

Information Science Network – ISN: social network analysis of scientific production of LIS field in Brazil

Renato Fabiano Matheus*

*rfmatheus@yahoo.com.br

Núcleo de Estudos em Tecnologias para Informação e Conhecimento – Netic (Brazil)

Federal University of Minas Gerais – UFMG, Information Science School (Brazil)

Introduction

This paper synthesizes previous results of the analyses of the network of co-authorship in scientific papers published by the Library and Information Science (LIS) research community in Brazil's main journals and annals of events. Data from 1991 to 2005 obtained from the Information Science Network (ISN) database – denominated RedeCI project – were analysed by means of statistical and network analysis techniques. Results show a low-density in the co-authorship network, few connections among different postgraduate programs, and a large number of sporadic authors. It is believed that this paper, and the RedeCI project, can stimulate further scientific network analysis studies, being a contribution to scientometric research. Finally, a general picture of the present reality of the Brazilian LIS research community is presented.

Methodology

Social network analysis is a methodology with a wide range of applications (Wasserman, Faust, 1994), many of which are used in the analysis of scientific production (Otte, Rousseau, 2002; Matheus, Silva, 2006). ISN (“RedeCI”) (Parreiras *et al.*, 2006) is a project that was conceived to give support to network analysis of data related to scientific collaboration. In order to achieve this goal, a database and a web site were set up (<http://ww.redeci.netic.com.br>) to allow the storage and retrieval of information about the network of authors of papers published since 1991 in the most important journals and annals of events in the LIS field in Brazil. The technological infrastructure adopted by RedeCI project is based on open software. Data about the scientific profile of the authors, such as the academic affiliation of professors, were obtained from their curriculum available at the Lattes platform (<http://lattes.cnpq.br/index.htm>), a governmental initiative to keep track of information about Brazilian scientific production. The selection of the most important publications was based on the classification of journals provided by the Capes evaluation system – named Qualis (<http://servicos.capes.gov.br/webqualis>). Data about authorship and the title of the selected papers were obtained from the journals' Web sites, from CDs of the events, and, in some cases, from a formatted file generated specifically for the RedeCI database. The resulting scientific network of co-authorship was

analysed using both UCINET (Borgatti, Everett, Freeman, 2002) and LOTKA (Rousseau; Rousseau, 2000) software. ISN adopts social network metrics, including centrality and power law degree distributions. Some statistical analysis, including numeric regression, is also used.

Network database

The following table contains information about data sources, their period and total numbers of papers and authors.

Table 1. Journals and annals of events in the RedeCI co-authorship network database

| Data Source* | Type (Event; Journal) | From | To | Nr. Paper | Nr. Author |
|--------------|-----------------------|------|------|-----------|------------|
| CINFORM | E | 2003 | 2005 | 127 | 195 |
| ENANCIB | E | 1994 | 2003 | 290 | 363 |
| CI | J | 1995 | 2005 | 337 | 469 |
| DGZ | J | 1999 | 2005 | 160 | 167 |
| EB | J | 1996 | 2005 | 72 | 96 |
| I&S | J | 1991 | 2005 | 219 | 256 |
| PCI | J | 1996 | 2005 | 156 | 195 |

Source: Parreiras *et al.*, 2006

* Data sources are the annals of events and main journals of the LIS community in Brazil, i.e.: CINFORM - *Encontro Nacional de Ensino e Pesquisa da Informação* (National Meeting of Teaching and Researching of Information); ENANCIB - *Encontro Nacional de Pesquisa em Ciência da Informação* (National Meeting of Research in Information Science); CI - *Ciência da Informação* (Information Science); DGZ - *DataGramaZero*; EB - *Encontros Bibli*; I&S - *Informação & Sociedade: Estudos* (Information and Society: studies); PCI - *Perspectivas em Ciência da Informação* (Perspectives in Information Science)

Table 2 shows the number of authors by type of publication. It is indicated if they have just one paper published (i.e., sporadic author) or more.

Table 2. Sporadic and prolific authors

| Type of publication | Nr. of Sporadic Authors (A) | Nr. of Prolific Authors (B) | Relation (A/B) |
|---------------------|-----------------------------|-----------------------------|----------------|
| Journals | 615 | 329 | 1,869 |
| Annals | 239 | 178 | 1,342 |
| Both | 854 | 507 | 1,684 |

Source: Parreiras *et al.*, 2006

Data analysis

Considering all the data available in the Rede CI, the scientific production was analysed using LOTKA

software. The results show that 72% of the papers have just one author, which indicates a low level of collaboration.

Table 3. Power Law Degree Distribution Estimate (Lotka's Law)

| Parameter | Value |
|-----------|--------|
| C | 0.7233 |
| Beta | 24.017 |

Source: Parreiras *et al.*, 2006

Centrality measures were made using UCINET for 949 authors that have at least one paper in co-authorship. Pearson correlation shows that different types of centrality measure provide different information about the network, with indicates that they should be used carefully.

Table 4. Pearson correlation for centrality measures

| Centrality Measure | Degree | Betweenness | Closeness |
|--------------------|--------|-------------|-----------|
| Degree | 1.000 | | |
| Betweenness | 0.356 | 1.000 | |
| Closeness | 0.178 | 0.275 | 1.000 |

Source: Parreiras *et al.*, 2006

Taking the group of authors with institutional academic affiliation, 116 teachers from 9 institutions were selected. Among them, 54 had papers in co-authorship with other teachers.

Table 5. Pearson correlation for centrality measures

| Institution* | Nr. Of Professors | Nr. of Professors With Co-authorship | % |
|--------------|-------------------|--------------------------------------|----|
| PUCCAMP | 8 | 4 | 50 |
| UFBA | 12 | 4 | 33 |
| UFF/IBICT | 18 | 14 | 78 |
| UFMG | 20 | 11 | 55 |
| UFPB | 7 | 0 | 0 |
| UFSC | 11 | 3 | 27 |
| UNB | 13 | 7 | 54 |
| UNESP | 11 | 5 | 45 |
| USP | 16 | 6 | 38 |
| Total | 116 | 54 | 47 |

Source: Silva *et al.*, 2006a

* Institutions that have postgraduate programs in Information Science in Brazil. See more at <http://www.capes.gov.br/>.

Comments

The bulk of the papers that had just one author (72%). The LIS postgraduate professors consist of a small group (116) scattered in nine academic centres across the country. A small community such as this is bound to produce few papers in co-authorships either in or between institutions (Silva *et al.*, 2006a). Beyond such group, there is a larger group of sporadic authors, indicating that postgraduate

students do participate in the scientific production of the field. Such results may indicate that LIS is not a mature area in Brazil. Future developments for the RedeCI project should include additional studies and developments such as: studies of different research branches (Silva *et al.*, 2006b); additional actor attributes; new importing schemes for data, as exemplified by OAI (Open Archive Initiative) protocol (<http://www.openarchives.org>); translation of RedeCI's web site; expansion of the network of papers to other knowledge areas; use of additional advanced network analysis techniques. This study also exemplifies an application of data about Brazilian researchers' profile made available through Lattes (<http://lattes.cnpq.br/index.htm>).

Acknowledgments

I want to thank Fernando Silva Parreiras, Antonio Braz de Oliveira e Silva, and Wladimir Cardoso Brandão for their partnership as coordinators of RedeCI (ISN). I thank my colleague Luiz Henrique Gomes for his careful review.

References

- Borgatti, S. P., Everett, M. G., Freeman, L. C (2002). *Ucinet for Windows: Software for Social Network Analysis*. Harvard, MA: Analytic Technologies.
- Matheus, R. F., Silva, A. B. de O. (2006). Análise de redes sociais como método para a Ciência da Informação, *DataGramaZero - Revista de Ciência da Informação*, 7, 2.
- Otte, E., Rousseau, R. (2002). Social network analysis: a powerful strategy, also for information sciences. *Journal of Information Science*, Thousand Oaks, 28, 6, pp. 441-453.
- Parreiras, F. S., Silva, A. B. de O., Matheus, R. F.; Brandão, W. C (2006). RedeCI: colaboração e produção científica em ciência da informação no Brasil. *Perspectivas em Ciência da Informação*, 11, 3, pp. .
- Rousseau, B., Rousseau, R (2000). LOTKA: A program to fit a power law degree distribution to observed frequency data. *Cybermetrics: International Journal of Scientometrics, Informetrics and Bibliometrics*, 4, 2, pp.1-13.
- Silva, A. B. de O., Matheus, R. F.; Parreiras, F. S.; Parreiras, T. S (2006a). Análise de redes sociais como metodologia de apoio para a discussão da interdisciplinaridade na ciência da informação. *Ciência da Informação*, 35, 1, pp. 72-93.
- Silva, A. B. de O., Parreiras, F. S., Matheus, R. F., Brandão, W. C. (2006b). Redes de co-autoria dos professores da Ciência da Informação: um retrato da colaboração científica dessa disciplina no Brasi. In: *Enancib*, 7, Marília, São Paulo, November, 2006.
- Wasserman, S., Faust, K (1994). *Social Network Analysis: methods and applications*. Cambridge: Cambridge University Press.