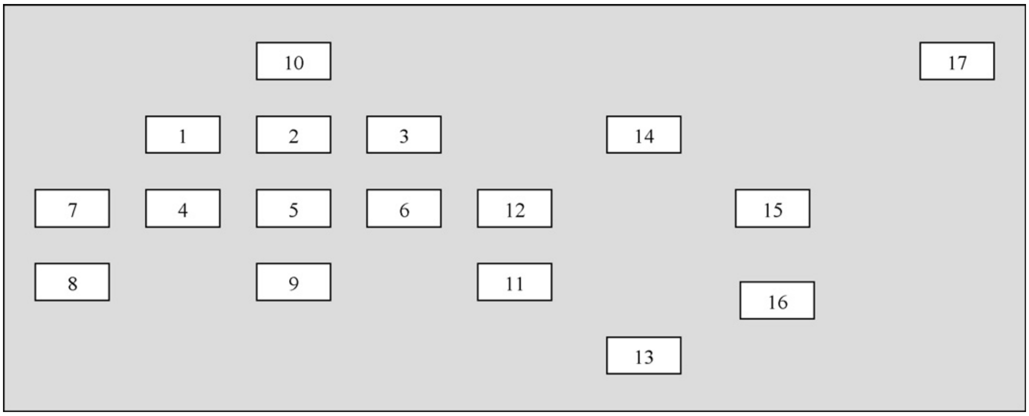

Fig. 10. The configuration of categories on the card sorting table.

4.4. The subject specialist classification

The subject specialist applied a *principle of division* based on four major categories:

i. properties;
ii. synthesis;
iii. understanding of mechanism; and
iv. method.

These four categories were in turn subdivided and a total of 17 categories was arrived at. A rectangular table used for the card sorting was applied in a way that a spatial representation was accomplished (Fig. 10).

The configuration of 17 piles of cards was arranged so that categories with a similar or connecting research foci were located in each others vicinity and each pile was assigned a label indicating the perceived research focus (Table 4).

All but three articles were contained in piles with a size of at least three, and three articles constituted one category each. The distribution of articles over categories is shown in Fig. 11.

The difference between the complete link size distribution (cf. Fig. 5) and the subject specialist's distribution of cards over categories is of a magnitude that clearly shows that two much deviating principles of classification have

Table 4
The subject specialist's labels

Cluster	Labels
1	Synthesis product
2	Synthesis reaction
3	Catalysis
4	Stereo selective synthesis-product
5	Stereo selective synthesis-reactions
6	Stereo selective synthesis-catalysis
7	Total synthesis
8	Total synthesis & medicinal chemistry
9	Stereo selective synthesis-racemization
10	Synthesis evaluation
11	Peptide synthesis
12	DNA templated organic synthesis
13	Stereo selective reaction mechanism
14	Reaction mechanism
15	DNA properties
16	Peptide structure
17	Nano

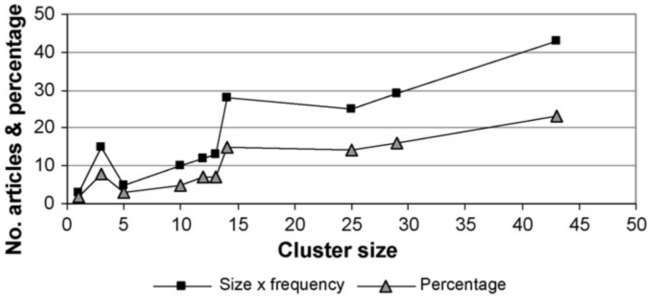

Fig. 11. The size–frequency distribution of card-sorting piles of 183 articles, and their percentages of a certain pile-size. Ten different sizes were found.

been applied. The two classifications were next compared with regard to the concentration of articles to categories. The concentration of articles to categories was assessed applying *Pratt's measure* of concentration which is of general use when one wants to see how concentrated or spread out items (here articles) are when partitioned into categories. It was originally suggested with the purpose of providing an index of concentration for rank–frequency distributions which permits comparisons of subject and journal concentration in various fields (Pratt, 1977). The starting point is the theoretical assumption that all articles are evenly distributed over *n* categories and the deviation from this norm is then measured. Pratt's measure is given as

$$C = \frac{2[((n+1)/2) - q]}{n-1}, \tag{4.4}$$

where *C* is the Pratt's measure of concentration, *n* the number of categories, and *q* is the sum of rank times frequency for each category, divided by the total number of articles (Pratt, 1977).

This measure will range between 0 and 1, where the most concentrated case (only one category) takes on the value of 1 and the "even" distribution the value of 0.[6] The value of *C* was 0.17 for the complete link classification and 0.57 for the subject specialist's, hence the complete link classification generated the more extreme classification. A final quantitative comparison of the two classifications was made with regard to the general agreement between the two classifications pair wise locations of articles in clusters. *Rand Index* is a measure of the similarity between two cluster solutions *Y* and *Y*′, of the same *N* data points. This similarity can be defined as $c(Y,Y')$ equal to the number of similar assignments of point-pairs normalized by the total number of point-pairs (Rand, 1971). Given two cluster solutions *U* and *V*, let *a* be the number of pairs of objects in the same cluster in *U* and in the same cluster in *V*, *b* be the number of pairs of objects in the same cluster in *U* but not in *V*, *c* the number of pairs of objects in the same cluster in *V* but not in *U* and *d* be the number of pairs of objects in different clusters in both cluster solutions. Then, $a+d$ is the total number of agreements and $c+b$ the total number of disagreements and Rand index $a+d/a+b+c+d$ (Yeung & Ruzzo, 2001). Rand index takes on values in the interval [0,1] and when the two partitions agree perfectly, the index is 1. A problem with this index is that values tend increase as the number of clusters increase (Yeung & Ruzzo, 2001). In order to correct for this Hubert and Arabie (1985) suggested an *adjusted Rand Index*, which has been recommended for general use (Everitt et al., pp. 182).

The results showed a Rand Index of 0.878, which may be somewhat misleading with concern to the larger number of clusters generated by the complete link clustering. A more comprehensible value of agreement, 0.102, was arrived at by applying the adjusted Rand Index. The contingency table (Table 5) details the agreement. The most striking deviation concerns the distribution of 43, 25 and 29 articles over three subject specialist classes, v2, v3 and v5. Over these three classes, approximately half of all articles are dispersed and the number of intersections with the complete link clusters is 53 out of 99. Several (14) complete link clusters are completely contained within subject specialist classes, constituting fractions.

[6] If one assumes that there are *n* categories and a total of *t* articles, in the even distribution there would be *t*/*n* articles in each category (ibid.).

Table 5
Contingency table where u*i* stands for the *i*th class in the ‘complete link distribution’, v*j* for the *j*th class in the ‘subject specialist distribution’

	v1	v2	v3	v4	v5	v6	v7	v8	v9	v10	v11	v12	v13	v14	v15	v16	v17
u1	0	4	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
u2	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0
u3	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0
u4	2	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
u5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0
u6	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0
u7	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0
u8	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
u9	0	0	0	0	1	0	3	1	0	0	0	0	0	0	0	0	0
u10	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
u11	0	0	0	0	5	1	0	0	0	0	0	0	0	0	0	0	0
u12	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0
u13	0	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
u14	0	0	1	0	1	0	0	0	0	0	0	0	0	3	0	0	0
u15	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0
u16	0	0	0	0	1	1	0	0	1	0	0	0	0	1	0	0	0
u17	0	0	0	1	1	5	0	0	0	0	0	0	0	0	0	0	0
u18	1	1	1	0	0	0	5	0	0	0	0	0	0	0	0	0	0
u19	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
u20	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
u21	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0
u22	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
u23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
u24	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
u25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0
u26	0	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
u27	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
u28	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
u29	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
u30	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
u31	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
u32	1	0	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0
u33	0	1	1	0	1	1	0	0	0	0	0	0	1	0	0	0	0
u34	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
u35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0
u36	0	0	2	1	0	1	0	0	0	0	0	0	0	0	0	0	0
u37	1	3	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0
u38	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
u39	0	0	0	0	3	1	0	0	0	0	0	0	0	0	0	0	0
u40	1	4	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
u41	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
u42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0
u43	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0
u44	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0

5. Final discussion

As stated in the methodology section, one may zoom in and out of a field and the analyst may choose a more or a less far-reaching study. In this study, explicit ‘close-up pictures’ of a field’s research fronts were developed. The findings in this study highlighted that despite a subject expert’s cumulated experience, he/she may not anticipate the many associations generated by the citing collective of peers. This was orally expressed by the trusted subject specialist and unexpected classifications may provide with topical and useful information. Hence, subjective evaluation may have its limitations as “. . .maps aggregate data about scientific fields in way that no individual expert, with his or her background and perspective would be able to do, and may not even recognize” (Rip, 1988). The strong deviations found between the two classifications are not surprising as the method applied aims at the identification of coherent 'cliques' and one may intuitively assume that such cognitive cores would be of a 'not too big' a size and strongly related to the time dimension. The subject specialist's classification, though structured and motivated, may to some extent be grounded in curricula of the discipline, and not reflect temporal currents in the research front. In a way, one may regard the two classifications extremes and reflecting strongly deviating principles of classification.

Reviewing the main features of the applied method, we can summarize findings and the consequences of its application:

i. It implies a strong reduction of the original document population.
ii. It retains the stronger links of normalized coupling strength and deletes the rest.
iii. It deletes all document relations that do not contribute to the generation of complete graphs.
iv. With regard to (iii), it may in some cases cut of continuing research themes.
v. It seems to do the job it was assigned to do; finding coherent and relevant 'cliques' in the research-front network.
vi. It seems not to comply with a subject specialist's a priori expectation of a field's structure.

One may assume that the high quality of clusters, both from a statistical as well as a subjective-judgement point of view, may be used for the identification and retrieval of groups of relevant and current documents in the research front. It may also be suggested that the proposed method be applied as a point of departure for an augmented analysis where links from 'cliques' are traced, and additional relevant documents are added step by step to a growing core. This should be executed in collaboration with a subject specialist. Presumably, 'cliques' containing articles with several strong ties

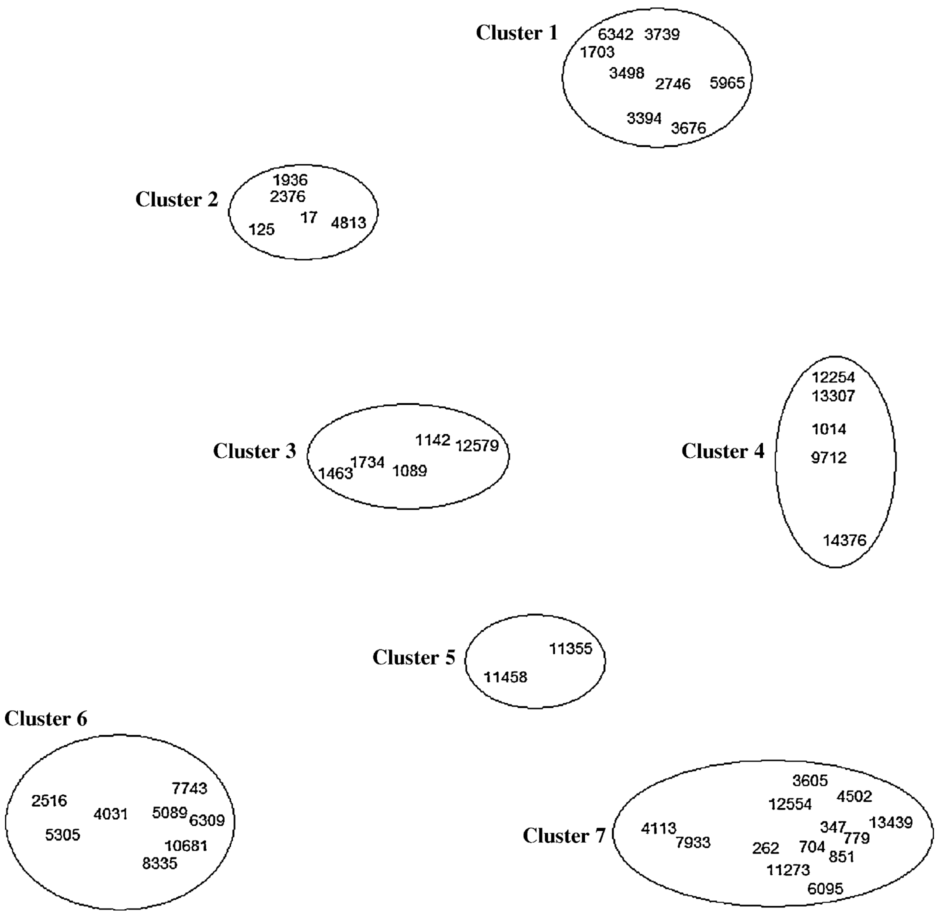

Fig. 12. Forty-six core documents from the study clustered by the complete links method and analyzed by MDS. Clusters were superimposed on the map and in perfect agreement with the MDS. The configuration was deemed completely relevant by the subject specialist. Kruskal's stress 0.03.

to a delimited number of other articles, mirror research-front scientists building on each others findings.[7] If so, one may hypothesize that clusters identified by the proposed method, to some extent, also mirror important indicators of social organisation like "...informal discussions, published collaborations, relationships with teachers, and the influence of colleagues upon the selection of research problems and techniques" (Crane, 1972, p. 41). This is, however, not covered for in the present study, but it may well be extended to cover also such aspects.

This study also raises the question of the shapes and configurations of research fronts. In that context, the author of this paper made relentless efforts to catch the overall structure of the research front(s), applying a variety of cluster methods, without success. These methods are based on the presumption that there is an inherent hierarchical structure in data. It may be hypothesized that such a structure is not immediately recognizable and that the associations between research front documents are better pictured as high density cliques in networks. An illustration of this is given in Fig. 12 where a set of core documents identified during this study has been mapped by both MDS and the complete link method. The complete link method was applied to identify complete graphs, rather than hierarchical structures, while MDS was used to depict the network structure. MDS may be the better method for the purpose of analyzing such networks, but is as for now limited in terms of the number of objects that can be analyzed. However, new methods are being developed which may facilitate this type of research. An example of this is Pajek, which is a program for the analysis of very large networks, not restricted to a maximal hundred vertices (cf. Batagelj & Mrvar, 2003). Conclusively, findings in this study suggest further research on the application of bibliographic coupling for research-front mapping.

References

Batagelj, V., & Mrvar, A. (2003). *Analysis and visualization of large networks*. Vol. 41, University of Ljublijana, Preprint series, p. 871.

Braam, R. R., Moed, H. F., & van Raan, A. J. F. (1991a). Mapping of science by combined cocitation and word analysis. 2. Dynamic aspects. *Journal of the American Society for Information Science*, *42*(4), 233–251.

Braam, R. R., Moed, H. F., & van Raan, A. J. F. (1991b). Mapping of science by combined cocitation and word analysis. 2. Dynamic aspects. *Journal of the American Society for Information Science*, *42*(4), 252–266.

Biglan, A. (1973). The difference of subject matter in different academic areas. *Journal of Applied Psychology*, *57*(3), 195–203.

Crane, D. (1972). *Invisible colleges: diffusion of knowledge in scientific communities*. Chicago & London: The University of Chicago Press.

Everitt, B. S., Landau, S., & Leese, M. (2001). *Cluster analysis* (4th ed.). London: Arnold.

Franklin, J. J., & Johnston, R. (1988). Co-citation bibliometric modelling as a tool for S&T and R&D management: Issues, applications, and developments. In A. F. J. van Raan (Ed.), *Handbook of quantitative studies of science and technology*. Amsterdam: North Holland.

Glänzel, W., & Czerwon, H. J. (1995). A new methodological approach to bibliographic coupling and its application to research-front and other core documents. In *Proceedings of the 5th international conference on scientometrics and informetrics* (pp. 167–176).

Glänzel, W., & Czerwon, H. J. (1996). A new methodological approach to bibliographic coupling and its application to the national, regional and institutional level. *Scientometrics*, *37*(2), 195–221.

Griffith, B., Small, H., Stonehill, J., & Dey, S. (1974). The structure of scientific literatures. II. Toward a macro- and microstructure for science. *Science Studies*, *4*, 339–365.

Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of Classification*, *2*, 193–218.

Jarneving, B. (2001). The cognitive structure of current cardiovascular research. *Scientometrics*, *50*(3), 365–389.

Johnsbaugh, R. (1997). *Discrete mathematics* (4th ed.). New Jersey: Prentice Hall International.

Kaufman, L., & Rousseeuw, P. J. (2005). *Finding groups in data: An introduction to cluster analysis*. Hoboken, New Jersey: John Wiley & Sons.

Kessler, M. M. (1960). *An experimental communication center for scientific and technical information*. Lincoln Laboratory: Massachusetts Institute for Technology.

Kessler, M. M. (1962). An experimental study of bibliographic coupling between technical papers. Massachusetts Institute for Technology, Lincoln Laboratory.

Kessler, M. M. (1963a). Bibliographic coupling between scientific papers. *American Documentation*, *14*(1), 10–25.

Kessler, M. M. (1963b). Bibliographic coupling extended in time: Ten case histories. *Information Storage and Retrieval*, *1*, 169–187.

Kessler, M. M. (1965). Comparison of the results of bibliographic coupling and analytic subject indexing. *American Documentation*, *16*(3), 223–233.

Martyn, J. (1964). Bibliographic coupling. *Journal of Documentation*, *20*, 236.

McCain, K. W. (1990). Mapping authors in intellectual space—A technical overview. *Journal of the American Society for Information Science*, *41*(6), 433–443.

McCain, K. W. (1986). Cocited author mapping as a valid representation of intellectual structure. *Journal of the American Society for Information Science*, *37*(3), 111–122.

Miller, G. A. (1969). A psychological method to investigate verbal concepts. *Journal of Mathematical Psychology*, *6*, 169–191.

Mubeen, M. A. (1995). Bibliographic coupling: an empirical study of economics. *Annals of Library Science and Documentation*, *42*(2), 41–53.

Noyons, E. C. M. (1999). *Bibliometric mapping as a science policy and research management tool*. Leiden: DSWO Press.

[7] This would presumably be verified by a large share of inter-citations.

Otte, E., & Rousseau, R. (2002). Social network analysis: A powerful strategy, also for the information sciences. *Journal of Information Science*, *28*(6), 441–453.

Persson, O. (1994). The intellectual base and research front of JASIS 1986–1990. *Journal of the American Society for Information Science*, *45*(1), 31–38.

Peters, H. P. F., Braam, R. R., & van Raan, A. F. J. (1995). Cognitive resemblance and citation relations in chemical engineering publications. *Journal of the American Society for Information Science*, *46*(1), 9–21.

Pratt, A. D. (1977). A measure of class concentration in bibliometrics. *Journal of the American Society for Information Science*, *28*, 285–292.

Rand, W. M. (1971). Objective criteria for the evaluation of cluster methods. *Journal of the American Statistical Association*, *66*(336), 846–850.

Rip, A. (1988). Mapping of science: Possibilities and limitations. In A. F. J. van Raan (Ed.), *Handbook of quantitative studies of science and technology*. Amsterdam: North Holland.

Sen, S. K., & Gan, S. K. (1983). A mathematical extension of the idea of bibliographic coupling and its applications. *Annals of Library Science and Documentation*, *30*(2), 78–82.

Sharabchiev, Y. T. (1988). Comparative analysis of 2 methods of cluster analysis of bibliographic citation. (In Russian). *Naucno-techniceskaja informacija*/2.

Sharada, B. A., & Sharma, J. S. (1993). A study of bibliographic coupling in linguistic research. *Annals of Library Science and Documentation*, *40*(4), 25–137.

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, *24*, 265–269.

Small, H., & Griffith, B. (1974). The structure of scientific literatures. I. Identifying and graphing specialities. *Science Studies*, *4*(1), 17–40.

Small, H., & Sweeney, E. (1985). Clustering the *Science Citation Index* using cocitations. I. A comparison of methods. *Scientometrics*, *7*(3–6), 391–409.

Tijssen, R. J. W. (1992). *Cartography of science: Scientometric mapping with multidimensional methods*. Leiden: DSWO Press, Leiden University.

Vladutz, G., & Cook, J. (1984). Bibliographic coupling and subject relatedness. *Proceedings of the ASIS annual meeting*, *47*, 204–207.

White, H. D., & Griffith, B. C. (1981). Author cocitation: A literature measure of intellectual structure. *Journal of the American Society for Information Science*, *32*, 163–171.

Yeung, K. Y., Ruzzo, W. L. (2001). Details of adjusted Rand Index and clustering algorithms. Supplement to the paper “An empirical study on principal component analysis for clustering gene expression data”. In Technical Report 20000 UW-CSE-00-11-01. Department of Computer Science and Engineering, University of Washington.