Journal of Informetrics 8 (2014) 369–383

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Optimizing SCImago Journal & Country Rank classification by community detection

Antonio J. Gómez-Núñez [a,*], Vladimir Batagelj [b], Benjamín Vargas-Quesada [c,e], Félix Moya-Anegón [d,e], Zaida Chinchilla-Rodríguez [d,e]

[a] CSIC, SCImago Research Group Associated Unit, Faculty of Communication and Documentation, Campus de Cartuja s/n, 18071 Granada, Spain
[b] University of Ljubljana, Faculty of Mathematics and Physics, Jadranska 19, 1000 Ljubljana, Slovenia
[c] University of Granada, Department of Information and Communication, Faculty of Communication and Documentation, Campus de Cartuja s/n, 18071 Granada, Spain
[d] CSIC, Institute of Public Goods and Policies, Albasanz 26-28, 28037 Madrid, Spain
[e] SCImago Research Group, Spain

ARTICLE INFO

Article history:
Received 11 June 2013
Received in revised form 14 January 2014
Accepted 16 January 2014
Available online 13 February 2014

Keywords:
Community detection
Clustering
SCImago Journal & Country Rank
Journal classification
Citation-based network

ABSTRACT

Subject classification arises as an important topic for bibliometrics and scientometrics, searching to develop reliable and consistent tools and outputs. Such objectives also call for a well delimited underlying subject classification scheme that adequately reflects scientific fields. Within the broad ensemble of classification techniques, clustering analysis is one of the most successful.

Two clustering algorithms based on modularity – the VOS and Louvain methods – are presented here for the purpose of updating and optimizing the journal classification of the SCImago Journal & Country Rank (SJR) platform. We used network analysis and *Pajek* visualization software to run both algorithms on a network of more than 18,000 SJR journals combining three citation-based measures of direct citation, co-citation and bibliographic coupling. The set of clusters obtained was termed through category labels assigned to SJR journals and significant words from journal titles.

Despite the fact that both algorithms exhibited slight differences in performance, the results show a similar behaviour in grouping journals. Consequently, they are deemed to be appropriate solutions for classification purposes. The two newly generated algorithm-based classifications were compared to other bibliometric classification systems, including the original SJR and WoS Subject Categories, in order to validate their consistency, adequacy and accuracy. In addition to some noteworthy differences, we found a certain coherence and homogeneity among the four classification systems analysed.

1. Introduction

Classification is a topic broadly covered in *Bibliometrics* and *Scientometrics* because of its significance in developing final bibliometric and scientometric outputs, mainly based on the scientific literature included in databases and repositories. Thus, the literature collected by these information and reference sources needs to be organized through an appropriate and consistent classification scheme. This is an essential objective not only for information retrieval purposes, but also for

* Corresponding author. Tel.: +34 958240923.
E-mail addresses: anxusgo@gmail.com (A.J. Gómez-Núñez), vladimir.batagelj@fmf.uni-lj.si (V. Batagelj), benjamin@ugr.es (B. Vargas-Quesada), felix.demoya@csic.es (F. Moya-Anegón), zaida.chinchilla@csic.es (Z. Chinchilla-Rodríguez).

1751-1577/$ – see front matter
http://dx.doi.org/10.1016/j.joi.2014.01.011

designing reliable and solid tools such as rankings, domain analysis or scientograms, all highly valuable in science policy design and science evaluation processes.

Normally, database subject classification schemes are constructed on the basis of a disciplinary structure that attempts to replicate the main fields and subfields of research and scientific knowledge recorded in the literature stored in databases. Then, scientific literature can be classified at journal or paper level. The most highly reputed scientific databases at present, namely Web of Science (Thomson Reuters, 2009) and Scopus (Elsevier, 2004), have very similar two-level hierarchical subject classification schemes consisting of subject areas at a high and wider level, and subject categories at a low and more specific level. In both databases, journals are assigned to one or more categories and their papers inherit the subject categories of the journals they belong to. In the case of the Web of Science (WoS), journal assignment is executed by ISI (currently, Thomson Reuters) staff, taking into account criteria such as journal titles or citation patterns (Pudovkin & Garfield, 2002).

The delimitation of scientific fields previous to developing disciplinary subject classification schemes may involve empirical and pragmatic techniques, or else automated procedures based on statistics and computerized methods. Of the latter, clustering analysis is a most valuable and popular method applicable in a wide variety of scientific fields, including *Library and Information Science*, *Psychology*, *Medicine* or *Biology*.

2. Related works

Many clustering algorithms and techniques have been developed to obtain optimal solutions for the classification problems mentioned above. Yet clustering methods have been most widely used by researchers dealing with information visualization techniques. Mapping the structure of scientific knowledge and research calls for a sound underlying classification of fields and subfields to be mapped. A total of 20 representative approaches for mapping science fields and their relations, working from Web of Knowledge and Scopus database literature, were compared and condensed by Klavans and Boyack (2009).

Clustering and mapping procedures may be conducted on different levels of aggregation, that is, using different units of analyses. At the journal level, numerous researchers have applied different cluster algorithms to journal–journal relation matrices, or networks based on citations, co-citations or bibliographic coupling. Chang and Chen (2011) applied the *minimum span clustering* (*MSC*) method to a citation square matrix of roughly 1600 SSCI journals. Leydesdorff, Hammarfelt, and Salah (2011) tried to merge a map of the humanities based on Thomson Reuters' A&HCI database in a global map of science previously developed (Rafols, Porter, & Leydesdorff, 2010), and used the k-core algorithm for mapping 25 specific A&HCI subject categories. Archambault, Beauchesne, and Caruso (2011) designed a scientific journal ontology aimed to simplify the output of bibliometric data and analysis. The new journal ontology was built on feedback from a previously existing journal classification whose categories were considered as "seeds" for the initial journal assignment. Three automatic classification procedures were executed, using either text or citation data from papers published in around 34,000 journals and conference proceedings from Scopus and WoS. However, the final solution was generated according to the iterative analysis of citation and references patterns between subject fields and journals.

Leydesdorff and Rafols (2012) collaborative work produced a study where a 9162 journal–journal citation matrix extracted from the 2009 volume of the SCI-Expanded was used to map interactive global journal maps. They compared several methods entailing different clustering algorithms to group journals into clusters. More recently, Börner et al. (2012) introduced a methodology to design an updated map of science requested by the University of California, San Diego (UCSD). A combination of text and link journal–journal similarity matrices based on Scopus and WoS data were used to build the map, after which journal clustering was executed on a filtered matrix derived from modified cosine similarities. Finally, the calculation of similarities among clusters, as well as their positions and relationships, made it possible to actually depict the UCSD map.

Lately, one research trend is to work with clustering algorithms for the analysis, validation, and improvement of classification schemes based on journals from various perspectives. The ECOOM research group of KU Leuven addresses this topic through several publications and different clustering algorithms (Ward clustering or Multi-level Aggregation Method, also known as Louvain method) applied to journal cross-citation and hybrid (text/citation) matrices (Janssens, Zhang, De Moor, & Glänzel, 2009; Zhang, Glänzel, & Liang, 2009; Zhang, Janssens, Liang, & Glänzel, 2010).

In contrast, by taking documents as the unit of analysis, Small (1999) developed a methodology to visualize and to obtain a hierarchical multidisciplinary map of science through a method combining fractional citation counting of cited papers, co-citation single-linkage clustering with limits on cluster size, and two-dimensional ordination according to a geometric triangulation process. Ahlgren and Colliander (2009) studied different document-document similarity approaches based on text, coupling and a combination of both as well as several methods to map and classify a set of 43 documents from the journal *Information Retrieval*. Complete-linkage clustering was applied to group articles and the final result of assignment was compared with an expert-based classification using an adjusted Rand Index. Similarly, Boyack et al. (2011) employed a combination of graph layout and average-link clustering to different text-based similarity-measure matrices constructed through relevant information from titles, abstracts, and MeSH subject headings of 2.15 million papers extracted from the Medline database. They compared and assessed nine similarity approaches through Jensen–Shannon divergence and concentration measures. Later on, Waltman and Van Eck (2012) faced an even more complex challenge by designing a detailed methodology to create a publication-level classification system using a multilevel clustering algorithm on a direct citation (disregarding the direction) network comprising nearly 10 million publications. In their opinion, their methodology' strength is rooted in transparency and simplicity, as well as its modest computing and memory requirements. Boyack, Small, and Klavans (2013) introduced the reference pair proximities as a new variable to improve the accuracy of co-citation clustering. To do so, they used a corpus of 270,521 Scopus full-text documents from 2007, comparing the results of traditional co-citation clustering approach to their new co-citation clustering, which yielded a manifest improvement in accuracy.

Generally, clustering procedures on networks and matrices involves complex calculations. This is more relevant when large datasets are being manipulated, since hardware and software requirements would be high. The fact that the visualization of clustered data should be clear and comprehensible is another matter to bear in mind. Software such as VOSViewer (Van Eck & Waltman, 2010) and Pajek (Batagelj & Mrvar, 1997; de Nooy, Mrvar, & Batagelj, 2012) are known to be good tools for network analysis and information visualization, especially when large networks have to be manipulated. VOSViewer moreover features its own classification algorithm, whereas Pajek integrates different clustering algorithms that can be run easily once a dataset is adapted to an appropriate format for the software.

3. Objectives

The main goal of this study is to optimize and update the journal classification results of the SCImago Journal & Country Rank (SJR) platform (SCImago, 2007) via clustering techniques. Using Pajek software, we ran two automatic classification algorithms to detect and extract communities (subject clusters) from a SJR journal network combining three citation-based measures. The set of automatically extracted communities represents the disciplinary structure of science and research recorded in SJR journals. The resulting cluster-based systems are compared to other classification systems, such as WoS Subject Categories and the original SJR Classification, to validate their consistency and accuracy by analysing the strengths and weakness of the results.

4. Material

Our data set, covering a total of 18,891 journals for a two-year time window (2009–2010), was gathered from SCImago Journal & Country Rank (SJR) database. In this set, only cited references going back from 2010 to the year 2000 were contemplated. All references were counted at paper level and later aggregated to journal level.

5. Methods

In order to clarify and favour a better understanding of the distinct procedures developed in performing our study, we present this section in seven stages covering and detailing the steps to be followed.

5.1. SJR journal classification: the starting point

The Scopus classification system, and by extension SJR original classification, constitutes an a priori two-level hierarchical classification system of an up-bottom nature. In its first level, the classification covers a total of 27 broad *subject areas* which, in turn, comprise a set of 308 specific subject categories at the second level. Then, journals recorded in the database are ascribed to one or several subject categories. Area and category tags were determined on the basis of All Science Journal Classification (ASJC). Each subject area normally includes a subject category taking the same tag, followed by '*miscellaneous*'. Journal assignment to categories was done on the basis of item adscription.

Then, SCImago Research Group introduced improvements based on journal scope analysis and the feedback from journal editors. The latter is an interesting source of information to take into account. Indeed, Archambault et al. (2011) claim to appraise feedback from researchers and practitioners using their journal ontology to persist in refining journal assignment. In our experience however, despite various attempts, vaster improvement of SJR journal classification was apparently needed to remove inconsistencies inherited from Scopus, allowing final users to customize the journal-sets of SJR subject categories and even generate tailored rankings (Jacsó, 2013). Previous work based on SJR journal reference analysis (Gómez-Núñez, Vargas-Quesada, Moya-Anegón, & Glänzel, 2011) was oriented to this end.

5.2. Journal citation-based relatedness measures: calculation and formatting

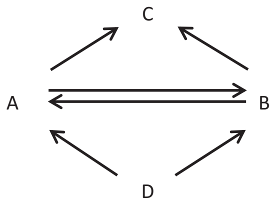
A plenty of publications dealing with classification and mapping of science and research have adopted text-based networks (Cantos-Mateos, Vargas-Quesada, Chinchilla-Rodríguez, & Zulueta, 2012; Liu, Hu, & Wang, 2011), citation-based networks (Leydesdorff & Rafols, 2012; Rafols & Leydesdorff, 2009) or combination of both (Glänzel, 2012; Janssens et al., 2009). Boyack and Klavans (2010) applied *Jensen–Shannon divergence* and *concentration* metrics to prove the accuracy of clustering solutions emerging from different citation-based mapping methods. The results revealed the best performance using the bibliographic coupling approach, followed closely by co-citation and direct citation. Waltman and Van Eck (2012) analysed advantages and disadvantages of three citation-based approaches, then choosing direct citation as the relatedness measure for constructing a publication-level classification. Their decision was primarily based on saving computer resources for processing a large data set of almost 10 million publications. Yet they argued that direct citation should provide strongest relatedness links between publication, as opposed to the more indirect mechanisms of co-citation and bibliographic coupling.

At any rate, they noted that the use of direct citations leads to a loss of information when citations of earlier publications or later publications were not contemplated.

In this work, we explore three citation-based approaches to express a degree of relatedness between journals, namely, direct citation, co-citation and bibliographic coupling. By doing so, we can combine both strengths and weakness from each measure. This approach might be considered more 'fair' and balanced, offsetting the inherent weaknesses of *direct citation*, *co-citation* and *bibliographic coupling* when used separately. In view of these important considerations, we constructed three journal networks, one for each citation-based measure. These networks were calculated at the document level, then aggregated to journals. For co-citation and bibliographic coupling calculation, references co-occurring were only counted once per paper, following the binary counting concept described by Rousseau and Zuccala (2004) instead of what Vargas-Quesada and Moya-Anegón (2007) named latent co-citation.

5.3. Citation-based measures combination

Once the three citation-based networks were generated, we combined them into a new one collecting pairwise journals and their relatedness strength as expressed by the sum of direct citation, co-citation and bibliographic coupling links. By doing so, we arrived at a final network based on raw data and containing what Persson (2010) named *Weighted Direct Citation* (WDC) links. Below we display the diagram used by Persson to integrate these three citation-based measures and calculate WDC. We introduced a small shift referring to both senses of the direct citation links, however.



Thus, we used this formula in combining citation based-measures:

$$c_{ij} = \mathrm{cu}_{ij} + \mathrm{cc}_{ij} + \max(\mathrm{ci}_{ij}, \mathrm{ci}_{ji})$$

where cu_{ij}, coupling; cc_{ij}, co-citation; ci_{ij}, direct citation from *i* to *j*; and ci_{ji}, direct citation from *j* to *i*.

Also, knowing that A, B, C and D are journals, we can adapt this formula according to Persson's diagram in this way:

$$c_{ij} = \mathrm{ABC} + \mathrm{DAB} + \max(\mathrm{AB}, \mathrm{BA})$$

5.4. Network normalization

At this point, the final network resulting from the aggregation of raw data links was normalized using *Geo normalization* formula as follows:

$$s_{ij} = \frac{c_{ij}}{\mathrm{sqrt}(c_i \times c_j)}, \quad c_i = \mathrm{sum}\{j : j! = i : c_{ij}\}$$

This similarity measure resembles the Cosine one. It divides elements of the matrix by the geometric mean of both diagonal elements (Batagelj & Mrvar, 2003). Thereby, raw data were corrected and relatedness values between pairwise journals were transformed to values ranging from 0 to 1. This avoids problems from misleading representations and overestimation of some science fields characterized by strong citation habits or covering large-size journals with a high power of attraction.

5.5. Clustering procedures

The next step in our methodology was to run clustering algorithms included in Pajek software on the normalized network. Pajek integrates several clustering methods in order to decompose networks by extracting different partitions such as islands, k-neighbours or block modelling. However, after several initial tests, we targeted on community detection algorithms: *VOS Clustering* (Waltman, Van Eck, & Noyons, 2010) and *Louvain Method* (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008). Both methods are grounded in modularity clustering proposed by Newman and Girvan (2004). Whereas the Louvain method optimizes modularity, VOS Clustering focuses on optimizing a quality function (Batagelj & Mrvar, 2011). For this experiment, we chose Louvain and VOS methods based on Multilevel Coarsening + Single Refinement. Moreover, we had to set up several options regarding *resolution parameters*, *random restarts*, *maximum number of levels in each iteration*, and *maximum number of repetitions in each level*. Here, we fixed just the same options for VOS and Louvain algorithms. Firstly, we introduced distinct values in *resolution parameter*, moving them from 10 to 20 in order to get different Pajek partitions depicting diverse solutions in decomposing the network and producing different sets of clusters or communities. Then, the remaining parameters were configured with default values.

By analysing certain relevant indicators for each parameterized clustering algorithm solution, basically, the *number of clusters generated* and the *number of journal per clusters*, we estimated that network decompositions providing between 250 and 300 groups would be interesting for our final journal classification objective. Here, some important issues were considered. Firstly, we took into account the 250 subject categories currently included in the WoS database, since this scientific information source is not only an international referent within bibliometric and scientometric fields, but also for scientists and researchers in general. Presently, SJR takes in 308 subject categories; therefore, we thought that a final set of 250–300 categories would provide a balanced and refined subject structure. This point was reinforced with the experience acquired in a previous work (Gómez-Núñez et al., 2011). There we noticed a regular behaviour in grouping journals, revealing a strong concentration of them in a few leading categories from a final set of 198 SJR categories. These leading categories are characterized by high attractiveness, especially when an iterative reference analysis method was used. As commented earlier, this behaviour may be derived from the citation habits of some scientific fields with an intense and well-defined citation practice, such as *Medicine* and allied science, or some social science subfields as *Economics* or *Education*.

Apart from the aforementioned indicators, we applied some others (see Results section) to VOS and Louvain partitions matching with different 10–20 resolution parameters and we proceeded to compare the results of both. Every partition was executed in Pajek and later saved to files, to be processed using spreadsheets and statistical software. Concretely, we selected VOS partitions referring to resolution parameter 15, while a resolution parameter 18 was appointed in the Louvain case. In this decision, we basically looked for similar partitions in terms of the final number of clusters generated by VOS and Louvain methods under the premise of obtaining comparable results and appraise the classification solutions of both clustering algorithms. Besides, we established a threshold to define a minimum cluster size of 10 journals, discarding all clusters that did not comply with this requirement.

5.6. Labelling

After executing automatic clustering techniques, we had to label the different subject groups or communities depicted by both algorithms and recorded in the selected partitions. To this end, we designed a multi-phase approach to resolve any discrepancies arising. We should explain at this point that our study proposes journal multi-assignment. Nevertheless, journal multi-assignment was due to labelling process; the clustering methods used led to journal single-assignment per cluster.

5.6.1. Labelling through SJR category tags

In a first approach, we took into consideration the citation frequencies from journals to former SJR categories. We counted how many times journals forming part of a cluster cited original categories from SJR. After that, frequencies were transformed into percentages and into weighted scores using the tf-idf formula by Salton and Buckley (1988) which we adapted in our particular case so that:

$$w_{i,j} = \mathrm{catf}_{i,j} \times \log\left(\frac{N}{\mathrm{cluf}_i}\right)$$

where $w_{\mathrm{i,j}}$, total weighted score; $\mathrm{catf}_{i,j}$, raw frequency of category '*i*' into cluster '*j*'; *N*, total number of clusters; and cluf_i, number of clusters containing category '*i*'.

Then, all the categories were ranked by tf-idf scores and only those categories amounting to at least 33% over the total set of references cited by journals forming distinct clusters were selected to delineate the topic of clusters. At this phase, clusters were labelled using from one to four SJR category tags. The cluster categories would hence be inherited by the journals they grouped. From a different standpoint, we can assert that journals were allocated to one up to four SJR categories. Although many research papers defend a single and exclusive assignment of journals to clusters or categories (Archambault et al., 2011; Thijs, Zhang, & Glänzel, 2013; Waltman & Van Eck, 2012), there are strong reasons to reconsider journal multi-assignment. Generally, most scientific journals do not cover a unique topic. This can be corroborated by simply having a look at journal scopes. In some cases, authors have an interest in publishing within journals outside of their expert field in order to attain higher prestige, visibility or impact. Moreover, current science often follows an interdisciplinary and collaborative model, with several fields involved in solving different problems, facing new challenges or contributing to a continuous development of science and research. Finally, we are aware that journal multi-assignment is carried out in original SJR journal classification, and we mean to keep taking this approach, but hopefully improving it.

5.6.2. Labelling through significant words of journal titles

This labelling approach was adopted in two particular cases:

(1) When using category tags, we found two clusters with exactly the same categories assigned, therefore representing two identical subject groups.
(2) In the whole labelling procedure, *Miscellaneous* and *Multidisciplinary* categories were rejected. After removing these categories, percentages and tf-idf scores were re-calculated. However, in some clusters the number of journals was

lower than the number of links pointing to SJR categories. This did not satisfy the condition of at least one link to category per journal.

In the two instances noted, we reconsidered the labelling approach for clusters by using a textual component, such as significant words extracted from journal titles. After counting them, the frequencies of most repeated words were taken so as to delineate the subject topic of clusters. To support the text-based labelling stage and fine-tune in denoting clusters, we used a *Voyeur Tools* platform that provides a set of online text analysis tools forming part of the *Hermeneuti*.ca collaborative project (Sinclair & Rockwell, 2009).

5.7. Validating classification proposals

In closing our methodological incursion, a validation of classifications generated by algorithms was desirable. There are different approaches aimed to this end. Expert assessment, while perhaps the best option, is very time- and cost-consuming. We opted for comparison with some other classification systems, particularly the original SJR classification, since the journal data set is involved. Yet we included a comparison with ISI Subject Categories, system behind the database of reference in bibliometric scope – WoS (and consequently, JCR + Arts & Humanities). To this end we prepared a combined list consisting of SCI + SSCI journals collected from the JCR 2010. Because JCR does not include journals of Arts & Humanities, an extra list of A&HCI journals of 2012 release was downloaded from Thomson Reuter's website and added. The final list of journals was integrated by 11,715 journals, 8005 pertaining to SCI, 2678 to SSCI and 1758 to A&HCI. Therefore, there is a certain level of overlapping: a total of 726 journals were covered by more than one index. Finally, the ISSN field served to generate matching between journals of SJR, WoS, VOS and Louvain classifications, leading to 9694 journals, 82.75% of the total set.

6. Results

6.1. Analysis of results derived from algorithm solutions

In an attempt to optimize and update SJR journal classification, we analysed and compared the results derived from VOS and Louvain clustering methods according to distinct indicators related to the proper performance of both algorithms, such as (1) *number of given clusters*, (2) *number of journals classified* after applying the threshold of 10 journals as the minimum cluster size, and (3) *mean number of journals per cluster*. Besides, we developed two indicators derived from the cluster labelling process: (4) *journal multi-assignment* and the (5) *mean number of categories per journal*. Again, we would like to remark that journal multi-assignment was a consequence of our labelling procedure and not due to VOS and Louvain performance, which carry out journal single assignment, being hard clustering techniques. As we pointed in the previous section, we projected around 250–300 journal subject groups to trace a basic and cohesive disciplinary structure in order to classify scientific journals. Therefore, we retained this premise during the parameterization of VOS and Louvain algorithms as well as when choosing final partitions, giving suitable results and better adapting to our final classification aim.

A simple glance at the figures presented evidences considerable parallelism in the distributions over the alternative resolution parameters of VOS and Louvain indicators. This might be viewed as normal, since both algorithms are grounded in the modularity clustering method proposed by Newman and Girvan (2004). If we focus on the (1) *number of clusters* offered by two clustering methods, the VOS algorithm is seen to need a resolution parameter lower than Louvain to arrive at a similar number of groups. Moreover, when a threshold of 10 journals as the minimum cluster size was set, the VOS algorithm presented a more balanced ratio between the clusters upholding this threshold and the clusters without doing so. Accordingly, the VOS partition with resolution parameter 15 returned just 270 clusters collecting ten or more journals from the total set of 848 clusters, equivalent to a ratio of 0.3184, or almost 32% of clusters satisfying the threshold. On the other hand, the Louvain partition with resolution parameter 18 produced a total set of 1064 clusters, with only 280 clusters reaching ten or more journals – a ratio amounting to 0.2632, meaning just slightly more than 26% of clusters with more than ten journals. Figs. 1 and 2 show the whole distribution of clusters according to the resolution parameter defined in the VOS and Louvain algorithms.

In the words of the authors of VOS, the resolution parameter included in their algorithm "helps to deal with the resolution limit problem of modularity-based clustering". They also claim that by introducing a sufficiently large value for the resolution parameter of their clustering technique, small clusters can always be determined, the number of clusters generated being larger when the resolution parameter is higher (Waltman et al., 2010). The final number of clusters is therefore directly proportional to the value of resolution parameter. Indeed, VOS and Louvain methods allowed us to classify the 18,891 journals forming part of the initial network explored through the set of clusters provided. Notwithstanding, many of these clusters were too small to form reliable or solid journal groups.

As we pointed out, only 31.84% of the total number of VOS clusters had a size higher than 10, while a mere 26.32% of clusters reached this threshold under the Louvain method. This phenomenon could be due to the use of citation and its derivatives as measurement units. Earlier on, we stated that some scientific fields report a strong concentration and power of attraction power of those citations linking to publications included in them. This could be attributed to marked citation habits occurring inside these fields. At any rate, the subject categories defining these fields are characterized by a great aggregation variance derived from the high quantity of citation received.

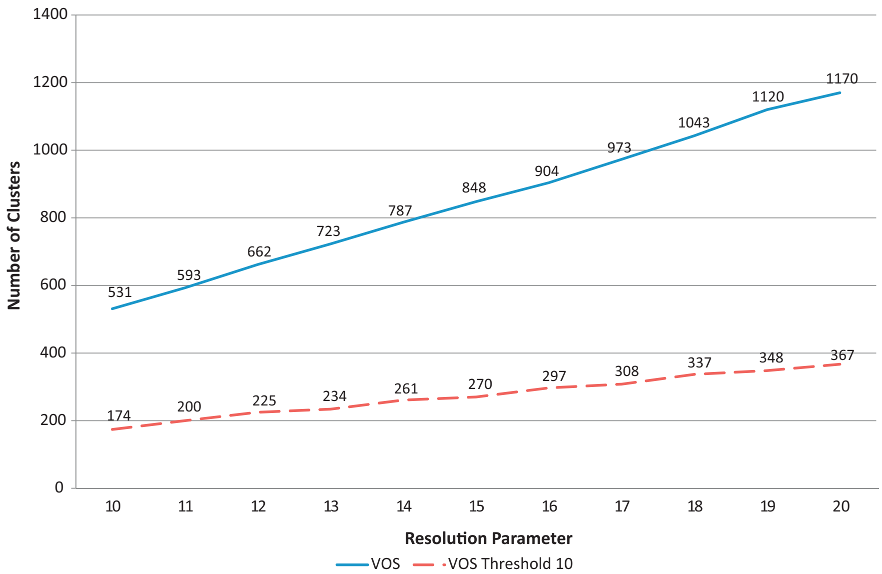


Fig. 1. VOS cluster distribution over the different resolution parameters tuned.

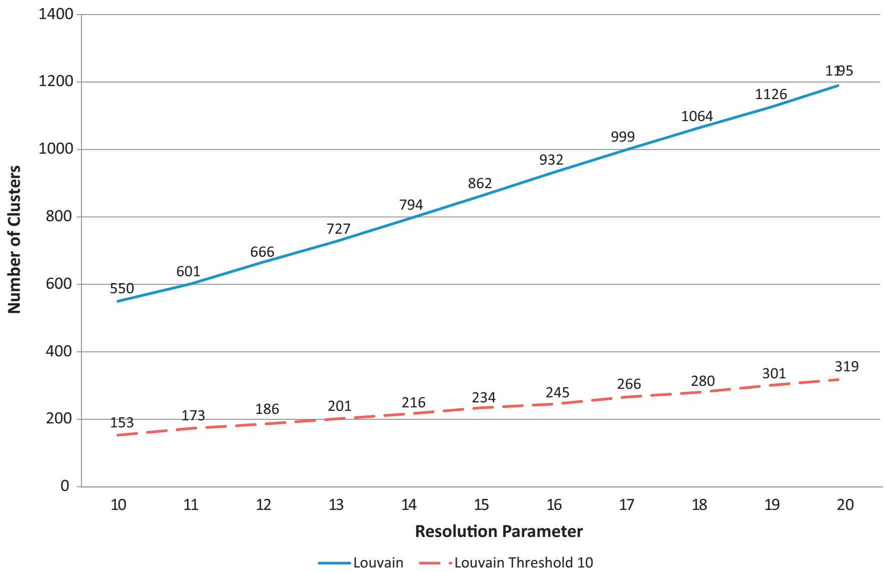


Fig. 2. Louvain cluster distribution over the different resolution parameters tuned.

By observing indicators related to the (2) *number of journals classified* and the (3) *mean number of journals per cluster*, we detected a general behaviour pointing to a better performance on the part of the VOS algorithm in classifying journals, that is, including journals in a particular cluster. In general, the mean of journals per cluster over the different resolution parameters returned by Louvain algorithm was higher. However, by examining the two partitions selected for our classification purpose, the mean number of journals per cluster was also a bit higher in favour of VOS resolution parameter 15. Fig. 3 shows the whole distribution of journals classified in VOS and Louvain clusters with regard to the distinct resolution parameters executed. Fig. 4 likewise shows the mean number of journals per cluster in two selected VOS and Louvain partitions.

Another example of the similitude of the results yielded by two clustering algorithms concerns the (4) *journal multi-assignment indicator* reflecting the number of journals assigned to one or multiple categories at once, and the (5) *mean number of categories per journal*. More details about the results for these two indicators in light of WoS and SJR results are offered in Table 1.

Table 1
Overall comparison among four classifications systems analysed. The number of classified journals in Louvain and VOS systems results from application of minimum cluster size threshold ($t \geq 10$).

	WoS	SJR	Louvain 18	VOS 15
Total set of journals	11,715	18,891	18,891	18,891
Number of classified journals	11,715	18,891	17,287	17,729
Number of categories	251	308	272	267
Mean number of journals per category	46.67	61.33	63.56	66.40
Mean number of categories per journal	1.54	1.61	1.48	1.50
Overlapping percentage	54.48%	60.73%	47.58%	49.89%
Journals changing their classification	–	–	5784	5874
Journals adding categories	–	–	3820	4192
Journals losing categories	–	–	4540	4603

	WoS Number of categories					SJR Number of categories					Louvain 18 Number of categories					VOS 15 Number of categories				
	1	2	3	4	+	1	2	3	4	+	1	2	3	4	+	1	2	3	4	+
Journal multi-assignment	6990	3432	986	261	46	12, 025	3893	1863	751	359	10, 806	4986	1237	256	0	10, 730	5251	1646	101	0
	59.7%	29.3%	8.4%	2.2%	0.4%	63.7%	20.6%	9.9%	4.0%	1.9%	62.5%	28.8%	7.2%	1.5%	0%	60.5%	29.6%	9.3%	0.6%	0%

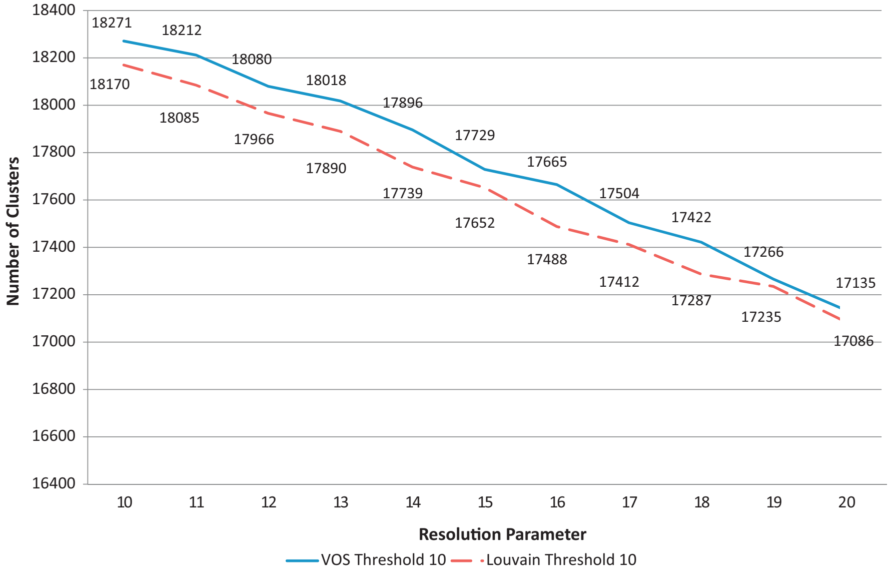


Fig. 3. Distribution of classified journals over the different resolution parameters tuned in VOS and Louvain clustering algorithms.

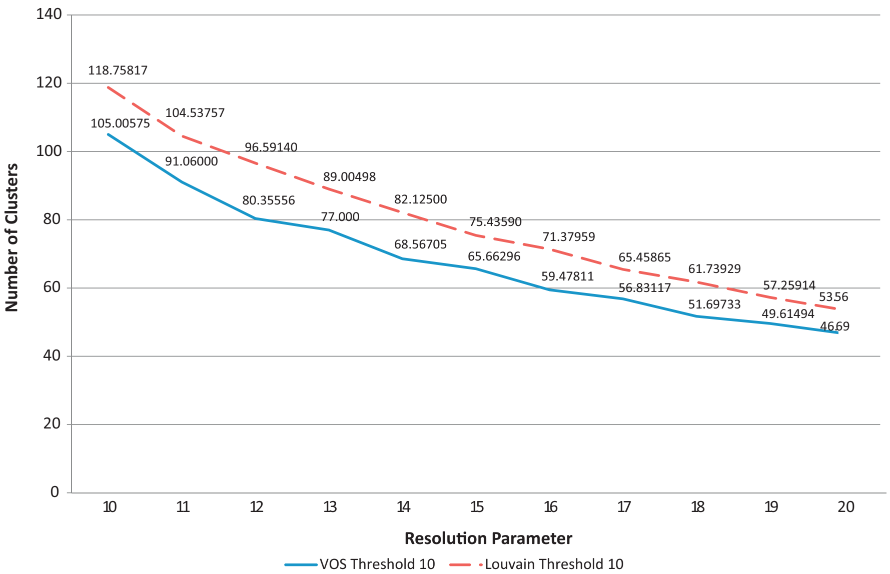


Fig. 4. Distribution of mean of journals per cluster over the different resolution parameters tuned in VOS and Louvain clustering algorithms.

6.2. Overall analysis and comparison among four different classification systems

Hitherto, we highlighted only the analysis of VOS and Louvain clustering algorithms on the basis of statistical data and indicators. At this moment, we detail results related to differences and resemblances in journal final classification obtained after applying both algorithms. To do so, we analyse and compare classifications originated by both clustering techniques, together with original SJR classification and WoS (ISI Subject Categories).

Table 1 captures overall data about the four classification systems compared. A closer look leads one to some important observations. Regarding the *total set of journals* included, the number of journals covered by SJR surpasses the WoS set by more than 1/3. Regarding the *number of classified journals* we can see that, after fixing threshold 10 in Louvain and VOS algorithms, the number of journals classified descended to 17,287 and 17,729, respectively. This is not a result of performance of algorithms, which were able to classify the whole set of 18,891 original SJR journals. Taking into account the huge effort to label small clusters, some even over 10-journal size, we decided that journals left out the final set would have to be classified separately using a different solution. One appropriate means could be the reference analysis method, described in a previous study (Gómez-Núñez et al., 2011), or the 'sibling journals' approach, involving those journals originally sharing former SJR categories, to extend their new cluster-based classification to journals under the threshold.

The next point to address is the final *number of categories* forming part of the classification system. At this point we should clarify why the number of clusters expressed in Figs. 1 and 2 for VOS and Louvain methods do not agree with the number of categories (subject clusters) displayed in Table 1. Figs. 1 and 2 collect the number of clusters generated by algorithms without labelling. On its side, Table 1 reflects the number of clusters after our labelling process. Our approach made it possible to have some clusters with different numbers and tags of categories assigned. For instance, cluster #82 in Louvain solution was labelled as 'Artificial Intelligence' + 'Information Systems' + 'Software' category tags, while cluster #90 was assigned to 'Artificial Intelligence' + 'Theoretical Computer Science' categories. The potential combinations of different number and tags of categories among the set of clusters is, therefore, the main reason behind the difference in the number of categories included in Table 1. The final number of categories in VOS and Louvain decreased meaningfully in comparison to the original SJR subject classification system, WoS being closer. This can be understood as an overall improvement, especially in the case of data referring to overlap (Table 1) and distribution of journals over categories (Table 2). After indicating the *number of classified journals* and the final *number of categories*, we can calculate the *mean number of journals per category*. The VOS, Louvain and SJR systems all surpass the value of 60 journals per category, the highest value being for VOS with a total of 66.4. The WoS system mean number only amounts to 46.67 journals per category, although it is true that WoS journal coverage is considerably reduced with respect to the other three systems.

Two further interesting observations from Table 1 concern *the mean number of categories per journal* and *overlapping percentage*. The latter was calculated by subtracting the number of records corresponding to journals covered by the system (or in other words, the set of journals under consideration) from the number of records referring to final journal multi-assignment (set of classified journals including multi-assigned journals), then multiplying by 100. Both indicators are correlated, and show the level of overlapping existing in the four classifications compared. The main difference resides in the fact that the mean number of categories per journal is expressed as per unit. The lowest level of overlapping was reached by Louvain system, followed closely by VOS. In both cases, overlapping levels are not over 50%. WoS and SJR systems surpass this level (the worst being SJR with 60.73%). In this sense, again VOS and Louvain methods evidenced better solutions than SJR and WoS. This point can be supplemented by comparing the *mean number of categories per journal* of VOS, which amounts to 1.50 categories per journal, and that of Louvain, just 1.48. Both figures also outperform the results of WoS and SJR, and show hardly significant differences in VOS and Louvain algorithms in terms of journal multi-assignment.

The following row displayed in Table 1 deals with the *journals changing their classification* from SJR to Louvain and VOS systems. In a first analysis, we can notice that figures for VOS and Louvain differ in just 90 journals. Thus, a total of 5874 VOS journals underwent changes with respect to their original classification, while a total of 5784 journals did it in Louvain. However, the distinct sizes of the whole set of journals classified under the two systems resulted in a higher ratio of journals altering their original classification in Louvain, with a 33.5%, in contrast to 33.1% of journals registering modifications in VOS. Regarding journals adding categories, there is a substantial difference of over 370 journals in VOS, amounting to 4192 journals under this indicator, while Louvain totals 3820 journals. In the meantime, the figures pointing to journals losing categories are much closer, with a difference of only 63 journals from VOS to Louvain system. We do not contemplate overlapping in our approach for calculating these three indicators. Accordingly, only those journals initially assigned to a given number of categories under SJR and later altering some of those given categories in VOS or Louvain are held to be journals changing their original classification. Instead, indicators referring to journals losing or adding categories point to journals firstly assigned to a given number of categories in SJR and later varying the number of assignments in VOS or Louvain.

Finally, Table 1 is displaying the figures related to *journal multi-assignment* in four classification systems compared. Here, the best assignment of journals to one category was for SJR, with 63.7% of the total set. The last place in the ranking was for WoS, with 59.7%. However, the four classification systems offered close percentages of journals assigned to one category. In aspiring to allow journal multi-assignment, the results obtained by Louvain and VOS can be judged as convenient, since they concentrated most journals in one or two categories. Louvain and VOS relative figures representing journal assignment executed in three and four categories outperform SJR and WoS systems by far. In addition, SJR and WoS systems made possible a journal assignment to more than four categories. Louvain and VOS solutions did not enable this kind of multi-assignment, giving rise to a more balanced classification system.

One last important issue to analyse among the four classification systems is the proper *distribution of journals over the set of subject clusters or categories* generated. Table 2 covers the top-20 categories regarding the number of journals included and expressed in raw data and percentage. We finally added a cumulative percentage in order to calculate the continuing aggregation of journals spread over categories. Here, the count of values and calculation of percentages were made considering journal overlap derived from journal multi-assignment in four classification systems. The slightest distribution of journals over the first ten categories resulted that of WoS. However, generally speaking, WoS, Louvain and VOS distributions are very similar. Admittedly, the SJR system achieved the worst distribution of journals over categories though having the largest set of classified journals. In many ways, this is a consequential effect of the remarkable concentration of journals on 'Medicine (miscellaneous)' category: 5.2% of the total set of SJR journals was included in this category.

A final consideration in view of Table 2 is how many really close or identical categories appear in 20-top ranking. A detailed analysis revealed that seven 20-top categories covered by the four classification systems were in all four systems. This signals a certain coherence and homogeneity, while changes in number and position of categories may imply, in the

Table 2
Top-20 categories of the four classification systems analysed.

WoS				SJR				Louvain resolution parameter 18				VOS resolution parameter 15			
Category	Num. of journals	% Journals	Cumul. %	Category	Num. of journals	% Journals	Cumul. %	Category	Num. of journals	% Journals	Cumul. %	Category	Num. of Journals	% Journals	Cumul. %
HISTORY	331	1.829	1.829	Medicine (miscellaneous)	1579	5.200	5.200	Sociology and Political Science	496	1.944	1.944	Electrical and Electronic Engineering	534	2.009	2.009
ECONOMICS	302	1.669	3.498	Education	524	1.726	6.926	Geology	417	1.634	3.579	Sociology and Political Science	480	1.806	3.816
BIOCHEMISTRY & MOLECULAR BIOLOGY	284	1.569	5.067	Sociology and Political Science	460	1.515	8.441	Literature and Literary Theory	415	1.627	5.205	Literature and Literary Theory	427	1.607	5.423
MATHEMATICS	276	1.525	6.592	Geography, Planning and Development	450	1.482	9.923	Geography, Planning and Development	393	1.540	6.746	Plant Science	393	1.479	6.901
PUBLIC, ENVIRONMENTAL & OCCUPATIONAL HEALTH	254	1.404	7.996	History	444	1.462	11.386	Electrical and Electronic Engineering	393	1.540	8.286	Geology	380	1.430	8.331
PHARMACOLOGY & PHARMACY	249	1.376	9.372	Electrical and Electronic Engineering	406	1.337	12.723	Psychiatry and Mental Health	372	1.458	9.744	Artificial Intelligence	371	1.396	9.728
ENGINEERING, ELECTRICAL & ELECTRONIC	247	1.365	10.737	Cultural Studies	375	1.235	13.958	Software	331	1.297	11.041	Education	369	1.389	11.116
NEUROSCIENCES	235	1.299	12.035	Social Sciences (miscellaneous)	374	1.232	15.190	Education	331	1.297	12.339	Software	352	1.325	12.441
MATHEMATICS, APPLIED	235	1.299	13.334	Economics and Econometrics	368	1.212	16.402	Hardware and Architecture	297	1.164	13.503	Psychiatry and Mental Health	341	1.283	13.724
PSYCHIATRY	233	1.288	14.621	Literature and Literary Theory	366	1.205	17.607	Religious Studies	297	1.164	14.667	Water Science and Technology	338	1.272	14.996
MATERIALS SCIENCE, MULTIDISCIPLINARY	219	1.210	15.831	Engineering (miscellaneous)	356	1.172	18.779	Applied Mathematics	283	1.109	15.776	Mathematics (general)	334	1.257	16.253
ENVIRONMENTAL SCIENCES	192	1.061	16.892	Psychology (miscellaneous)	351	1.156	19.935	Geochemistry and Petrology	282	1.105	16.882	Economics and Econometrics	318	1.197	17.449
LANGUAGE & LINGUISTICS	192	1.061	17.953	Public Health, Environmental and Occupational Health	335	1.103	21.039	Cultural Studies	264	1.035	17.916	Agronomy and Crop Science	312	1.174	18.623
SURGERY	186	1.028	18.981	Plant Science	327	1.077	22.116	Economics and Econometrics	252	0.988	18.904	Paleontology	309	1.163	19.786
CLINICAL NEUROLOGY	185	1.022	20.003	Language and Linguistics	327	1.077	23.193	History	250	0.980	19.884	History	300	1.129	20.915
PLANT SCIENCES	185	1.022	21.026	Psychiatry and Mental Health	325	1.070	24.263	Mechanical Engineering	245	0.960	20.844	Geography, Planning and Development	283	1.065	21.980
ONCOLOGY	181	1.000	22.026	Animal Science and Zoology	315	1.037	25.301	Civil and Structural Engineering	242	0.949	21.793	Mechanical Engineering	279	1.050	23.030
PHILOSOPHY	178	0.984	23.009	Mathematics (miscellaneous)	306	1.008	26.308	Rehabilitation	240	0.941	22.734	Developmental and Educational Psychology	278	1.046	24.076
EDUCATION & EDUCATIONAL RESEARCH	177	0.978	23.987	Cardiology and Cardiovascular Medicine	273	0.899	27.207	Mathematics (general)	235	0.921	23.655	Language and Linguistics	274	1.031	25.107
CELL BIOLOGY	174	0.961	24.949	Agricultural and Biological Sciences (miscellaneous)	269	0.886	28.093	Language and Linguistics	222	0.870	24.525	Cultural Studies	269	1.012	26.120

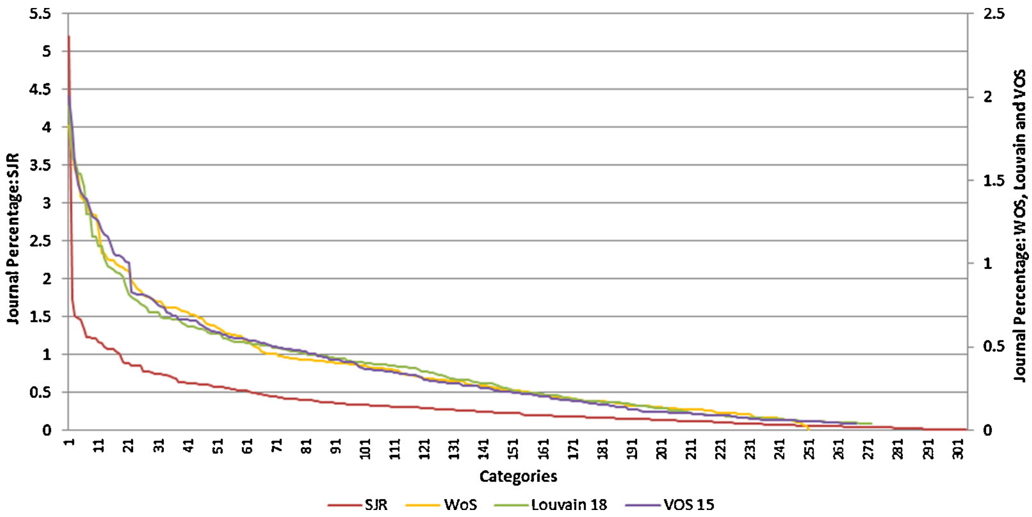


Fig. 5. Distribution of journals over categories in SJR, WoS, Louvain and VOS systems.

case of algorithm systems, a refinement of original SJR classification, a matching of a considerable number of categories may be a symptom of stability and consistency. The seven categories matching in WoS are: (1) 'HISTORY'; (2) 'ECONOMICS'; (3) 'MATHEMATICS'; (4) 'ENGINEERING, ELECTRICAL & ELECTRONIC'; (5) 'PSYCHIATRY'; (6) 'LANGUAGE & LINGUISTICS'; (7) 'EDUCATION & EDUCATIONAL RESEARCH'. The correspondence of these categories in SJR is: (1) 'History'; (2) 'Economics and Econometrics'; (3) 'Mathematics (miscellaneous)'; (4) 'Electrical and Electronic Engineering'; (5) 'Psychiatry and Mental Health'; (6) 'Language and Linguistics'; (7) 'Education'. Finally, the set of categories in the Louvain and VOS systems was identical to that of SJR, except for the category (3) 'Mathematics (miscellaneous)' which was labelled as 'Mathematics (general)'.

Apart from Table 2, Fig. 5 depicts the *distribution of journals over the set of subject clusters or categories* generated for the four classification systems in order to facilitate a better analysis and observation of journal assignment over categories. The *x-axis* denotes the rank order of the various subject categories (owing to the different numbers and names of categories of the four classification systems), while the *y-axis* covers the percentage of journals assigned to a determined category. We would like to emphasize that due to the similarity shown by the distributions of four classification systems and with the aim of improving their visualization two scales for measuring the percentage of journals over categories (*y-axis*) have been introduced in Fig. 5. Thus, on the left-hand side of the figure we have placed the main *y-axis* displaying the scale for SJR distribution, whereas on the right-hand side we have placed an alternative scale for WoS, Louvain and VOS distributions. Although the four systems presented similar distributions, SJR classification shows the most skewed distribution, characterized by a high concentration of journals in the 'Medicine (miscellaneous)' category as aforementioned. WoS, Louvain and VOS solutions presented very similar behaviour regarding the final distribution of journals over the different set of categories obtained. In comparison with the original classification of SJR, there is a more even distribution curve which describes a lower concentration of journals over their respective categories. In spite of this fact, the Louvain, VOS and SJR distributions tend to progressively level; but it is noteworthy that the sets of categories in Louvain and VOS systems are lower than in SJR (respectively, 36 and 41 categories less than SJR), which could favour a higher concentration of journals. This would confirm that Louvain and VOS systems outperform in the distribution of journals as opposed to SJR. Moreover, the similar representations of four systems substantiate our earlier argument about the coherence and homogeneity revealed by the four classification systems.

The final master tables covering the new classification of SJR journals proceeding from VOS and Louvain clustering methods can be accessed through the following links:

- VOS Classification: http://www.ugr.es/local/benjamin/vos15_classification.pdf
- Louvain Classification: http://www.ugr.es/local/benjamin/louvain18_classification.pdf

7. Discussion and conclusions

A wide variety of research studies have approached the problem of science classification for mapping, knowledge organization, information retrieval or bibliometric and scientometric purposes. Up to date, some authors have underlined the non-existence of a classification system as an international standard in bibliometric fields (Archambault et al., 2011; Gomez & Bordons, 1996; Waltman & Van Eck, 2012) Different levels of aggregation, the distinct systems adopted for organizing information, or the varying degrees of specialization or multidisciplinarity of several scientific databases would be sufficient reasons to impede the construction of an international classification system for bibliometric ends. In this work, however, we propose a methodology to update and refine the SJR journal classification system. The proposed method, based on clustering and bibliometric techniques, can be applied to other systems as well.

Another topic commonly addressed by the scientific literature on classification is the adequacy and possibility of developing automatic classification systems to avoid, as far as possible, human intervention. Early works were developed by authors such as Luhn (1957) in Information Retrieval scope at the end of 1950s; but interest remained strong in the 1960s (Garland, 1982), and furthered with the advance and development of scientific databases, bibliometric indicators, science mapping, etc. up to the present. Some research reviewed here tried to avoid human intervention, but concluded it was not possible to do so completely. Waltman and Van Eck (2012) ensured that human involvement was minimized to the choice of certain values in parameters. Archambault et al. (2011) asserted that human intelligence and expertise originates more useful and flexible classification schemes, yet at the same time they can be considered inadequate and biased systems: "From the outset, we decided that it would also be necessary to use expert judgement to finalize the work. In the end, it took substantially more work than initially expected, with alternating iterations using an algorithmic approach followed by manual fine-tuning".

In accordance with the corpus of literature and experience gained from our previous studies, we surmised that a classification system based on a fully automatic approach had not been devised to date. Many options can be enriched by expertise and human learning. Relevant stages emanating from automatic classification implementations, such as labelling in clustering approaches, are very complicated to conduct without human involvement. Decisions as to labelling based on significant words or citation links, single or multiple assignments, definition of thresholds, etc., are indeed complicated. Human expertise and guidance can moreover become very helpful during these tasks. In this work, we avoided human intervention as much as possible. After examining the final results, however, we believe that a mixed approach could be very realistic and convenient. There is no doubt that the clustering algorithms used here work fine in classifying journals. This is clearly evident when results of our tests are checked. However, having run our algorithms and having labelled the set of clusters, we have found that some were termed through adjacent and close categories. In some cases, these categories came from original tags of SJR system, while others resulted from our text-based approach.

For instance, in VOS system we obtained categories such as 'Anatomy' or 'Anatomy and Morphology' respectively covering 18 and 15 journals. In turn, using the Louvain system we found that the categories 'Women's Reproductive Health' and 'Women's and Children's Health', included 10 and 28 journals, respectively. In light of our knowledge of the SJR database, we opine that we could group categories covering very close knowledge domains, by checking journals inside them, to obtain a VOS final category named 'Anatomy and Morphology' consisting of 33 journals, as well as a Louvain final category termed 'Women's and Children's Health' embracing 38 journals. These examples may be extended to approximately two dozen categories in both algorithm classifications.

After analysing and comparing clustering methods introduced in this work, we should emphasize the similarity of final results from VOS and Louvain clustering solutions in relation to the facets studied, evidenced by our figures and tables. Still, the same value for the resolution parameter produces a higher number of clusters under the Louvain method, signalling a finer granularity. According to the initial objectives pursued, this could be an important criterion to consider in selecting one or the other algorithm. Altogether, consideration of the several points analysed makes it hard to decide which one of the clustering algorithms analysed is better suited to our journal classification aim. Both VOS and Louvain clustering solutions provide a good performance in classifying SJR journals deriving from the extensive journal citation-based measurement network. A particular analysis of journals assigned to clusters of specific and well-known knowledge field for authors (such as Library and Information Science) and/or cluster validation techniques based on expert opinions or statistical methods to validate the number and the goodness of clusters generated (Rand Index, Silhouette, Entropy, etc.) might be useful for selecting a final clustering solution.

In comparison with the original SJR journal classification, we detected an especially marked improvement regarding the distribution of journals over categories and the final number of categories available in the new solutions based on VOS and Louvain methods. The original SJR classification scheme includes 308 categories, where 29 have less than 10 journals assigned, while remaining categories cover more than 10. This means that almost all 18,891 journals fit in just 275 categories. Thus, a less skewed distribution of journals over categories is achieved by means of VOS and Louvain solutions. Besides, journal multi-assignment is reduced, and 'Miscellaneous' categories are removed, so that overlapping is minimized for both algorithm solutions. Comparison with WoS Subject Categories showed a certain consistency among the several classification systems, both in the number of journals distributed over categories and in the number of categories appearing together in top-20 categories.

A final but equally relevant issue arises with regards to the large, leading Multidisciplinary journals such as Science, Nature or PNAS, which are not included in any cluster of a size higher than 10. This might be due to their special features, e.g. a citation pattern characterized by a vast quantity of citations emitted and received. By looking at the whole set of clusters (including those below threshold 10), we discovered that Science, Nature and PNAS were allocated in different singletons. This finding suggests it is necessary to look for an alternative method to classify Multidisciplinary journals. One reasonable choice may be to classify these journals on the basis of the papers published in them. Multidisciplinarity could be ascribed to journals publishing on a broad spectrum of topics, overcoming the limitations of journal multi-assignment previously defined.

This work corresponds to a succession of several studies (Gómez-Núñez et al., 2011; Gómez-Núñez, Vargas-Quesada, & Moya-Anegón, 'unpublished results') concerned with optimizing and boosting SJR journal classification system and the related subsequent journal assignment. We can articulate future research by testing new clustering algorithms and automatic techniques (factor analysis) as well as different units of analysis (papers) and measures (text-approaches). We do not wish to proclaim any one of these classification proposals as definitive or exclusively appropriate. Rather, we believe it necessary to keep working in an effort to combine several techniques and process units of analysis that will lead us closer to a consensus among the scientific community, the ultimate aim being to develop an optimal SJR classification.

Acknowledgments

The authors thank the reviewers for their helpful and relevant comments.

References

Ahlgren, P., & Colliander, C. (2009). Document–document similarity approaches and science mapping: Experimental comparison of five approaches. *Journal of Informetrics*, *3*(1), 49–63. http://dx.doi.org/10.1016/j.joi.2008.11.003

Archambault, É., Beauchesne, O. H., & Caruso, J. (2011). Towards a multilingual comprehensive and open scientific journal ontology. In E. C. M. Noyons, P. Ngulube, & J. Leta (Eds.), *Proceedings of the 13th international conference of the international society for scientometrics and informetrics* (pp. 66–77).

Batagelj, V., & Mrvar, A. (1997). *Pajek – Program for large network analysis*. Retrieved from. http://pajek.imfm.si/doku.php

Batagelj, V., & Mrvar, A. (2003). *Density based approaches to network analysis: Analysis of Reuters terror news network*. Retrieved from http://www.cs.cmu.edu/∼dunja/LinkKDD2003/papers/Batagelj.pdf.

Batagelj, V., & Mrvar, A. (2011). *Pajek: Program for analysis and visualization of large networks: Reference manual. Version 3.13*. Retrieved from. http://pajek.imfm.si/lib/exe/fetch.php?media=dl:pajekman.pdf (02.09.13)

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, (October (10)), P10008. http://dx.doi.org/10.1088/1742-5468/2008/10/P10008

Börner, K., Klavans, R., Patek, M., Zoss, A. M., Biberstine, J. R., Light, R. P., et al. (2012). Design and update of a classification system: The UCSD map of science. *PLoS ONE*, *7*(7), e39464. http://dx.doi.org/10.1371/journal.pone.0039464

Boyack, K. W., & Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology*, *61*(12), 2389–2404, doi:10.1002/asi.

Boyack, K. W., Newman, D., Duhon, R. J., Klavans, R., Patek, M., Biberstine, J. R., et al. (2011). Clustering more than two million biomedical publications: Comparing the accuracies of nine text-based similarity approaches. *PLoS ONE*, *6*(3), e18029. http://dx.doi.org/10.1371/journal.pone.0018029

Boyack, K. W., Small, H., & Klavans, R. (2013). Improving the accuracy of co-citation clustering using full text. *Journal of the American Society for Information Science and Technology*, *64*(9), 1759–1767. http://dx.doi.org/10.1002/asi.22896

Cantos-Mateos, G., Vargas-Quesada, B., Chinchilla-Rodríguez, Z., & Zulueta, M. A. (2012). Stem cell research: Bibliometric analysis of main research areas through KeyWords Plus. *Aslib Proceedings*, *64*(6), 561–590. http://dx.doi.org/10.1108/00012531211281698

Chang, Y. F., & Chen, C. (2011). Classification and visualization of the social science network by the minimum span clustering method. *Journal of the Association for Information Science and Technology*, *62*(12), 2404–2413, doi:10.1002/asi.

de Nooy, W., Mrvar, A., & Batagelj, V. (2012). *Exploratory social network analysis with Pajek revised and expanded* (2nd ed.). Cambridge: Cambridge University Press.

Elsevier. (2004). *Scopus*. Retrieved from. http://www.scopus.com/home.url (30.05.13)

Garland, K. (1982). An experiment in automatic hierarchical document classification. *Information Processing & Management*, *19*(3), 113–120.

Glänzel, W. (2012). Bibliometric methods for detecting and analysing emerging research topics. *El Profesional de la Informacion*, *21*(2), 194–201. http://dx.doi.org/10.3145/epi.2012.mar.11

Gomez, I., & Bordons, M. (1996). *Coping with the problem of subject classification diversity*, *35*(2), 223–235.

Gómez-Núñez, A.J., Vargas-Quesada, B., & Moya-Anegón, F. ('unpublished results'). A new SJR journal classification through a combination of citation measures.

Gómez-Núñez, A. J., Vargas-Quesada, B., Moya-Anegón, F., & Glänzel, W. (2011). Improving SCImago Journal & Country Rank (SJR) subject classification through reference analysis. *Scientometrics*, *89*(3), 741–758. http://dx.doi.org/10.1007/s11192-011-0485-8

Jacsó, P. (2013). The need for end-user customization of the journal-sets of the subject categories in the SCImago Journal Ranking database for more appropriate league lists. A case study for the Library & Information Science field. *El Profesional de la Informacion*, *22*(5), 459–473. http://dx.doi.org/10.3145/epi.2013.sep.12

Janssens, F., Zhang, L., De Moor, B., & Glänzel, W. (2009). Hybrid clustering for validation and improvement of subject-classification schemes. *Information Processing & Management*, *45*(6), 683–702. http://dx.doi.org/10.1016/j.ipm.2009.06.003

Klavans, R., & Boyack, K. W. (2009). Toward a consensus map of science. *Journal of the American Society for Information Science and Technology*, *60*(3), 455–476. http://dx.doi.org/10.1002/asi.20991

Leydesdorff, L., Hammarfelt, B., & Salah, A. (2011). The structure of the arts & humanities citation index: A mapping on the basis of aggregated citations among 1,157 journals. *Journal of the Association for Information Science and Technology*, *62*(12), 2414–2426, 10.1002/asi.

Leydesdorff, L., & Rafols, I. (2012). Interactive overlays: A new method for generating global journal maps from Web-of-Science data. *Journal of Informetrics*, *6*(2), 318–332. http://dx.doi.org/10.1016/j.joi.2011.11.003

Liu, G.-Y., Hu, J.-M., & Wang, H.-L. (2011). A co-word analysis of digital library field in China. *Scientometrics*, *91*(1), 203–217. http://dx.doi.org/10.1007/s11192-011-0586-4

Luhn, H. P. (1957). A statistical approach to mechanized encoding and searching of literary information. *IBM Journal of Research and Development*, *1*(4), 309–317. http://dx.doi.org/10.1147/rd.14.0309

Newman, M., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, *69*(2), 026113. http://dx.doi.org/10.1103/PhysRevE.69.026113

Persson, O. (2010). Identifying research themes with weighted direct citation links. *Journal of Informetrics*, *4*(3), 415–422. http://dx.doi.org/10.1016/j.joi.2010.03.006

Pudovkin, A. I., & Garfield, E. (2002). Algorithmic procedure for finding semantically related journals. *Journal of the American Society for Information Science and Technology*, *53*(13), 1113–1119.

Rafols, I., & Leydesdorff, L. (2009). Content-based and algorithmic classifications of journals: Perspectives on the dynamics of scientific communication and indexer effects. *Journal of the American Society for Information Science and Technology*, *60*(9), 1823–1835. http://dx.doi.org/10.1002/asi.21086

Rafols, I., Porter, A. L., & Leydesdorff, L. (2010). Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*, *61*(9), 1871–1887. http://dx.doi.org/10.1002/asi

Rousseau, R., & Zuccala, A. (2004). A classification of author co-citations: Definitions and search strategies. *Journal of the American Society for Information Science and Technology*, *55*(6), 513–529.

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, *24*(5), 513–523.

SCImago. (2007). *Scimago Journal & Country Rank (SJR)*. Retrieved from. http://www.scimagojr.com/ (15.04.11)

Sinclair, S., & Rockwell, G. (2009). Voyeur tools: See through your texts. In *Hermeneuti. ca – The rhetoric of text analysis*. Retrieved from. http://hermeneuti.ca/ (01.09.13)
Small, H. (1999). Visualizing Science by Citation Mapping. *Journal of the Association for Information Science and Technology*, *50*(1973), 799–813.
Thijs, B., Zhang, L., & Glänzel, W. (2013). Bibliographic coupling and hierarchical clustering for the validation and improvement of subject-classification schemes. In *14th international conference on scientometrics and informetrics* 15–19 July, 2013, Vienna (Austria), (pp. 237–250). Vienna: International Society of Scientometrics and Informetrics. Retrieved from. http://www.mtakszi.hu/kszi_aktak/
Thomson Reuters. (2009). *ISI web of knowledge*. Retrieved from. http://thomsonreuters.com/web-of-science/ (01.09.13)
Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. http://dx.doi.org/10.1007/s11192-009-0146-3
Vargas-Quesada, B., & Moya-Anegón, F. (2007). *Visualizing the structure of science*. New York: Springer.
Waltman, L., Van Eck, N. J., & Noyons, E. C. M. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, *4*(4), 629–635. http://dx.doi.org/10.1016/j.joi.2010.07.002
Waltman, L., & Van Eck, N. J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the Association for Information Science and Technology*, *63*(12), 2378–2392. http://dx.doi.org/10.1002/asi
Zhang, L., Glänzel, W., & Liang, L. (2009). Tracing the role of individual journals in a cross-citation network based on different indicators. *Scientometrics*, *81*(3), 821–838. http://dx.doi.org/10.1007/s11192-008-2245-y
Zhang, L., Janssens, F., Liang, L., & Glänzel, W. (2010). Journal cross-citation analysis for validation and improvement of journal-based subject classification in bibliometric research. *Scientometrics*, *82*(3), 687–706. http://dx.doi.org/10.1007/s11192-010-0180-1