ENTERPRISE CONTENT MANAGEMENT SYSTEMS POTENTIAL TO SUPPORT HUMAN CAPITAL MANAGEMENT INITIATIVES

Miguel Neto
Instituto Superior de Estatística e Gestão de Informação, Lisboa, Portugal
mneto@isegi.unl.pt

Carlos A. Fernandes
Instituto de Telecomunicações, Lisboa, Portugal
carlos.fernandes@lx.it.pt

Abstract: Enterprise Content Management Systems (ECMS) and the Enterprise Information Portals (EIP) they support are being increasingly referred in the literature as one interesting technological solution to help organizations in their knowledge management initiatives. This paper seeks to explore the hidden potential of enterprise information portals delivered by ECMS to support knowledge management initiatives, namely human capital management, through the usage of social network analysis on research results and co-authorship/co-work relationships that may suggest ways to more effectively utilize knowledge capital and other organizational resources. For that purpose in this paper we will present field research results on human capital assets management based on EIP data repositories using social network analysis. This evaluation will be made through the use of social network analysis techniques applied to authorship data from papers published in international journals with refereeing covering the last twenty years of research activities from a Portuguese leading research institution in the field of telecommunications - Instituto de Telecomunicações (IT).

Keywords: Enterprise Content Management Systems; Enterprise Information Portals; Knowledge Management; Social Network Analysis; R&D Activities.

1. Introduction

The growing interest on EIP delivered by ECMS as a technological solution to support organizations in their knowledge management initiatives received more recently an important contribute by the aggregation in this platform of solutions to facilitate social networking in what has been referred to as Enterprise 2.0 with the intent of translating the Web 2.0 or social web software adoption by the organizations.

The modern ECMS are based on data repositories where we can find information not only about the organization and its objectives, but also, and more important, about the people and the activities they undertake as well as the relations they establish between them along the way. Among these we can include projects and project teams' composition, documents and documents

co-authorship, functional and physical proximity, etc. In fact, looking at the organization as a network of actors who establish relations between them to achieve their personal and institutional goals through their activities we can find in the ECMS data repositories valuable information about how knowledge is created and shared and also gained a deeper knowledge on the human capital assets, namely on the way they are organized and interact with each other over time. In this context, we can find in the literature arguments to support the use of social network analysis and its metrics as one of the most interesting research techniques to study this kind of problems.

In this paper we will present field research results on evaluating knowledge management and human capital assets based on ECMS data repositories using social network analysis. Furthermore we will evaluate how social network analysis results confirm the traditional approach for R&D evaluation validity - paper authorship productivity. The research is based on data covering the last twenty years of research activity from a Portuguese leading research institution in the field of telecommunications - Instituto de Telecomunicações (IT) - www.it.pt.

In contemporary society, collaboration in research and scientific publication is very common in most areas of academic science. In fact, as Li-Chun et al. (2006) pointed out there is evidence that cooperation among researchers is increasing in all scientific fields. As Hara et al. (2003) highlighted, in science and technology, the effect of this cooperation is significant when addressing complex problems in the contemporary world of "rapidly changing technology, dynamic growth of knowledge and highly specialized areas of expertise".

In this context we believe organizations should look not only to productivity indicators such as the number of papers published in reference journals but also to the collaboration processes behind it and what they represent in terms of knowledge creation, management, and sharing. This paper presents a study, using a social network analysis approach to examine the structure of co-authorship collaboration within IT research community from 1990 to 2009; it uses the most common measures of macro (whole network) and micro (actor centered) structures of this collaboration (Li-Chun, Y et al., 2006).

The data used to build the social network was obtained from IT ECMS backend database. By 31 December 2009, there were around 1.500 papers published in international journals with refereeing involving IT researchers, covering the above-mentioned period. Based on the knowledge network constructed, we can analyze specific paths through which knowledge sharing occurred and by which knowledge capital was nurtured within the IT research community.

2. Social network analysis

Social Network Analysis (SNA) is a methodology that is finding increasing applications outside the social sciences and to date it has been applied to areas as diverse as business organization, electronic communication, health and psychology (Clark, 2006).

Information and intellectual capital stocks are some of the most important resources that flow through research communities; in this context SNA is often applied to identify both information flows and bottlenecks. In theory, it should help to identify strategies that encourage participating actors to share knowledge through the improvement of information flows in the existing social system rather then seeking to create new ones. Social network studies usually take one of two approaches (Clark 2006).: The first approach uses a closed group, for example, to look at the relationships among people working in an institution, whereas the other approach focuses on one person and seeks to understand their *egonet* (self-centered) social relationships surrounding them.

The existence of informal social networks within organizations has long been recognized as important and the unique working relationships among scientific and technical personnel have been well documented by both academics and practitioners. The growing interest in knowledge sharing methods led to increased attention being paid to SNA as a tool for mapping the nature and membership of informal networks.

Managerial social network studies have sought to establish the extent of these informal networks within organizations and assess how the informal organization compares with the formal structures prescribed by management (Cross et al. 2001, 2002a, b; Cross and Parker 2004). Critically, they observed that the patterns of collaboration and communication revealed in informal networks are significantly different from the formal organizational structures implemented by managers. Thus, formal organizational structures fail to reflect accurately the true nature of social relationships and the dynamics and dependencies between staff, which can compromise efficient knowledge exchange within the organization.

The interest in the nature and scale of scientific collaboration, including co-authorship, is growing; especially in the ways that knowledge creation and sharing processes unfold (Barabási et al, 2002; Barnett et al., 1988; Katz and Martin, 1997; Moody, 2004; Newman, 2001). One of the possible approaches to assess scientific collaboration is based on the assumption that co-authorship creates a social network of researchers that develops over time (Hara et al., 2003). If one maps the network of co-authorship, using SNA, it is possible to infer the structure of the collaboration that is taking place between the network members.

In networks of research communities, information and intellectual capital are among the most important resources; their flows and bottlenecks, within the communities, can be studied by SNA (Neto et al., 2008). In theory, these studies should provide a basis for thinking about how a community is organized and what actions might be appropriate to create and develop an environment in which collaboration research is encouraged and ideas shared (Vidgen et al., 2007) providing a deeper knowledge then the traditional R&D publication metrics.

As Cronin (1996) points out, although the study of co-authorship relationships is only one of the possibilities for measuring formal and informal collaborations of scientists, it assumes particular relevance because it is fundamental in leveraging scientific activity; it can also be used as one of the tools to develop a reward system

for academics. This type of metric assumes an important contemporary role, because of the institutional pressure on researchers to increase their productivity by publishing scientific work and the fierce competition for the finite space available in scientific journals. "Publish or perish" dictates success or failure in the competition for funds and other resources (Piette, M. J. & Ross K. L., 1992).

In fact, in recent decades there has been a growing interest in the nature and scale of scientific collaboration and studies into co-authorship have taken two different approaches (Acedo et al. 2006). The first approach attempts to analyze the reasons why authors collaborate and the consequences of such a decision. The second approach is based on the idea that co-authorship creates a social network of researchers. In this paper we carry out an exploratory analysis of co-authorships in IT, a research institution in the telecommunications field. Following the second approach since amongst the wide data available in the ECMS database we will focus our attention in papers published on international journal with refereeing over the last 20 years.

3. Instituto de Telecomunicações

Instituto de Telecomunicações (IT) is a Portuguese private, not-for-profit organization, of public interest, a partnership of six institutions with experience and traditions in research and development in the field of Telecommunications:

* Instituto Superior Técnico (IST);
* Universidade de Aveiro (UA);
* Faculdade de Ciências e Tecnologia da Universidade de Coimbra (FCTUC);
* Universidade da Beira Interior (UBI);
* Portugal Telecom Inovação, S.A. (PTIn);
* Nokia Siemens Networks.

IT's mission is to create and disseminate scientific knowledge in the field of telecommunications. Scientific expertise, which is the basis for their main research and education activities, spans through the following wide areas:

* Wireless Communications;
* Optical Communications;
* Networks and Multimedia; and
* Basic Sciences and Enabling Technologies.

IT is actively involved in fundamental and applied research both at the national and international level. Simultaneously it is committed to fostering higher education and training, by hosting and tutoring graduate and postgraduate students. IT also plays a role in society by providing public awareness initiatives, knowledge transfer to industry, and consulting services, which are offered on a non-competing basis.

IT is organized around three sites: Aveiro at the University Campus, Site II of the University of Coimbra, and at Lisbon in IST. In addition, there is an

external laboratory in Porto, shared by the Faculty of Engineering and the Faculty of Sciences of the University of Porto, and three delegations: Leiria, at the Leiria Polytechnic (IPL), Lisbon, at ISCTE-Instituto Universitário de Lisboa and Covilhã, at the University of Beira Interior.

IT members are organized into research groups within the scientific areas referred above, with specific scientific backgrounds, covering the whole range of IT areas of R&D. The large majority of permanent researchers have a PhD degree and the research teams further include PhD students, MSc students and graduated collaborators. Advanced laboratory facilities are available in most scientific areas to support applied research, which is carried in the framework of national and international projects in cooperation with similar research institutions worldwide. The high level of scientific research is attested by a solid number of peer reviewed international publications.

Considered an important instrument for R&D information management IT Web information portal, and the underlying ECMS, has been an ongoing project for the last few years with new developments and increased functionalities being added along the way. In general terms we can say that all the information concerning IT human resources and the different activities they undertake are registered and stored in a central database which feeds the public web information portal.

Without going in further detail, since that is not the purpose of the present work, we may say that this is only part of the present IT Web portal usage since the adopted portal development strategy relies on a decentralized information management policy where each human resource information is directly managed by the individual through a private Web back-office, with a formal and explicit policy on rules and procedures for this process.

One of the most important IT information outputs, in terms of R&D results, is the real time Web dynamic publication of key IT indicators (http://www.it.pt/indicators_p.asp) where we can find a human resources profile summary and more important statistical information about the scientific output and participation in national and international projects along the years. In the first case - scientific output - this includes:

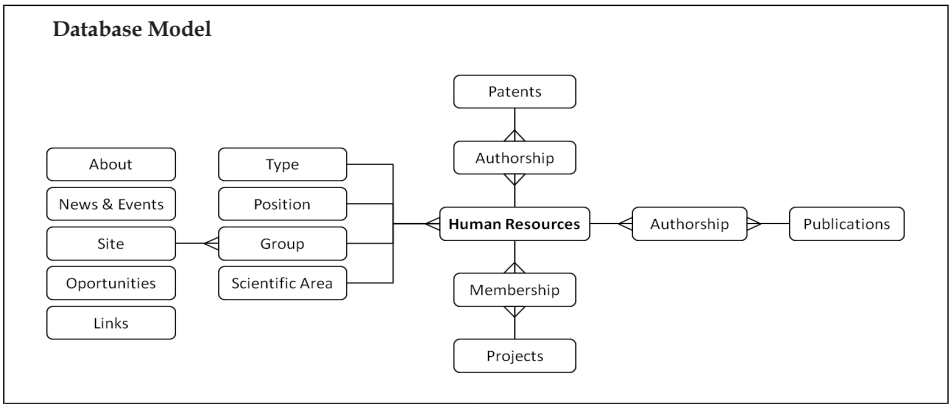
- "Books" - Fully authored books, Editor of books, and book chapter contributions
- "Journals" - Papers in International refereed journals
- "Conferences" - Communications in International and National refereed conferences
- "MSc" - Concluded MSc Theses (includes 2nd cycle dissertations since 2007)
- "PhD" - Concluded PhD Theses

Year	Books	Journals	Patents	Conferences	MSc Theses	PhD Theses
1994	3	53	n.a	136	9	4
1995	1	60	n.a	149	20	12
1996	3	48	n.a	130	25	8
1997	10	60	n.a	242	21	5
1998	5	72	n.a	225	27	6
1999	11	60	n.a	242	17	6
2000	6	42	n.a	193	13	5
2001	11	67	n.a	263	11	8
2002	29	76	2	189	15	7
2003	23	[88]	2	323	23	8
2004	22	95	2	277	42	14
2005	32	117	4	389	29	13
2006	40	128	17	341	33	9
2007	28	152	10	520	78	16
2008	46	190	18	407	160	25
2009	67	197	17	520	194	22

Table 1: IT Scientific output

4. IT Enterprise Content Management System Social Network Analysis

Due to the nature of relational databases, all the data registered concerning IT human resources and R&D activities includes information about co-authorship in the case of publications or team membership in the case of projects or patents are linked in the database. We propose that this enterprise information repository could play an important role in terms of supporting a knowledge management initiative since with a SNA approach we can bring to surface valuable information about how the research community evolves over time. This analysis could identify if the researchers work together bridging the scientific areas boundaries, what is the impact of physical distance on collaboration, etc.



Site Map

Home

About IT

- IT Organisation
- History
- Sites
- Funding
- Key IT Indicators
- Annual Reports
- Press Kit

Scientific Areas

- Wireless communications
- Optical communications
- Networks and multimedia
- Basic sciences and enabling technologies

Research and Educational Opportunities

- Post-docs
- Other research positions
- Short courses

News and Events

- News
- Coming events
- Past events and seminars

Links

Contacts

- Addresses
- How to reach IT

Site Map

Intranet Logon

Figure 1: IT backend database model and EIP site map

In Figure 1 we can see IT's database structure schematic representation and the public Web site map supported by it. From this database and for the present work we analyzed only part of the available data. In fact the analysis carried out is performed only on the data concerning papers published in international journals with refereeing over the last twenty years (from 1 January 1990 until 31 December 2009, although IT was only created in 1992). Through the "author(s)"

data on journals papers it was possible to ascertain and quantify the co-authorship relationships and their unit(s), making it possible to build the relationship matrix for the SNA approach described above.

The data concerning these journals papers and the related authorship information necessary to support the SNA was pre-processed in order to produce information on December 31 in five different moments: year 1990; year 1995; year 2000; year 2005; and year 2009. It must be noted that the data is cumulative, e.g., year 2005 includes all the information since 1 January 1990 through 31 December 2005.

This data set supported the construction of five social networks where the nodes were the IT researchers in the different moments and included identification data, scientific research area, site location, and number of journal papers authorship. The ties represented the co-authorship relationships with the tie strength representing the number of co-authorship for two given connected nodes.

Figure 2 graphically illustrates the evolution of the social network in terms of actors, the relations they have established and their strength. The nodes represent the social network actors, in this case the IT research community. The size of the nodes indicates the number of papers published in international journals and the lines indicate the co-authorship relationships, with the line size indicating the strength of the tie; the latter being the number of co-authorships two connected nodes share. In order to facilitate the comparison between years, the colors remain unchanged over time, meaning for instance that the Wireless Communications scientific area is always red. The social networks maps and the centrality measures were obtained by using NetDraw (Borgatti 2002), a free software tool used for SNA. The nodes aligned on the left of the maps represent researchers who publish alone.

Since the data collection process took place within the last five years, there may be some bias in the early years under analysis due to the fact that the process only received input from those researchers active over this more recent period. Figure 2 also confirms the literature referred research findings where the informal organization reality doesn't follow the formal structure. From the 2009 network map we can see that many researchers from different scientific areas work together. Nevertheless, it is also apparent (Figure 3) that there are sub-groups who under the "all periods of analysis" (1990 - 2009) only contributed to their area of expertise by publishing activities in international journals within their scientific area research network (dotted line).

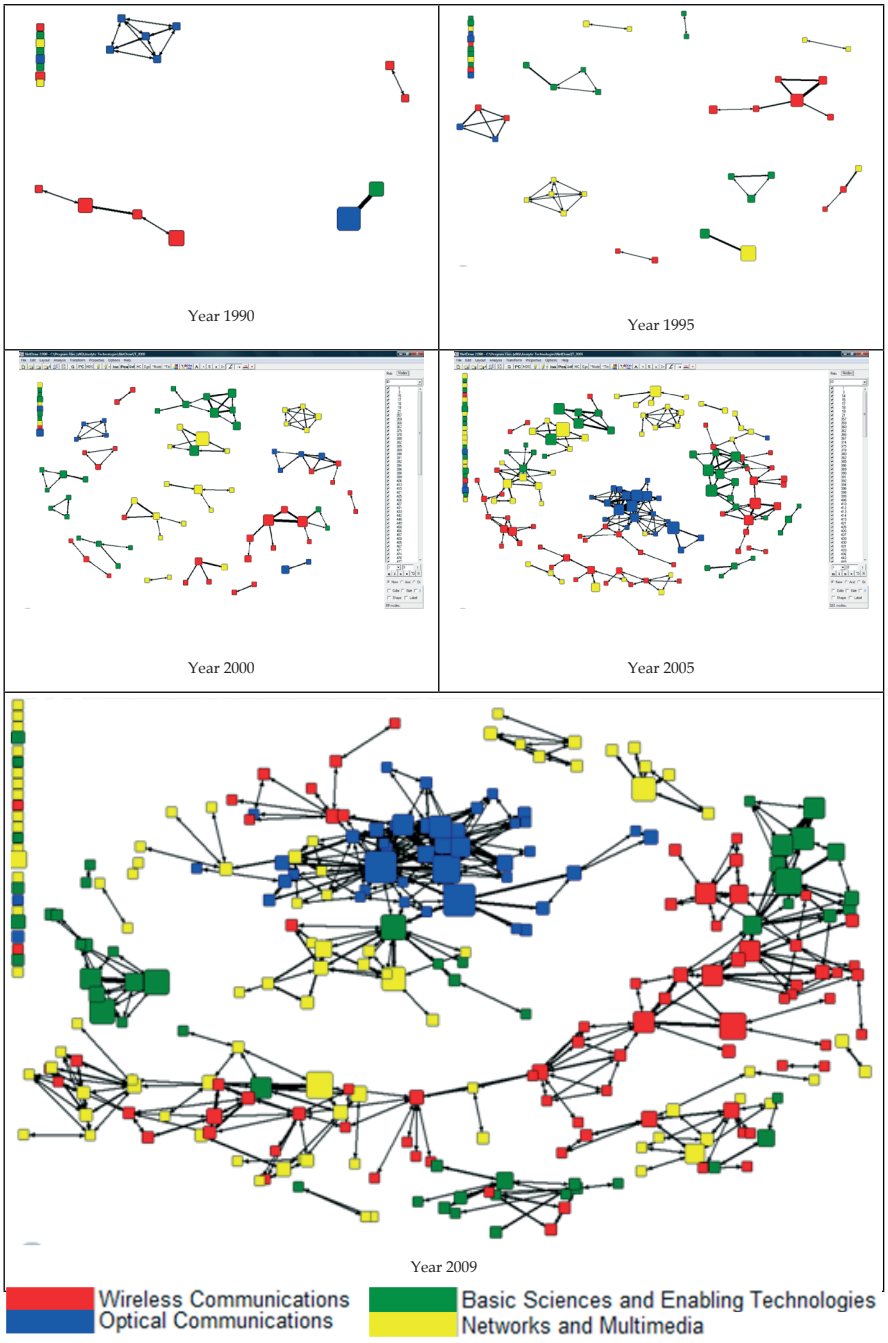


Figure 2: IT research community co-authorship social network evolution over time

In the following maps also a spatial analysis dimension was added to the network map trough the use of different node symbols representing the three IT site locations where the researchers work (Lisboa, Coimbra e Aveiro).

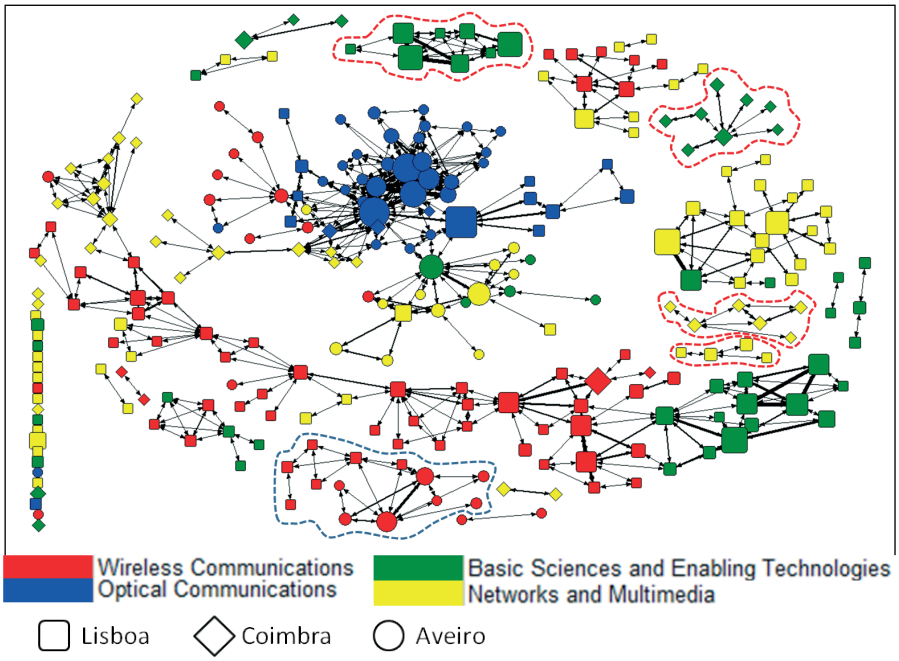


Figure 3: 2009 IT research community co-authorship social network

Analysing the last year's network data in terms of centrality measures we can achieve a better insight into how the network nodes relate with each other. In the following figures we highlight the network nodes in terms of centrality measures. Centrality is a structural attribute of nodes in a network and is a measure of the contribution of network position to the importance, influence or prominence of an actor in a network. Centrality translates the extent to which an actor (in this case the researcher) occupies a central position in the network in one of the following ways (Kilduff and Tsai 2003):

- having many ties to other actors - degree of centrality;
- being able to reach many other actors - closeness centrality;
- connecting other actors who have no direct connections - betweenness centrality.

Centrality		Closeness		Betweenness		Productivity	
ID	Value	ID	Value	ID	Value	ID	Value
3541	32	3541	52707	3541	1830,5	3541	69
5485	23	2822	52740	4142	1280,3	3003	69

5557	20	3003	52746	3074	1273,3	5485	66
6055	18	5557	52748	118	1189,8	5557	58
2822	16	5485	52751	3857	1137,5	110	53
3841	15	4434	52751	2822	1105,7	3976	51
5580	14	3841	52752	3770	654,9	13904	50
1365	12	6055	52756	3043	638,0	3557	49
3770 3857 3723 1596	11	6047	52757	7	622,7	4450	49

Table 2: IT 2009 top 10 summary data (note: ID - fictitious node identification)

From summary data in Table 2, we can confirm that although the most productive researcher was the one that obtain the highest scores in the calculated centrality measures but this correlation cannot be generalized to the entire population. This is particularly relevant in what concerns the "betweenness" role since the corresponding top 10 researchers have a bounding capital in maintaining the network that is far superior to their productivity capacity. In fact nine out of the ten people with highest betweenness centrality records aren't in the productivity top ten. Yet, some caution must be exercised when ranking IT researchers by productivity, since the typical number of papers per year is very diverse according to the different areas of research.

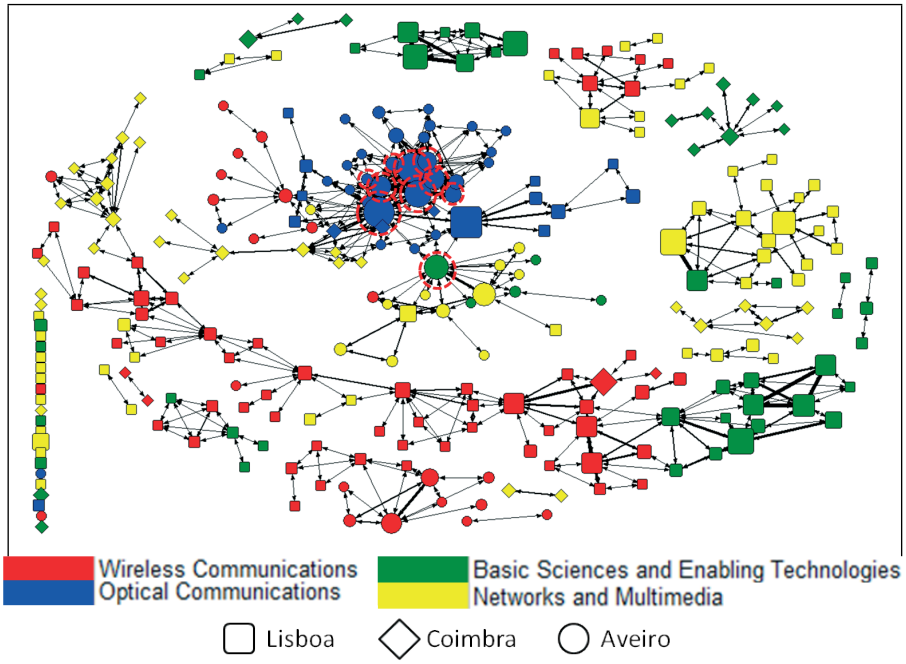


Figure 4: 2009 top ten IT degree of centrality researchers

In Figure 4 we present the 2009 IT co-authorship social network with the indication of the top 10 degrees of centrality nodes (red dotted circles), a measure defined as the number of linked incidents for a node, which can be interpreted as the extent to which a given actor has direct ties to numerous other network actors. The degree of centrality does not take into account the strength of direct links between the two actors; it only reveals the number of people with whom any one scientist has collaborated. This represents access to information and can be considered as a hub or a connector in this network.

Figure 5 points out the top 10 betweenness centrality nodes (red dotted circles), which is seen as the number of geodesic paths that go through a node, expressed as a measure of centrality. This reflects the capacity of an author to connect with other authors within the network, i.e., it is a measure of an author's ability to perform a "broker" role within the network (Acedo et al., 2006). In this 2009 network map we can clearly identify two sub-networks (red dotted line) that would be greatly fragmented if these top 10 betweenness centrality nodes weren't present.

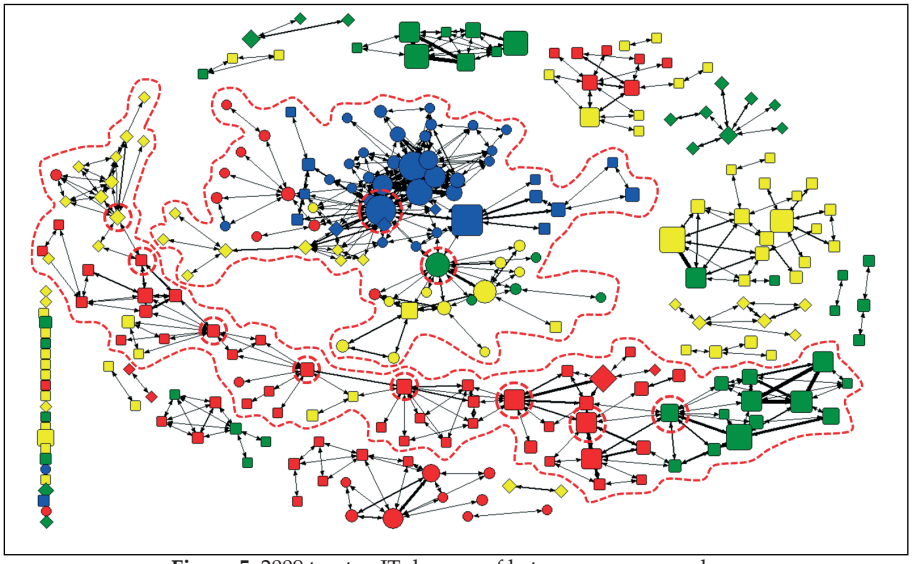


Figure 5: 2009 top ten IT degrees of betweenness researchers

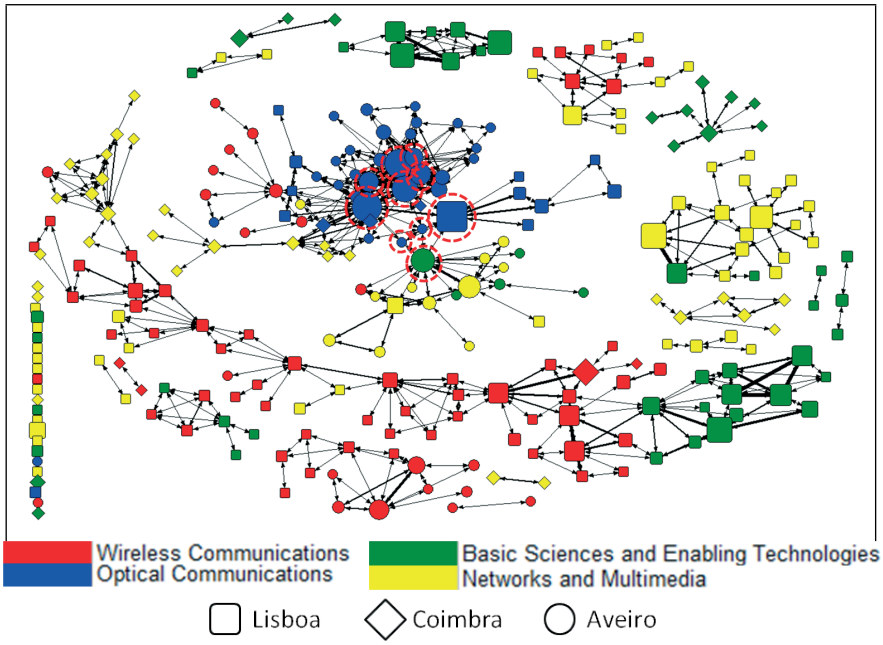


Figure 6: 2009 top ten IT degrees of closeness researchers

Figure 6 points out the top 10 closeness centrality nodes (red dotted circles), which defines paths to the other actors. This represents the capability to monitor the information flow in the network and therefore what is happening in the network.

Finally in the co-authorship network presented in Figure 7 the node size represents the researchers productivity in terms of number of papers published in international journals with refereeing.

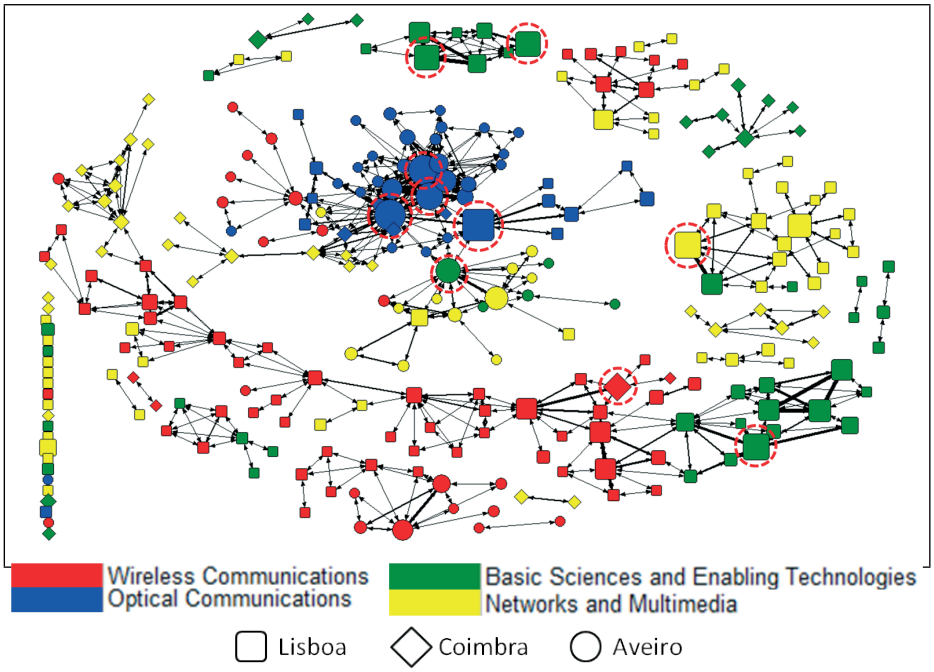


Figure 7: 2009 top ten IT authorship productivity researchers

5. Conclusions and future work

This work is based on part of the information available in the IT Enterprise Content Management System backend database that supports the Web information portal and covered only papers published in international journals, provides valuable insight into the characteristics of the research network that grew inside IT, enabling managers to identifying weaknesses in a more systematic manner and acting to correct them. The results of the presented analysis are far more informative than the usual paper count. The first evidence from the study is that the institution is no longer a mere sum of individual contributors as tended to be at its onset, but has gradually evolved to profit from the created synergies, shared labs, etc. It is interesting to see in Figure 2 the positive collaboration network evolution with the initial isolated clusters becoming more and more connected over time which can be a result from the IT active policy to promote cooperation and synergies amongst researchers.

It becomes evident the crucial role that some researchers play in terms of maintaining the collaboration network and the knowledge creation and sharing processes (e.g. role played by the researchers with the highest degrees of betweenness). The managers might look at these researchers when assigning strategic funding, as potential drivers to engage in strategic projects or initiatives.

It is also relevant to point out that despite all the communication technologies available today, we can still identify the strong impact of physical distance in

collaboration patterns. In Figure 3 (red dotted line) it becomes evident that in the time span under analysis (20 years) we have researchers that publish only with researchers within their scientific area and never published with people from different site locations. This happens practically in all scientific areas, but the cases within the "Basic Sciences and Enabling Technologies" might need attention, since the work in this area should be by definition subsidiary to the other scientific areas.

It is our intention to extend the analysis to the remaining data, either in the publication field (books published and edited, book chapters contributions, communications in refereed conferences, MSc Theses, PhD Theses and students final year projects) and in team activities (projects, patents, etc.). Another important aspect in future research is to refine the analysis to get a deeper understanding of the network dynamics and the real value of productivity indicators for knowledge management initiatives and to include in the analysis other variables such as journals impact factor or papers citation records. Also the co-authorship ties that exist outside IT boundaries will deserve increased attention, for example considering the external researchers who collaborate with IT as a group or research area for the SNA purposes.

The continuation of this analysis will provide a deeper knowledge on how the network evolves over time, how different it is from the formal hierarchical structure, and identifying the actors with the most significant connections. Perhaps, more importantly, it could also highlight those in danger of becoming isolated from the network. Finally, another aspect to be researched is the role that research projects have in social network construction and how public funding policies affect this process.

In conclusion, the presented authorship analysis based on social network analysis tools, profiting from the structured database of papers and authors available at IT, proves to be a valuable resource for institution managers to monitor, diagnose and propose strategic measures to leverage the research and the institution.

6. References

Acedo, F. J. et al., 2006. Co-Authorship in Management and Organizational Studies: An Empirical and Network Analysis. *Journal of Management Studies*, Vol. 43, N.º 5, pp. 957-983.

Barabási, A. L. et al., 2002. Evolution of the Social Network of Scientific Collaboration. *Physica A*, N.º 311, pp. 590-614.

Barnet, A. H. et al., 1988. The Rising Incidence of co-authorship in economics: further evidence. *Review of Economics and Statistics*, Vol. 70 N.º 3, pp. 539-543.

Borgatti, S. P., 2002. *Netdraw: Graph Visualization Software. Harvard Analytic Technologies* [http://www.analytictech.com/Netdraw/netdraw.htm] last accessed 31 March 2008.

Clark, L., 2006. *Network Mapping as a Diagnostic Tool*. Centro Internacional de

Agricultura Tropical, La Paz, Bolivia.

Cronin, B., 1996. Rates of return of citation. *Journal of Documentation*, Vol. 52, no 2, pp. 188-197.

Cross, R. and Parker, A., 2004. *The Hidden Power of Social Networks*. Harvard Business School Press, Boston, USA.

Cross, R. et al., 2001. Knowing what we know: supporting knowledge creation and sharing in social networks. *Organizational Dynamics*, Vol. 30, N.º 2, pp. 100-120.

Cross, R. et al., 2002a. Making invisible work visible: using social network analysis to support strategic collaboration. *California Management Review*, Vol. 44, N.º 2, pp. 25-46.

Cross, R. et al., 2002b. Six myths about informal networks and how to overcame them. *MIT Sloan Management Review*, Vol. 43, N.º 3, pp. 66-75.

Hara, N., et al., 2003. An Emerging View of Scientific Collaboration: Scientists' Perspectives on Collaboration and Factors that Impact Collaboration. *Journal of the American Society for Information Science and Technology*, Vol. 54, N.º 10, pp. 952-965.

Katz, J. & Martin, B. R., 1997. What is research collaboration?. *Research Policy*, N.º 26, pp. 1-18.

Kilduff, M. and Tasi, W., 2003. *Social Networks and Organizations.* SAGE Publications, London, UK.

Li-chun, Yin et al., 2006. Connection and Stratification in research collaboration: An analysis of the COLLNET network. *Information Processing and Management*, N.º 42, pp. 1599-1613.

Moody, J., 2004. The Structure of a social science collaboration network: disciplinary cohesion from 1963 to 1999. *American Sociological Review*, N.º 69, pp. 213-238.

Neto, M.C. et al., 2008. Social Network Analysis Applied to Knowledge Creation and Transfer in the Portuguese Agricultural R&D Field: an exploratory study. *Proceedings of the 9th European Conference on Knowledge Management*. Southampton, UK. pp. 585-592.

Newman, M. E. J., 2001. Scientific collaboration networks. *Physical ReviewE*, N.º 64.

Piette, M. et al., 1992. A study of the publication of scholarly output in economics journals. *Eastern Economic Journal*, Vol. 18, N.º 4, pp. 429-436.

Piette, M. J. and Ross, K. L., 1992. An analysis of determinants of co-authorship in economics, *Journal of Economic Education*, N.º 23, pp. 277-283.

Vidgen, R. et al., 2007. What sort of community is the European Conference on Information Systems? A social network analysis 1993-2005. *European Journal of Information Systems*, N.º 16, pp. 5-19.