

An Exploratory Social Network Analysis of Academic Research Networks

Sergio L. Toral¹, Nik Bessis², M.R. Martinez-Torres¹, Florian Franc³, Federico Barrero¹ and Fatos Xhafa⁴

¹University of Seville, Spain

²University of Derby, United Kingdom

³University of Toulouse, France

⁴Polytechnic University of Catalonia (UPC), Spain

¹toral@esi.us.es, ¹rmtorres@us.es, ¹fbarrero@esi.us.es, ²n.bessis@derby.ac.uk, ³ffranc@etud.insa-toulouse.fr,

⁴fatos@lsi.upc.edu

Abstract – For several decades, academics around the world have been collaborating with the view to support the development of their research domain. Having said that, the majority of scientific and technological policies try to encourage the creation of strong inter-related research groups in order to improve the efficiency of research outcomes and subsequently research funding allocation. In this paper, we attempt to highlight and thus, to demonstrate how these collaborative networks are developing in practice. To achieve this, we have developed an automated tool for extracting data about joint article publications and analyzing them from the perspective of social network analysis. In this case study, we have limited data from works published in 2010 by England academic and research institutions. The outcomes of this work can help policy makers in realising the current status of research collaborative networks in England.

Keywords-component; Knowledge networks, Social network analysis, Bibliometrics, Collaborative research networks

I. INTRODUCTION

Co-operations between different scientific disciplines, different organizational units, and external actors seem to be a common and increasing phenomenon of academic reality [1]. Obviously, collaboration has always been part of academic life, but the context of an increasingly globalised research environment has encouraged academic institutions to strength their external and international dimension. Both of them are considered essential to remain competitive and to drive economic growth. Several reasons can explain this trend. Research collaboration is important in order to meet the big global challenges confronting science. In fact, an increasing number of topics require today an interdisciplinary treatment, being necessary the participation of department or organization belonging to different fields of knowledge [2], [3]. From the researchers point of view, group collaboration show that people who already have written a paper previously together are much more likely to succeed in future collaborations, as they have already paid the start-up costs of getting to know each other's languages, approaches and methodologies [4]. Furthermore, attracting and retaining links with the best scientific talents ensures that researchers and institutions stay at the centre of global innovation networks.

In general, all the actors involved in academic world, researchers, managers, politicians and policy makers, are

broadly in agreement about the benefits of reinforcing external collaborations. The main benefits for Higher Education institutions derive from being visible and attracting reputed researchers as well as potential research students. Researchers also consider more exciting working people and groups that have different skills and viewpoints. Politicians and policy makers are usually more worried about the competitiveness and sustainability of the domestic research system and the domestic economic growth.

The most extended technique for analyzing collaborations is co-authorship analysis [5]. Much of the previous work in this area has used co-authorship analysis to assess collaboration among researchers or the structure of scientific collaborations [6], [7]. In this paper, we propose the use of co-authorship analysis to assess the collaboration among academic institutions in a certain geographical environment using social network analysis techniques. As a case study, academic institutions of England are considered, and thus, the extracted results have been analyzed using social network analysis techniques. The remainder of this paper is presented as follows: section II details previous research in the field of co-authorship analysis and social network analysis. Section III describes the methodology, including the set of considered data, retrieval of information and the analysis of this information. Obtained results are presented in section IV as well as their discussion and implications. Finally, the conclusions can be found in section V.

II. RELATED WORK

Co-authorship networks represent a class of social networks typically used to determine the structure of scientific collaborations and the status of individual researchers. These networks are usually analyzed using bibliometric methods. Although they are somewhat similar to the much studied citation networks [8], co-authorship implies a much stronger link than citation, which can occur without the authors knowing each other.

Regarding co-authorship, several previous works agree that collaboration of individual scientists and that of institutions or of even higher levels of aggregation have to be clearly distinguished [9], [10]. In this context, institutional collaboration can, in turn, be studied in two important aspects: collaboration between different research institutions disregarding their organizational type, and collaboration

between different sectors such as university, industry, and government [11]. This work is focused on the first one, sometimes called inter-institutional collaboration. An analysis of scientific collaboration from an inter-institutional point of view can contribute to define the ranking position of academic institutions in a specific geographical area or in a specific knowledge domain [12]. Most studies of inter-institutional collaboration are restricted to national or regional analyses [13], [14]. For instance, a study of domestic inter-institutional collaboration in Canada, Australia, and the UK has concluded that research cooperation decreases exponentially with the distance separating the collaborative partners [15]. As a difference, international collaboration is not only determined by distance, but also by other factors such as the country size and political and economic reasons, as well as certain aspects of mobility and migration at the individual level [11]. Besides, there are also strong influences of historical, cultural and linguistic proximities on co-operation patterns at the national level [16].

In general, scientific collaboration is accepted as a basically positive phenomenon and is unanimously recognized as exerting a significant influence on the performance of individual researchers and institutions, in terms of both effectiveness and efficiency [17]. Several studies conclude that collaborations contribute to scientific productivity [18]. In particular, international collaboration has been commented for producing real and remarkable results in the scientific performance of research groups [19], [20].

The application of social network analysis to a co-authorship networks scenario has become increasingly common during the last decade [21]. They share with other social networks global topological properties such as small world-property, long-tail degree distribution and a scaling law for the clustering coefficient [22]. In a similar vein, several previous works have analyzed the relationships among researchers in specific geographical areas [23] or specialities [24].

In particular, they model researchers as nodes of the social networks and extract some conclusions from the links of collaborations. As a difference, some other works model institutions as nodes, visualizing inter-university and international collaboration networks. For instance, a microanalysis of inter-institutional co-authorship networks comprising universities, government and private companies located in Madrid, Spain, is proposed in [25]. This study was later extended to the rest of Spanish Universities in [21]. Authors concluded that Spanish inter-university collaboration patterns appear to be influenced by both geographic proximity and administrative and political affiliation, which is in line with some other previous studies.

In this work, academic institutions are also modeled as the nodes of the co-authorship network in a specific geographical area (England). However, three partitions of the global network are going to be considered: England Universities, other England academic institutions and external (outside of England) academic institutions. The purpose of this research consists of identifying similarities

and dissimilarities in the collaboration policies of England institutions, and their impact on the basis of their overall performance.

III. METHODOLOGY AND DATA

A social network can be represented as a graph $G = (V, E)$ where V denotes a finite set of nodes and E denotes a finite set of edges such that $E \subseteq V \times V$. Some network analysis methods are easier to understand when graphs are conceptualized as matrices [26], [27]:

$$M = (m_{i,j})_{n \times n} \quad \text{where } n = |V|, \quad m_{i,j} = \begin{cases} 1 & \text{if } (v_i, v_j) \in E \\ 0 & \text{otherwise} \end{cases}$$

In the context of co-authorship, collaborations can be modeled using this kind of networks. As the purpose of this work is about analyzing patterns of collaborations among institutions, nodes of the network represent those academic institutions in which authors are affiliated, and edges are set whenever institutions share a common work.

Data has been extracted from databases contained on the Web of Science. More specifically, records corresponding to England academic institutions during the year 2010 were downloaded. The resulting network is shown in Figure 1. This network contains 9344 nodes corresponding to England and foreign academic institutions collaborating through almost 213.000 papers. Non-England academic institutions (6693), shown as white nodes in Figure 1, England academic institutions except Universities (2518), and England Universities (133), both of them shown as black nodes in Figure 1.

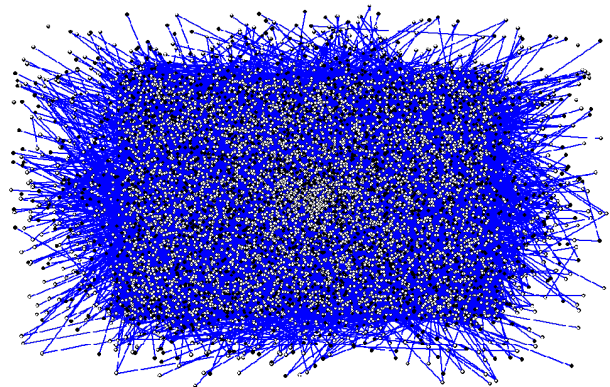


Figure 1. Collaboration Network of England academic institutions in 2010.

Using the three extracted partitions, the following networks will be considered:

- Complete network (CN), including the whole data set of 9344 nodes.
- England academic institutions network (EN), which only consider collaborations among England institutions (excluding foreign institutions).

- England Universities plus foreign institutions network (UFN), focused on the international collaborations of England Universities.

Several features of these networks can be calculated through Social Network Analysis (SNA) techniques:

- Density: it is defined as the number of lines in a simple network, expressed as a proportion of the maximum possible number of lines. The main problem of this definition is that it does not take into account valued lines higher than 1 and it depends on the network size. A different measure of density is based on the idea of the degree of a node, which is the number of lines incident with it [28]. A higher degree of nodes yields a denser network, because nodes entertain more ties, and the average degree is a non-size dependent measure of density.
- Closeness centralization: it is an index of centrality based on the concept of distance. The closeness centrality of a node is calculated considering the total distance between one node and all other nodes, where larger distances yield lower closeness centrality scores. The closeness centralization is an index defined for the whole network, and it is calculated as the variation in the closeness centrality of vertices divided by the maximum variation in closeness centrality scores possible in a network of the same size [28].
- Brokerage roles: A broker is a middle node in a directed triad (a set of three vertices and the lines among them). Different types of brokerage roles can be distinguished considering mediation between different types of nodes. In the context of this study, brokerage roles among nodes of the three extracted partitions can be considered separately.
- Neighbours: number of nodes adjacent to each node.
- Clustering coefficient: It measures whether first degree neighbor of a particular node interact with each other. Basically, clustering coefficient is a measure of local cohesiveness through the neighbor interactions of a node.
- Structural holes: they refer to the extent a node performs a bridging role among a set of nodes that are not directly linked [26].

IV. RESULTS

Table 1 details the set of indicators that have been measured in this study. *DegreeCN* refers to the degree of each node of the complete collaboration network. Degree value considers multiple lines, that is, multiple collaborations between the same institutions. Consequently, the degree shows not only multiple collaborations with other institutions but also repeated collaborations with the same institutions. *ClosenessCN* is the closeness centrality of a node, which measures multiple collaborations with other institutions but without considering multiple lines. *ClosenessUEN* and *closenessUFN* are the closeness centrality of nodes of the England academic institutions network and the England Universities plus foreign institutions network, respectively. *CoordCN* is the number of brokerage roles developed by

each node but among other nodes belonging to the same partition. As a difference, *ItinUEN* and *ItinUFN* refer to the number of brokerage roles developed by each node but among other nodes belonging to the different partitions. In the case of *ItinUEN*, partitions considered are Universities and the rest of England academic institutions whole in the case of *ItinUFN*, partitions considered are Universities and foreign academic institutions. *Nlines1neighCN*, *Nlines1neighUEN*, *Nlines1neighUFN* represent the number of lines with adjacent nodes for the three considered networks. It is a measure of the intensity of collaboration. Finally, *SHolesCN*, *SHolesUEN*, *SHolesUFN* measures the extent each node behaves as a structural hole in each of the three considered networks, while *CC_CN*, *CC_UEN* and *CC_UFN* are the corresponding clustering coefficients.

Table 1. Set of measured indicators

Indicator	Description
DegreeCN	Degree of each node (complete network)
ClosenessCN	Closeness centralization (complete network)
ClosenessUEN	Closeness centralization (England academic institutions network)
ClosenessUFN	Closeness centralization (Universities and foreign institutions network)
CoordCN	Number of brokerage roles among nodes belonging to the same partition (complete network)
ItinUEN	Number of brokerage roles among nodes belonging to the different partitions (England academic institutions network)
ItinUFN	Number of brokerage roles among nodes belonging to the different partitions (Universities and foreign institutions network)
Nlines1neighCN	Number of lines with adjacent nodes (complete network)
Nlines1neighUEN	Number of lines with adjacent nodes (England academic institutions network)
Nlines1neighUFN	Number of lines with adjacent nodes (Universities and foreign institutions network)
SHolesCN	Structural holes (complete network)
SHolesUEN	Structural holes (England academic institutions network)
SHolesUFN	Structural holes (Universities and foreign institutions network)
CC_CN	Clustering coefficient (complete network)
CC_UEN	Clustering coefficient (England academic institutions network)
CC_UFN	Clustering coefficient (Universities and foreign institutions network)

The set of indicators shown in Table 1 have been measured for the list of 133 England Universities corresponding to one of the extracted partitions. Those Universities with a zero degree value have been excluded, leading to the total of 115 Universities.

First of all, a ranking of England Universities has been used to test the influence of the listed indicators in the overall performance of Universities. For this purpose, the ranking of the Complete University Guide (<http://www.thecompleteuniversityguide.co.uk/league-tables/rankings>) has been chosen. In particular, this guide offers a ranking of England Universities attending to their average quality of their research. Table 2 shows the correlation values of this ranking with the set of extracted indicators. Notice that a negative value means a high correlation with the performance of Universities in terms of research, as usually rankings are ordered from lower to higher values.

Table 2. Correlation of the average quality of England Universities research with the set of extracted indicators.

	Ranking		Ranking
DegreeCN	-,577**	Nlines1neighUEN	-,698**
ClosenessCN	-,526**	Nlines1neighUFN	-,606**
ClosenessUEN	-,434**	SholesCN	,180
ClosenessUFN	-,528**	SholesUEN	,570
CoordCN	-,646**	SholesUFN	,189
ItinUEN	-,493**	CC1_CN	,438**
ItinUFN	-,474**	CC1_UEN	,570**
Nlines1neighCN	-,606**	CC1_UFN	,517**

** Correlation is significant at the 0,01 level (2-tailed).

According to Table 2, performance of Universities is positively correlated with external collaborations, both in terms of internal and international collaborations. It is also positively correlated with the scope and intensity of these collaborations. On the other hand, performance is not significantly correlated with the role of Universities as structural holes and negatively correlated with their clustering coefficient.

To obtain the different patterns of collaborations of England Universities, a factor analysis has been applied to the extracted data set. Factor Analysis is a way to fit a model to multivariate data, estimating their interdependence. It addresses the problem of analyzing the structure of inter-relationships among a number of variables by defining a set of common underlying dimensions, the factors, which are not directly observable, segmenting a sample into relatively homogeneous segments [29]. Because each factor may affect several variables in common, they are known as "common factors". Each variable is assumed to be dependent on a linear combination of the common factors, and the coefficients are known as loadings. Factor analysis can be used for either exploratory or confirmatory purposes: exploratory analyses do not set any a priori constraints on the estimation of factors or the number of factors to be extracted, while confirmatory analysis does. In our case, we have developed an exploratory analysis as we did not know the

number of underlying dimensions. That is to say, a decision must be made about the number of factors to be extracted. There are several criteria for doing this, being the most extensive the eigenvalue and percentage of variance criterion. The eigenvalue criterion considers a number of factors equals to the number of eigenvalues higher than 1. The percentage of variance criterion considers all factors accounting for about 70% of the variance of the original variables [29]. Table 3 details the eigenvalues and the percentage of variance explained for the proposed case study. Both criteria are satisfied for a number of factors equals to three.

Table 3. Total variance explained.

Factor	Eigenvalues		
	Total	% of Variance	Cumulative %
1	8,868	55,422	55,422
2	3,753	23,457	78,879
3	1,840	11,499	90,378
4	,572	3,574	93,952
5	,335	2,097	96,049
...
...
15	,000	,001	100,000
16	5,66E-005	,000	100,000

Table 4. Rotated factor loadings with Varimax rotation.

	Factor		
	F1	F2	F3
DegreeCN	,967	,142	-,152
ClosenessCN	,367	,844	-,165
CoordCN	,798	,318	-,345
ItinUEN	,949	,080	-,080
ItinUFN	,947	,067	-,054
ClosenessUEN	,249	,911	-,087
ClosenessUFN	,373	,842	-,162
CC1_UEN	-,335	,151	,747
CC1_UFN	-,196	-,174	,919
SHolesCN	-,059	-,970	-,005
SHolesUEN	-,030	-,945	,096
SHolesUFN	-,066	-,971	,009
Nlines1neighCN	,926	,191	-,219
Nlines1neighUEN	,847	,318	-,346
Nlines1neighUFN	,923	,193	-,221
CC1_CN	-,149	-,226	,930

Once the number of factors has been determined, the next step is to interpret them according to the factor loadings matrix. The estimated loadings from an unrotated factor analysis fit can usually have a complicated structure. The

goal of orthogonal factor rotation is to find a parameterization in which each variable has only a small number of large loadings, i.e., is affected by a small number of factors. The rotated factor analysis ensures that factors represent unidimensional constructs while preserving the essential properties of the original loadings. The most popular of these techniques is the varimax rotation, which seeks rotated loadings that maximize the variance of the squared loadings in each column of the factor loading matrix [29].

Factor loadings with varimax rotation are shown in Table 4. Each row represents the factor loadings of each variable. Moving horizontally from left to right across the five loadings in each row, the highest loading has to be identified. All the variables associated in this way with the same factor are hypothesized to share a common meaning that the analyst should discover.

According to Table 4, several factors can be distinguished; these are discussed next.

F1. Collaboration intensity: This factor characterizes those Universities with the highest collaboration ratings, both in terms of internal and international collaboration. The high value of factor loadings for the degree, brokerage roles and number of lines with the first neighbor confirms this pattern of behavior. This group is composed by 15 Universities, most of them occupying the first positions of the ranking used to test their research performance.

F2. Scope of collaboration: This group is defined by the high value of closeness centrality and the low value of structural holes (negative factor loadings). That means this group exhibit a good connectivity with other institutions but with lower intensity than the previous group. This group is integrated by 55 Universities.

F3. Absence of collaboration: This group is defined by the high value of their clustering coefficient. The clustering coefficient depends on the connectivity of neighbours. In this case, its high value can be interpreted as nodes with a low number of neighbours, which is also justified by the low value of factor loadings corresponding to the degree or the brokerage roles. Up to 33 Universities are included in this group.

Obtained results show that there is a small group of top Universities which concentrates the majority of collaborations. Actually, they accumulate about 68% of external collaborations of all England Universities. In the case of international collaborations, this percentage increases till 84,7%. Figure 2 illustrates the network of England Universities, where the area of nodes is proportional to their degree. This figure visually highlights the dominant group of Universities in terms of collaborations with other institutions. The second group of Universities is responsible of almost the rest of collaborations. Although they collaborate with a good number of other institutions, the intensity of this collaboration is still far from those of top Universities. Finally, the third group does not exhibit any research orientation nor collaboration strategy.

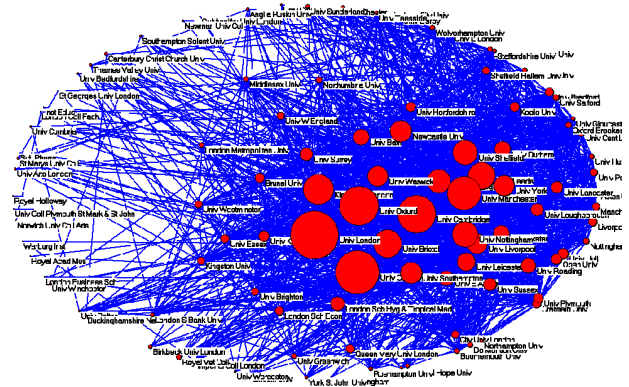


Figure 2.England Universities network.

From the viewpoint of Universities, there is a clear interest in promoting external collaborations, as this activity have a significant impact in their performance and ranking position. In this sense, it is important to know their current situation related to internal and international collaborations. From the viewpoint of policy makers, the priority should be to increase the list of top Universities to improve the national research system performance.

Although this study is focused on collaborations in 2010, several future works can improve obtained results. First, the evolution of the categorization of Universities can be analyzed over time, by extracting the factors through several years. Another important issue is analyzing the quality of collaborations. In this study collaborations are all of them considered the same way. However, not all the journals have the same impact. Consequently, the extracted network can be modified adding a value associated to the strength of edges.

V. CONCLUSION

This study has analyzed the impact of external collaborations on University performance and has extracted three profiles of Universities regarding their external collaborations. Obtained results show a reduced top list of Universities, responsible of the majority of collaborations, a high group of Universities with lower intensity in their external collaboration and a third group of Universities with no external collaboration orientation. This categorization provides valuable information to University managers and policy makers about the current position of their Universities.

REFERENCES

- [1] R. Guimera, B. Uzzi, J. Spiro, L. Amaral, "Team assembly mechanisms determine collaboration network structure and team performance", *Science*, Vol. 308, pp. 697-702, 2005.
- [2] N. Bessis. *Grid Technology for Maximizing Collaborative Decision Management and Support: Advancing Effective Virtual Organizations*, IGI Publishing, 2009
- [3] M. Bordons, M. A. Zulueta, F. Romero, S. Barrigón, "Measuring interdisciplinary collaboration within a university: The effects of the multidisciplinary research programme", *Scientometrics*, Vol. 46, no. 3, pp. 383-398, 1999

- [4] J. N. Cummings and S. Kiesler, "Who collaborates successfully?: prior experience reduces collaboration barriers in distributed interdisciplinary research", In Proceedings of the ACM 2008 Conference on Computer supported Cooperative Work, 2008.
- [5] X. Liu, J. Bollen, M. L. Nelson, H. Van de Sompel, "Co-authorship networks in the digital library research community", *Information Processing & Management*, Vol. 41, Iss. 6, pp. 1462-1480, 2005.
- [6] I. Farkas, I. Derenyi, H. Jeong, Z. Neda, Z. N. Oltvai, E. Ravasz, et al. "Networks in life: scaling properties and eigenvalue spectra", *Physica A*, Vol. 314, pp. 25-34, 2002.
- [7] M. A. Rodríguez and, A. Pepe, "On the relationship between the structural and socioacademic communities of a coauthorship network", *Journal of Informetrics*, Vol. 2, Iss. 3, pp. 195-201, 2008.
- [8] E. Garfield, *Citation indexing—Its theory and application in science, technology, and humanities*. New York, NY: John Wiley & Sons, 1979.
- [9] G. Laudel, What do we measure by co-authorships? *Research Evaluation*, Vol. 11, no. 1, pp. 3-15, 2002.
- [10] J. S. Katz, B. R. Martin, What is research collaboration? *Research Policy*, Vol. 26, pp. 1-18, 1997.
- [11] W. Glänzel, and A. Schubert, *Analysing Scientific Networks through co-authorship*, H.F. Moed et al. (eds.), *Handbook of Quantitative Science and Technology Research*, Kluwer Academic Publishers, pp. 257-276, 2004.
- [12] C. Olmeda-Gómez, A. Perianes-Rodríguez, M.A. Ovalle-Perandones, Visualization of scientific co-authorship in Spanish universities: From regionalization to internationalization, *Aslib Proceedings: New Information Perspectives*, Vol. 61, no. 1, pp. 83-100, 2009.
- [13] D. Hicks, J. S. Katz, The changing shape of British industrial research, STEEP Special Report N° 6, SPRU, 1997.
- [14] I. Gómez, M. T. Fernández, A. Méndez, Collaboration patterns of Spanish scientific publications in different research areas and disciplines. In M.E.D. Koenig and A. Bookstein (Eds.), *Proceedings of the Biennial Conference of the International Society for Scientometrics and Informetrics* (pp. 187-196). Learned Inf., Medford, NJ, 187-196, 1995.
- [15] J.S. Katz, "Geographical proximity and scientific collaboration", *Scientometrics*, Vol. 31, no. 1, pp. 31-43, 1994.
- [16] D. deB. Beaver, "Reflections on scientific collaborations (and its study): Past, present and prospective", *Scientometrics*, Vol. 52, no. 3, pp. 365-377, 2001.
- [17] G. Abramo, C. A. D'Angelo and F. Di Costa, "Research collaboration and productivity: is there correlation?", *Higher Education*, Vol. 57, pp. 155-171, 2009.
- [18] S. Lee, & B. Bozeman, "The impact of research collaboration on scientific productivity", *Social Studies of Science*, Vol. 35, no. 5, pp. 673-702, 2005.
- [19] M. J. Martín-Sempere, J. Rey-Rocha, & B. Garzon-Garcia, "The effect of team consolidation on research collaboration and performance of scientists. Case study of Spanish University researchers in Geology", *Scientometrics*, Vol. 55, no. 3, pp. 377-394, 2002.
- [20] F. Barjak, & S. Robinson, "International collaboration, mobility and team diversity in the life sciences: Impact on research performance", *Social Geography Discussion*, Vol. 3, pp. 121-157, 2007.
- [21] C. Olmeda-Gómez, A. Perianes-Rodríguez, M. A. Ovalle-Perandones, V. P. Guerrero-Bote, F. de Moya Anegón, "Visualization of scientific co-authorship in Spanish universities: From regionalization to internationalization", *Aslib Proceedings: New Information Perspectives*, Vol. 61, no. 1, pp. 83-100, 2009.
- [22] T. Velden, A. U. Haque, C. Lagoze, "A new approach to analyzing patterns of collaboration in co-authorship networks: mesoscopic analysis and interpretation", *Scientometrics*, Vol. 85, pp. 219-242, 2010.
- [23] V. Larivière, Y. Gingras, and E. Archambault, "Comparative analysis of networks of collaboration of Canadian researchers in the natural sciences, social sciences and the humanities", *Scientometrics*, Vol. 68 No. 3, pp. 519-33, 2006.
- [24] H. Hou, H. Kretschmer, and Z. Liu, "The structure of scientific collaboration networks in scientometrics", *Scientometrics*, Vol. 75 No. 2, pp. 189-2002, 2008.
- [25] C. Olmeda-Gómez, A. Perianes-Rodríguez, M. A. Ovalle-Perandones, & F. Moya-Anegón, "Comparative analysis of university-government-enterprise co-authorship networks in three scientific domains in the region of Madrid, 1995-2003", *Information Research*, Vol. 13, no. 3, paper 352, 2008, [Available at <http://InformationR.net/ir/13-3/paper352.html>]
- [26] W. Nooy, A. Mrvar, & V. Batagelj, *Exploratory Network Analysis with Pajek*, New York, Cambridge University Press, 2005.
- [27] M. R. Martínez-Torres, S. L. Toral, F. Barrero, "The role of Internet in the development of Future Software Projects", *Internet Research*, Vol. 20, no. 1, pp. 72-86, 2010.
- [28] S. L. Toral, M. R. Martínez-Torres, & F. Barrero, "Virtual Communities as a resource for the development of OSS projects: the case of Linux ports to embedded processors", *Behavior and Information Technology*, Vol. 28, no. 5, pp. 405-419, 2009.
- [29] A.C. Rencher, *Methods of Multivariate Analysis*. 2nd ed. Wiley Series in Probability and Statistics, John Wiley & Sons, 2002.