

Análise da colaboração científica departamental através de redes de coautoria

Andréa Sabedra Bordin

Professora Assistente da Universidade Federal do Pampa (UNIPAMPA).Doutoranda do Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento (EGC) na Universidade Federal de Santa Catarina (UFSC)

Alexandre Leopoldo Gonçalves

Professor do Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento (PPGEGC) na Universidade Federal de Santa Catarina (UFSC).Bolsista CNPq

José Leomar Todesco

Professor do Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento (PPGEGC) na Universidade Federal de Santa Catarina (UFSC). Bolsista CNPq

<http://dx.doi.org/10.1590/1981-5344/1796>

O entendimento da rede de colaboração de um departamento de Pós-graduação permite a identificação de uma série de itens relacionados ao processo de colaboração científica, tais como: nível de colaboração da comunidade envolvida, distância média entre os colaboradores, autores que mais colaboram, grupos que colaboram com mais intensidade, etc. Através da produção bibliográfica informada ao sistema Coleta de Dados da CAPES, referente ao período de 2005-2012, redes de coautoria foram criadas e analisadas através de métricas de análise de rede social. Foi encontrado um índice médio de colaboração de 3,4 autores por publicação, redes de coautoria com baixa densidade, um maior número de componentes nos primeiros anos do programa e autores com alta produtividade, porém, com baixa centralidade de grau. Acredita-se que a análise e discussão dos resultados apresentados nesse artigo alcançaram o objetivo macro de diagnosticar a produção bibliográfica do programa, em termos de colaboração científica, e espera-se que seja uma ferramenta útil para tomada de decisão em nível organizacional e individual.

Palavras-chave: *Colaboração; Rede de coautoria; Análise de rede social.*

Departmental scientific collaboration analysis through coauthorship networks

The understanding of the collaboration network of a department of postgraduate allows the identification of a number of items related to the process of scientific collaboration, such as level of cooperation of the community involved, average distance between collaborators, authors who most collaborate, groups collaborating with more intensity etc. Through the bibliographic production informed to CAPES system for the period 2005-2012 coauthorship networks were created and analyzed using social network analysis. It was found an average of 3.4 collaborative authors per publication, coauthorship networks with low density, a larger number of components in the early years of the program and authors with high productivity, but with low degree centrality. It is believed that the analysis and discussion of the results presented in this paper have achieved the goal of diagnosing macro bibliographic production of the program in terms of scientific collaboration and is expected to be a useful tool for decision making in organizational and individual level.

Keywords: *Collaboration; Coauthorship networks; Social network analysis.*

Recebido em 08.10.2013 Aceito em 07.03.2014

1 Introdução

A colaboração científica é uma das características da ciência moderna e inúmeras são as suas vantagens. O indicador mais evidente de colaboração científica e mais utilizado nos estudos sobre esse tema é a coautoria, embora alguns autores defendam que a colaboração científica se manifesta de várias formas (coorientação de trabalho acadêmico, escrita conjunta de projeto, etc.) e algumas delas não são nem ao menos formalizadas (KATZ; MARTIN,1997). Segundo Milojevic (2010), a colaboração científica geralmente é estudada através de questionários e pesquisas (*surveys*), métodos bibliométricos (Glänzel e Schubert, 2004) e análise de redes sociais (BARABÁSI *et al.*, 2002; NEWMAN, 2001).

O objeto de estudo do método de análise de rede social (*social network analysis*) é uma rede social composta de nodos ou atores e arestas ou relacionamentos entre os atores. Uma rede de colaboração científica é um tipo de rede social no qual os atores são os pesquisadores e os relacionamentos podem ser manifestados pelas diversas formas de colaboração existentes no contexto científico. A coautoria, como mencionado anteriormente, é o indicador mais utilizado e, por isso, esse tipo de rede é chamada de rede de coautoria (*co-authorship network*).

As métricas de análise de redes sociais são baseadas essencialmente na teoria de grafos e permitem analisar a estrutura e as relações da rede como um todo, subgrupos de atores e atores individualmente dentro da rede. No contexto de uma rede de coautoria, os resultados dessa análise são importantes para o entendimento de diversos aspectos relacionados à colaboração, tais como: Qual o nível de colaboração de toda a comunidade de pesquisa? Quem são os grupos que trabalham isoladamente? Quais os autores que mais colaboram? As respostas a essas questões podem fornecer insumos para a tomada de decisão em um contexto científico.

Este artigo analisa a produção bibliográfica do programa de pós-graduação em Engenharia e Gestão do Conhecimento (PPGEGC) da Universidade Federal de Santa Catarina, no período de 2005 a 2012. O PPEGC foi autorizado, em meados de 2004, em nível *stricto sensu* (mestrado e doutorado) e tem direcionado suas pesquisas e sua formação no conhecimento, percebido como produto, processo e resultado de interações sociais e tecnológicas entre agentes humanos e tecnológicos. O programa possui três áreas de concentração, que se articulam de forma interdisciplinar: gestão do conhecimento, engenharia do conhecimento e mídia e conhecimento.

A análise da produção foi realizada através de métricas de análise de redes sociais. Para isso, foram criadas redes de coautoria da produção científica do período analisado. Os objetivos da análise foram identificar o nível de colaboração científica da comunidade do programa ano a ano, identificar grupos de pesquisadores que publicam conjuntamente e isoladamente e identificar os pesquisadores centrais na rede. Espera-se que o método apresentado e a discussão dos resultados sirvam como insumo para tomada de decisão em nível organizacional e individual pela comunidade acadêmica.

O artigo apresenta, a seguir, uma revisão de análise de redes sociais, os resultados da aplicação das métricas na produção departamental e a discussão dos resultados.

2 Análise de redes sociais

A noção de rede social e métodos de análise de rede social têm atraído muito interesse nas últimas décadas. Muito desse interesse pode ser atribuído ao foco nos relacionamentos entre as entidades sociais e nos padrões e implicações desses relacionamentos (WASSERMAN; FAUST, 1994). Através dos métodos de análise de rede social, é possível identificar aspectos, tais como, a) padrões de relacionamento entre os

atores de uma rede; b) a conectividade entre os mesmos; c) a formação de *clusters*; d) a evolução da rede ao longo do tempo; e e) o fluxo de comunicação, informação e conhecimento dentro da rede.

Uma rede pode ser representada por um grafo $G = (V, E)$, formado por Vértices (V) e Arestas (E). Cada vértice ou nodo representa um ator e cada aresta representa a relação existente entre dois atores integrantes da rede. No caso de uma rede de coautoria, os vértices representam os autores e as arestas representam os artigos produzidos em parceria. De modo geral, uma rede pode ser direcionada ou não direcionada e as arestas podem ser valoradas ou não valoradas. No caso da rede de coautoria, a rede é não direcionada e valorada, porque a relação de coautoria é assíncrona e a valoração se dá em função do número de artigos publicados em conjunto.

Uma rede pode ser analisada segundo três escopos: de rede, na qual a rede como um todo é analisada; individualmente, na qual cada ator é analisado; e de grupo, em que são analisadas as formações de subgrupos.

a)escopo de rede: no escopo de rede, três medidas são muito utilizadas: densidade, diâmetro e distância.

Segundo Scott (2000), a densidade é um dos conceitos mais utilizados em teoria dos grafos. Ela descreve o nível geral de ligações entre os pontos de um grafo. Um grafo "completo" é aquele em que todos os pontos são adjacentes um ao outro, ou seja, cada ponto é ligado diretamente a todos os outros pontos. Tal completude é muito rara, mesmo em redes muito pequenas, e o conceito de densidade é uma tentativa de resumir a distribuição total de linhas, a fim de medir o quão longe este estado de completude o gráfico está. Quanto mais pontos estão ligados uns aos outros, mais denso será o gráfico. No contexto de uma rede de coautoria, a densidade reflete o percentual do total da rede com o qual um ator foi coautor de um artigo (FISCHBACH; PUTZKE; SCHODER, 2011).

O diâmetro é uma medida que indica o maior grau de separação entre dois nodos da rede unidos pelo caminho mais curto (geodésico). De maneira geral, ela está relacionada à quantidade de tempo que a informação leva para passar através de toda a rede. No contexto de uma rede de coautoria, ela indica a possibilidade de colaboração mais longa.

A distância é outra medida e denota o menor caminho entre dois nodos. Em uma rede de coautoria, a distância de um autor que publica com outro autor é 1. Dois autores que não publicaram conjuntamente, mas compartilham um autor em comum, possuem distância 2 e, assim, segue. A distância foi utilizada em um famoso experimento conduzido por Travers e Milgram (1969), no qual eles demonstraram que o envio de uma mensagem de um remetente qualquer para um destinatário qualquer tem, na média, 6 intermediários, ou seja, 6 graus de separação.

b)escopo individual: no escopo individual, existem algumas métricas de centralidade que procuram descrever as propriedades de localização de

um ator na rede. Estas medidas levam em consideração as diferentes maneiras com que um ator interage e se comunica com o restante da rede. Os atores mais importantes ou mais proeminentes estão normalmente localizados em posições estratégicas dentro da rede (WASSERMAN; FAUST, 1994).

A centralidade de um ponto pode ser local ou global. A centralidade local está preocupada com a importância de um ponto focal na sua vizinhança, enquanto centralidade global diz respeito à proeminência do ponto dentro de toda a rede.

A centralidade local é medida através da centralidade de grau (*degree centrality*), enquanto que a centralidade global é medida através da centralidade de intermediação (*betweenness centrality*) e centralidade de proximidade (*closeness centrality*).

Na centralidade de grau, o grau de um vértice v corresponde ao número de arestas incidentes ou ao número de vértices adjacentes a ele. Segundo Freeman (1979), a centralidade de grau reflete a posição e o papel do nodo em termos de popularidade e atividade. Além disso, nodos com alto grau de centralidade podem ser identificados como líderes informais do grupo (KRACKHARDT, 2010). Em redes valoradas, nas quais a aresta possui um peso, como é o caso da uma rede de coautoria, a centralidade de grau pode levar em conta o valor ou peso da aresta. Em uma rede, vários atores podem possuir o mesmo grau de centralidade.

Em uma rede de coautoria, esse grau indica o total de autores da rede que publicaram em parceria com um determinado autor. O grau de centralidade varia de 0 (publicação sem parceria com outros autores) até $n-1$ (publicação com todos os demais autores da rede excluindo a si próprio), no qual n é o número total de autores da rede.

Para a centralidade de proximidade, Freeman (1979) propôs uma medida de centralidade global baseada na proximidade (*closeness*) dos pontos. Esta medida indica a proximidade de um determinado ator em relação aos demais atores da rede, sendo definida pela soma das distâncias geodésicas entre um determinado vértice e todos os outros vértices do grafo (distância geodésica é o menor caminho – número de arestas – que liga dois vértices de uma rede). Em uma rede de coautoria, um autor com uma centralidade de proximidade alta pode indicar uma maior possibilidade de estabelecer parcerias de publicação na rede por estar mais próximo em relação a todos os outros autores (SOUZA; BARBASTEFANO; LIMA, 2012).

Já na centralidade de intermediação, Freeman (1979) acrescenta um conceito adicional de centralidade global, denominado de intermediação (*betweenness*). Este conceito mede o quanto um determinado ponto se encontra "entre" os vários outros pontos no grafo, ou seja, atribui importância a um ator em função do fluxo que passa por ele para interligar outros dois atores da rede, através do menor caminho possível. Em uma rede de coautoria, um autor com alto valor de centralidade de intermediação indica que um número significativo das parcerias estabelecidas na rede envolve, de forma direta ou indireta, as publicações relacionadas a esse ator (SOUZA; BARBASTEFANO; LIMA, 2012).

c)escopo de grupo: uma das maiores preocupações de analistas de redes sociais é a identificação de subgrupos de atores dentro de uma rede. Subgrupos são subconjuntos de atores, entre os quais existem laços fortes, diretos, intensos, frequentes ou positivos (WASSERMAN; FAUST, 1994).

A identificação de componentes ou subgrafos dentro de uma rede é uma das técnicas para analisar uma rede do ponto de vista de grupo de atores. Um componente é formalmente definido como um subgrafo conectado ao máximo. Um subgrafo é dito conectado, quando todos os seus pontos se conectam através de um ou mais caminhos e não possuem outras conexões fora do subgrafo. Quando um subgrafo é máximo, é impossível que um nodo adicione outro nodo sem destruir a qualidade de conectividade.

Componentes de um grafo são subgrafos que estão conectados dentro do grafo, mas desconectados entre os subgrafos. Se um grafo contém um ou mais pontos "isolados", esses pontos também são chamados de componentes. Um componente que possui um conjunto de nós que estão ligados independentemente da direção de laços é denominado componente fraco. Quando o componente contém uma fração significativa de todos os nodos ele é denominado de componente gigante.

Em uma rede de coautoria, a presença de mais de um componente na rede indica a existência de grupos que publicam isoladamente

3 Procedimentos metodológicos

Para realizar essa pesquisa, os seguintes procedimentos foram realizados:

a)*coleta de dados da produção*: os dados de produção utilizados nesse estudo foram extraídos dos relatórios gerados pelo sistema Coleta de Dados CAPES. Para cada ano base de produção coletada (2005 a 2012), foram extraídos relatórios no formato *pdf* de produção bibliográfica do tipo trabalho em anais, artigo em periódico, livro ou capítulo de livro. Ao total, foram gerados 24 relatórios;

b)*extração e normalização dos dados*: em cada relatório os dados relacionados a produção, tais como título, tipo de produção, nome do evento/periódico, nome, abreviação e tipo do(s) autor(es), entre outros, estavam em um formato semiestruturado, ou seja, identificados com um rótulo. Com isso, foi possível desenvolver um *software (parser)* para extrair os dados das produções necessários para a criação das redes de coautoria. Os dados extraídos foram o nome e abreviação dos autores de cada produção.

Esses dados foram armazenados em um arquivo com formato exigido para leitura em um *software* de análise de redes sociais. O formato é dependente do *software* a ser utilizado, mas, geralmente, consiste de uma relação de nodos (vértices) ou autores, com um identificador e um rótulo, no qual o identificador é número sequencial

referente a um determinado autor e o rótulo é do nome do autor. Além disso, o formato exige a relação de todas as arestas da rede, que nesse caso de rede de coautoria consiste na relação de combinações de coautoria das produções, juntamente com o número de ocorrências de uma coautoria. Foram gerados arquivos de rede para cada ano base de produção e um arquivo de rede com toda a produção do período.

Nessa etapa, foi necessário corrigir o nome e abreviações dos autores, pois muitos estavam incorretos, o que causaria a duplicação de autoria. Esse trabalho foi feito manualmente e demandou um tempo considerável;

c) *análise dos dados*: a análise dos dados de produção foi feita através de algumas medidas bibliométricas, tais como, número de publicações, número de autores, número autores por publicação (índice de colaboração) e principalmente através de métricas de análise de rede social e visualização de redes.

A análise aconteceu em três níveis: o nível global ou de rede forneceu estatísticas da rede como um todo; o nível individual forneceu estatísticas ou métricas para os autores da rede; o nível de grupo identificou subgrupos de autores na rede. Além dessas análises, foram criadas visualizações das redes de coautoria.

A análise de rede social foi feita através do *software* de visualização e análise exploratória de dados de rede Gephi. (BASTIAN; HEYMAN; JACOMY, 2009).

4 Resultados

Foram analisadas as produções bibliográficas informadas ao sistema Coleta de Dados da CAPES relativas ao ano base de 2005 até 2012, totalizando oito anos. A seguir, serão apresentados os resultados em cada nível de análise e a visualização de alguns aspectos das redes.

a) *nível global*: ao total, foram 2182 produções bibliográficas, divididas em anais, artigos, livros e capítulos produzidos por 1488 autores diferentes.

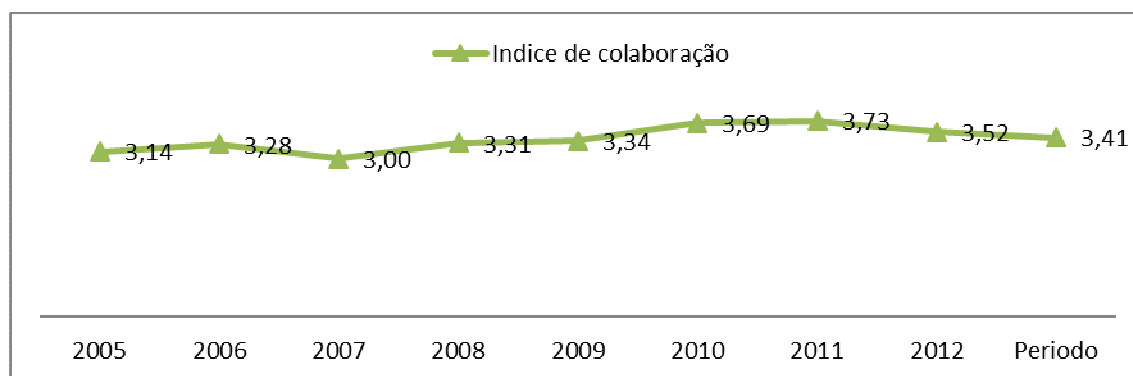
Para as medidas bibliométricas, a Tabela 1 exibe algumas medidas, tais como: o número de produções, o número de produções coautoradas, o percentual de produções individuais, o número de autores e a densidade da rede em cada ano. Como pode ser observado, o ano com maior número de produções bibliográficas foi 2009. O percentual de publicações individuais, obtido através da diferença entre o número de produções e o número de produções coautoradas, teve uma tendência decrescente ao longo de oito anos, com exceção do ano de 2009. O índice de colaboração (média de autores que colaboraram em uma publicação), que é visualizado na Figura 1, manteve-se estável ao longo dos anos, com uma média de 3,4 autores por publicação.

Tabela 1 - Medidas bibliométricas e de rede da produção bibliográfica do PPGEGC

Ano	Anais	Artigos	Livros	No. de produções	No. de produções coautoradas	Percentual de produções individuais	No. de autores/nodos rede	Densidade
2005	134	33	15	182	157	13,7	247	0,02 (2%)
2006	161	39	24	224	206	8,0	306	0,02 (2%)
2007	138	36	38	212	194	8,5	281	0,02 (2%)
2008	158	38	28	224	211	5,8	319	0,02 (2%)
2009	216	99	96	411	377	8,0	425	0,01 (1%)
2010	121	89	81	291	281	3,4	368	0,02 (2%)
2011	166	71	68	305	295	3,3	362	0,02 (2%)
2012	176	105	52	333	321	3,6	394	0,02 (2%)
Período	1270	510	402	2182	2042	6,4	1488	0,01 (1%)

Fonte: Dados da pesquisa.

Figura 1 – Número médio de autores por produção bibliográfica



Fonte: Dados da pesquisa.

Nas medidas de análise de rede social, o tamanho de uma rede social é dado pelo número de nodos ou nesse caso, coautores e dá uma indicação da probabilidade de interação entre os nodos, ou seja, quanto maior a rede maior a probabilidade de interação entre os nodos. Entretanto, quanto maior a rede, menor a possibilidade de um nodo estar conectado com o conjunto total de nodos da rede.

Como pôde ser observado na Tabela 1, o número de autores/nodos nas redes de cada ano apresentou um crescimento médio no período em torno de 8%, com um significativo aumento no ano de 2009.

O grau de conectividade de uma rede é dada pela sua densidade, cujo valor máximo é 1. Ela indica a proporção de conexões existentes em relação ao número total de conexões possíveis na rede. A Tabela 1 mostra as densidades das redes de coautoria anualmente e no período inteiro. Observa-se que todas as redes são esparsas, com um grau de densidade baixo em torno de 2%, o que já era esperado, devido à provável existência de vários componentes isolados dentro de cada rede.

b) *nível de grupo*: uma rede pode não ser inteiramente conectada, ou seja, ela pode estar dividida em um número de subgrafos chamados

componentes. Entre os nodos de componentes diferentes não existe um caminho possível. Em nível de grupo, a análise de rede foi feita através da identificação desses componentes. A identificação de componentes indica a presença de grupos publicando isoladamente.

A Tabela 2 mostra o número de componentes encontrados em cada rede de coautoria, o número de nodos do componente gigante, o percentual de representação do componente gigante encontrado em cada rede, a densidade, o diâmetro e a distância média entre dois nodos nesse componente. Destaca-se que redes mais fragmentadas, ou seja, com um maior número de componentes, são encontradas nos primeiros anos do programa, com exceção do ano 2006 e, nos quatro últimos anos (2009 a 2012), o número de componentes vai decrescendo. As redes dos últimos quatro anos possuem componentes gigantes (o maior componente de uma rede), que representam percentualmente a maior parte da rede ($\geq 94\%$), restando aos outros componentes uma representação muito pequena perante toda a rede, o que indica uma pequena quantidade de grupos trabalhando isoladamente e, conseqüentemente, uma coesão e colaboração maior entre a comunidade acadêmica do programa. A Figura 2 ilustra essa situação, ao exibir a rede de coautoria do ano de 2012, na qual foram encontrados três componentes pequenos e um componente gigante, que compreende 95% da rede.

A densidade do componente gigante de cada rede, entre 2% e 3%, mostra um valor superior ao encontrado na densidade de todas as redes (vide tabela 1). Porém, o nível de conectividade ainda permanece baixo, indicando que existe espaço para mais conectividade ou colaboração em relação ao total de possibilidades (nodos) existentes.

O diâmetro do componente gigante de cada rede mostra que o caminho mais longo foi encontrado na rede de 2007. Nos dois últimos anos, foi encontrado um diâmetro de oito, indicando que existe uma cadeia máxima de oito autores conectando dois coautores.

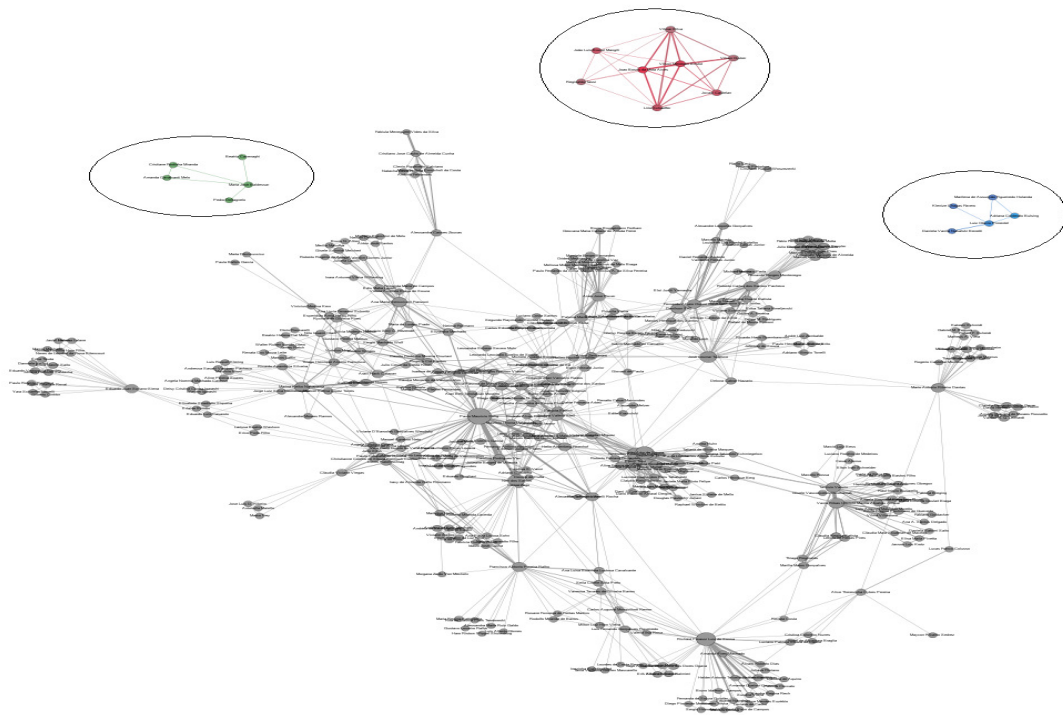
A distância média entre dois nodos do componente gigante de cada rede indica que a possibilidade de colaboração direta entre dois autores é, em média, de quatro autores.

Tabela 2 - Número de componentes e medidas do componente gigante de cada rede

Ano	No. nodos rede	No. componentes	No. nodos comp. gigante	%	Densidade comp. gigante	Diâmetro comp. gigante	Distância média
2005	247	11	192	77,7	0,03 (3%)	9	4,2
2006	306	7	269	87,9	0,02 (2%)	7	3,6
2007	281	13	232	82,6	0,02 (2%)	11	5,2
2008	319	11	266	83,4	0,02 (2%)	10	4,8
2009	425	6	406	95,5	0,02 (2%)	8	4,1
2010	368	5	346	94,0	0,02 (2%)	7	3,6
2011	362	5	346	95,6	0,02 (2%)	8	3,7
2012	394	4	376	95,4	0,02 (2%)	8	3,7
Período	1488	4	1473	99,0	0,01 (2%)	7	3,4

Fonte: Dados da pesquisa.

Figura 2 - Rede de coautoria do ano 2012

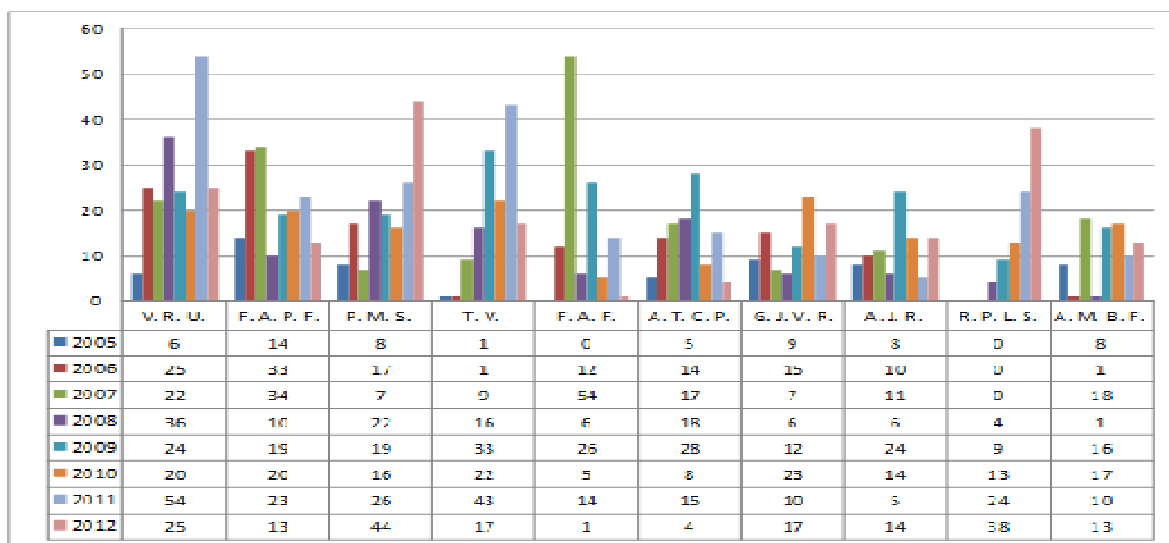


Fonte: Dados da pesquisa.

c) nível individual: no nível individual, foram identificados os autores com maior número de publicações e os autores mais centrais à rede do período 2005-2012.

Como medidas bibliométricas, a Figura 3 mostra os 10 autores com maior número de produções bibliográficas no período e a distribuição do número de produções dos autores ao longo de oito anos. Todos os autores foram ou são docentes do PPGECC.

Figura 3 - Autores com maior número de produções bibliográficas no período 2005-2012



Fonte: Dados da pesquisa.

Nas medidas de análise de rede social, para se identificar os autores mais centrais à rede, foram usadas três métricas de centralidade: grau, intermediação e proximidade.

A centralidade de grau (*degree centrality*) foi calculada para todos os atores na rede do período. A tabela 4 mostra os autores com as dez maiores centralidades de grau, juntamente com a sua área de concentração no programa. A centralidade de grau indica que esses são os autores que mais colaboraram, publicando conjuntamente com outros autores. Essa medida leva em conta o número de coautores que colaboraram com um determinado autor, juntamente com número de publicações que os mesmos fizeram conjuntamente. A Figura 3, quando comparada à Tabela 4, mostra que os últimos cinco autores (A.T.C.P., G.J.V.R., A.J.R., R.P.L.S e A.M.B.F) com maior número de publicações não estão presentes na Tabela 4, o que indica que esses autores não possuem muitas publicações em conjunto com outros autores. Ao contrário, quatro autores presentes na Tabela 4 (M. K. N., F. J. S., N. S. e J. L.T.) não estão na figura 3, o que indica que os autores colaboram com mais intensidade, publicando conjuntamente com outros autores.

Outro ponto que pode ser observado na Tabela 4 é que do quarto colocado para o quinto existe uma significativa diferença na centralidade de grau, em torno de 37%, e que os quatro primeiros autores com maior centralidade de grau também são os quatro primeiros autores com maior número de publicações. Todos os autores da Tabela 4 foram ou são docentes do programa.

Tabela 3 - Autores com maior centralidade de grau

Autor	Centralidade de grau	Área de concentração
V. R. U.	517	Mídia e conhecimento
P. M. S.	497	Gestão do conhecimento
F. A. P. F.	428	Mídia e conhecimento
T. V.	400	Mídia e conhecimento
M. K. N.	291	Mídia e conhecimento
F. J. S.	269	Mídia e conhecimento
N. S.	263	Gestão do conhecimento
J. L. T.	256	Engenharia do conhecimento
F. A. F.	255	Gestão do conhecimento
A. M. B. F.	243	Gestão do conhecimento

Fonte: Dados da pesquisa.

A centralidade de intermediação (*betweenness centrality*) é uma medida de centralidade que pode ser calculada somente em grafos conectados. No caso das redes de coautoria, foi escolhido o componente gigante de cada rede. Nesse contexto, um nodo ou coautor, que tem um alto grau de centralidade de intermediação, pode ser considerado uma "ponte" entre distintos grupos de pesquisa, uma vez que está no caminho mais curto entre outros nodos (FISCHBACH; PUTZKE; SCHODER, 2011).

A Tabela 5 mostra os cinco autores com maior centralidade de intermediação na rede, do período 2005-2012. Os valores foram normalizados entre 0 e 1. Todos os autores da Tabela 5 são docentes do programa e podem ser vistos como aqueles que produzem conjuntamente com diversos grupos. Considerando os resultados apresentados na Tabela 5, observa-se que os autores que estão nas posições intermediárias, em relação a todos os outros autores, são o primeiro e o segundo autores na tabela, com um grau em torno de 18%.

Tabela 4 - Autores com maior centralidade de intermediação

Autor	Centralidade de intermediação	Área de concentração
F. A. P. F.	0,19	Mídia e conhecimento
P. M. S.	0,17	Gestão do conhecimento
G. J. V. R.	0,10	Gestão do conhecimento
A. T. C. P.	0,08	Mídia e conhecimento
V. R. U.	0,07	Mídia e conhecimento

Fonte: Dados da pesquisa.

A centralidade de proximidade (*closeness centrality*) também só pode ser calculada em grafos conectados. Ela indica a proximidade de um determinado ator em relação aos demais atores da rede. Autores que têm um alto grau de centralidade de proximidade podem alcançar todos os outros autores, através uma cadeia menor de coautores, em relação a autores com um baixo grau de centralidade (FISCHBACH; PUTZKE; SCHODER, 2011).

A Tabela 6 mostra os cinco autores com maior centralidade de proximidade na rede, do período 2005-2012. Todos os autores da tabela 6 são docentes do programa. Esses são os autores mais próximos (índice aproximado de 50%) de todos os outros autores na rede.

Tabela 5 - Autores com maior centralidade de proximidade

Autor	Centralidade de proximidade	Área de concentração
F. A. P. F.	0,48	Mídia e conhecimento
P. M. S.	0,45	Gestão do conhecimento
G. J. V. R.	0,45	Gestão do conhecimento
N. S.	0,43	Gestão do conhecimento
S. S. L. F.	0,41	Mídia e conhecimento

Fonte: Dados da pesquisa.

5 Discussão

A seguir, serão apresentados e discutidos os principais pontos resultantes da análise:

1) a média em torno de 3,4 coautores por publicação se manteve estável durante todos os anos do período analisado. De acordo com Katz e Martin (1997), o aumento no índice de coautoria é uma evidência do aumento de colaboração. Persson, Glänzel e Danell (2004) destacam que

publicações com mais autores tendem a ter um respaldo científico maior em termos de aceitação e citação. Nesse sentido, apesar da média estável encontrada, entende-se que, para que haja uma maior colaboração nesse programa, deve ser estimulado o aumento de coautores por publicação;

2) o primeiro ano apresenta o maior percentual de publicações individuais (13,7%), o que já seria esperado nos primeiros anos de um programa de Pós-graduação, no qual a comunidade acadêmica não se conhece o suficiente para colaborar, publicando conjuntamente. Como consequência da maturidade do programa, observa-se que o percentual de publicações individuais decresce no decorrer dos anos (com exceção do ano de 2009).

Em relação ao número de componentes encontrados em cada rede, nota-se um grande número de componentes nos primeiros quatro anos (11, 7, 13 e 11), revelando a presença de grupos que publicaram isoladamente, o que também se atribui aos primeiros anos de existência do programa e ao fato do programa ser interdisciplinar, o que fez com que pequenos grupos (componentes) de diferentes áreas de conhecimento publicassem de forma isolada;

3) por outro lado, nos últimos quatro anos, observa-se um decréscimo no número de componentes, o que indica uma existência menor de grupos que publicam isoladamente e, conseqüentemente, uma maior consciência de trabalho colaborativo e interdisciplinar entre as áreas de concentração do programa. Além disso, os componentes gigantes encontrados nas redes dos últimos quatro anos possuem o maior percentual de representação em relação a todos os nodos da rede, o que corrobora com o argumento de maior consciência de trabalho colaborativo, por parte da comunidade acadêmica do programa;

4) a densidade encontrada no componente gigante das redes de coautoria de todos os anos é esparsa (entre 1% e 2%), o que pode indicar um espaço ou oportunidade de colaboração maior. Entretanto, cabe salientar que os coautores das redes são docentes, discentes e participantes externos ao programa e que um percentual em torno de 60% dos autores são participantes externos ao programa, o que diminui a capacidade de colaboração, visto que um participante externo tende a não ser conhecido ou acessível pela maioria dos docentes e discentes do programa. Essa situação é, mais uma vez, atribuída ao fato do programa ser interdisciplinar, quando é esperado que exista uma maior participação externa da comunidade do programa, uma vez que as suas interações de pesquisa se dão nas várias áreas do conhecimento;

5) os dez autores com maior centralidade de grau foram ou são professores do programa, o que já era esperado, no contexto de uma rede de coautoria de produção de um programa de Pós-graduação. Esses autores representam as três áreas de concentração do programa, na seguinte proporção: cinco autores da área de Mídia e Conhecimento, quatro autores da área de Gestão do Conhecimento e um autor da área de Engenharia do Conhecimento. Como o número total de docentes do programa (37 docentes) está equilibrado nas três áreas de concentração, essas centralidades de grau indicam que a área com menor

representação, em termos de atores com alto grau de colaboração, é a área de Engenharia do Conhecimento;

6) apesar de serem listados os cinco autores com maior grau de centralidade de intermediação, 46 autores apresentaram algum grau de intermediação, ou seja, o caminho mais curto para chegar a algum autor passa por eles. Comparando com a quantidade de autores no componente gigante da rede do período (1473), percebe-se que esse grau de centralidade segue uma distribuição do tipo *Power Law*, na qual somente alguns autores possuem os maiores graus em relação a maioria. Tal fato já foi observado por Newman (2004), nos estudos de redes de coautoria;

7) foram apresentados os cinco autores com maior centralidade de proximidade, porém, observou-se que todos os autores, até a décima quinta posição no *ranking*, são docentes do programa, representando 50% do corpo docente, o que indica que os autores mais próximos de toda a rede são os docentes do programa. Segundo Lima (2011), a alta proximidade em relação aos demais nós significa, para o autor, que ele está estruturalmente bem posicionado na rede e que possui mais chances do que os demais autores de expandir sua rede, ou seja, de aumentar o número de autores para coautoria.

6 Conclusão

Este artigo analisou a produção bibliográfica do PPGE GC, no período de 2005 a 2012. Foram analisadas produções do tipo anais publicados em eventos, artigos publicados em revistas e livros e capítulos de livros. A análise, feita através de métricas de análise de redes sociais e posterior discussão dos resultados, não considerou o aspecto qualitativo das produções bibliográficas e não fez uma correlação com a avaliação trienal do PPGE GC pela CAPES.

Acredita-se que a análise e discussão dos resultados apresentados neste artigo, alcançaram o objetivo macro de diagnosticar a produção bibliográfica do programa em termos de colaboração científica e espera-se que seja uma ferramenta útil para tomada de decisão em nível organizacional e individual.

No âmbito organizacional, incentivar a colaboração entre os membros da comunidade docente e discente do programa pode aumentar a densidade das redes de colaboração; incentivar a colaboração na comunidade vinculada a uma determinada área de concentração pode ajudar a equilibrar a distribuição do número de pesquisadores com maior centralidade de grau.

No âmbito individual, os resultados podem ajudar novos integrantes da comunidade acadêmica do PPGE GC a identificar membros expoentes dentro da comunidade, com os quais possa ser interessante trabalhar, visto que um novo membro frequentemente tem esse tipo de dificuldade. Além disso, pode ajudar membros de outras comunidades acadêmicas a identificar possíveis parcerias para trabalhos conjuntos, propiciando, assim, a interdisciplinaridade que está na gênese do PPGE GC.

Foram encontrados poucos trabalhos que analisam a produção científica de um departamento de pós-graduação, através do método de análise de redes sociais. Alguns deles são: Katerndahl (2012); Lima(2011), Rossoni(2009) e Silva et. al (2006). Logo, espera-se, também, que o método empregado neste artigo sirva como de referência para a comunidade do PPGEGC e outras comunidades acadêmicas, em estudos similares.

Referências

BARABÁSI, A. L. *et al.* Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, v. 311, n. 3/4, p. 590-614, 2002.

BASTIAN, M.; HEYMANN, S.; JACOMY, M. Gephi: an open source software for exploring and manipulating networks. In: INTERNATIONAL AAAI CONFERENCE ON BLOGS AND SOCIAL MEDIA, 3., 2009, San Jose, California. Proceedings... Menlo Park: AAAI Press, 2009. p. 361-362.

FISCHBACH, K.; PUTZKE, J.; SCHODER, D. Co-authorship networks in electronic markets research. *Electronic Markets*, v. 21, n. 1, p. 19-40, 2011.

FREEMAN, L. Centrality in social networks: conceptual clarification. *Social Networks*, v. 1, n. 3, p. 215-239, 1979.

GLÄNZEL, W.; SCHUBERT, A. Analysing scientific networks through co-authorship. In: MOED, H. F.; GLÄNZEL, W.; SCHMOCH, U. (Eds.). Handbook of quantitative science and technology research: the use of publication and patent statistics in studies of S&T systems. Dordrecht: Kluwer Academic Publishers, 2004. p. 257-276.

KATERND AHL, D. Evolution of the research collaboration network in a productive department. *Journal of Evaluation in Clinical Practice*, v. 18, n. 1, p. 195-201, 2012.

KATZ, J. S.; MARTIN, B. R. What is research collaboration? *Res. Policy*, v. 26, n. 1, p. 1-18, 1997.

KRACKHARDT, D. Social networks. In: LEVINE, J. M.; HOGG, M. A. (Eds.). *Encyclopedia of group processes and intergroup relations*. Los Angeles: SAGE, 2010. p. 817-821.

LIMA, M. Y. Coautoria na produção científica do PPGeo/UFRGS: uma análise de redes sociais. *Ci. Inf, Brasília*, v. 40, n. 1, p. 38-51, jan./abr. 2011.

MILOJEVIC, S. Modes of collaboration in modern science: beyond power laws and preferential attachment. *Journal of the American Society for Information Science and Technology*, v. 61, n. 7, p. 1410-1423, 2010.

NEWMAN, M. E. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*, v. 98, n. 2, p. 404, jan. 2001.

NEWMAN, M. E. Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences of the United States of America*, v. 101, n. suppl 1, p. 5200-5205, April 2004.

PERSSON, O.; GLÄNZEL, W.; DANELL, R. Inflationary bibliometric values: the role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, v. 60, n. 3, p. 421-432, 2004.

ROSSONI, L.; FILHO, E. G.; CO, S. Cooperação entre programas de pós - graduação em administração no Brasil: evidências estruturais em quatro áreas. *RAC*, Curitiba, v. 13, n. 2, p. 366-390, jul./ago. 2009.

SCOTT, J. *Social network analysis: a handbook*. 2. ed. London: SAGE Publications, 2000.

SILVA, A. B. O. *et al.* Estudo da rede de coautoria e da interdisciplinaridade na produção científica com base nos métodos de análise de redes sociais: avaliação do caso do Programa de Pós-Graduação em Ciência da Informação - PPGCI / UFMG. *Encontros Bibli (UFSC)*, Florianópolis, n. esp., p. 179-194, 2006.

SOUZA, C. G.; BARBASTEFANO, R. G.; LIMA, L. S. Redes de colaboração científica na área de química no Brasil: um estudo baseado nas coautorias dos artigos da revista Química Nova. *Química Nova*, São Paulo, v. 35, n. 4, p. 671-676, 2012.

TRIVERS, J., MILGRAM, S. An experimental study of the small world problem. *Sociometry*, v. 32, p. 425-443, 1969.

WASSERMAN, S.; FAUST, K. *Social network analysis: methods and applications*. Structural analysis in social the social sciences series. Cambridge: Cambridge University Press, 1994.