Ranking Strategy for Graduate Programs Evaluation

Giseli Rabello Lopes, Mirella M. Moro, Roberto da Silva, Eduardo M. Barbosa, and José Palazzo Moreira de Oliveira

***Abstract*—The demand for quality assessment criteria and associated evaluation methods in academia is increasing and has been the focus of many studies in the last decade. This growth arises due to the pursuit of academic excellence and support for the decision making of funding agencies. The high pressure from such scenario requires quality criteria objectively defined. In this paper, we develop an assessment procedure for graduate programs evaluation based on the internal collaborations among their research groups. These collaborations are evaluated through analysis on co-authorships networks based on novel metrics of social interaction. Furthermore, our procedure is easily reproduced and may be customized for evaluating any set of research groups. Our experiments show that the ranking provided by our metrics are according to the based (which is the official ranking defined by a national agency).**

***Index Terms*—Measurement; Quality Assessment; Social Networks.**

I. Introduction

The demand for quality assessment criteria in academia is increasing and has been the focus of studies in the last decade [1], [2], [3], [4], [5], [6]. This growth arises due to the pursuit of excellence in major areas of research. It is also motivated by other factors such as the competition for grants and the decision making of funding agencies. The high pressure on such scenario requires that quality criteria be objectively defined, preferably, by using a procedure that is easily reproduced. This last feature is important because the result of any ranking strategy may be questioned after its publication. Hence, it is desirable to enable anyone to reproduce the procedure in order to *double-check* the results.

Note that the quality of research groups has been used to assess other related features and the respective measures of quality have been used to define rankings. Examples include ranking journals and conferences based on the quality of their editorial boards, ranking universities based on the quality of their researchers and professors, and ranking research project proposals based on the quality of their researchers proponents. For all those rankings, the trend is to employ bibliometric techniques, especially citation statistics [6], [7], [8], [9].

This research is partially supported by MCT/CNPq projects, grants 57.3871/2008-6 (INWeb) and 475.530/2010-1 (Universal), and by the authors' scholarships and individual research grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil.

G. R. Lopes and J. Palazzo M. de Oliveira are with the Instituto de Informática, Universidade Federal do Rio Grande do Sul (UFRGS), Brazil (e-mail: {grlopes, palazzo}@inf.ufrgs.br).

M. M. Moro and E. M. Barbosa are with Departamento de Ciência da Computação, Universidade Federal de Minas Gerais (UFMG), Brazil (e-mail: {mirella, emb}@dcc.ufmg.br).

R. da Silva is with the Instituto de Física, Universidade Federal do Rio Grande do Sul (UFRGS), Brazil (e-mail: rdasilva@if.ufrgs.br).

ISBN: 978-0-9803267-4-1

Scientific research is commonly developed through collaboration involving different groups of researchers, for example in the areas of natural sciences and computer science. Therefore research communities have employed social network analysis (SNA) to understand their own interconnections, evolution and behaviors [10], [11]. More importantly, SNA allows to analyze the collaborations among researchers as well as quantify such scientific interaction behaviors, which in turn are two interesting facets for quality evaluation purposes.

Specifically, SNA assumes that the importance of relationships among the interaction units is a central point to the evaluation and analysis of social interaction. Such a perspective includes theories, models, and applications that are expressed in terms of relational concepts. Some fundamental concepts used on SNA include actors and relational ties [12], [13]. Actors are social entities that have social linkages modeled by the social network. Actors are linked to other actors by relational ties. The range and the kind of these ties can be quite extensible. Thus, a critical and defining feature of a Social Network is the presence of relational information.

In the aforementioned academic context, developing new methods in SNA is important because a particular unit on a network is not a single individual anymore, but an *entity* consisting of individuals and links among them. Also, since SNA is inherently an interdisciplinary effort, its concepts are developed as part of advances in social theory, empirical research and formal mathematics and statistics [13], [14].

For the academic context, one example of social network is a co-authorships network, in which actors represent researchers (authors) and relational ties represent their co-authored publications (i.e., the presence of at least one co-authored paper between two researchers determines a relational tie between them). For defining such a network, we need access to the researchers' publications. With the advance of digital libraries, this task is relatively easy, since we can use as data sources DBLP (Digital Bibliography & Library Project), Google Scholar, CiteSeer, BDBComp, ISI-JCR, among others.

In this paper, we develop an assessment procedure for graduate programs evaluation based on the internal collaborations among their groups of researchers. These collaborations are evaluated through analysis on co-authorships networks. Specifically, we propose new analysis to infer quality of researchers groups as social efficiency and highest eigenvalue. These measures adequately quantify the quality based on desired features in co-authorships networks, such as low number of non-collaborative researchers and large number of connected researchers with high density of collaborations. We present experiments that rank Brazilian Computer Science graduate programs through their faculty scientific behavior. In our experiments, the results are discussed and compared to an official classification scheme adopted by CAPES, a Brazilian funding agency dedicated to human resources qualification.

The remainder of this paper is organized as follows. Section II describes some related work. Section III overviews traditional measures used in SNA. Section IV introduces our new procedure to assess graduate programs and new measures to infer quality in co-authorships networks. Section V discusses the experiments and obtained results. Finally, section VI presents our main conclusions and summaries.

II. Related work

Previous work on quality assessment within the academic context aims to evaluate researchers, journals, conferences, institutions, among others. For example, some approaches are developed to evaluate other subjects considering the group of researchers involved, such as [8], [9]. Nonetheless, most approaches in the academic quality assessment consider bibliometric techniques, especially citation statistics. A thorough discussion about the advantages and disadvantages of using citation statistics is presented in [15]. Examples of such proposals based on citation statistics includes those to assess researchers [1], [3], [8], institutions [6] and publication venues [4], [5], [9].

One of the most popular indicators that consider bibliometric techniques is the h-index proposed by Hirsch [3]. The h-index is calculated based on the number of published papers in n years and the number of citations of each them. This number indicates that a researcher has h-index, if h of their N_p papers have at least h citations each one and the others $(N_p - h)$ papers have $\leq h$ citations each one. Then, one researcher with h-index has h published papers and each of them received at least h citations. Other measures for determining the quality/impact based on h-index (called h-type indices [2]) have been proposed. Some examples of them specifically introduced to evaluating research groups or institutions can be found at [7], [16], [17].

Combining different criteria to obtain the quality/impact assessment is also used [18]. Note that the advantage of using *one* unique criterion is the simplicity of the evaluation process. However, it is not easy to discover one unique facet of the problem that generates satisfactory results. On the other hand, using multiple criteria focuses on obtaining results that are better substantiated (even though, sometimes, not even using a handful of criteria can get to a good result).

In general, research is done through collaboration involving different groups of researchers. Research communities have employed social network analysis to understand their own interconnections, evolution and behaviors [10], [11], [19]. Examples vary from consolidated communities such as physics [20] and mathematics [21] to relatively new ones, such as information retrieval [22]. Then, analyzing the collaborations among researchers as well as quantifying these scientific interactions behavior define interesting criteria to the overall quality analysis.

Our work is inserted in the context of quality assessment in academia and it aims to evaluate graduate programs based on the internal collaboration of their groups of researchers. Unlike the aforementioned related work, ours does not consider citation statistics. The insight of our work is to explore other facet that can be used to infer quality/relevance of graduate programs: *the internal collaboration of research groups*. Therefore, differently from all previous work, we introduce social network analysis techniques to assess graduate programs. The related work in social networks uses analysis to comparative purposes and to understand the behavior of interactions but they are not intended to infer quality or construct a ranking of graduate programs such as our case.

III. Traditional Social Networks Analysis

This section presents some of the most traditional measures used in Social Networks Analysis [12], [20], [23], [24], [25]. The traditional way to represent a network is a graph $G := (\aleph, \xi)$, with nodes (vertices) $n \in \aleph$ and edges (links, linkages, connections) $e \in \xi$. Also, the nodes represent the network actors and the edges their relational ties. In our case, the actors represent the researchers and the relational ties are the co-authorships between researchers. For the equations presentation, the total number of nodes in the analyzed social network is presented as N; and $e(n_i, n_k)$ returns 1 when there is an edge between the indicated nodes (n_i and n_k) and 0 (zero) otherwise. Next, we overview some of the SNA metrics.

Degree Centrality. The concept of degree centrality presumes that a node that has many connections is considered important, while a node without connections is considered irrelevant. This degree reflects the direct relational activity of node [23]. The degree centrality of a node is calculated as the number of direct ties (edges) that involve a given node. Equation 1 presents the calculation of degree centrality of a node n_i, named $dc(n_i)$. If the network is an undirected graph, i.e. the connection between two nodes is not directional, the metric is just called degree. If the network is a directed graph, i.e. the connection between two nodes is directional, the metric is categorized into in-degree and out-degree according to the direction of the relationships being analyzed.

$$dc(n_i) = \sum_{k=1}^{N} e(n_i, n_k) \tag{1}$$

Density. The density of a network is defined based on the degree centrality. The density is calculated as the number of edges divided by the number of all possible edges of this network. A network entirely connected has density 1. The number of all possible edges will change according to the kind of graph describing the network. This concept is not useful when multiple edges are allowed or when the edges are weighted, because no total number of possible connections can be evaluated. If the network is an undirected graph where only one edge is allowed, the possible number of connections between each two nodes is 1 and equation 2 can be used to calculate the density (named as d) of the graph G representing the network. In this equation, the total number of edges is calculated by sum of the degree centrality ($dc(n_i)$) of all nodes and the number of all possible edges is calculated for a directed network as $(N(N-1)/2)$ where N represents the total number of nodes.

$$d(G) = \frac{2\sum_{i=1}^{N} dc(n_i)}{N(N-1)} \quad (2)$$

Clustering Coefficient. The clustering coefficient of a node n_i, named $cc(n_i)$, is calculated as the number of edges between neighbors of the node divided by the total number of all possible edges between the neighbors of node n_i. The clustering coefficient of a node aims to determine the density of edges established between the neighbors of a node. The concept of "transitivity" is applied describing symmetry of interaction among triples of nodes. Three nodes n_1, n_2, n_3 are transitive if n_1 is connected to n_2 and n_2 is connected to n_3 then vertex n_1 is connected to n_3. The transitive among triples of nodes is calculated as the number of triples that are transitive divided by the number of triples that have the potential to be transitive (paths of length 2). The overall clustering coefficient of a network (equation 3), named occ, is calculated as the average of clustering coefficient of all its nodes. For the clustering coefficient of a network, we can use a weighted clustering coefficient (equation 4), named wcc, that is calculated by a weighted average of the clustering coefficient of all the nodes, each one weighted by its degree (number of neighbors).

$$occ(G) = \frac{1}{N}\sum_{i=1}^{N} cc(n_i) \quad (3)$$

$$wcc(G) = \frac{\sum_{i=1}^{N} dc(n_i)cc(n_i)}{\sum_{i=1}^{N} dc(n_i)} \quad (4)$$

Giant Coefficient. The giant coefficient is calculated based on the size of the giant component of a network. The giant component, also known as main component, is the connected component with the largest number of nodes. The main component (MC) is a subgraph of graph G representing the network. The giant coefficient (equation 5), named as gc, is calculated as the size of the giant component divided by the total number of nodes of the network being analyzed. This value represents the percentage of nodes that are part of the giant component.

$$gc(G) = \frac{1}{N}\sum_{i=1}^{N} mc(n_i) \quad (5)$$

where:

$$mc(n_i) = \begin{cases} 1, & \text{if } (n_i \in MC) \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

As an example of the measures calculation, Fig. 1 shows a simple, generic network and the respective values for several measures. The values include the metrics density, clustering coefficient (overall and weighted) and giant coefficient as well as the new measures (social efficiency and highest eigenvalue) to be discussed in the next section (IV).

IV. Quality Measures to Rank Grad Programs

In this section we introduce the new procedure to assess quality of graduate programs as well as the new measures and analysis for assessing quality in co-authorships networks.

Measures	Values
Density	0.06
Overall Clustering Coefficient	0.26
Weighted Clustering Coefficient	0.18
Giant Coefficient	0.45
Social Efficiency	0.68
Highest Eigenvalue (binary)	2.32

Fig. 1. Examples of a social network and the respective measures values.

A. Overall Procedure

Fig. 2 describes the general procedure for assessing the quality of graduate programs. Given a list of graduate programs (remember we want to rank them at the end), it gets the list of their researchers (lines 1-2), and the set of publications for each of those researchers (lines 3-5). Once we have the researchers and their publications, the next step is to match their names and their publications, looking for their co-authorship relations (line 6). The list of researchers and their co-authorships will then define the social network (line 7), which in turn will be analyzed with SNA metrics (line 8). Finally, a ranking of the programs will be the output of the procedure (line 10).

Procedure QualityAssessment
Input: A list of graduate programs G
1. **for all** $g_i \in G$ **do**
2. R_i = Get the list of researchers that are currently members of g_i
3. **for all** $r_j \in R_i$ **do**
4. P_j = Get the list of publications of r_j from digital library
5. **end for**
6. E_i = Find the co-authorship relations within P
7. SN_i = Materialize the social network $SN(R_i, E_i)$
8. Apply SNA metrics to SN_i
9. **end for**
10. **return** $Ranking(SN)$

Fig. 2. Procedure for assessing quality of graduate programs.

B. New Measures

The new measures are based on the desired features for high quality research groups. The first one is the Social Efficiency that is based on the necessity of collaborative behavior in the group and the necessity for non-existant "social inefficient" individuals. The second one is the highest Eigenvalue analysis proposed to infer quality based on the necessity of large number of good researchers and the high density of collaborations. The measures are defined as follows.

Social Efficiency. The social efficiency aims to measure the percentage of nodes that contribute to the network connections. The social efficiency, named as se, is presented in equation 7, defined as one minus the social inefficiency. The social inefficiency is calculated by the number of nodes without edges in the network being analyzed, divided by the total number of nodes of the network, as presented in equation 8.

$$se(G) = 1 - si(G) \quad (7)$$

where:

$$si(G) = \frac{1}{N}\sum_{i=1}^{N} E(n_i) \quad (8)$$

$$E(n_i) = \begin{cases} 1, & \text{if } (dc(n_i) = 0) \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

Highest Eigenvalue. Spectral properties of graphs have pointed out interesting ways to describe their topologies [26]. However, these properties could also be used to measure the interaction level of persons in a graph. The highest eigenvalue of adjacency matrix of a graph (or even of valued-adjacency matrix, where the edges describe the level of interaction among pairs of authors) has two features that help to classify the interaction level of the group defined by its topology: larger matrix dimensions and large density of edges lead to an enlarge of highest eigenvalue.

Our hypothesis is that two features define high quality groups of researchers: (i) large number of good researchers; and (ii) high density of edges (which means a high level of joint publications, or even a good communication among researchers of the same group, and not only of these researchers with members of external groups).

Based on such hypothesis, we propose that two interesting metrics to quantify quality of groups of researchers are the highest eigenvalue of adjacency matrix and of valued-adjacency matrix. As we will see in Section V, our results do indeed agree with CAPES classification, which is used as baseline.

C. *Measures application to assessment and ranking purposes*

Based on the hypothesis that a good behavior is a high cooperative interaction among researchers, the co-authorship networks of top graduate programs must be as connected as possible. For evaluating such behavior and trying to infer quality, we use the measures presented in the Sections III and IV.

Besides evaluating graduate programs, the measures are used to build rankings ordered by quality. For all measures, the higher its value, the greater the quality of the analyzed network. This is the case of the following measures: density, clustering coefficient, giant coefficient, social efficiency and highest eigenvalue. These measures must be in descending order for the defining ranking. More details about the results of their application for quality ranking purposes are discussed in the next section.

V. Experiments

In this section, we present experiments that rank Brazilian Computer Science graduate programs through collaboration interactions among their group of researchers. First, we overview the classification from CAPES, which is employed as our baseline. Then, we describe the complete dataset and discuss the experimental results.

A. *Baseline*

With these experiments, we want to show that our new metrics (social efficiency and highest Eigenvalue) are appropriate to assess the quality of graduate programs. We have already discussed (on Related Work) that one single facet is not enough to qualify the graduate programs. Then, it is very reasonable that we do not use any of them as baseline for comparing our results. Hence, in this section we present the CAPES (http://www.capes.gov.br/) evaluation as baseline.

TABLE I
QUALITY EVALUATION OF BRAZILIAN COMPUTER SCIENCE GRADUATE PROGRAMS PERFORMED BY CAPES.

Graduate Program	CAPES Classification	Graduate Program	CAPES Classification
COPPE/UFRJ	7	UFCG	4
PUC/RIO	7	UFES	4
UFMG	7	UFPR	4
UFPE	6	UFRJ	4
UFRGS	6	UFRN	4
UNICAMP	6	UFSC	4
USP/SC	6	UFSCAR	4
UFF	5	UNB	4
USP	5	UNISINOS	4
PUC/PR	4	PUC/MG	3
PUC/RS	4	UCPEL	3
UFAM	4	UFG	3
UFBA	4	UFPA	3
UFC	4		

CAPES is the Brazilian federal agency for qualification of human resources and for official evaluation of university graduate programs. The graduate programs are classified into five levels (from 3 to 7). The classification is performed by experts and it is based on a series of criteria such as program proposal, faculty, students and thesis, intellectual production (publications), and social insertion. The agency has a complex set of equations in order to ponder all those facets. The top quality levels are 7 and 6 and they represent the graduate programs with performances comparable to top international programs. The evaluation is performed considering a period of 3 years. Table I shows the latest CAPES evaluation of 27 Computer Science graduate programs performed in year 2010 (evaluating the period of 2007 until 2009). This set of programs includes all those from levels 7, 6 and 5, and a selected set of levels 4 and 3. This table is the baseline of our experiments.

B. *Dataset Description*

The dataset used in these experiments includes the researchers of the 27 Brazilian graduate programs in Computer Science (from Table I) and their publications. Although we consider such programs, our procedure may be applied in programs from any of the other fields, such as physics and medicine.

In total, we have considered 732 researchers (i.e., the faculty members included in the lastest CAPES evaluation). Their publication data was extracted from DBLP (http://www.informatik.uni-trier.de/~ley/db) on August 03, 2010. We considered only the papers of published in conference proceedings or in journals indexed by DBLP until 2009 (since CAPES evaluation considers the publications from 2007 to 2009). Note that we need those publications in order to specify the co-authorship relations among the researchers. Also, recent studies [27], [28] discuss the coverage of computer science sub-fields by DBLP, which has reached the approximate value of 67%, covering up to 96% of some sub-fields. Nonetheless, DBLP is widely applied to obtain Computer Science publications (even though some exception results can be motivated by the limited coverage of some specific sub-field).

C. Evaluation and Results

In these experiments, we have followed the procedure from Fig. 2 for the graduate programs aforementioned. We have materialized the co-authorship social networks and some of them (one for each of the CAPES levels) are illustrated in Fig. 3. In this visual representation, researchers are represented by numbered points and the pairs of those who have published at least one paper together are linked by lines. A quick view of the figure is enough for realizing that the higher levels of CAPES classification programs have more connected researchers (i.e. higher collaborations interactions) while the others have lower collaboration/social behavior. Furthermore, the differences between the programs of levels 7 and 6 to the others are normally more apparent.

Now, in order to evaluate the results against the baseline, we need some measure that enables to compare rankings. One common way to do so is to employ the Spearman's coefficient to evaluate rank correlation. However, since CAPES ranks the programs by levels, all programs classified in a level can be considered as tied. Therefore, we use the variation of Spearman's coefficient that deals with tied ranks, as proposed in [29]. In summary: the higher the Spearman's coefficient value, the higher the correlation between the rankings being compared. Furthermore, the significance level of the obtained results is also evaluated. The statistical significance threshold of 0.01 is used and the results are summarized in Table II - which presents the results in descending ordered according to the Spearman's coefficient.

The results of Spearman's coefficient obtained by the traditional metrics of SNA (density, overall clustering coefficient and weighted clustering coefficient) were not satisfactory. The worst result of Spearman's coefficient was obtained by the density measure. Those three measures generate rankings that are *not* correlated to the CAPES classification (tested by a significance threshold of 0.01). This probably happened because considering the ideal network as a network totally connected can be very ambitious in the academic context. Indeed, this can depreciate the graduate programs with higher number of researchers, which even with a large number of connections can obtain a low density value (because the number of possible connections is very high). On the other hand, graduate programs with a small number of researchers may get an unfair advantage, because the number of possible connections is very restricted. Then, these three measures were not adequate to quantify quality for ranking graduate programs.

The highest eigenvalue measure was calculated for two adjacency matrices representing the weights of the analyzed social networks: (i) binary adjacency matrix, whose weights are binary (1 for presence and 0 for the absence of co-authored papers between pairs of researchers); and (ii) valued-adjacency matrix, whose weights represent the number of co-authored papers between pairs of researchers.

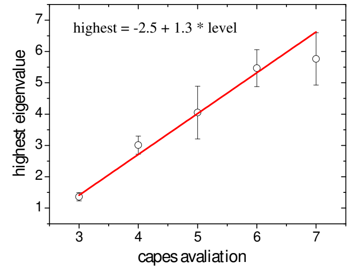


Fig. 4. Highest eigenvalue by CAPES evaluation.

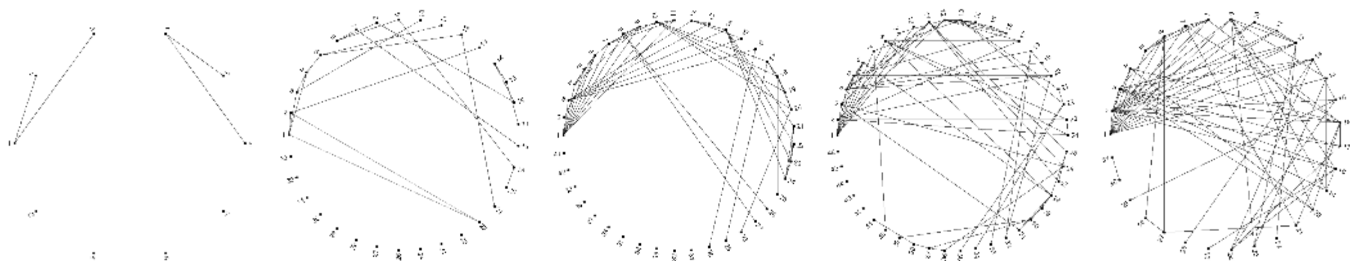
The measures giant coefficient, social efficiency, the difference between giant coefficient and social inefficiency, and highest eigenvalue (of both adjacency matrices, binary and valued) presented adequate results for ranking graduate programs. The correlations with CAPES classification was verified by a significance threshold of 0.01. All those measures obtained good values of Spearman's coefficient. Specifically, the best result (approximately 0.807) was obtained by the highest eigenvalue measure using binary adjacency matrix. For automated measures of quality assessment, the main desired features are the simplicity and objectivity. Then, among the presented measures, the social efficiency and highest eigenvalue are more adequate since they are very simple to calculate and both obtain satisfactory results.

A complementary analysis was performed to closely investigate the best results of the highest eigenvalue using binary adjacency matrix. This analysis was performed in order to determine a function for representing the variation of highest eigenvalue by CAPES level (presented in Fig. 4). Specifically, we separated the results of highest eigenvalues by the CAPES levels (from 3 to 7). We calculated the average and the standard error of values for each level, as illustrated by the results in Fig. 4. Note that the average is represented by the circle and the standard error interval by the lines around the circle. Then, we plotted the highest eigenvalue by CAPES classification and found a fit line determined as $highest = -2.5 + 1.3 * level$. This analysis shows that there is a linear behavior of the highest eigenvalues (determining the quality of internal collaboration) and the CAPES level.

VI. Conclusion

In this paper, we proposed a new procedure for quality assessment of graduate programs. We considered the hypothesis that research groups that are internally collaborative have more chance of achieving success and excellence in research than groups without social activity. Moreover, we proposed new measures to adequately quantify the quality for generating ranks.

We performed experiments using a real dataset of Brazilian graduate programs. The analysis showed that the researchers in top programs have the tendency to present a collaborative behavior. We also established a comparative analysis using the official evaluation performed by CAPES as baseline. The results showed evidences that one important facet of quality



(a) (b) (c) (d) (e)

Fig. 3. Examples of Social Networks modeling the internal collaborations among researchers of graduate programs. The programs were classified by CAPES at: (a) Level 3, (b) Level 4, (c) Level 5, (d) Level 6, and (e) Level 7.

TABLE II
SPEARMAN'S COEFFICIENT RESULTS OF THE DIFFERENT QUALITY MEASURES.

#	Quality Measures	Spearman's coeff.	Significance
1	Highest Eigenvalue (binary adjacency matrix)	0.807	Correlated
2	Giant Coefficient - Social Inefficiency	0.736	Correlated
3	Highest Eigenvalue (valued-adjacency matrix)	0.732	Correlated
4	Giant Coefficient	0.707	Correlated
5	Social Efficiency	0.682	Correlated
6	Weighted Clustering Coefficient	0.386	Not correlated
7	Overall Clustering Coefficient	0.325	Not correlated
8	Density	0.248	Not correlated

assessment of graduate programs is indeed to analyze the internal collaboration. Moreover, our new measures were more appropriate to assess quality for rankings. We also emphasized that our proposal of highest eigenvalue corroborates the excellence rank of the CAPES classification.

An interesting future work is to study the collaborations with external individuals. Some individuals identified as "social inefficient" may have collaborative behavior with external researchers.

REFERENCES

[1] L. Egghe, "An improvement of the h-index: the g-index," *ISSI Newsletter*, vol. 2, no. 1, pp. 8–9, 2006.

[2] ——, "The hirsch-index and related impact measures," *Annual Review of Information Science and Technology*, vol. 44, pp. 65–114, 2010.

[3] J. E. Hirsch, "An index to quantify an individual's scientific research output," *Proceedings of the National Academy of Sciences*, vol. 102, no. 46, pp. 16 569–16 572, Nov. 2005.

[4] F. Habibzadeh and M. Yadollahie, "Journal weighted impact factor: A proposal," *Journal of Informetrics*, vol. 2, no. 2, pp. 164–172, Apr. 2008.

[5] J. Nicolaisen and T. F. Frandsen, "The reference return ratio," *Journal of Informetrics*, vol. 2, no. 2, pp. 128–135, Apr. 2008.

[6] J. Ren and R. N. Taylor, "Automatic and versatile publications ranking for research institutions and scholars," *Commun. ACM*, vol. 50, no. 6, pp. 81–85, 2007.

[7] J.-F. Molinari and A. Molinari, "A new methodology for ranking scientific institutions," *Scientometrics*, vol. 75, pp. 163–174, 2008.

[8] R. da Silva et al., "Statistics for ranking program committees and editorial boards," *CoRR*, vol. abs/1002.1060, 2010.

[9] S. Yan and D. Lee, "Toward alternative measures for ranking venues: a case of database research community," in *Procs. of ACM/IEEE Joint Conference on Digital Libraries*, 2007, pp. 235–244.

[10] Y. Ding, "Scientific collaboration and endorsement: Network analysis of coauthorshipand citation networks," *Journal of Informetrics*, vol. 5, no. 1, pp. 187 – 203, 2011.

[11] C. Wang et al., "Mining advisor-advisee relationships from research publication networks," in *SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining*, Washington, DC, USA, 2010, pp. 203–212.

[12] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.

[13] S. Wasserman and K. Faust, *Social Network Analysis: methods and applications*. Cambridge University Press, 1994.

[14] J. A. Barnes, "Class and committees in a norwegian island parish," *Human Relations*, vol. 7, no. 1, pp. 39–58, Feb. 1954.

[15] R. Adler et al., "Citation statistics," Technical Report, Jun. 2008, available at: http://www.mathunion.org/Publications/Report/CitationStatistics.

[16] M. da Luz et al., "Institutional h-index: The performance of a new metric in the evaluation of brazilian psychiatric post-graduation programs," *Scientometrics*, vol. 77, pp. 361–368, 2008.

[17] A. F. J. van Raan, "Comparison of the hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups," *Scientometrics*, vol. 67, pp. 491–502, 2006.

[18] M. A. M. Souto et al., "An ontological approach for the quality assessment of computer science conferences," in *Procs. of Intl. Workshop on Quality of Information Systems*, Auckland, New Zealand, 2007.

[19] G. V. Menezes et al., "A geographical analysis of knowledge production in computer science," in *Procs. of Intl. Conf. on World Wide Web*, 2009, pp. 1041–1050.

[20] M. E. J. Newman, "Scientific collaboration networks. i. network construction and fundamental results," *Physical Review E*, vol. 64, no. 1, p. 016131, 2001.

[21] A.-L. Barabási, *Linked: The New Science of Networks*. Basic Books, 2002.

[22] A. F. Smeaton et al., "Analysis of papers from twenty-five years of sigir conferences: What have we been doing for the last quarter of a century," *ACM SIGIR Forum*, vol. 36, no. 2, pp. 49–53, 2003.

[23] L. C. Freeman, "Centrality in social networks: Conceptual clarification," *Social Networks*, vol. 1, pp. 215–239, 1979.

[24] B. Hoser et al., "Semantic network analysis of ontologies," in *The Semantic Web: Research and Applications*, 2006, pp. 514–529.

[25] P. V. Marsden, "Egocentric and sociocentric measures of network centrality," *Social Networks*, vol. 24, no. 4, pp. 407–422, 2002.

[26] M. Najiminaini, L. Subedi, and L. Trajkovic, "Analysis of internet topologies: a historical view," in *IEEE Int. Symp. Circuits Syst.*, May 2009, pp. 1697–1700.

[27] F. Reitz and O. Hoffmann, "An analysis of the evolving coverage of computer science sub-fields in the dblp digital library," in *Procs. of European Conf. on Research and Advanced Technology for Digital Libraries*, 2010, pp. 216–227.

[28] A. H. F. Laender et al., "Assessing the research and education quality of the top Brazilian Computer Science graduate programs," *SIGCSE Bulletin*, vol. 40, no. 2, pp. 135–145, 2008.

[29] S. Siegel and N. J. Castellan, *Nonparametric Statistics for the Behavioral Sciences*, 2nd ed. New York: McGraw-Hill, 1988.