Manuscrito del autor publicado en "Aslib Proceedings, vol. 61, nº 1, 2009, p.83-100"

Visualization of scientific co-authorship in Spanish universities: from regionalization to internationalization

Olmeda-Gómez Carlos, Antonio Perianes-Rodriguez and Ma Antonia Ovalle-Perandones.
Department of Library and Information Sciences,
Carlos III University of Madrid, Madrid, Spain
Vicente P. Guerrero-Bote
Department of Information and Communication, University of Extremadura, Badajoz, Spain, and
Felix de Moya Anegón
Institute of Public Goods and Policies, CSIC, Madrid, Spain

Abstract
Purpose – To visualize the inter-university and international collaboration networks generated by Spanish universities based on the co-authorship of scientific articles.
Design/methodology/approach - Formulation based on a bibliometric analysis of Spanish university production from 2000 to 2004 as contained in Web of Science databases, applying social network visualization techniques. The co-authorship data used were extracted with the total counting method from a database containing 100,710 papers.
Findings – Spanish inter-university collaboration patterns appear to be influenced by both geographic proximity and administrative and political affiliation. Inter-regional co-authorship encompasses regional sub-networks whose spatial scope conforms rather closely to Spanish geopolitical divisions. Papers involving international collaboration are written primarily with European Union and North and Latin American researchers. Greater visibility is attained with international co-authorship than any other type of collaboration studied.
Research limitations/implications - Impact was measured in terms of journals rather than each individual article. The co-authorship data were taken from the Web of Knowledge and were not compared to data from other databases.
Practical implications - The data obtained may provide guidance for public policy makers seeking to enhance and intensify the internationalization of scientific production in Spanish universities.
Originality – The Spanish university system is in the midst of profound structural change. This is the first article to describe Spanish university collaboration networks using social network visualization techniques, covering an area not previously addressed.
Keywords Universities, Research co-authorship, Proximity, Social networks, Spain
Paper type Research paper

Introduction
In the two decades running from 1985-1986 to 2007-2008 the Spanish university system has undergone significant institutional and organizational change. The period is characterized by pronounced expansion, intensification and diversification of the university undergraduate offering and a growing demand for university services. The system is presently immersed in a process midway between reform and overhaul.

Spanish scientific production in institutions of higher education has been specifically analyzed in a number of articles, PhD theses and research reports. Particularly prominent in this regard is the line of work that analyzes specific aspects of nation-wide university production, such as excellence (Moya Anegón *et al*., 2004), citation (Camí, 2004), interdisciplinarity (Rovira Pato, 2006) or university performance (Ramos *et al.*, 2007; Gómez *et al*., 2007).

The present paper discusses a study, from the vantage of institutional aggregation, of internationally visible scientific output produced by Spanish universities, using social network analysis and visualization techniques (Otte and Rousseau, 2002; Freeman, 2000).

The study is designed to respond to questions such as:

- What degree of national cooperation can be detected in co-authorship practices among Spanish university researchers?
- Is inter-university collaboration in Spain structured or hierarchized?
- What is the impact of collaborative production? And at the international level?

Related papers

Scientific collaboration is one of the key social mechanisms in contemporary research. Greater intensity and breadth of co-authorship have been detected by bibliometric studies conducted internationally (Glänzel, 2001; Persson, Glänzel and Danell, 2004; Wagner and Leydesdorff, 2005), in the United States (Hill *et al*., 2007), the European Union (European Commission, 2003) and Spain (Moya Anegón *et al*., 2007), and the causes of this rise have been researched over the years (Laband and Tollison, 2000; Beaver, 2001).

An analysis of scientific collaboration from a structural standpoint contributes to a better understanding of the topology and laws governing network dynamics. The early attempts to analyze this type of network had a dual focus: on the one hand they purposed to define large networks from an analysis of individual co-authorship patterns (|Newman, 2001a, 2001b) to determine their statistical properties (Newman, 2004) and characterize them as "small worlds" (Nascimento *et al*., 2003), and on the other to study their dynamics and evolution (Barabási *et al.*, 2002).

The application of social network analysis to bibliographic co-authorship networks has become increasingly common in informetrics. Nagpaul (2002) analyzed an inter-institutional network comprising 50 elite research centres in India. In an analysis of co-authorship networks in social science, Moody (2004) concluded that co-authorship is more common in specialities where the division of research work is most readily identified, such as those in which quantitative methodology is used. Acedo *et al.* (2006) showed that the most influential authors of management and organizational studies are linked by co-authorship bonds. Lariviere *et al.* (2006) applied social network visualization techniques to analyze collaboration among Canadian researchers engaging in natural and social science and the humanities. They found both geographic distance and the language used by the researchers, in a bilingual country such as Canada, to be determinants in the resulting network configurations. Hou *et al*. (2008), in turn, described the microstructure of the collaboration networks

of authors publishing in *Scientometrics*, identifying the most influential components and authors on the basis of co-authorship.

Materials and methods

The source of the data used to formulate the bibliometric indicators was the Web of Science (WOS) and, more specifically, three of its databases:

1. Science Citation Index Expanded (SCI- Expanded), specializing in mathematics and medicine.
2. Social Science Citation Index (SSCI), specializing in social science.
3. Arts and Humanities Citation Index (AHCI).

Data

All the papers for the years 2000 to 2004, inclusive, containing the word Spain in the address field, were retrieved from the above databases. Records were retrieved from each complete database using the online version of the WOS. The initial results of the searches performed with the above criteria yielded a total of 151,600 papers of all types, published by all manner of Spanish research institutions.

Data standardization

The data in certain fields of the database had to be standardized prior to analysis, especially the names of the authors' affiliations. The address field usually comprises four levels: the main organization, a department within the organization, the city and country. In many cases, only three levels are listed, excluding the department or institutional level. The country is generally highly standardized and the city can be standardized using postal codes. Many variations can be found at all these levels. This was one of the problems that had to be solved, for it directly affected the identification of relationships among institutions and organizations in the same and different autonomous regions. In order to correctly match organization sites to universities, the variations in the names of each institution were identified, adopted and allocated to the respective university and region using semi-automatic procedures. Previously, an authority file had been created in which the admissible variations in an address were referred to the file adopted as the accepted entry. After this process, the papers attributed to the university sector were grouped in a sub-set of 100,710 texts of all types (P): articles, congress abstracts, reviews, letters, editorials and book reviews. Only the papers classified as "articles" (Pa) were retrieved from sub-set P, for a total of 88,753.

All the information from *Journal Citation Reports* JCR-SCI and JCR-SSCI for the period 2000-2004 was added to these databases. The information gathered on each of the journals included: bibliographic identification, number of papers published per year, subject category and impact index by year.

Data processing

The variables considered to classify the bibliographic data were: time, geography, sector and institution. The analysis focused on a sector at the national level, namely the production attributed to Spain's 70 private and public universities registered in the Ministry of Education and Science's National Register of Universities, Schools and Training in February 2008.

Distribution by time
Articles were dated on the grounds of the year of publication of the journal issue in which they appeared. This information, typically found in all bibliographic references, can be used to date bibliometric analyses. The immediate aim was to group data by year to detect year-by-year variations in any of the bibliometric indicators used for the study.

Geographic distribution
Spanish autonomous regions constituted the unit used for the geographic distribution of papers. With this approach, inter-regional comparisons and comparisons of each region to the country as a whole could be drawn. Production in Ceuta and Melilla, Spain’s two autonomous cities, was included in the production for Andalusia.

Counting, indicators and graphics
The set of indicators used to quantify the results of collaborative scientific production involving Spanish universities was based on whole, total or standard publication counting (Gauffriau *et al.*, 2007). The indicator P represents the number of papers of whatever type authored by at least one Spanish university.

While the databases contain data on the institutional affiliations of the authors of the articles listed, the records contain no further information, no data on the disciplines dealt with in their papers or their nationality, age, sex, administrative rank or status within their organizations. In the present study, the use of the term “co-authorship" should be understood to mean at the institutional level, i.e. authors' institutional affiliation (Melin and Persson, 1996). This method is not perfect, for when attributions are based on authors’ institutional addresses, if, in internationally co-authored articles, the original journal omits the address of one of the authors, co-authorship cannot be attributed either to the institution or the country in question.

Production indicators were broken down by type of collaboration as discussed below. “Non collaboration” refers to papers whose institutional authorship can be attributed to a single Spanish university. “Inter-university collaboration” is understood to mean papers signed by at least two different Spanish universities. “National collaboration” covers papers involving inter-university collaboration and signed as well by at least one national institution of whatever nature: private enterprise, hospital or public research body. “International collaboration” means that at least one of the authors is affiliated with a foreign (non-Spanish) institution.

When the total counting method is employed, as in this case, any given paper may be attributed to two or even three categories, depending on the type of aggregation used. The drawback to this counting method is multi-attribution, for each author receives full credit (Egghe *et al.*, 2000) and as a result the production summations are greater than the actual number of papers. By way of clarification, the two tables (Tables I and III) have two rows of totals: with (R) and without (NR) repetitions.

When applying visualization techniques to bibliometric co-authorship networks, one aspect to be borne in mind is the graphic representation of the direction of the relationship or link established by collaborating universities, and the effectiveness of that collaboration. The existence of collaboration between two countries, institutions or persons implies reciprocity, but provides no insight into the degree of dependence of one or the other. The degree of dependence may vary among organizations, for collaboration may not be symmetric. Confirmation or reciprocity is an important property of links in network analysis. Confirmation is not defined simply by the existence of the link, but by the degree to which the value of reciprocity is the same in the various nodes in the network (Tichy *et al.*, 1979).

Such dissimilarity in the degree of collaboration between universities is represented by computing the asymmetric collaboration rate and mapping the inter-university collaboration network, in which asymmetry is denoted by the different direction of the points on the arrows between nodes. This indicator, borrowed from the affinity index used to measure asymmetric relations between two countries (Zitt *et al.*, 2000), is adapted here to estimate asymmetric collaboration between universities.

It is calculated from formulas used to measure the direction of cooperation between any two nodes, as follows:

$$TCA(Insti_1 \rightarrow Insti_2) = \frac{COL(Insti_1 \leftrightarrow Insti_2)}{COL(Insti_1 \leftrightarrow total_{class})} \times 100$$

$$TCA(Insti_2 \rightarrow Insti_1) = \frac{COL(Insti_2 \leftrightarrow Insti_1)}{COL(Insti_2 \leftrightarrow total_{class})} \times 100$$

The expected impact is obtained for each periodical appearing in the JCR from the Impact Factor (IF). The expected impact factor used here as an indicator is calculated on the basis of the following premises: each scientific paper automatically inherits the IF, defined in the JCR, of the journal where it is published. Each paper is assigned the IF corresponding to the year of publication and, wanting that, the factor for the closest year available. This is subsequently normalized with a procedure that accommodates comparative terms. A normalization procedure based on typification (Braun *et al.*, 1985) generates IF values that conserve their variability while harmonizing the scales of the various subject categories. This yields the optimal reference point on which the domain analyzed should be positioned, whereas in other types of calculations the resulting value is given as a range.

The TIF is found with the following formula:

$$tif_{jc} = \frac{if_{jc} - \overline{if_c}}{\sigma if_c}$$

where *if* is the impact factor for journal *j* in JCR category *c* and *tif* is the normalized impact factor of journal *j* in JCR category *c*. The values found with this function may be positive. The *tif* values for different categories can be compared. Nonetheless, since negative values are difficult to understand and use additively, a *tif* scale corrector is proposed, as follows:

$$tifn_{jc} = m + \left(tif_{jc}/k\right)$$

In this expression, m and k are two constants whose values are chosen in keeping with the objectives of the study. In the present case the values used are m=1 and k=3 to ensure that the values generated conserve their variability, are positive, allow comparison among different categories, and ensure that if an article has an average IF it has a value of 1 and that the normalized IF is assigned to each paper.

The Kamada-Kawai (Kamada and Kawai, 1989) graphic representation algorithm included in Pajek (Batagelj and Mrvar, 1997) network analysis software is used to position Spanish universities on the inter-university collaboration network, together with the findings for total collaboration without repetitions.

This algorithm designs the node network assuming that the links between them behave like springs, for which there is an ideal spring *length*, corresponding to the distance between nodes, and a *force* acting on the spring. The nodes can be positioned in two- or three-dimensional space and the system as a whole is made to evolve in a way that the energy on the springs declines. According to Vargas-Quesada and Moya Anegón (2007), to avoid computational problems, evolution is calculated in this algorithm for each node individually, rather than for the whole. In other words, all the nodes remain unchanged except the one accumulating the greatest energy, which is allowed to evolve until its energy drops to below a certain threshold, at which point a new threshold is established. Subsequently, a second node, the one that now has the highest accumulated energy, is allowed to evolve to below the established threshold and again a new threshold is defined. This process is repeated until none of the nodes in the network accumulates more than the threshold energy. After applying the algorithm, the distance between nodes is readily visible because it is closely correlated to the physical distance.

The network generated from the raw co-authorship data (Leydesdorff and Vaughan, 2006) can be subsequently enhanced by adding notations to the nodes (name), or using different node sizes (to indicate production values, for instance) or colours (sectors, autonomous regions or countries). This type of graph can include more than three dimensions, thereby increasing the number of variables that can be added and giving rise to "hyper-varied" representations.

The international collaboration map is built as an adapted heliocentric map (Moya Anegón *et al.*, 2005) using the following methodology:

1. A list of neighbours is generated based on the number of articles co-authored by the university with each country.
2. The impact obtained for the articles written in collaboration with each country is normalized by applying the following function for normalizing the scale on the map:

$$Col\ \frac{tifn_{country} - 1}{tifn_{university} - 1}$$

   Where $tifn_{country}$ is the normalized typified impact factor for the publications co-produced by the university and the country and $tifn_{university}$ is the normalized impact factor for university production.

3. The networks are depicted on the basis of value similarity, yielding links with identical thicknesses but variable lengths.

Results

Since 1978, Spain has been organized into towns/cities, provinces, seventeen autonomous regions and two autonomous cities, Ceuta and Melilla. This arrangement has brought radical change to the system for governing science and technology policy, in which the regions have acquired a decisive role in controlling, financing and guiding research activities (Cruz Castro *et al.*, 2004).

The data on university production by type of collaboration and autonomous region are given in Table I. Nationwide, six of every ten articles are co-authored, and international collaboration is involved in three of every ten. National collaboration is the most prevalent form of co-authorship, while collaboration exclusively among university professors is the least common of all the types analyzed.

Take in Table I. Spanish university production by type of collaboration and autonomous region

Regionally speaking, the researchers most prone to co-authorship are found in Aragon, the Balearic Isles, Cantabria, Catalonia, Madrid, Navarre and Valencia. The figures for researchers working out of Andalusia, Asturias, the Canary Islands, Castile-La Mancha, Castile-Leon, Extremadura, Murcia, Basque Country and Rioja are below the national average (64 per cent). The highest percentages of university researchers working alone, without research partners, are also found among this second list of regions: Extremadura (48 per cent), Murcia (43 per cent), Andalusia (42 per cent), Galicia (42 per cent) and the Basque Country (40 per cent). Lastly, attention is drawn to the numerous links with foreign researchers established by scientists in the Balearic Isles (40 per cent) and Catalonia (38 per cent), the two regions with the highest percentages of papers co-authored internationally in the five-year period studied. On the other extreme, with figures below the national average, are: La Rioja (19 per cent), Castile-La Mancha (25 per cent), Murcia (25 per cent), Asturias (28 per cent) Extremadura (29 per cent) and Castile-León (30 per cent).

The four highest producers are also the university regions most prone to produce papers with other authors in their region, but the percentages vary (Figure 1). Researchers in Catalonia sign 70 per cent of their national production with institutions lying within their region, while the figure comes to 61 per cent in Madrid, 55 per cent in Valencia and 54 per cent in Andalusia.

By contrast to these regions where endogenous links prevail, in others collaboration with institutions outside their geographic area is clearly the norm. This is the case of Castile-La Mancha (93 per cent of regional production), La Rioja (89 per cent) and Extremadura (89 per cent). Cantabria (62 per cent), Asturias (61 per cent), Navarre (60 per cent) and Aragon (58 per cent) occupy intermediate positions.

Take in Figure 1. Spanish university intra/ inter-regional production (%) and autonomous regions

The map of Spanish universities' inter-university research collaboration network is shown in Figure 2. In this figure, the size of the nodes is proportional to the volume of their inter-university scientific production. The spatial distribution of the nodes reveals the existence of a series of interconnected, region-wide collaboration sub-networks. The region of Madrid comprises a large nucleus, with three large-scale nodes representing its Complutense, Autonomous and Polytechnic Universities, with the first two occupying the most central positions, by the number of its neighbours or degree. The Complutense University of Madrid is surrounded by a small constellation of satellite universities: small private institutions (Francisco de Vitoria, San Pablo CEU, Europea, Pontificia de Comillas), recently created public universities such as King Juan Carlos University or specialist universities such as the National Distance University. The direction of the links indicates dependence, as far as collaboration is concerned, on researchers working out of the Complutense University, which is the partner of choice in all cases.

Take in Figure 2. Inter-university research collaboration network, Spain 2000-2004

A similar situation is found in the Catalonian sub-network. Universities created in the early 1990s, namely the Universities of Lleida and Girona and Ramón Llull, International, Rovira Virgili and Pompeu Fabra Universities, flank the older institutions: the Central, Autonomous and Polytechnic Universities. Like their Madrilenian counterparts, they co-author most of their production with professors and researchers from the more senior institutions. The Public University of Navarre and the Universities of Zaragoza, Rioja and Balearic Isles are positioned near the Catalonian sub-network. The private University of Navarre, in turn, is peripheral, forming part of no specific group. The Valencian sub-network exhibits a peculiarity which is not, however, wholly exceptional. While the universities located in the provinces of Valencia and Castellon (the University of Valencia, the Catholic University of Saint Vincent Martyr and the Polytechnic, Jaume I and Cardinal Herrera CEU Universities) maintain close ties, the professors and researchers from the University of Alicante, which administratively speaking forms part of the same region, collaborate more intensely with universities that are geographically closer, in the region of Murcia: Catholic University of San Antonio, University of Murcia and Polytechnic University of Cartagena. The Andalusian sub-network has two nuclei. The first, headed by the University of Seville, is surrounded by the Universities of Cordoba, Cadiz and Huelva and Pablo de Olavide University. The other, located on the lower right side of the map, has the University of Granada in the centre, circled by the Universities of Malaga, Jaen and Almeria. Geographic proximity also appears to be a predominant factor in the spatial configuration of the Galician universities. In this case the University of Santiago de Compostela acts as the nucleus, with connections to Coruña and Vigo. In Castile-Leon, which has seven universities, two sub-sets form: one consisting of the Universities of Salamanca and Leon, and the other comprising the Universities of Valladolid and Burgos and the Miguel de Cervantes European

University. The exception in this region is the Catholic University of Avila, which “orbits” the Madrilenian universities, specifically the Autonomous University of Madrid. Lastly, mention must be made of the closely collaborating Universities of Cantabria and Oviedo, located in two adjacent regions in northern Spain, that appear in the centre of the map, near the Universities of Las Palmas de Gran Canaria and La Laguna, Extremadura and Castile-La Mancha.

The indicator proposed to comparatively measure expected visibility (TIFN) shows that the visibility of co-authored papers, regardless of type, declines steadily and more steeply in the last two years of the study. The highest visibility is consistently found for papers involving international collaboration (see Table II).

Take in Table II. Variations in visibility by type of co-authorship (TIFN)

Table III shows the countries with which Spanish university researchers chiefly collaborate. The list includes the 25 preferred countries, with which Spanish researchers co-authored over 450 articles in the period, and which account for 86.5 per cent of the total production of internationally co-authored articles produced by Spanish universities.

Take in Table III. Production of articles involving international collaboration and mean impact by country

Over 85 per cent of internationally co-authored articles in which Spanish universities participate involve partners in the European Union and 30 per cent carry the names of authors in North America. The United States, England and France are the partners of choice and, among the Latin American countries, Argentina, Mexico and Chile prevail. The figures on the mean impact of international papers clearly show the countries with which greatest visibility is attained. The heliocentric map in Figure 3 graphically represents production involving international collaboration and the visibility reached with such production, country by country.

Take in Figure 3. Spanish university research: international co-authorship network, 2000-2004

The aim of such maps is to show the international co-publication preferences of Spanish university researchers and, at the same time, the way that these relations affect visibility, based on the expected impact of such production. The main characteristic of this graph is that it contains a central node, which in this case represents Spanish university production of articles involving international collaboration. The spheres representing the articles produced with authors from the respective countries “orbit” around the main node, at a greater or lesser distance. Here the size of the spheres is proportional to the volume of the articles co-authored with each country. The countries of choice are, in descending order, the United States,

France, England and Germany. The distance to the centre is inversely proportional to the impact attained. In terms of impact, then, the countries with a higher mean TIFN are closer to the centre. Therefore, the graph shows that the mean score attained for articles involving co-authorship with researchers in the United States, Denmark or Austria, for instance, is higher than for papers co-authored with Argentina, Mexico, Chile, Poland or Russia. The shaded area in the ellipse separates the countries with mean impact values higher than the figure recorded for Spain (1.120), from those with a lower impact value, which lie outside that area.

Discussion

From an institutional perspective, the proportion of university production involving national and international collaboration is on the rise, whereas the percentage of non-collaborative papers is declining, in accordance with patterns observed in other countries.

Public universities collaborate more intensely than private institutions and long-standing universities are more active in this regard than those recently founded. Researchers working out of universities in Catalonia, in particular Barcelona, are the ones most prone to collaborate with other colleagues. As a rule, the older universities have the most heterogeneous departmental compositions, cover a larger number of specialities and have more PhDs as tenured professors.

Universities located in regions where the primary sector is the predominant economic activity, such as Andalusia, Castile-La Mancha, Extremadura, Rioja and Murcia, tend to collaborate with colleagues from other regions more than with researchers in other universities in their own region. The fact that scientists from more recently founded universities, whose production is lower and whose researchers are younger, seek to collaborate with more active nodes on the network, where they appear as satellites, is the reflection of an attempt to work with more productive researchers or form part of networks with greater influence and visibility.

Regional differences in the percentage of the various types of collaboration are the result of a number of factors, including the existence of more than one university and the diversity of institutions located in each region, both, in turn, a result of the institutional configuration of Spanish scientific and technological policy (Sanz-Menéndez, 1997). Equally important are regional disparities in socio-economic conditions, the relative level of development and the weight of the various R&D actors in each region.

A network having a single component, sub-divided into several sub-groups, with a series of peripheral universities belonging to those sub-groups and a third group consisting in all other universities, would appear to adjust to the coherent core/periphery structure model described in the literature on social networks (Everett and Borgatti, 1999), from the representational standpoint, at least. The results also suggest that geographic proximity plays an important role in the spatial configuration of Spanish inter-university collaboration networks.

Bibliometric studies have shown geographic proximity to have a beneficial effect on the intensity and frequency of scientific collaboration. Sylvan Katz, analyzing

collaboration among universities in Canada, Australia and the United Kingdom, generated the mathematical expression that relates the distance separating two universities to the number of their joint endeavours, observing that the latter decline with increasing distance (Katz, 1994). This same author, working with Smith (Smith and Katz, 2000), concluded that "50% of institutional collaboration [among higher education institutions in the United Kingdom] occurred within a radius of from 60-80 km. For institutions outside greater London the radius was 80-100 Km." (p. 5) Liang and Zhu (2002), in a study of inter-regional co-authorship of scientific articles produced in China, determined that geographic proximity is one of the major factors affecting inter-regional research. More recently, Okubo and Zitt (2004) explored co-authorship patterns among certain French regions and adjacent regions in other European Union countries. One of their conclusions was that the "Other regions in neighboring countries have more chance to figure among preferred partners [by French authors] than randomly chosen other EU regions" (p. 224). The data reported in the present paper, while collected with a different methodology and approach, corroborate these previous findings, at least as regards the resulting network that illustrates the links among professionals working in institutions within the same sector.

The reasons underlying that effect have been explained by a number of authors (Kraut, *et al.,* 1988; Katz and Martin, 1997; Laudel, 2001). At the individual level, proximity facilitates identification of the most suitable partners, problem definition, project planning and verification of interpersonal and intellectual compatibility. Since collaboration is based on interpersonal contact, geographic proximity among researchers enhances the possibility of meetings or attendance at courses, conferences and seminars. Physical proximity affords opportunities to discover common interests, exchange ideas, verify compatibilities and discuss the possibility of working together, all within a framework of face-to-face encounters. Researchers' proven preference for such types of relationships explains the influence of proximity on the shape of the Spanish network.

Other types of factors should not be overlooked, however, such as being under the aegis of the same regional authority. Since the regions are responsible for managing the universities located within their boundaries, joint projects can often qualify for financial incentives for cooperation only if all the partners are located in the same region (Sanz Menéndez and Cruz-Castro, 2005).

When collaborating internationally, Spanish university researchers establish links with colleagues from countries with larger scientific systems, measured in number of publications (Luukkonen, *et al.*, 1992), rather than with researchers who are geographically closer. The bonds between Spanish university researchers and their counterparts in the European Union corroborates a trend observed in the European Union as a whole (Mattsson *et al.,* 2008). The explanation for the ties with Latin American countries lies in social, historic and linguistic affinities, as well as in the attempt to generate an Ibero-American higher education area.

Conclusions

The purpose of this paper is to describe and analyze collaboration patterns among professors and researchers working out of Spanish universities, based on data on co-

authorship of academic studies. The trends identified, in particular with regard to type of collaboration, corroborate and confirm the findings for Spain as a whole. The network charted on the grounds of an analysis of collaboration among peers working in comparable, institutionally similar organizations appears to be based on geographic proximity. Its strong regional component gives rise to regional sub-networks that conform very closely to Spain's "federalist" geopolitical structure (Moreno, 2007).

From the methodological standpoint, the use of social network visualization techniques with the algorithm proposed proves to be ideal for graphing co-authorship network configurations. Geographic proximity plays an important role in Spanish scientists' co-authorship behaviour when they seek national partners, while the size of the respective scientific systems is the main criterion in international collaboration, even where the highest visibility is not necessarily attained.

The present results suggest further lines of possible ongoing research. Specifically, studies might address other geographic aspects of collaboration among Spanish university researchers, such as analyses of networks involving institutions pertaining to other national sectors, or the mapping of institutional networks by subject categories or areas.

References

Acedo, F.J., Barroso, C., Casanueva, C. and Galán, J.L. (2006), "Co-authorship in management organizational studies: an empirical and network analysis", *Journal of Management Studies*, Vol. 43 No. 5, pp. 957-982.

Barabasi, A. L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., Vicsek, T. (2002), "Evolution of the social network of scientific collaborations", *Physica A*, Vol. 311 No. 3-4, pp. 590-614.

Batagelj, V. and Mrvar, A. (1997), "Pajek. Program package for large networks analysis", available at: http://vlado.fmf.uni-lj.si/pub/networks/pajek/sunbelt.97/pajek.htm (accessed 23 June 2008).

Beaver, D.D. (2001), "Reflections on scientific collaboration (and its study): past, present and future", *Scientometrics*, Vol. 52 No. 3, pp. 365-377.

Braun, T., Glänzel, W. and Schubert, A. (1985), *Scientometrics Indicators*, World Scientific, Philadelphia.

Camí, J. (2004), "Dinámica de intercitación y cocitación entre las universidades españolas: análisis bibliométrico de los documentos citados producidos en las universidades y de los documentos citantes en el período 1994-2000, a partir de las bases de datos National Citation Report del ISI", Dirección General de Universidades, Madrid.

Cruz Castro, L., Sanz Menéndez, L. and Romero, M. (2004), "Convergencia y divergencia en las políticas de ciencia y tecnología de los gobiernos regionales", working paper 04-02, Unidad de Políticas Comparadas, Consejo Superior de Investigaciones Científicas, Madrid, October, available at: http://www.iesam.csic.es/doctrab2/dt-0402.pdf (accessed 11 July 2008).

Egghe, L., Rousseau, R. and Van Hooydonk, G. (2000), "Methods for accrediting publications to authors or countries: consequences for evaluation studies", *Journal of the American Society for Information Science*, Vol. 51 No.2, pp. 145-157.

European Commission (2003), *Third European Report on Science and Technology Indicators 2003: Towards a Knowledge-based Economy*, Office for Official Publications of the European Communities, Luxembourg.

Everett, M.G. and Borgatti, S.P. (1999), "Peripheries of cohesive subsets", *Social Networks*, Vol. 21 No. 4, pp. 397-407.

Freeman, L.C. (2000), "Visualizing social networks", *Journal of Social Structure*, Vol. 1 No. 1 , pp.1-15.

Gauffriau, M., Larsen, P.O., Maye, I., Roulin-Perriard, A and Von Ins, M. (2007),"Publication, cooperation and productivity measures in scientific research", *Scientometrics*, Vol. 73 No. 2, pp.175-214.

Glänzel, W. (2001), "National characteristics in international scientific co-authorship relations", *Scientometrics*, Vol. 51 No. 1, pp. 69-115.

Gómez, I., Bordons, M., Fernández, M.T. and Morillo, F. (2007), "Structure and research performance of Spanish universities", in Torres-Salinas, D. and Moed, H.F. (Eds), *Proceeedings of ISSI 2007: 11th International Conference of the International Society for Scientometrics and Informetrics*, CINDOC, Madrid, pp. 335-345.

Hill, D., Rapoport, A.I., Lehming, R.F. and Bell, R.K. (2007), "Changing U.S. output of scientific articles: 1998-2003", Special report, National Science Foundation, Arlington, May.

Hou, H., Kretschmer, H and Liu, Z. (2008), "The structure of scientific collaboration networks in Scientometrics", *Scientometrics*, Vol. 75 No. 2, pp. 189-2002.

Kamada T. and Kawai S. (1989), "An algorithm for drawing general undirected graphs", *Information Processing Letter*, Vol. 31 No. 1, pp.7-15.

Katz, J.S. (1994), "Geographical proximity and scientific collaboration", *Scientometrics*, Vol. 31 No. 1, pp. 31-43.

Katz, J.S. and Martin, B.R. (1997), "What is research collaboration?", *Research Policy*, Vol. 26 No. 1, pp. 1-18.

Kraut, R.E., Galegher, J. and Egido, C. (1988), "Relationships and task is scientific research collaboration", *Human-Computer Interaction*, Vol. 3 No. 1, pp. 31-58.

Laband, D.N. and Tollison, R.D. (2000), "Intellectual collaboration", *The Journal of Political Economý*, Vol. 108 No. 3, pp. 632-662.

Lariviere, V., Gingras, Y. and Archambault, E. (2006), "Comparative analysis of networks of collaboration of Canadian researchers in the natural sciences, social sciences and the humanities", *Scientometrics*, Vol. 68 No. 3, pp. 519-533.

Laudel, G. ((2001), "Collaboration, creativity and rewards: why and how scientists collaborate", *International Journal Technology Management*, Vol. 22 No. 7/8, pp. 762-780.

Leydesdorff, L. and Vaughan, L. (2006), "Co-occurrence matrices and their applications in information science: extending ACA to the web environment", *Journal of the American Society for Information Science & Technology*, Vol. 57 No. 12, pp. 1616-1628.

Liang, L. and Zhu, L. (2002), "Major factors affecting China´s inter-regional research collaboration: regional scientific productivity and geographical proximity", *Scientometrics*, Vol. 55 No. 2, pp. 287-316.

Luukkonen, T., Persson, O. and Sivertsen, G. (1992), "Understanding patterns of international scientific collaboration", *Science, Technology & Human Values*, Vol. 17 No. 1, pp. 101-126.

Mattsson, P., Laget, P., Nilsson, A. and Sundberg, C-J. (2008), "Intra-EU vs. extra-Eu scientific co-publication patterns in EU", *Scientometrics*, Vol. 75 No. 3, pp. 555-574.

Melin, G. and Persson, O. (1996), "Studying collaboration using co-authorships", *Scientometrics*, Vol. 36 No. 3, pp. 363-377.

Moody, J. (2004), "The structure of a social science collaboration network: disciplinary cohesión from 1963 to 1999", *American Sociological Review*, Vol. 69 No. 2, pp. 213-238.

Moreno, L. (2007), "Federalization in multinational Spain", in Burgess, M. and Pinder, J. (Eds.), *Multinational Federation*, Routledge, London, pp. 86-107, available at: http://www.iesam.csic.es/doctrab2/dt-0704.pdf (accessed 11 July 2008).

Moya Anegón, F., Chinchilla Rodríguez, Z., Corera Alvarez, E., Gómez Crisóstomo, M.R., González Molina, A., Muñoz Fernández, F. and Vargas Quesada, B. (2007), "Indicadores bibliométricos de la actividad científica española. (1990-2004)", Fundación Española para la Ciencia y la Tecnología, Madrid.

Moya Anegón, F., Chinchilla Rodríguez, Z., Corera Alvarez, E., Hassan Montero, Y., Herrero Solana, V., Muñoz Fernández, F.J., Ruiz de Osma, E., Faba Pérez, C., Guerrero Bote, V., Zapico Alonso, F. (2004), "Análisis de la excelencia científica en la investigación universitaria española (1998-2002)", Dirección General de Universidades, Madrid.

Moya Anegón, F., Vargas Quesada, B., Chinchilla Rodríguez, Z., Corera Alvarez, E., Herrero Solana, V. and Muñoz Fernández, F. (2005), "Domain analysis and information retrieval through the construction of heliocentric maps based on ISI-JCR category cocitation", *Information Processing and Management*, Vol. 41 No. 6, pp. 1520-1533.

Nagpaul, P.S. (2002), "Visualizing cooperation networks of elite institutions in India", *Scientometrics*, Vol. 54 No. 2, pp. 213-228.

Nascimento, M.A, Sander, J. and Pound, J. (2003). "Analysis of SIGMOD's co-authorship graph", *ACM SIGMOD Record*, Vol. 32 No. 3, pp. 8-10.

Newman, M.E.J. (2001a), "Scientific collaboration networks I. Network construction and fundamental results", *Physical Review E*, 64, pp. 016131-1-016131-8.

Newman, M.E.J. (2001b), "Scientific collaboration networks II. Shortest paths, weighted networks and centrality", *Physical Review E*, 64, pp. 016132-1-016132-7.

Newman, M.E.J. (2004), "Coauthorship networks and patterns of scientific collaboration", *Proceedings of the National Academy of Sciences of the United States of America*, 101 (suppl. 1), pp. 5200-5205.

Okubo, Y. and Zitt, M. (2004), "Searching for research integration across Europe: a closer look at international and inter-regional collaboration in France", *Science and Public Policy*, Vol. 31 No. 3, pp. 213-226.

Otte, E. and Rousseau, R. (2002), "Social network analysis: a powerful strategy, also for the information Sciences", *Journal of Information Science*, Vol. 28 No. 6, pp. 441-453.

Persson, O., Glänzel, W. and Danell, R. (2004), "Inflationary bibliometric values: the role of scientific collaboration and the need for relative indicators in evaluative studies", *Scientometrics*, Vol. 60 No. 3, pp. 421-432.

Ramos, R., Royuela,V. and Suriñach, J. (2007), "An analysis of the determinants in Economics and Business publication by Spanish universities between 1994 and 2004", *Scientometrics*, Vol. 71 No. 1, pp.117-144.

Rovira Pato, L.l. (2006), "Análisis de la actividad interdisciplinar en base a los documentos ISI de las universidades españolas: estudio de las aplicaciones para su óptima evaluación", Dirección General de Universidades, Madrid.

Sanz Menéndez, L. (1997), *Estado, Ciencia y Tecnología en España: 1939-1997*. Alianza Universidad, Madrid.

Sanz Menéndez, L. and Cruz Castro, L. (2005), "Explaining the science and technology policies of regional governments", working paper 05-10, Unidad de Políticas Comparadas, Consejo Superior de Investigaciones Científicas, Madrid, September, available at: http://www.iesam.csic.es/doctrab2/dt-0510.pdf (accessed 11 July 2008).

Smith, D. and Katz, J.S. (2000), "Collaborative approaches to research. HEFCE fundamental review of research policy and funding. A joint project with the Higher Education Policy Unit (HEPU), University of Leeds and the Science Policy Research Unit (SPRU) University of Sussex", available at: http://www.sussex.ac.uk/Users/sylvank/pubs/collc.pdf (accessed 11 July 2008).

Tichy, N.M., Tushman, M.L. and Fombrun, Ch. (1979), "Social network analysis for organization", *The Academy of Management Review*, Vol. 4 No.4, pp. 507-519.

Vargas-Quesada, B. and Moya Anegón, F. (2007), *Visualizing the structure of science*, Springer, Berlin.

Wagner, C.S. and Leydesdorff, L. (2005), "Mapping the network of global science: comparing international co-authorship from 1990 to 2000", *International Journal of Technology and Globalization*, Vol. 1 No. 2, pp.185-208.

Zitt, M., Bassecoulard, E. and Okubo, Y. (2000), "Shadows of the past in international cooperation. Collaboration profiles of the top five producers of science", *Scientometrics*. Vol. 47 No. 3, pp. 627-657.

Corresponding author
Carlos Olmeda-Go´mez can be contacted at: carlos.olmeda@uc3m.es

Received
Revised
Accepted

Table I. Spanish university production by type of collaboration and autonomous region

Region	P	P with coll. (%)	Type of collaboration			
			P w/o coll. %	P, inter-university %	P, national %	P, international %
Andalusia	148840	8534 (57.51)	6306 (42.49)	2285 (15.40)	3904 (26.31)	4630 (31.20)
Aragon	3528	2376 (67.35)	1152 (32.65)	408 (11.56)	1180 (33,45)	1196 (33.90)
Asturias	3253	2006 (61.67)	1247 (38.33)	406 (12.48)	1078 (33.14)	928 (28.53)
Balearic I.	1034	668 (64.60)	366 (35.40)	139 (13.44)	246 (23.79)	422 (40.81)
Canary I.	2757	1732 (62.82)	1025 (37.18)	410 (14.87)	835 (30.29)	897 (32.54)
Cantabria	1466	1052 (71.76)	414 (28.24)	222 (15.14)	539 (36.77)	513 (34,99)
Castile-La Mancha	1444	841 (58.24)	603 (41.76)	345 (23,89)	480 (33.24)	361 (25.00)
Castile-Leon	5161	3161 (61,25)	2000 (38.75)	864 (16.74)	1569 (30.40)	1592 (30.85)
Catalonia	23178	16767 (72,34)	6411 (27,66)	2243 (9.68)	7958 (34.33)	8809 (38.01)
Extremadura	1651	857 (51.91)	794 (48.09)	286 (17.32)	378 (22.90)	479 (29.01)
Galicia	7880	4545 (57.68)	3335 (42.32)	1294 (16,42)	2120 (26.90)	2425 (30.77)
Madrid	20356	13602 (66.82)	6754 (33.18)	2897 (14.23)	7023 (34.50)	6579 (32.32)
Murcia	2855	1624 (56.88)	1231 (43.12)	571 (20.00)	889 (31.14)	735 (25.74)
Navarre	2810	1889 (67.22)	921 (32.78)	364 (12.95)	1030 (36.65)	859 (30.57)
Basque Country	4086	2434 (59.57)	1652 (40.43)	661 (13.73)	1110 (27.17)	1324 (32.40)
Rioja	326	203 (62.27)	123 (37.73)	89 (27.30)	138 (42.33)	65 (19.94)
Valencia	12091	7847 (64.90)	4244 (35.10)	1867 (15.44)	3868 (31.99)	3979 (32.91)
Total R*	108716	70138 (64.51)	38578 (35.49)	15251 (14.03	34345 (31.59)	35793 (32.92)
Total NR*	100710	62479 (62.02)	38231 (37.96)	9814 (9.74)	28802 (28.60)	33677 (33.44)

Notes: * Total production with repetitions ** Total production without repetitions

Table II. Production of articles involving international collaboration and mean impact by country

COUNTRY	Total articles with international co-authorship, Pa R*	Mean TIFN of articles with international co-authors**
USA	7465	1.187
France	4747	1.143
England	3845	1.153
Germany	3697	1.156
Italy	3526	1.152
Netherlands	1541	1.164
Switzerland	1223	1.202
Argentina	1204	1.075
Canada	1108	1.176
Belgium	1110	1.155
Portugal	1082	1.095
Russia	1102	1.126
Mexico	1072	1.060
Sweden	915	1.172
Scotland	939	1.135
Brazil	857	1.082
Japan	806	1.168
Poland	767	1.086
Denmark	630	1.227
Austria	615	1.185
Chile	523	1.071
Finland	521	1.186
Greece	504	1.151
China	468	1.120
Australia	457	1.113
All others (107)	6398	
Total Spain (R)	47122	1.120
Total Spain (NR)***	33677	1.097

Notes: * Total university production (Pa) involving international co-authorship, with repetitions
** Mean impact of articles involving international co-authorship
*** Total university production (Pa) involving international co-authorship, without repetitions

Table III. Variations in visibility by type of co-authorship (TIFN)

Year	TIFN co-authorship			
	No collaboration	Intra-sectoral	National	International
2000	1.05	1.05	1.07	1.14
2001	1.05	1.05	1.07	1.13
2002	1.05	1.06	1.07	1.13
2003	1.03	1.04	1.06	1.10
2004	1.03	1.04	1.06	1.11
Mean	1.04	1.05	1.07	1.12

Figure 1. Spanish university intra/inter-regional production (%) and autonomous regions

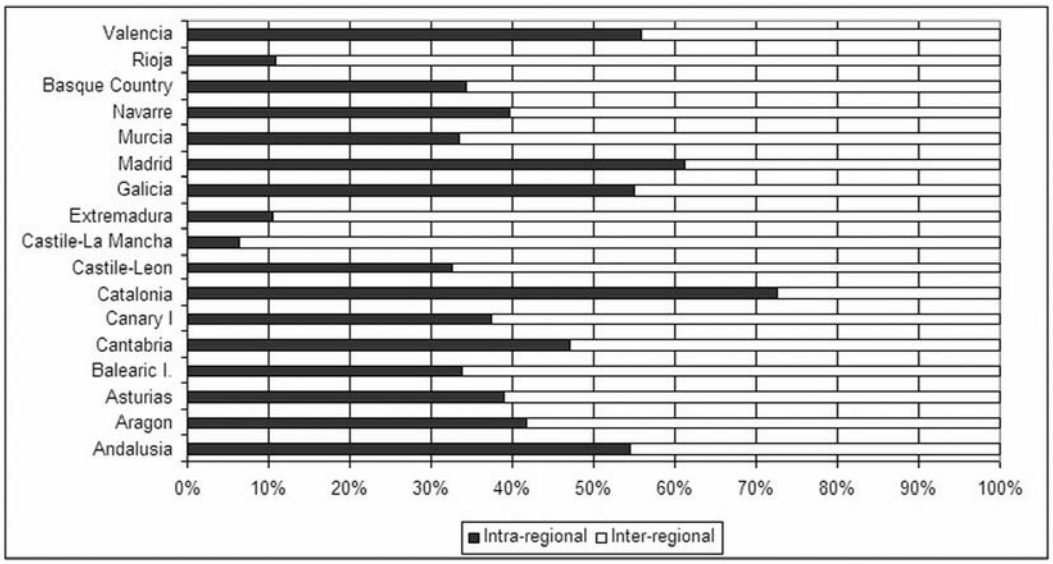

Figure 2. Inter-university research collaboration network. Spain 2000-2004

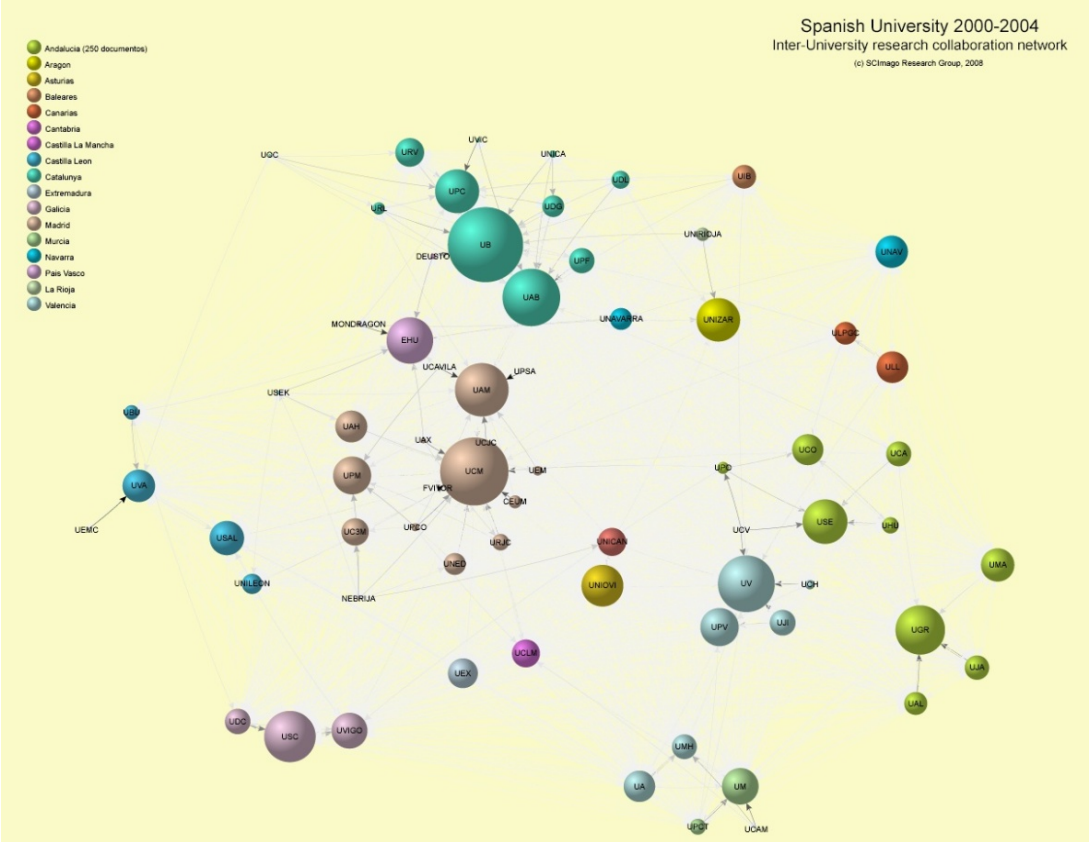

Figure 3. Spanish university research: international co-authorship network, 2000-2004