Acta Scientiarum. Technology
ISSN: 1806-2563
eduem@uem.br
Universidade Estadual de Maringá
Brasil

de Oliveira, Sandra Cristina; Ferreira, Taiane de Paula; Barbero Brigantini, Beatriz
A comparative study on the reliability of co-authorship networks with emphases on edges and nodes
Acta Scientiarum. Technology, vol. 38, núm. 3, julio-septiembre, 2016, pp. 353-360
Universidade Estadual de Maringá
Maringá, Brasil

Acta Scientiarum

http://www.uem.br/acta
ISSN printed: 1806-2563
ISSN on-line: 1807-8664
Doi: 10.4025/actascitechnol.v38i3.26348

A comparative study on the reliability of co-authorship networks with emphases on edges and nodes

Sandra Cristina de Oliveira*, Taiane de Paula Ferreira and Beatriz Barbero Brigantini

*Faculdade de Ciências e Engenharia, Universidade Estadual Paulista "Julio de Mesquita Filho", Avenida Domingos da Costa Lopes, 780, 17602-496, Tupã, São Paulo, Brazil. *Author for correspondence. E-mail: sandra@tupa.unesp.br*

ABSTRACT. A scientific co-authorship network may be modeled by a graph G composed of k nodes and m edges. Researchers that make up this network may be interpreted as its nodes and the link between these agents (co-authored papers) as its edges. Current work evaluated and compared the reliability measure of networks with two emphases: 1) On nodes (perfectly reliable edges) and 2) On edges (perfectly reliable nodes). Specifically, the reliability of a fictitious co-authorship network at a given time t was analyzed taking into account, first, the reliability of nodes (researchers) equal and different, and, second, the reliability of edges (co-authorship relations), equal and different. Additionally, centrality measures of nodes were obtained to identify situations where the insertion of an edge significantly increased the reliability of the network. Results showed that the reliability of the co-authorship network focusing on edges is more sensitive to changes in individual reliabilities than the reliability of the network focusing on nodes. Additionally, the use of centrality measures was viable to identify possible insertions of edges or co-authorship relations to increase the reliability of the network in the two approaches.

Keywords: scientific co-authorship, theory of graphs, simulation, centrality measures.

Um estudo comparativo sobre confiabilidade de redes de coautoria com enfoques nas arestas e nos vértices

RESUMO. Uma rede de coautoria científica pode ser modelada por um grafo G composto por k vértices e m arestas. Os pesquisadores que compõem essa rede podem ser interpretados como seus vértices, e as relações entre esses agentes (trabalhos em coautoria) como suas arestas. Neste trabalho avaliou-se e comparou-se a medida da confiabilidade de redes considerando dois enfoques: nos vértices (arestas perfeitamente confiáveis) e nas arestas (vértices perfeitamente confiáveis). Especificamente, foi analisada a confiabilidade de uma rede de coautoria fictícia em um dado tempo t, considerando primeiramente as confiabilidades dos vértices (pesquisadores) iguais e distintas, e, em seguida, as confiabilidades das arestas (relações de coautoria) iguais e distintas; e, foram obtidas medidas de centralidade de vértices que auxiliaram na identificação de situações onde a inserção de uma aresta aumentava significativamente a confiabilidade desta rede. Os resultados mostraram que a confiabilidade da rede de coautoria com enfoque nas arestas é mais sensível às alterações das confiabilidades individuais do que a com enfoque nos vértices. Além disso, o uso de medidas de centralidade se mostrou viável na identificação de possíveis inserções de arestas ou relações de coautoria visando o aumento da confiabilidade da rede para os dois enfoques.

Palavras-chave: coautoria científica, teoria dos grafos, simulação, medidas de centralidade.

Introduction

Physical, biological or social systems characterized by a huge set of well-defined items that interact dynamically are networks. The maintenance of the functionality of a network requires information of its structure, functions and characteristics.

Since a network structure may be represented by a graph G composed of k nodes and m edges, the theory of graphs is basic to determine the properties, which refer to topological aspects of the network (Brigantini, Oliveira, & Braga Junior, 2014).

The network reliability is given by the probability of this network to remain operating even when one or more subsets of components (edges and/or nodes) are removed (Barlow & Proschan, 1981). Highly reliable networks are strong structures. Moreover, a network is more reliable than another if the probability of one network is disconnected is smaller than that of the other (Oliveira, Ferreira, Brigantini, & Uehara, 2014).

Social networks are structures composed of people, organizations, territories or others, connected among themselves by one or several types

of relationships (friendship, family, commercial etc.) through which information, knowledge, interests, values and aims are shared (Lyra & Oliveira, 2011).

Co-authorship networks may be included in this presented context. In fact, they are made up of researchers and shared tasks. Co-authorship networks are symmetrical in the sense that researcher A is a collaborator of researcher B at a given time t in the exact number of times that researcher B is a collaborator of A. Co-authorship of products obtained by scientific activities, with a special mention to scientific publications, indicates collaboration.

A co-authorship network is called highly reliable when is likely to continue producing science. If this probability decreases the reliability of the network also decreases.

The analysis of co-authorship networks is one of the most visible and accessible ways to identify scientific collaboration relationships within the academic milieu. Those networks have been extensively employed by the social network analysis (SNA) method which is specifically focused on the inter-relationship of agents (researchers). It aims at understanding the social mechanisms behind the connections (co-authorship, citations and others) and the manner they make easy the flow of information and knowledge among agents (Oliveira et al., 2014).

Several studies on SNA-employing co-authorship show that collaborations among researchers increased in all areas of knowledge and reveal the importance of co-authorship in the establishment of scientific production (Newman, 2004; Lee & Bozeman, 2005; Maia & Caregnato, 2008; Hayashi, Hayashi Lima, 2008; Yan & Ding, 2009; Abbasi, Altmann, & Hwang, 2010; Goldenberg et al., 2010; Souza & Barbastefano, 2011; Abbasi, Altmann, & Hossain, 2011; Barbastefano, Souza, Costa, & Teixeira, 2013; Abbasi, Wigand, & Hossain, 2014; among others).

On the other hand, network structural analysis precisely investigates the roles and positions of agents in the network and employs the concepts of the theory of graphs that provide the mathematical basis for computer calculations and use more specific approaches in data analyses. Recent studies on the reliability of co-authorship networks may be included since they have mainly emphasized combinatory distributions (Lyra & Oliveira, 2011; Brigantini, Oliveira, & Braga Junior, 2014; Oliveira et al., 2014).

Given this context, studies on the structural analyses of co-authorship networks underscoring reliability measures coupled to a methodology of SNA are relevant to maintain the working of the networks and their activities and for the emergence of strategies in survival and competiveness.

A network of scientific co-authorship may be modeled by graph G composed of k nodes and m edges. Researchers of the network are its nodes and connections or links among the agents (represented by common or co-authored publications) are the edges.

Current paper evaluate and compare the reliability measure of networks with two emphases: 1) On nodes (unreliable nodes or nodes prone to failure, or rather, one or more researchers leaving the network, and perfectly reliable edges); and 2) On edges (unreliable edges or edges prone to failure, or rather, one or more co-authorship relations among researchers may not exist anymore, and perfectly reliable nodes). Specifically, the reliability of a fictitious co-authorship network at a given time t was analyzed taking into account, first, the reliability of nodes (researchers) equal and different, and, second, the reliability of edges (co-authorship relations), equal and different. Additionally, centrality measures of nodes were obtained to identify situations in which the insertion of an edge may increase significantly the reliability of the network with regard to influence on nodes and on edges.

Material and methods

In a co-authorship network each researcher is represented by a node and two nodes are linked by one edge when the represented researchers have at least one publication in common. Then, the reliability of the co-authorship network is the probability of this network to be active at time t, even though one or more faults cause the removal of one or more subsets of nodes or edges of the graph.

Current essay comprises a network of fictitious scientific co-authorship modeled by graph G of Figure 1 with $k = 6$ nodes or researchers (each researcher i is denoted by $i = 1,2,3,...,6$, respectively) and $m = 6$ edges or co-authorship relations (each relation j is denoted by $j = 1,2,3,...,6$, respectively).

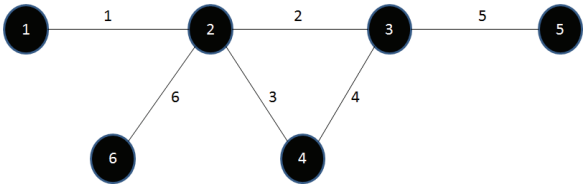

Figure 1. Graph G with six nodes or researchers (numbers in white) and six edges or co-authorship relations (numbers in black).

Calculation of the network reliability

Let a network be modeled by a simple undirected graph $G = (V, E)$ with k nodes and m edges. In order for the network to function (or to be active) at time t, every pair of nodes should be connected by at least one path. In this work were considered two situations:

Unreliable edges and perfectly reliable nodes (emphasis on edges)

Let´s suppose that the nodes are perfectly reliable and only the edges tend to be faulty. Therefore, each edge j ($j = 1,2,...,m$) has a probability of operation (reliability of edge j) denoted by p_j. There are instances in which all the edges of a graph that models the network have the same reliability, simply denoted by p. Further, edges are independent two by two. In other words, the failure of one does not imply the failure of the other.

In order for the reliability of a network (probability of graph G continues connected, even given the failure of one or more edges) may be calculated, the probability of each operating stage of the network must first be determined by Equation 1:

$$\prod_{j\in E'} p_j \prod_{j\in(E\setminus E')}(1-p_j) \quad (1)$$

in which E is the set of edges of graph G and E' is the set made up by the operating edges of graph G. When the edges of graph G that models the network have the same reliability p, the network reliability is given as Equation 2:

$$p_{R_G} = \sum_{j=k-1}^{m} S_j (1-p)^j p^{m-j} \quad (2)$$

in which G is the graph that models the network with k nodes and m edges; S_j is the number of connected sub-graphs of G with j edges (Kelmans, 1966).

When the edges of the graph that models the network have different reliabilities p_j, the network reliability p_{R_G} is calculated similarly as Equation (2), or rather, when the connected sub-graphs of G with j edges are obtained, the probability of each operating state of the network should be calculated and results added.

Unreliable nodes and perfectly reliable edges (emphasis on nodes)

Let´s suppose that the edges are perfectly reliable and only the nodes tend to be faulty. Therefore, each node i ($i = 1,2,...,k$) has a probability of operation (reliability of node i) denoted by p_i. There are instances in which all the nodes of a graph that models the network have the same reliability, simply denoted by p. Further, nodes are independent two by two. In other words, the failure of one does not imply the failure of the other.

In order for the reliability of a network (probability of graph G continues connected, even given the failure of one or more nodes) may be calculated, the probability of each operating stage of the network must first be determined by Equation 3:

$$\prod_{i\in V'} p_i \prod_{i\in(V\setminus V')}(1-p_i) \quad (3)$$

in which V is the set of nodes of graph G and V' is the set made up by the operating nodes of graph G. When the nodes of graph G that models the network have the same operating probability p, the network reliability is given as Equation 4:

$$p_{R_G} = \sum_{i=2}^{k} S_i p^i (1-p)^{k-i} \quad (4)$$

in which G is the graph that models the network with k nodes and m edges; S_i is the number of connected sub-graphs of G with i nodes (Goldschmidt, Jaillet, & Lasota, 1994).

When the nodes of the graph that models the network have different reliabilities p_i, the network reliability p_{R_G} is calculated similarly as Equation (4), or rather, when the connected sub-graphs of G with i nodes are obtained, the probability of each operating state of the network should be calculated and results added.

Centrality measures

Centrality measures are employed to verify the relevance of a node with regard to the others in a network. Through centrality measures, nodes may be ordered according to their relative importance.

Different centrality measures are used for different types of relevance (position, flux, influence and others). In this work it was employed two centrality measures that assess the importance of nodes in a network according to their structural position (Newman, 2010):

Closeness measure relates total distance of a node to other nodes of the network, or rather, it indicates the access velocity of a node to another in the network and shows the nodes that need improvement. Closeness measure of node i (v_i) is calculated by Equation 5:

$$C_p(v_i) = \sum_{j=1}^{k} d_{v_i v_j} \quad (5)$$

in which $d_{v_i v_j}$ represents the least distance between node i (v_i) and node j (v_j); k is the number of nodes in the network. The most central item of the network has the lowest rate of $C_p(v_i)$, or rather, the item that communicates with the highest speed with the other items of the network due to its structural position.

- Information degree measure gives relevance to a node due to the number of direct bonds that it establishes with other nodes of the network. In other words, it evaluates direct interference (or immediate effect for time $t+1$) of a node in the other by the number of measurement unit paths originating from a node. The calculation of the information degrees measure of node i (v_i) is given by Equation 6:

$$C_g(v_i) = d(v_i), 1 \le i \le k \quad (6)$$

in which k is the number of nodes in the network.

Results and discussion

Perfectly reliable nodes and unreliable edges (emphasis on edges or co-authorship relations)

Graph G of Figure 1 is approached for emphasis on edges, or rather, taking into consideration nodes, or totally reliable researchers, and edges, or relations of co-authorship prone to failure:

Let p, $j = 1,2,3,...,6$ be the reliability of each edge or the co-authorship relation j, such that $p_1 = 0.6$; $p_2 = 0.8$; $p_3 = 0.9$; $p_4 = 0.7$; $p_5 = 0.5$ and $p_6 = 0.8$. In practice, these reliabilities may be obtained by the frequentist probability of each edge or co-authorship relation, or rather, it is the number of co-authored publications of a pair of researchers with regard to total number of publications of this same pair of researchers for the specific network (Brigantini, Oliveira, & Braga Junior, 2014).

Four connected sub-graphs may be formed from graph G of Figure 2: three sub-graphs with five edges and one sub-graph with six edges. Due to the configuration of graph G, it is not possible to form connected sub-graphs with four or less edges. The reliability of the network modeled by graph G in time t may be given by $p_{R_G} = p_A + p_B + p_C + p_D$, in which, replacing the respective reliabilities, according as Equations 7 at 10:

$$p_A = p_1 \times p_2 \times p_3 \times p_4 \times p_5 \times p_6 = 0.12096 \quad (7)$$

$$p_B = p_1 \times (1-p_2) \times p_3 \times p_4 \times p_5 \times p_6 = 0.03024 \quad (8)$$

$$p_C = p_1 \times p_2 \times (1-p_3) \times p_4 \times p_5 \times p_6 = 0.01344 \quad (9)$$

$$p_D = p_1 \times p_2 \times p_3 \times (1-p_4) \times p_5 \times p_6 = 0.05184 \quad (10)$$

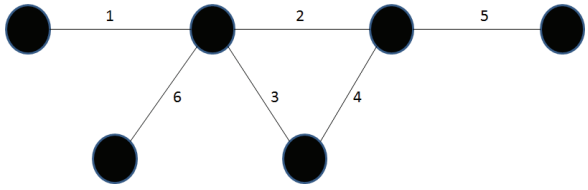

Figure 2. Graph G from Figure 1 with emphasis on edges or co-authorship relations.

Therefore, the reliability of the co-authorship network with emphasis on edges or co-authorship relation is given by Equation 11:

$$p_{R_G} = 0.12096 + 0.03024 + 0.01344 + 0.05184 = 0.21648 \quad (11)$$

Since all edges or co-authorship relations of graph G have the same reliability p, the reliability of the co-authorship network may be expressed by (2). If $S_5 = 3$ and $S_6 = 1$, then the reliability of the network is given by Equation 12:

$$p_{R_G} = \sum_{j=5}^{6} S_j p^j (1-p)^{m-j} = S_5 p^5 (1-p)^1 + S_6 p^6 (1-p)^0 = 3p^5(1-p) + p^6 \quad (12)$$

Table 1 shows results of simulations for different rates of p. One may observe the behavior of the network reliability as the reliability of each edge or the co-authorship relation (value of p) increases. Therefore, due to the configuration of the researchers´ network and the co-authorship relation, it may be verified that the probability of failure of edge or co-authorship relation above 0.6 ($p \le 0.4$) causes the reliability of the network in t to be close to zero.

Table 1. Reliability of the co-authorship network modeled by graph G (edges or co-authorship relations with equal reliabilities) for different rates of p.

Values of p	Network reliability p_{R_G}	Values of p	Network reliability p_{R_G}
0.9	0.708588	0.4	0.022528
0.8	0.458752	0.3	0.005832
0.7	0.268912	0.2	0.000832
0.6	0.139968	0.1	0.000028
0.5	0.062500		

Perfectly reliable edges and unreliable nodes (emphasis on nodes or researchers)

In the case of focus on nodes or researchers, or rather, taking into consideration the edges or relations of totally reliable co-authorship and the nodes or researchers prone to failure, graph G of Figure 1 may be thus approached.

Let p_i, $i = 1,2,3,...,6$ be the reliability of each node or researcher i, such that $p_1 = 0.6$; $p_2 = 0.7$; $p_3 = 0.8$; $p_4 = 0.9$; $p_5 = 0.5$ and $p_6 = 0.8$. In a real situation, reliabilities may be obtained by the frequentist probability of each node or researcher, i.e., it is the number of publications of a researcher for the network of scientific co-authorship with regard to total number of publications of this same researcher for that network and for other ends (Oliveira et al., 2014).

Twenty-six connected sub-graphs may be formed from graph G of Figure 3, with one sub-graph with six nodes, four sub-graphs with five nodes, seven sub-graphs with four nodes, eight sub-graphs with three nodes and six sub-graphs with two nodes. The reliability of the co-authorship network modeled by graph G in time t is given by $p_{R_G} = p_A + p_B + p_C + p_D + p_E$, where, replacing the respective reliabilities, according as Equations 13 to 17:

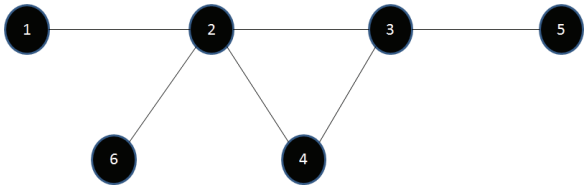

Figure 3. Graph G from Figure 1 with emphasis on nodes or researchers.

$$p_A = p_1 \times p_2 \times p_3 \times p_4 \times p_5 \times p_6 = 0.12096 \quad (13)$$

$$p_B = (1-p_1)\times p_2 \times p_3 \times p_4 \times p_5 \times p_6 + p_1 \times p_2 \times p_3 \times \times (1-p_4)\times p_5 \times p_6 + p_1 \times p_2 \times p_3 \times p_4 \times (1-p_5)\times p_6 + \quad (14)$$
$$+ p_1 \times p_2 \times p_3 \times p_4 \times p_5 \times (1-p_6) = 0.24528$$

$$p_C = (1-p_1)\times p_2 \times p_3 \times (1-p_4)\times p_5 \times p_6 + (1-p_1)\times \times p_2 \times p_3 \times p_4 \times (1-p_5) + (1-p_1)\times p_2 \times p_3 \times p_4 \times p_5 \times \times (1-p_6) + p_1 \times p_2 \times (1-p_3)\times p_4 \times (1-p_5)\times p_6 + p_1 \times \times p_2 \times p_3 \times (1-p_4)\times (1-p_5)\times p_6 + p_1 \times p_2 \times p_3 \times \times (1-p_4)\times p_5 \times (1-p_6) + p_1 \times p_2 \times p_3 \times p_4 \times (1-p_5)\times \times (1-p_6) = 0.18704 \quad (15)$$

$$p_D = (1-p_1)\times (1-p_2)\times p_3 \times p_4 \times p_5 \times (1-p_6) + (1-p_1)\times \times p_2 \times (1-p_3)\times p_4 \times (1-p_5)\times p_6 + (1-p_1)\times p_2 \times p_3 \times \times (1-p_4)\times (1-p_5)\times p_6 + (1-p_1)\times p_2 \times p_3 \times (1-p_4)\times p_5 \times \times (1-p_6) + (1-p_1)\times p_2 \times p_3 \times p_4 \times (1-p_5)\times (1-p_6) + p_1 \times \times p_2 \times (1-p_3)\times (1-p_4)\times (1-p_5)\times p_6 + p_1 \times p_2 \times (1-p_3)\times \times p_4 \times (1-p_5)\times (1-p_6) + p_1 \times p_2 \times p_3 \times (1-p_4)\times (1-p_5)\times \times (1-p_6) = 0.07444 \quad (16)$$

$$p_E = (1-p_1)\times (1-p_2)\times p_3 \times (1-p_4)\times p_5 \times (1-p_6) + + (1-p_1)\times (1-p_2)\times p_3 \times p_4 \times (1-p_5)\times (1-p_6) + + (1-p_1)\times p_2 \times (1-p_3)\times (1-p_4)\times (1-p_5)\times p_6 + + (1-p_1)\times p_2 \times (1-p_3)\times p_4 \times (1-p_5)\times (1-p_6) + + (1-p_1)\times p_2 \times p_3 \times (1-p_4)\times (1-p_5)\times (1-p_6) + + p_1 \times p_2 \times (1-p_3)\times (1-p_4)\times (1-p_5)\times (1-p_6) = = 0.01996 \quad (17)$$

Therefore, the reliability of the co-authorship network with emphasis on nodes or researchers is given by: $p_{R_G} = 0.12096 + 0.24528 + 0.18704 + 0.07444 + 0.01996 = 0.64768$. If all the nodes of graph G have the same reliability p, the reliability of the co-authorship network may be expressed by (4). Since $S_2 = 6$, $S_3 = 8$, $S_4 = 7$, $S_5 = 4$ and $S_6 = 1$, then the network reliability is given by Equation 18:

$$p_{R_G} = \sum_{i=2}^{6} S_i p^i (1-p)^{6-i} = p^6 + 4p^5(1-p)^1 + + 7p^4(1-p)^2 + 8p^3(1-p)^3 + 6p^2(1-p)^4 \quad (18)$$

Table 2 shows the results of simulations for different rates of p. One may observe the behavior of the network reliability in proportion to the increase of the reliability of each node or researcher (value of p). Therefore, owing to the configuration of this network and the relationships of existing co-authorship, the probability of the node´s or researcher´s failure above 0.8 ($p \leq 0.2$) causes the reliability of the network in time t to be close to zero.

According to the configuration of the fictitious co-authorship network of Figure 1 and the relationships of the existing co-authorship, when the two situations for different rates of p (emphasis on edges and then emphasis on nodes) are compared, one may see that as p decreases, the reliability of the

network becomes close to zero more quickly in the first case (edges or unreliable co-authorship relations) than in the second (nodes or unreliable researchers), as Figure 4 shows. Therefore, the reliability measure of network p_{R_G} with emphasis on edges or co-authorship relation is more sensitive to changes of individual reliabilities p than that of the reliability of the network with emphasis on nodes or researchers.

Table 2. Reliability of the co-authorship network modeled by graph *G* (nodes or researchers with equal reliabilities) for different values of p.

Values of p	Network reliability p_{R_G}	Values of p	Network reliability p_{R_G}
0.9	0.819882	0.4	0.328192
0.8	0.677888	0.3	0.239058
0.7	0.568498	0.2	0.139328
0.6	0.482112	0.1	0.045802
0.5	0.406250		

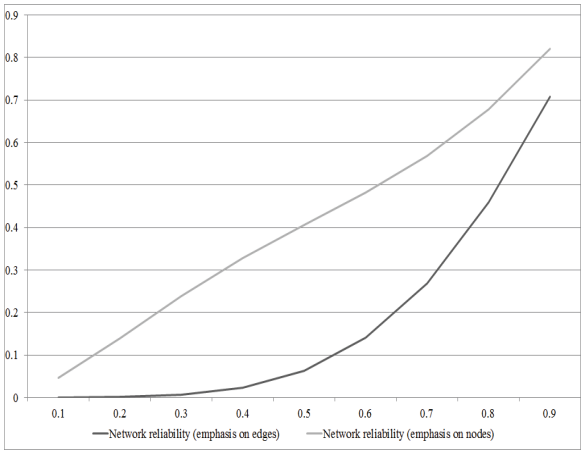

Figure 4. Comparison between reliabilities of the co-authorship network modeled by graph *G* (p_{R_G}) shown in Table 1 and 2 for different rates of p.

So that the reliability of the fictitious co-authorship network could be increased and the importance of centrality measures in this context could be introduced, within all possible insertion options of a new edge (or co-authorship relation) in graph *G* of Figure 1, nine non-isomorph graphs[1], described in Table 3, could be produced.

Since the nine graphs of Table 3 have equal reliabilities p (for emphasis on edges and on nodes), a simulation was conducted for different rates of p whose results of reliability calculated for *G*1, *G*2,..., *G*9 are represented in Tables 4 and 5, below. Results were then compared with the reliabilities of the network modeled by graph *G*, calculated in Tables 1 and 2, to identify which graph(s) provided the highest increase in reliability of *G* within all possibilities of insertion of an edge (link between nodes or researchers who were not previously extant).

Table 3. Non-isomorph graphs obtained from *G* (Figure 1) introducing an edge or co-authorship relation.

Graph	Link	Graph	Link
*G*1	Researchers 1 and 6	*G*6	Researchers 1 and 5
*G*2	Researchers 4 and 6	*G*7	Researchers 1 and 4
*G*3	Researchers 1 and 3	*G*8	Researchers 2 and 5
*G*4	Researchers 4 and 5	*G*9	Researchers 5 and 6
*G*5	Researchers 3 and 6		

In the case of unreliable edges (Table 4), it should be noted that as from the insertion of an edge or co-authorship relation in graph *G*, the graphs *G*6 and *G*9 had respectively the highest increase in reliability. Therefore, the insertion of an edge or co-authorship link between nodes or researchers '1' and '5' and between researchers '5' and '6' provided an average 32.87 (*G*6) and 31.22% (*G*9) higher reliability of the network modeled by graph *G*.

In the case of unreliable nodes (Table 5), one notes that as from the insertion of an edge or co-authorship relation in graph *G*, the graphs *G*8, *G*6 and *G*9 had respectively the highest increase in reliability. Therefore, the insertion of an edge or co-authorship link between nodes or researchers '2' and '5' and between researchers '1' and '5', or similarly, between researchers '5' and '6' increased, on an average, by 22.26 (*G*8) and 20.33% (*G*6 and *G*9) the reliability of the network modeled by graph *G*.

Centrality measures may be employed to verify how a node or researcher of a co-authorship network is relatively more important with regard to the others and to indicate among which nodes or researchers a new edge may be inserted for maximum reliability within the network.

Two centrality measures of nodes, previously given, were taken into account in the analysis of the fictitious co-authorship network modeled by graph *G*: Closeness measure and Information degree measure.

The former refers to the path that the node has to go to reach the others and the latter refers to the direct links that the node does to the others. The node with the lowest closeness measure is the most central of the network, or rather, that which communicates more rapidly with the other nodes. In the case of the information degree measure, the node that has the highest rate (or degree) is that which has the best direct contact with the other nodes.

[1]Isomorph graphs have the same structure; in other words, they have the same number of nodes and edges, albeit a different pattern.

Table 4. Reliability of non-isomorph graphs obtained from *G* introducing an edge (edges with equal reliabilities) for different values of p.

Values of p	*G*1	*G*2	*G*3	*G*4	*G*5	*G*6	*G*7	*G*8	*G*9
0.9	0.850306	0.785352	0.803066	0.785352	0.791257	0.862115	0.7617321	0.7794468	0.856211
0.8	0.642253	0.563610	0.602931	0.563610	0.576717	0.668467	0.5111808	0.5505024	0.655360
0.7	0.430259	0.364712	0.410091	0.364712	0.379838	0.460512	0.3042067	0.3495856	0.445386
0.6	0.251942	0.208397	0.245722	0.208397	0.220838	0.276826	0.1586304	0.1959552	0.264384
0.5	0.125000	0.101563	0.125000	0.101563	0.109375	0.140625	0.0703125	0.0937500	0.132813
0.4	0.049562	0.039731	0.050790	0.039731	0.043418	0.056934	0.0249856	0.0360448	0.053248
0.3	0.013997	0.011105	0.014677	0.011105	0.012296	0.016378	0.0063423	0.0099144	0.015188
0.2	0.002163	0.001702	0.002317	0.001702	0.001907	0.002573	0.0008832	0.0014976	0.002368
0.1	0.000078	0.000061	0.000086	0.000061	0.000069	0.000095	0.0000289	0.0000532	0.000086

Table 5. Reliability of non-isomorph graphs obtained from *G* introducing an edge (nodes with equal reliabilities) for different values of p.

Values of p	*G*1	*G*2	*G*3	*G*4	*G*5	*G*6	*G*7	*G*8	*G*9
0.9	0.819963	0.827253	0.887031	0.892863	0.827982	0.900153	0.827253	0.900882	0.900153
0.8	0.678912	0.699392	0.769024	0.781312	0.703488	0.801792	0.699392	0.805888	0.801792
0.7	0.572467	0.603337	0.663019	0.675367	0.612598	0.706237	0.603337	0.715498	0.706237
0.6	0.491328	0.525888	0.570816	0.577728	0.539712	0.612288	0.525888	0.626112	0.612288
0.5	0.421875	0.453125	0.484375	0.484375	0.46875	0.515625	0.453125	0.53125	0.515625
0.4	0.348928	0.371968	0.391936	0.387328	0.385792	0.410368	0.371968	0.424192	0.410368
0.3	0.260667	0.273897	0.284859	0.279567	0.283158	0.292797	0.273897	0.302058	0.292797
0.2	0.155712	0.160832	0.165184	0.162112	0.164928	0.167232	0.160832	0.171328	0.167232
0.1	0.052363	0.053173	0.053911	0.053263	0.053902	0.054073	0.053173	0.054802	0.054073

Table 6 shows the centrality measures calculated for the nodes of graph *G* of Figure 1. The table shows that, according to the closeness measures and information degree measures, the most central node or researcher of graph *G* is '2', or rather, it has the highest access speed and the greatest influence on the other nodes or researchers. Therefore, if the node or researcher '2' were to be removed from the graph (for any reason), the network become less connected and consequently with decreased reliability, since certain pathways were eliminated.

Table 6. Centrality measures of nodes for the graph *G*.

Nodes of the graph *G*	Closeness measures	Information degree measures
1	10	1
2	6	4
3	7	3
4	8	2
5	11	1
6	10	1

On the other hand, nodes or less central researchers of graph *G* are nodes '5', '1' and '6', respectively. Taking into consideration results in recent studies on the reliability of co-authorship network (Lyra & Oliveira, 2011; Oliveira et al., 2014; Brigantini, Oliveira & Braga Junior, 2014), one must pinpoint links of nodes '1' and '5' or '5' and '6' if the aim is to make the network more reliable through the insertion of an edge or relation between nodes or researchers.

Simulation results in Table 4 (emphasis on edges) showed that the greatest increases of reliability of the fictitious network of researchers modeled by graph *G* were precisely obtained with the insertions of the edges between nodes or researchers '1' and '5' and between researchers '5' and '6', respectively. The above corroborated totally results obtained by centrality measures.

Conclusion

Studies on the reliability of scientific co-authorship network identify which networks are reliable from different approaches (edges and/or nodes) according to the participation of researchers and the intensity of extant co-authorship relations.

The advantage of an approach with emphasis on nodes is that it is possible, in a future research, to measure the importance of each researcher in the network through the calculus of the network reliability conditioned on its absence. Although the approach considering both edges and nodes is desired there are some reasons for not considering it. First, because it may be intractable. And second, the extra information obtained with the addition of the edge to the approach considering just the nodes may be negligible. The same logic may be considered for an approach focused on edges.

The results of this work showed that the measure of reliability of the co-authorship network with emphasis on edges or relations of co-authorship is more sensitive to changes in individual reliabilities of the network with emphasis on nodes or researchers.

The example provided showed that the calculation of reliability of a co-authorship network may be stressing when executed manually or by computer. The employment of centrality measures may be considered a feasible alternative. However, the use of other centrality measurements and the execution of simulations for more trust-worthy results are recommended besides the employment of these measurements.

Acknowledgements

The authors thank CNPq for financial support (Ed. 18/2012 – Proc. 406626/2012-0) and funding the scientific initiation scholarships (Pibic/CNPq – Proc. 28151/2013 and Fapesp – Proc. 2012/01690-8).

References

Abbasi, A., Altmann, J., & Hossain, L. (2011). Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network measures. *Journal of Informetrics, 5*(4), 594-607.

Abbasi, A., Altmann, J., & Hwang, J. (2010). Evaluating scholars based on their academic collaboration activities: two indices, the RC-index and the CC-index, for quantifying collaboration activities of researchers and scientific communities. *Scientometrics, 83*(1), 1-13.

Abbasi, A., Wigand, R. T., & Hossain, L. (2014). Measuring social capital through network analysis and its influence on individual performance. *Library & Information Science Research, 36*(1), 66-73.

Barbastefano, R. G., Souza, C., Costa, J. S., & Teixeira, P. M. (2013). Impactos dos nomes nas propriedades de redes sociais: Um estudo em rede de coautoria sobre sustentabilidade. *Perspectivas em Ciência da Informação, 18*(3), 78-95.

Barlow, R. E., & Proschan, F. (1981). *Statistical theory of reliability and life testing*. New York City, NK: Holt, Rinehart and Winston.

Brigantini, B. B., Oliveira, S. C., & Braga Junior, S. S. (2014). Classical statistical inference for the reliability of a co-authorship network with emphasis on edges. *American International Journal of Contemporary Research, 4*(3), 19-31.

Goldenberg, J., Libai, B., Muller, E., & Stremersch, S. (2010). The evolving social network of marketing scholars. *Marketing Science, 29*(3), 561-567.

Goldschmidt, O., Jaillet, P., & Lasota, R. (1994). On reliability of graphs with node failures. *Networks, 24*(4), 251-259.

Hayashi, M. C. P. I., Hayashi, C. R. M., & Lima, M. Y. (2008). Análise de redes de coautoria na produção científica em educação especial. *Liinc em Revista, 4*(1), 84-103.

Kelmans, A. (1966). Connectivity of probabilistic networks. *Automatic Telemekhania, 3*, 98-116.

Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science, 35*(5), 673-702.

Lyra, T. F., & Oliveira, C. S. (2011). Um estudo sobre confiabilidade de redes e medidas de centralidade em uma rede de co-autoria. *Pesquisa Operacional para o Desenvolvimento, 3*(2), 160-172.

Maia, M. F. S., & Caregnato, S. E. (2008). Coautoria como indicador de redes de colaboração científica. *Perspectivas em Ciência da Informação, 13*(2), 18-31.

Newman, M. E. J. (2004). Coauthorship networks and patterns of scientific collaboration. *PNAS, 101*(1), 5200-5205.

Newman, M. E. J. (2010) *Networks*: An introduction. New York City, NK: Oxford University Press.

Oliveira, S. C., Ferreira, T. P., Brigantini, B. B., & Uehara, J. K. (2014). Inferência estatística clássica para a confiabilidade de rede de coautoria com enfoque nos vértices. *Perspectivas em Ciência da Informação, 19*(4), 202-225.

Souza, C. G., & Barbastefano, R. G. (2011). Knowledge diffusion and collaboration networks on life cycle assessment. *The International Journal of Life Cycle Assessment, 16*(6), 561-568.

Yan, E., & Ding, Y. (2009). Applying centrality measures to impact analysis: A coauthorship network analysis. *Journal of the American Society for Information Science and Techonology, 60*(10), 2107-2118.

Received on January 20, 2015.
Accepted on October 22, 2015.