

Network of collaboration among PC members of Brazilian computer science conferences

Ana L.C. Bazzan · V.F. Argenta

Received: 14 September 2010 / Accepted: 2 May 2011 / Published online: 17 May 2011
© The Brazilian Computer Society 2011

Abstract The structure, dynamics, and importance of the social network of collaboration among scientists has been already studied, sometimes yielding counter-intuitive conclusions. In this paper we investigate the role played by people who served as PC (Program Committee) members in the network formed by members of the Brazilian computer science community and their co-authors. Some characteristics of such network are compared with those reported in similar studies involving other scientific collaboration networks. As a result, we show that apart from the evidence of Milgram's phenomenon (six degrees of separation), there is no other community with completely similar patterns (among those used for comparison). This is probably due to the unique characteristics of the target network. For instance, their members do not necessarily interact with each other in terms of co-authorship since they belong to different sub-areas of computer science. There are strong evidences that the clusters in this network are connected by non-Brazilian members. Moreover, nodes with high degrees have little connection to Brazilian authors.

Keywords Complex networks · Social networks

1 Introduction

The distance between people in social networks has been the object of various studies, as e.g. the Erdős number

that refers to collaboration among mathematicians, and the Bacon number, to mention only two that deal with co-authorship of scientific papers and participation in the same movie. After these studies, there have been some attempts to measure collaboration and other properties of social networks that relate to social networks of researchers in various areas.

One problem normally faced by such studies is to gather data that refer to a significant number of people that compose a given social network. In this paper we use data related to members of program committees of several conferences in different areas of computer science. This study differs from others as it involves people that would not necessarily interact among themselves given that they work in different areas. Also, these data refer primarily to more senior people, a characteristic that is expected when dealing with PC (program committee) members of conferences. In fact this is, to the best of our knowledge, the first study that looks at social networks of a discipline with such a bias.

The data used here relate to conferences sponsored by the Brazilian Computing Society. However, these data by no means involve Brazilian researchers only. On the contrary, a high number of people from many areas of computer science serve as PC members or are themselves authors, given that the majority of the conferences accept papers only in English and/or have established themselves in the international scientific community.

The Brazilian Computing Society provides a submission system for conferences and events it sponsors or supports. Similarly to EasyChair, EDAS, and others, the JEMS¹—Journal and Event Management System—aims at supporting the computer science community in Brazil in handling and managing submission of papers related to conferences

A.L.C. Bazzan (✉) · V.F. Argenta
Instituto de Informática, Universidade Federal do Rio Grande do Sul, Caixa Postal 15064, CEP 91501-970 Porto Alegre, RS, Brazil
e-mail: bazzan@inf.ufrgs.br

V.F. Argenta
e-mail: vfargenta@inf.ufrgs.br

¹<https://jems.sbc.org.br/>.

in areas as diverse as microelectronics and artificial intelligence. Therefore the data stored at JEMS refer to people who do not necessarily interact with each other in terms of co-authorship, thus making the study of this social network different from, for instance, the communities reported by M. Newman in [13] (e.g. mathematicians, physicists, or biologists) or in [7] (database community). Moreover, as mentioned, JEMS involves not only the Brazilian computer science community but also a large number of non-Brazilian authors and, especially, non-Brazilian reviewers.

The goal of the present paper is to investigate the properties of part of the social network associated with the community involved with JEMS in terms of collaboration (co-authorship), namely the members of the program committees (called TPC members in JEMS jargon). We restrict ourselves to TPCs because they include a large proportion of non-Brazilians and also a large number of researchers that are well-known in the computer science community. The senior character of these members is notorious and hence, makes the study different from those that we mention in the next section.

In order to reach that goal, we analyze the co-authorship network using data from DBLP, a well-known source of bibliographic information in computer science. The DBLP server provides bibliographic information on major computer science journals and proceedings. Initially the server was focused on DataBase systems and Logic Programming (hence the acronym DBLP), but it now covers major fields of computer science.

DBLP makes its data available in XML files that can be parsed. These data refer to publications (paper, book chapter, etc.) title, name of the publication where the given paper has appeared, name of author(s), etc. For our purposes we basically use the latter as we are interested in building the co-authorship social network.

Working with authors names (both in DBLP as well as in JEMS) of course leads to the problem of name disambiguation (name authority control problem), also mentioned in similar studies such as [7].

To tackle this problem we have used a very simple preprocessing step (manually fixing or removing inconsistencies in JEMS and using the DBLP default disambiguation). However, we remark that, recently, steps have been taken toward better handling this problem. For instance, in [9] the authors evaluated the use of social networks to solve the author name disambiguation problem in digital libraries. They introduce a set of match functions that combine traditional similarity functions with strength and importance of connections between authors in their social network. Experimental results showed that the use of social networks significantly improves the performance of syntax based similarity functions.

The present paper is organized as follows. In the next section we present basic concepts related to complex networks.

Section 3 discusses related works. In Sect. 4 we provide details about the method used as well as the data and setup. Section 5 presents and discusses the main results obtained so far. Section 6 summarizes the main conclusions and the future steps of this research.

2 Concepts related to complex networks

The research on complex networks has originated around works on percolation and random graphs (e.g. [8]), but has experienced a boom more recently when the underlying concepts were applied to problems related to the world-wide web, sociology, biology, and other sciences. In many cases one is interested in discovering and/or investigating structural or topological characteristics of such networks, by means of quantitative description of their properties.

Three particular topological models have been proposed and are widely used: random graphs (the Erdős and Rényi model) [8]; the small world of Watts and Strogatz [14]; and the Barabási–Albert model [2]. Many surveys exist for those interested in the details of these models. We refer the reader to [1] and to [6] (the latter is especially useful for issues related to metrics).

For our purpose it suffices to define some basic concepts; we restrict ourselves to the case of unweighted, undirected graphs. A graph G is a set $\mathcal{N}(G)$ of n vertices (or nodes) and a set of edges; each edge is identified by a pair of vertices. The degree of a vertex i (k_i) is given by the number of edges connected to it. The average degree is given by Eq. 1.

$$\langle k \rangle = \frac{1}{n} \sum_i k_i \quad (1)$$

Most networks of interest are sparse but two non-adjacent vertices i and j may be connected through a sequence of length m of edges. Many measurements related to the characterization of complex networks are based on the length of the connecting paths, especially the shortest ones. Some networks may have relatively few edges (in relation to n) thus $\langle k \rangle$ may be small and many vertices may be isolated. Moreover, a high number of small clusters exist. As more edges are added, clusters are connected to each other. At some point, most of the vertices are connected into a giant cluster. The proportion of vertices in the giant cluster can also be important in the determination of the characteristics of the network.

Regarding characterization of complex networks, the quantification of their properties is fundamental. Therefore a variety of measurements have been proposed (see [6] for a survey). For our purposes, we concentrate on those based on *degree*, *distance* (shortest and longest paths), *clustering*, and *centrality* (closeness and betweenness).

Degree was already defined. Be i, j vertices in $\mathcal{N}(G)$. For each i , define $d_{i,j}$ as the geodesic or *shortest* distance between i and each j to which i can be connected. Then the average geodesic distance is

$$\langle d \rangle = \frac{1}{n-1} \sum_{i \neq j} d_{i,j} \tag{2}$$

Another important measure relates to clustering level i.e. how tightly vertices are connected. There are at least two ways to measure clustering. Here we use the one given in Eq. 4 as, at least for now, we want to assign the same weight to each vertex.

In an undirected graph, the clustering coefficient of a node (vertex) i , C_i , measures how many edges are actually present in the neighbors N_i of i ($|E(N_i)|$), in comparison to a fully connected graph i.e. how close the neighborhood of i is of being a clique. Thus, in Eq. 3 $|E_{\max}(N_i)| = \frac{k_i \times (k_i - 1)}{2}$, with k_i as defined before.

$$C_i = \frac{|E(N_i)|}{|E_{\max}(N_i)|} \tag{3}$$

$$\langle C \rangle = \frac{1}{n} \sum_i C_i \tag{4}$$

Regarding centrality, we have measured the betweenness and closeness of each node i . Both relate to distance, at vertex level. Vertices that have short geodesic distances, have higher closeness as this is defined as the reciprocal of mean geodesic distance, as given in Eq. 5. Note that only the reachable nodes (from i) are considered in the computation. Closeness scores higher for more central vertices and can be regarded as a measure of how long it will take for information to spread from a given vertex to other reachable vertices in the network.

$$C_i = \frac{n-1}{\sum_{i \neq j} d_{i,j}} = \frac{1}{\langle d \rangle} \tag{5}$$

Betweenness measures how many shortest paths between any i and j pass through a vertex u . The more paths u participates, the more important it is. Thus, if $\sigma(i, u, j)$ is the number of shortest paths between i and j via u , and $\sigma(i, j)$ is the total number of shortest paths between i and j , we have

$$B_u = \sum_{i,j} \frac{\sigma(i, u, j)}{\sigma(i, j)} \tag{6}$$

These metrics just presented were used to evaluate the data from JEMS and DBLP (Sect. 5).

3 Related work

Grossman and Ion Batagelj and Mrvar, Melin and Peterson, Barabási and colleagues and, especially, Newman pioneered the study of social networks involving scientific collaboration. For a review please see [13]. By collaboration we mean here the connection that stems from the fact that two researchers have co-authored a paper.

Newman [11, 12] has used Medline and databases that include publications on physics, and, to a lesser extent, computer science, allowing the comparison among these disciplines. He has shown some differences in the patterns of collaboration in these areas. He has found that these collaboration networks form small worlds in which two authors randomly selected are typically separated by a short path, and has demonstrated the presence of clustering. These conclusions were somehow counter-intuitive if one considers the sizes of the sets involved (authors, publications, etc.).

Other authors have investigated the characteristics of social networks of single disciplines such as database [7] and evolutionary computation [5].

The former authors, working with DBLP data from 1968 to 2003 have also reported a small diameter and increasing trend in the level of collaboration, which the authors link to the increasing pressure for publication.

Barabási and colleagues [3] have focused on a database containing all relevant journals in mathematics and neuroscience for an 8-year period (1991 to 1998). Results indicated that this network is scale-free. Moreover, they were able to infer the dynamics and the structural mechanisms that govern the evolution and topology of this system. For instance, they have concluded that the diameter is decreasing.

More recently, Bird et al. [4] have compared different areas of computer science. However, their focus is on the investigation of similarities and differences among these areas (interdisciplinarity, connectivity, size, etc.), as well as on the characteristics of the authors themselves (dominance, assortivity), and of the main conferences of each area, for comparison among them.

4 Data, setup, and methods

As mentioned before, for our study we use data from JEMS as well as DBLP. JEMS stores various datasets such as names of conferences whose submissions it manages, submitted papers, names of authors, and names of people in the program committees, the so-called TPC. For this study we have downloaded the JEMS data as in January 2010. The DBLP data were downloaded in February 2010.

Overall the data involve the following figures:

- number of JEMS entries (TPC members): roughly 8K (7600)
- number of DBLP authors involved: roughly 80K (83,079)
- number of DBLP publications involved: roughly 800K

For each TPC member in JEMS, we have looked at his or her record at DBLP and retrieved how many co-authors this person has (direct co-authorship). We have done this for all people involved in program committees of JEMS.

Direct co-authorship between two vertices i and j means that distance $d_{i,j} = 1$. Following this, we have looked at further distances of relationship i.e. co-authors of co-authors of each TPC member, and so on. Finally, we have investigated standard properties of the resulting social network.

Next, we show and discuss some of the results.

5 Results

5.1 Degree

The first measure relates to number of direct co-authors i.e. the degree. It was found that the average degree $\langle k \rangle$ or number of co-authors per author is $\langle k \rangle = 3.13$, but there is a significant deviation in the overall distribution. Table 1 shows the 10 TPC members that have the highest number of co-authors (again, only publications listed in DBLP were used). At this stage we should also remark that the graph is not completely connected, as discussed later.

It is possible to see that only one author of this list appears in the list of the top 10 authors in terms of number of publications recorded in DBLP (E. Bertino).² The most connected person in this social network is Ian T. Foster, who co-authors papers (260 of them) with 508 different persons. For this author, we give in Table 2 not only the number of immediate co-authors (i.e. distance $d = 1$) but also the number of authors he reaches with distance 2, 3, etc. We see that he reaches the highest number of people with distance 5 (around 21K) but given that the universe of people we considered in DBLP is around 80K, this universe is not reached even with distance 10. This is a particular characteristic of this dataset, as will be clearer when we discuss centrality properties.

As mentioned before, the average degree (number of co-authors) is 3.13 but the distribution varies significantly. The number of authors without co-authors is 60, while around 55K authors have one co-author. The number of authors with 2, 3, ..., 10 co-authors is given in Table 3; of course at least one author has more than 500 co-authors, as seen in Table 1.

The complete plot of degree of collaboration versus number of authors with this degree is given in Fig. 1, where we

²This list can be found at <http://www.informatik.uni-trier.de/~ley/db/indices/a-tree/prolific/index.html>.

Table 1 Top 10 TPC members in JEMS in terms of collaboration

Number of direct co-authors ($d = 1$)	Name of author
508	Ian T. Foster
380	Ying Zhang
345	Elisa Bertino
315	Jack Dongarra
313	Carl Kesselman
294	Ewa Deelman
292	Edward A. Fox
286	Geoffrey Charles Fox
273	Howard Jay Siegel
254	Giannong Cao

Table 2 Collaboration network of the most connected TPC (authors reached with distance $d = 1$ to $d = 10$)

Degree d	Number of co-authors of Ian T. Foster	Accumulated number
1	508	508
2	2521	3029
3	8038	11,067
4	17,191	28,258
5	21,135	49,393
6	12,732	62,125
7	4376	66,501
8	1446	67,947
9	852	68,799
10	84	68,883

see that this distribution exhibits a power-law tail with exponent -1.85 . With 95% confidence, this exponent lies between -1.79 and -1.92 . We have performed an ANOVA analysis. Hereby, the significance of F is smaller than 0.05 thus permitting to reject the null hypothesis. Also, we remark that $R^2 = 0.944$.

5.2 Distance and clustering

We now discuss the characteristics of the resulting social network when seen as graph properties related to distance.

The giant component (largest subset of interconnected vertices in the graph) has 80,804 vertices, yielding a ratio of 0.973 to the whole graph. This may look counter-intuitive given that TPC members in JEMS stem from different areas of computer science. However, we remark that the whole graph also includes co-authors of those TPC members. This way, we can conclude that these co-authors have a role of connecting TPC members.

Table 3 Number of authors with a given number of co-authors (up to 10 co-authors only)

Number of co-authors ($k = 1$ only)	Number of authors
0	60
1	55,893
2	13,465
3	4884
4	2250
5	1271
6	740
7	499
8	349
9	272
10	197

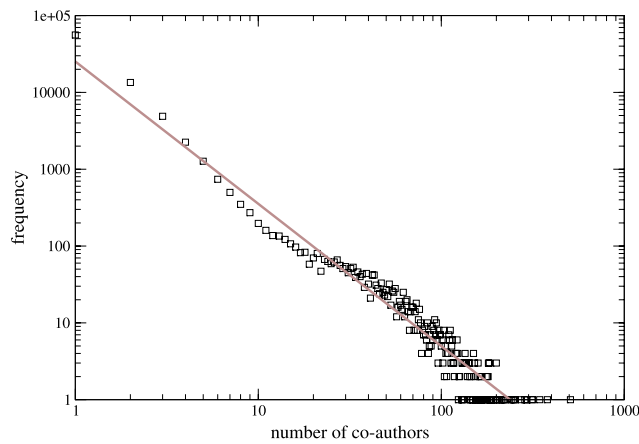


Fig. 1 Number of authors with a given number of co-authors: complete distribution

It would be interesting to have a plot of this measure along time as in [7] but unfortunately we do not have JEMS data of this kind.

Another important measure for our purpose is how authors are clustered. Considering all vertices, the global clustering coefficient $\langle C \rangle$ —computed here as the average over all C_i 's—yields $\langle C \rangle = 0.51$. The clustering coefficient of the giant component is roughly the same, as expected given that the giant component is almost the size of the whole graph. In [7] this coefficient reaches 0.63 (in 2003) but it was computed for the largest subgraph only, and refers, as put by the authors “to a tight community of people working on database only”.

We notice that a large number of authors (11K) have individual clustering coefficient near 1. This happens because many have at most 2 neighbors, who are themselves also interconnected. Most of them are non-JEMS people, with no or little connection with JEMS TPCs. In fact these vertices

Table 4 Clustering coefficient of top 10 TPC members in JEMS

Name of author	Clustering coefficient	Betweenness	Closeness
Foster	8.5E-05	2.86E-02	0.241127
Zhang	3.9E-04	2.13E-02	0.226683
Bertino	4.5E-03	7.62E-03	0.234095
Dongarra	5.0E-03	1.92E-02	0.222461
Kesselman	1.3E-02	7.73E-03	0.230369
Deeman	1.7E-02	1.81E-02	0.226285
E.A. Fox	4.4E-03	1.76E-02	0.227568
G.C. Fox	9.6E-03	1.71E-02	0.230322
Siegel	2.0E-03	1.48E-02	0.220378
Cao	2.9E-03	1.56E-02	0.229101

tend to belong to small sub-graphs of the network so that they form a clique.

The clustering coefficient of the JEMS TPC members who have high degrees (i.e. those appearing in Table 1) are given in Table 4. These low figures—that are in sharp contrast with e.g. values reported in [7] for the Database community—show that these TPC members (and their direct co-authors) have little connection with Brazilian authors, even if they have a large number of co-authors.

Next we give some figures that refer to the geodesic of all pairs of authors. In the JEMS social network, two TPC members may be related by their co-authors. There are several paths that may relate two given TPC members. In our case, we have measured the diameter of the social graph, i.e. the maximum of the pairwise distances in the giant component. This diameter is 17, a number similar to that measured for the database community in [7], which is 20. We remark that there are several paths with diameter 17.

Figure 2 depicts the distribution of sizes related to the shortest paths. For instance 36% of them have size 6. Less than 0.3% of these paths have size higher than 10. More interesting, we were also able to verify the phenomenon described by Milgran [10], namely that the average distance in the network is nearly six (6.07 in our case).

5.3 Centrality

As mentioned, regarding centrality, we have measured the betweenness and closeness of each node. Regarding betweenness we noted that only 2.3% of the authors have a value superior to 10^{-2} ; only around 10% have betweenness above 10^{-3} . This means that very few nodes have an importance in the task of connecting other nodes in the network. They participate in a relatively high number of paths between any two nodes.

As for closeness, 96% are between 0.1 and 0.2. We have also measured closeness centrality and found that 97% are smaller than 10^{-3} .

Table 5 Summary of the results for 5 communities. Results for *a*, *b* from [11]; for *c* from [7]; for *d* from [5]

	Medline ^a	Physics ^b	Database ^c	EC ^d	TPC JEMS/ JEMS + DBLP
Total authors	1,520,251	52,909	32,689	5492	7600/83,079
Mean collaborators per author	18.1	9.7	3.93	4.2	3.13
Size giant component	1,395,693	44,337	18,542	3686	80,804
Perc. giant component	92.6%	85.4%	57%	67.1%	97%
Clustering coefficient	0.066	0.43	0.63	0.798	0.51
Diameter	24	20	20	18	17
Mean distance	4.6	5.9	≈ 6	6.1	6.07

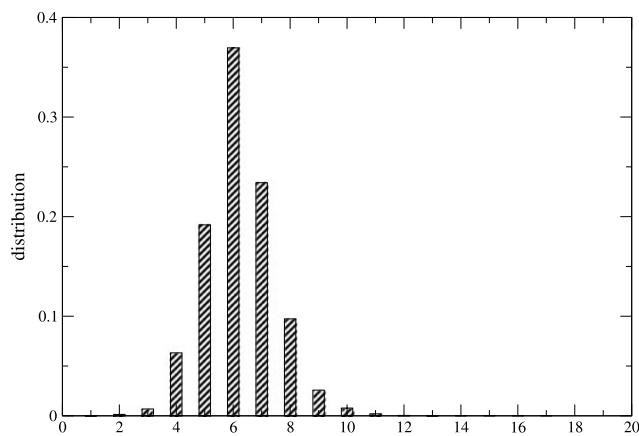


Fig. 2 Distribution of shortest paths

Table 4 lists clustering and also the betweenness and closeness of the authors with high degree. Results show that these people have relatively high betweenness and closeness scores. For sake of comparison, the top 3 Brazilians (regarding degree) have betweenness around 10^{-3} and closeness around 0.21.

From these measures we cannot say that there is a particular role played by the top 10. Their centrality scores are relatively high but so are those of some other authors. In fact, in order to assess the importance and centrality of the top 10 in the JEMS community, we would probably need to visualize the connections as well as the clusters. Given the number of nodes, this task cannot be done without making some simplifications. In any case, the top 10 play an important role in the conferences (managed by JEMS) themselves.

Another conclusion of this study is that although there are evidences that the top 10 TPC members (and possibly their non-Brazilian co-authors) act as liaison elements between clusters of the JEMS network, they are not well connected with the Brazilians in terms of co-authorship. This possibly indicates that there is still much work to be done in terms of the internationalization of this community. This is especially the case if we consider the *general* picture i.e. the majority of the Brazilians present in this network.

5.4 Comparison with other communities

In order to perform a comparison with investigations mentioned in Sect. 3, Table 5 summarizes the main figures. We can draw the following conclusions.

Similarly to the EC community, the JEMS TPC community is small when considered by itself (i.e. without the co-authors taken from DBLP). However, the similarities between these two only concerns geodesic: both the mean distance and the diameter are very similar. But given that these two metrics do not differ significantly for all five communities (with the possibly exception of Medline), these metrics are not really much informative.

In terms of average degree, the highest similarity is with the Database community but this may be explained by the fact that the main source of data for both studies was DBLP (though covering different time periods).

If one looks at the structure of the graphs, the JEMS community is similar to Medline in the sense that the giant component is almost the complete graph in both cases, which is a good indication that the community is connected. However the clustering coefficients for both communities are completely different.

Contrarily to studies mentioned in [13], the distribution of degree of collaboration with number of authors follows a power-law tail—see Fig. 1. This has been similarly reported for the Database community [7]. Also, among similar plots reported by Newman, the one that most approximates a power-law tail is exactly the one for the computer science community (NCSTRL).

In all cases (again, with the possible exception of Medline), the mean distance between pairs is around 6. As mentioned in [13], there is something special about the structure of such networks, namely that these authors are good at connecting themselves. In the case of JEMS in particular, our results show that non-Brazilian researchers were fundamental in reducing this distance.

6 Conclusion and future work

In this paper we have investigated the properties of the social network defined by researchers that have served as PC members in the Brazilian computer science conferences. This community is composed by non-Brazilian as well as Brazilian members, who tend to be senior researchers. In order to establish the connections between the nodes of this network, we have gathered data about co-authorship from DBLP.

It was found that contrarily to most of the studied communities (with exception of the Database one), in the JEMS TPC (plus DBLP) community, a power law in the distribution of number of co-authors was observed. Moreover, this community does not completely fit any of the other patterns we use as comparison, probably due to its own characteristics.

Other main findings were that most connected nodes are non-Brazilian TPC members. These however play an important role in the network by acting as connectors between Brazilian members. For instance, the giant component is almost the size of the whole graph meaning that non-TPC members (those who are co-authors of TPC members) are important connectors. Perhaps a lesson to be taken from the measures presented in Sect. 5 is that actions must be taken in order to extend the ties with the most connected nodes (Table 1). A concrete measure could be to invite these people (and probably their direct co-authors as well) to take part as external referees regarding main actions by the SBC (such as the “Computer Science Challenges in Brazil”), so that opportunities for further international collaborations could be recognized and/or fostered.

Regarding further comparisons, it remains to be seen how the JEMS dataset compares with data representative for the whole computer science community. However the literature does not report any broad attempt to do so. In [11] only data from NCSTRL were used, while authors in [4] use DBLP data but do not fully report macroscopic network properties. Therefore, in the future we intend to perform such comparison, possibly involving data from ACM digital library as well.

Also, we plan to focus on the Brazilian conferences and Brazilian authors and TPC only (possibly using data from the Brazilian Computer Society Digital Library), in order to be able to compare with the present results and thus derive a clearer picture about the role of the key people in the community. Similarly, it would be interesting to draw conclusions grounded on quantitative studies about the degree

of internationalization of members of this network. So far this was not possible because the pre-processing of the data from JEMS has excluded the affiliations.

Finally, we intend to carry out an investigation related to the detection of community structure within the JEMS graph. Intuitively, the JEMS TPC members should present a clear structuring in sub-communities (e.g. database, network, artificial intelligence, etc.). It is expected that most connections are to be found between vertices inside such communities, with a few connections between vertices of different communities. In order to verify this and possibly find out other interesting structural characteristics, visualization tools and metrics related to the hierarchical clustering of the network are key.

Acknowledgments We would like to thank Prof. L. Granville for making the JEMS data available to us. Also, we are grateful to the anonymous reviewers for their suggestions.

References

1. Albert R, Barabási A-L (2002) Statistical mechanics of complex networks. *Rev Mod Phys* 74(1):47–97
2. Barabasi AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–512
3. Barabási AL, Jeong H, Nédá Z, Ravasz E, Schubert A, Vicsek T (2002) Evolution of the social network of scientific collaborations. *Physica A* 311(3–4):590–614
4. Bird C, Barr E, Nash A, Devanbu P, Filkov V, Su Z (2009) Structure and dynamics of research collaboration in computer science. In: *Proceedings of the ninth SIAM international conference on data mining*. SIAM, Philadelphia, pp 826–837
5. Cotta C, Merelo J-J (2006) The complex network of EC authors. *SIGEVolution* 1(2):2–9
6. da Costa FL, Rodrigues FA, Travieso G, Boas PRV (2007) Characterization of complex networks: A survey of measurements. *Adv Phys* 56(1):167–242
7. Elmacioglu E, Lee D (2005) On six degrees of separation in DBLP-DB and more. *SIGMOD Rec* 34(2):33–40
8. Erdős P, Rényi A (1959) On random graphs, I. *Publ Math (Debr)* 6:290–297
9. Levin FH, Heuser CA (2009) Evaluating the use of social networks in author name disambiguation in digital libraries. In: *SBBD*, pp 46–60
10. Milgram S (1967) The small world problem. *Psychol Today* 2
11. Newman MEJ (2001) Scientific collaboration networks. I. network construction and fundamental results. *Phys Rev E* 64:2001
12. Newman MEJ (2001) Scientific collaboration networks. II. shortest paths, weighted networks, and centrality. *Phys Rev E* 64:2001
13. Newman MEJ (2003) The structure and function of complex networks. *SIAM Rev* 45:167–256
14. Watts DJ, Strogatz SH (1998) Collective dynamics of ‘small-world’ networks. *Nature* 393(6684):397–498