Small World Phenomenon and Author Collaboration: How Small and Connected Is the Digital Library World?

Monica Sharma and Shalini R. Urs

International School of Information Management, University of Mysore, Mysore, India
{monica,shalini}@isim.ac.in

Abstract. We present here the findings of our research to study the "Small World Phenomenon" and the scientific collaboration of authors in the Digital Library domain by using Social Network Analysis metrics. Co-authorship network of prolific authors is created and Social Network Analysis is carried out using UCInet.

1 Introduction

Using Social Network Analysis, it is possible to study the structural features of academic communities by examining their publications [1][3]. Liu et al. [2] studied the co- authorship network in Digital Library which focused on the DL conferences. The present work is an attempt to study the network characteristics of Digital Library as reflected in author collaboration and to study the small world phenomenon as expressed in Social Network Metrics.

2 Data and Analysis

Publication count and coauthor count of authors from World of Science (WoS) are used for identifying the front-runners in the DL domain. A search on 'digital libraries' yielded 1838 records. Eliminating authors with less than 5 papers gave us a set of 52 authors. DBLP record count and coauthor count of these 52 authors were also recorded. The 11 authors who scored high on all the three parameters were identified as the core members of the DL community. Co authorship network of these 11 authors with their coauthors (1053) was constructed using UCInet (Fig 1.).

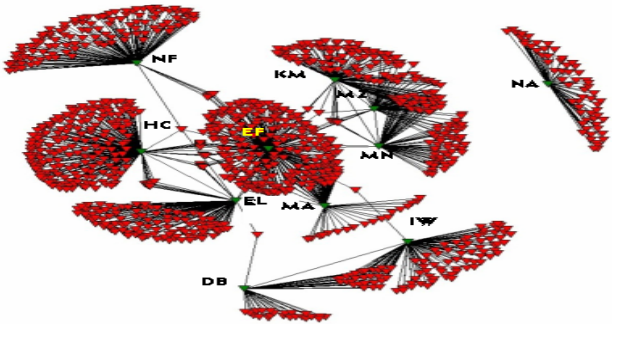

Fig. 1. Co authorship Network of authors in the field of Digital Library

D.H.-L. Goh et al. (Eds.): ICADL 2007, LNCS 4822, pp. 510–511, 2007.

Different Parameters such as Betweenness centrality, Degree centrality, Clique, etc. were studied. All the measures strongly suggested that Edward A. Fox and Hsinchun Chen are the most active members in terms of co authorship and record count. The network forms 2 components, Giant (94% authors) and Small component (6%). Table 1 shows various metrics of Social Network Analysis of top 5 authors under study (Values in parenthesis are of study carried out by Liu et al. [2]). It can be interpreted from the results of both the studies that Edward A. Fox (maximum number of cliques) and Hsinchun Chen hold an important position in the network in terms of centrality.

These measures indicate that they play a powerful role in the network and have a great influence over the flow and dissemination of resources in the network. The results indicate that the network is highly connected and that Digital Library community has significantly high international collaboration.

Table 1. Results of Social Network Analysis of co authorship network

Author		Degree	Betweenness	Closeness	Size	Ties	2Step reach
Edward A. Fox	(EF)	261(55)	342150.5(83163.9)	1.540(0.251)	261	194	72.05
Hsinchun Chen	(HC)	209(59)	167245.3(89250.9)	1.532(0.259)	209	56	52.95
Kurt Maly	(KM)	134	116338.6	1.525(0.212)	134	210	42.11
Ee-Peng Lim	(EL)	123	105330.0	1.531	123	50	53.04
Norbert Fuhr	(NF)	121	87432.5	1.514	121	0	11.79
Ian H. Witten	(IW)	97(18)	85283.3	1.507	97	56	11.5

Geodesic or the shortest path is the path between a pair of nodes which involve minimum number of nodes in between which connect the 2 nodes. Elmacioglu and Lee [1] studied the 6 degrees of separation in database community (DB) and found that average distance of coauthors in DB world is 6 whereas the network of Liu et al. [2] work yielded average distance of 3.6, which is approximately equal to our value. In our network the Average Geodesic between the reachable pairs is 3.5. The results reveal that compared to DB community, DL community is a small world and highly connected suggesting that dissemination of information can be fast in the network. This could be attributed to the fact that DL community is a small community.

3 Conclusion

The digital library community is indeed a small world lead by Edward A. Fox and Hsinchun Chen, having a significantly high international collaboration with an average geodesic distance of 3.5 between two pairs.

References

1. Elmacioglu, E., Lee, D.: On Six Degrees of Separation in DBLP-DB and More. ACM SIGMOD Record 34(2), 33–40 (2005)
2. Liu, X., et al.: Co-Authorship Networks in the Digital Library Research Community. Infor. Proces. & Mgt.-An International Journal 41(6), 1462–1480 (2005)
3. Newman, M.E.J.: Co authorship Networks and Patterns of Scientific Collaboration. PNAS 101, 5200–5205 (2004)