

Redes de Interação na Ciência: mapeamento de comunidades de autores

Roberto Mario Lovón Canchumani

Doutor; Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil;
rlovon@gmail.com

Resumo: Este artigo apresenta os resultados de uma proposta de identificação de comunidades de autores realizado na Universidade Federal do Rio de Janeiro (UFRJ). O objetivo foi identificar a tendência de agrupamento de autores da instituição ao longo do tempo, assim como também observar características mais específicas das principais comunidades identificadas. O trabalho investigou 44.233 registros recuperados da base institucional SIGMA.UFRJ, de artigos publicados em periódicos no período de 2001 a 2012. Os dados foram organizados e analisados com auxílio dos softwares *Excel* e *Gephi*. As métricas de grau e modularidade/transitividade mostraram que, ao longo do período estudado, novas colaborações entre os autores da UFRJ foram estabelecidas, formando novos agrupamentos/comunidades de autores. As maiores comunidades de autores identificadas foram detalhadas, revelando suas principais especificações.

Palavras-chave: Comunidades de Autores. Redes de Colaboração. Coautoria.

1 Introdução

O desenvolvimento de pesquisas no campo de redes complexas tem apresentado, nas últimas décadas, significativos avanços na descrição e análise dos traços topológicos e dinâmicos de redes de tipo social, biológica ou tecnológica (BARABÁSI et al., 2002). Nesse cenário, um dos aspectos que tem concentrado particular atenção tem sido a existência de subconjuntos de nós fortemente entrelaçados, às vezes conectados e outras vezes desconectados da rede. Neste último caso, nos deparamos com agrupamentos ou comunidades (*clusters*) que se encontram ligadas ou não à rede (FORTUNATO, 2010).

Essas comunidades cumprem papel relevante nas propriedades de estruturas complexas. Identificar e analisar a natureza das comunidades aí ocultas é uma importante tarefa para revelar a organização informal e a natureza dos fluxos de informação que acontecem dentro desses sistemas complexos

(PALLA et al., 2005). É importante destacar que tais comunidades são compostas por diferentes subconjuntos de nós, os quais se caracterizam por estarem densamente ligados mais entre si do que em relação ao resto da rede. A identificação dessas comunidades facilita a compreensão das redes.

Considerando que a comunidade científica atual se estrutura cada vez mais em um sistema integrado, cooperativo ou de rede, torna-se relevante identificar a natureza, a intensidade e as características específicas das interações entre seus membros (RIVELLINI; RIZZI; ZACCARIN, 2006; RODRIGUEZ, 1995).

Quando realizadas no contexto dos estudos da Informetria, essas análises buscam a representação dos laços colaborativos a partir das informações contidas na produção científica, mais frequentemente a partir da comunicação em formato de artigos em periódicos. Nesse âmbito, algumas abordagens têm permitido identificar comunidades científicas, nas quais se adotam funções específicas para descrever características dos grupos a serem determinados (CANCHUMANI; LETA; FIGUEIREDO, 2017; BARBOSA et al., 2011; PERIANES-RODRIGUEZ; OLMEDA-GÓMEZ; MOYA-ANEGON, 2010).

Este trabalho apresenta os resultados de uma proposta de mapeamento de comunidades de autores realizado na Universidade Federal do Rio de Janeiro (UFRJ). O objetivo do estudo foi identificar a tendência de agrupamento de autores da instituição ao longo do tempo, assim como também observar características mais específicas das principais comunidades de autores identificadas.

Cabe mencionar que uma primeira versão deste estudo foi apresentada no 6º Encontro Brasileiro de Bibliometria e Cientometria (6º EBBC) (CANCHUMANI, 2018), realizado na cidade do Rio de Janeiro em julho de 2018. O presente artigo é um desdobramento desse trabalho, sendo, portanto, uma versão mais abrangente e detalhada dos resultados apresentados no 6º EBBC.

2 Metodologia

Foi realizado um levantamento da produção bibliográfica, envolvendo artigos publicados em periódicos no período de 2001 a 2012, na base de dados institucional SIGMA.UFRJ. O SIGMA.UFRJ foi criado em 1998 com a finalidade de ser um ambiente virtual e interativo de representação das atividades de ensino, de pesquisa e de extensão realizadas na UFRJ, construído sobre uma base de dados corporativa, na qual eram correntemente registradas essas atividades e evidenciadas suas múltiplas inter-relações.

Ressalta-se que até o ano de 2012, as informações da base SIGMA.UFRJ foram utilizadas nos processos de avaliação dos programas de pós-graduação da instituição, já que a coleta de dados era realizada nessa base, para posteriormente ser exportada para a Coleta CAPES. Em 2013, por decisões da Pró-Reitoria de Pós-Graduação da UFRJ, essa coleta deixou de ser realizada para fins avaliativos. A partir de então, a base se encontra inativa. Daí porque o período de análise se estende até 2012, último ano em que as informações da base, como já mencionado, foram atualizadas e utilizadas nos processos de avaliação dos programas de pós-graduação da UFRJ. Convém ressaltar também que essa base informacional ainda não tinha sido utilizada/explorada para estudos desta natureza.

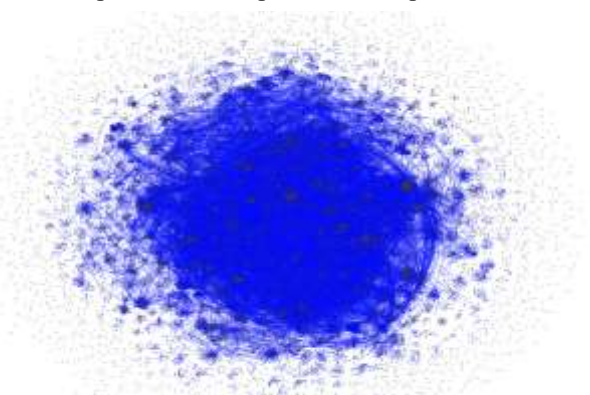
A busca na mencionada base considerou os seguintes critérios: Produção Intelectual => artigos completos publicados em periódicos => autores. A partir desses critérios foi possível extrair 44.233 registros da produção da UFRJ de artigos publicados em periódicos no período de 2001 a 2012. Os registros com suas informações foram exportados para matrizes multidimensionais, um processo que usou ferramentas computacionais, como software para cruzamento de dados (*Excel*) e geração/visualização de redes/mapas (*Gephi*).

A interação entre autores foi analisada a partir das informações de coautoria e as métricas de grau e modularidade/transitividade.

3 Resultados

A população identificada na produção acadêmica no período de 2001 a 2012 é constituída por 28.248 autores com vínculo formal com a UFRJ, responsáveis pelos 44.233 artigos publicados no período. A Figura 1 mostra o grafo que representa o total de autores ($n=28.248$). Cada ponto do grafo representa um autor. Uma aresta na cor azul indica pelo menos uma publicação realizada em colaboração entre dois pesquisadores.

Figura 1 - Mapa de interação dos autores com vínculo formal com a UFRJ, de artigos publicados em periódicos no período 2001-2012



Fonte: Elaborado pelo autor (2018).

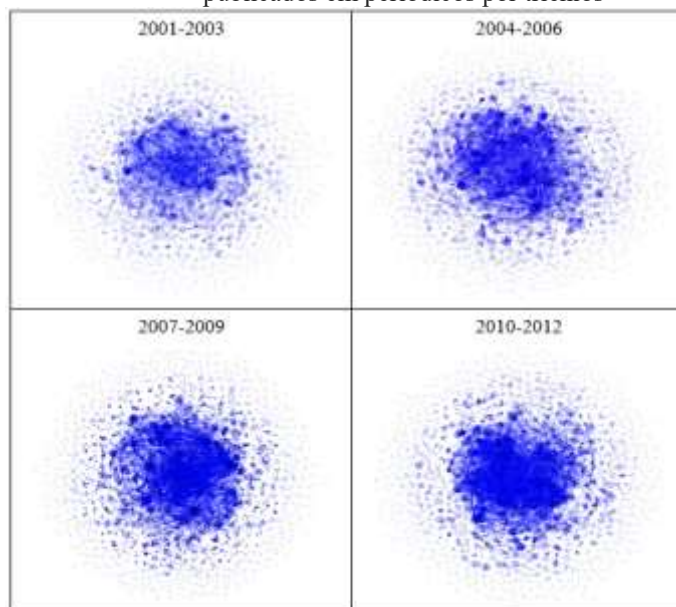
Nota: O total de autores é 28.248, e as coautorias estão representadas por arestas na cor azul.

Os nós que aparecem visualmente isolados, mais externamente no grafo, correspondem a autores de artigos sem colaboração ou subconjunto de autores (geralmente formado por poucos pesquisadores) desligados da grande rede interligada.

Nas redes de interação de coautoria, em média, cada autor da UFRJ mantém colaboração com 6,73 pesquisadores. Observa-se que no grafo existem vários subgrupos, ou seja, conjuntos/agrupamentos de nós bem próximos uns dos outros, de autores que formam comunidades (*clusters*). Podemos notar, visualmente, que autores desses subgrupos colaboram mais entre si.

Um olhar por períodos (triênios), Figura 2, mostra que ao longo dos triênios a densidade do grafo aumenta. Essa simples representação visual possibilita observar o incremento da atividade colaborativa na UFRJ ao longo do tempo.

Figura 2 - Mapa de interação dos autores com vínculo formal com a UFRJ, de artigos publicados em periódicos por triênios



Fonte: Elaborado pelo autor (2018).

Nota: O total de autores é 28.248, distribuídos por triênios, e as coautorias estão representadas por arestas na cor azul.

Verificou-se que no triênio 2001-2003, o grau médio foi de 4,33, ou seja, cada autor colaborou, em média, com pouco mais de quatro autores, e a transitividade, isto é, a tendência que os autores têm de se agruparem (coeficiente de *clustering*) foi de 0,76. Em 2004-2006, o grau médio foi de 4,87 e o coeficiente de *clustering* foi de 0,77. No período seguinte, 2007-2009, o grau médio foi de 5,12 e a transitividade foi de 0,78. Já o último triênio apresentou grau médio de 5,57 e transitividade de 0,79.

Essas medidas indicam que, ao longo dos períodos, novas colaborações foram estabelecidas formando novos agrupamentos/comunidades. Em outras palavras, ao longo dos triênios, a tendência de todos os nós (autores) se agruparem aumenta.

A fim de verificar a rede de autores com participação mais intensa em termos de colaboração na instituição ao longo do tempo, identificamos a formação de diferentes subgrupos de pesquisadores que formam comunidades (*clusters*) para cada triênio, focalizando nossa atenção na maior comunidade de cada um desses períodos.

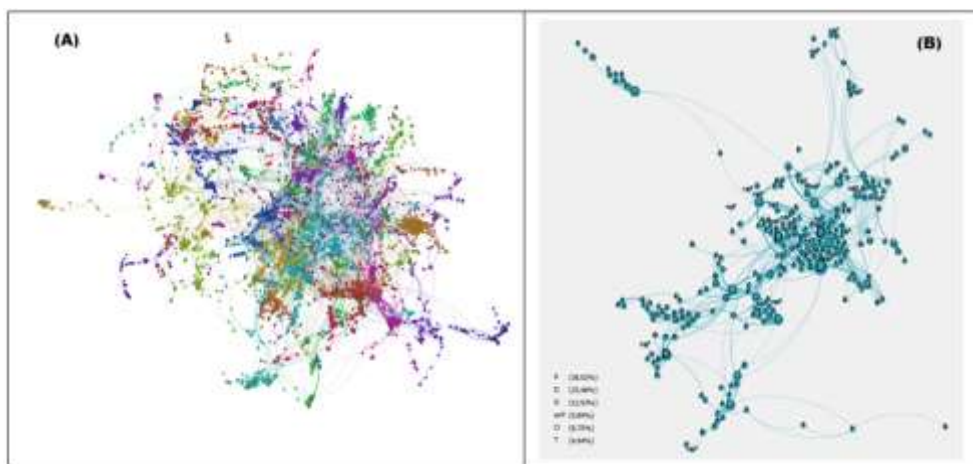
3.1 Comunidades de autores no triênio 2001-2003

A rede de coautoria da UFRJ, no período 2001-2003, é composta por 5.695 autores. A identificação de agrupamentos dessa população, aplicando o algoritmo Blondel et al. (2008), encontrou um total de 40 comunidades, envolvendo 3.736 nós (autores), identificadas na Figura 3-A através das cores.

Os nós (autores) estabeleceram entre si 11.304 relações (colaborações), representadas na rede por arestas. O grau médio foi de 6,05, ou seja, em média, estes autores participaram de publicações com pouco mais de seis autores e a transitividade (coeficiente de *clustering*) de 0,76.

A maior comunidade de autores identificada no período de 2001-2003 é constituída por 236 autores (o que representa 7,04% do total de autores das 40 comunidades) e, 689 relacionamentos, arestas (Figura 3-B). O grau médio é de 5,24, ou seja, cada autor colabora, em média, com mais de cinco pesquisadores. O coeficiente de transitividade apresentado é de 0,79.

Figura 3 - (A) Grafo das comunidades de autores da UFRJ de publicações em periódicos identificadas no período 2001-2003. (B) Grafo da maior comunidade de autores de artigos em periódicos da UFRJ no período 2001-2003



Fonte: Elaborado pelo autor (2018).

A classificação de autores dessa comunidade por tipo de vínculo mostra que a maioria é composta por alunos de pós-graduação - P (38,02%), seguida por docentes - D (25,48%) e alunos de graduação - G (12,93%). Alunos egressos da pós-graduação (exP) representam 9,89% e os autores com outros tipos de

vínculo (O) 8,75%, podendo ser professor egresso, técnico egresso ou egresso da graduação. Já os técnicos (T) representam 4,94%.

Do total de autores dessa comunidade, identificamos os 20 que apresentam as maiores medidas de centralidade de grau, ou seja, são autores que mais interagem, tem maior número de colaborações (Tabela 1).

Verifica-se que dos 20 autores identificados, 16 são docentes. Os outros quatro autores têm os seguintes tipos de vínculo: egresso da pós-graduação (exP), aluno de pós-graduação (P), técnico (T) e um que possui outros tipos de vínculo (O).

No que se refere às áreas dos autores, verificamos que dez deles atuam na grande área de Ciências Biológicas, dois nas Ciências da Saúde e um nas Ciências Humanas. Os outros autores identificados atuam em duas áreas do conhecimento, assim distribuídos: quatro deles em Ciências Biológicas e Ciências da Saúde; dois em Ciências Agrárias e Ciências da Saúde e; um autor em Ciências Biológicas e Ciências Humanas.

Tabela 1 - Vínculo e grande área dos autores com maiores medidas de centralidade de grau da maior comunidade identificada das publicações da UFRJ em 2001-2003

Id	Vínculo*	Grau	Áreas de Atuação
330	D	46	Ciências Biológicas; Ciências da Saúde
2139	D	34	Ciências Biológicas, Ciências da Saúde
342	D	26	Ciências Biológicas
339	exP	24	Ciências Biológicas, Ciências da Saúde
343	O	20	Ciências Biológicas
344	D	19	Ciências Biológicas
2140	D	19	Ciências da Saúde
268	D	17	Ciências Biológicas
406	D	16	Ciências Biológicas
2042	D	15	Ciências Agrárias, Ciências da Saúde
2320	D	15	Ciências da Saúde, Ciências Biológicas
445	D	14	Ciências Biológicas
540	D	14	Ciências Biológicas
2156	T	14	Ciências da Saúde
223	D	13	Ciências Biológicas
412	D	12	Ciências Biológicas, Ciências Humanas
1924	D	12	Ciências da Saúde, Ciências Agrárias
338	P	11	Ciências Humanas
341	D	11	Ciências Biológicas
541	D	11	Ciências Biológicas

Fonte: Elaborado pelo autor (2018).

* Tipos de Vínculo: D (Docente); P (Discente de Pós-graduação); exP (Egresso Discente de Pós-graduação); T (Técnico); O (Outros).

Uma visão detalhada da rede em torno do grupo de autores mais centrais revela algumas informações importantes (Figura 4). O tamanho do nó (e do Id) é proporcional ao número de colaborações (grau do nó) e cada tipo de autor é representado por uma determinada cor.

Figura 4 - Grafo dos autores mais centrais da maior comunidade de autores de publicações da UFRJ identificada em 2001-2003

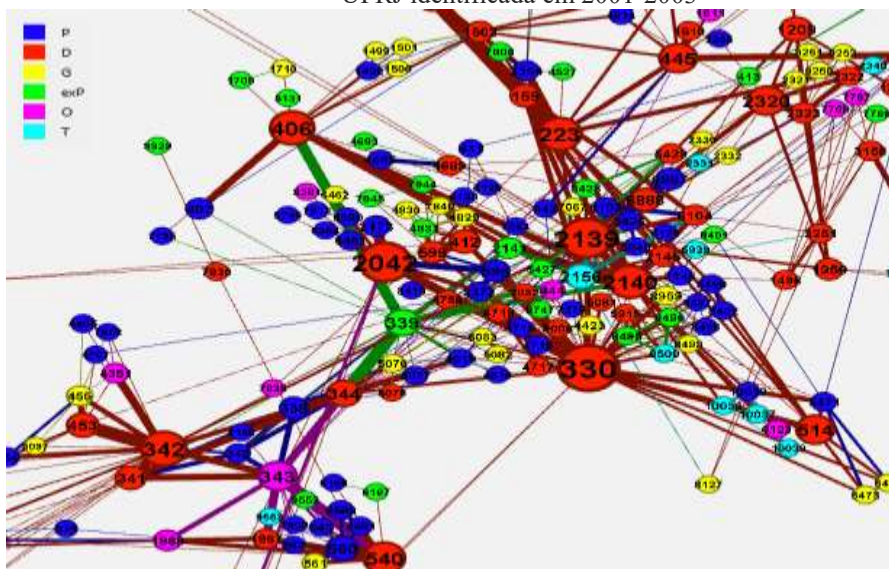
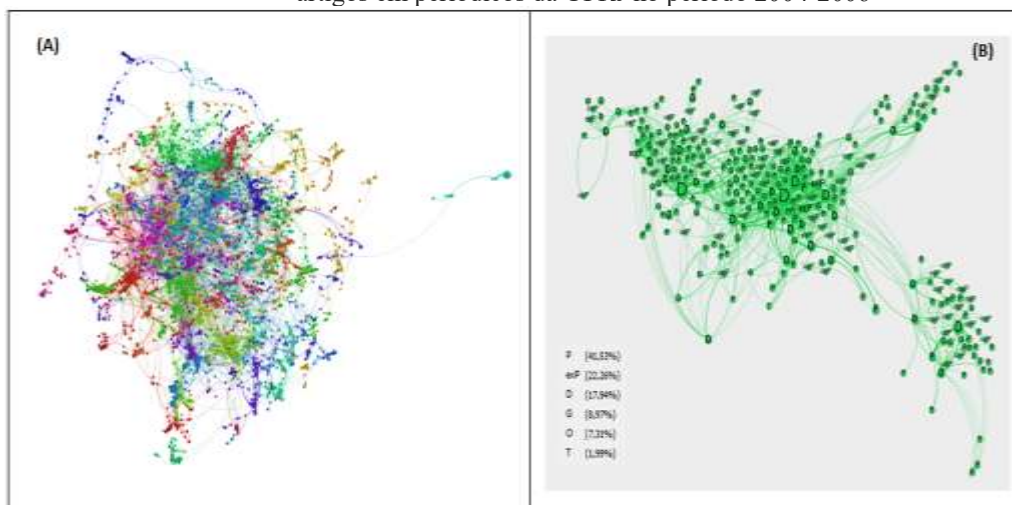


Figura 5 - (A) Grafo das comunidades de autores da UFRJ de publicações em periódicos identificadas no período 2004-2006. (B) Grafo da maior comunidade de autores de artigos em periódicos da UFRJ no período 2004-2006



Fonte: Elaborado pelo autor (2018)

O grau médio da rede foi de 6,45 e a transitividade (coeficiente de *clustering*) de 0,77. Esses dados mostram uma expansão do número de subgrupos (agrupamentos), resultante do maior nível de interação entre autores em relação ao triênio anterior.

A maior comunidade de autores, identificada no período de 2004-2006, é constituída por 301 autores (o que representa 6,81% do total de autores das 45 comunidades), 65 autores a mais que no triênio anterior. Esses autores estabeleceram 1.127 relacionamentos, arestas (Figura 5-B). O grau médio é de 7,48, ou seja, cada autor colabora, em média, com mais de sete pesquisadores. O coeficiente de transitividade apresentado é de 0,80.

A classificação de autores dessa comunidade por tipo de vínculo mostra que a maioria é composta por alunos de pós-graduação - P (41,53%), seguida por alunos egressos da pós-graduação - exP (22,26%) e docentes - D (17,94%). Alunos de graduação (G) representam 8,97% e autores com outros tipos de vínculo (O) 7,31%. Já o número de técnicos (T) participantes nessa comunidade representam 1,99%.

Os 20 autores que apresentam as maiores medidas de centralidade de grau são mostrados na Tabela 2.

Tabela 2 - Vínculo e grande área dos autores com maiores medidas de centralidade de grau da maior comunidade identificada das publicações da UFRJ em 2004-2006

Id	Vínculo*	Grau	Áreas de Atuação
56	D	70	Ciências Biológicas
233	D	58	Ciências Biológicas
54	D	45	Ciências Biológicas
593	D	44	Ciências Biológicas
55	D	39	Ciências Biológicas
57	D	38	Ciências Biológicas
182	D	33	Ciências Biológicas
600	D	31	Ciências Biológicas
558	D	30	Ciências Biológicas
442	D	25	Ciências Biológicas
3389	D	25	Ciências da Saúde
53	exP	24	Ciências Biológicas
337	D	24	Engenharias, Ciências Biológicas
395	D	23	Ciências Biológicas
244	D	22	Ciências Biológicas
397	D	22	Ciências Biológicas
833	P	21	Ciências Biológicas
4797	D	21	Ciências da Saúde
1047	D	20	Ciências Biológicas
6096	exP	20	Ciências Biológicas

Fonte: Elaborado pelo autor (2018).

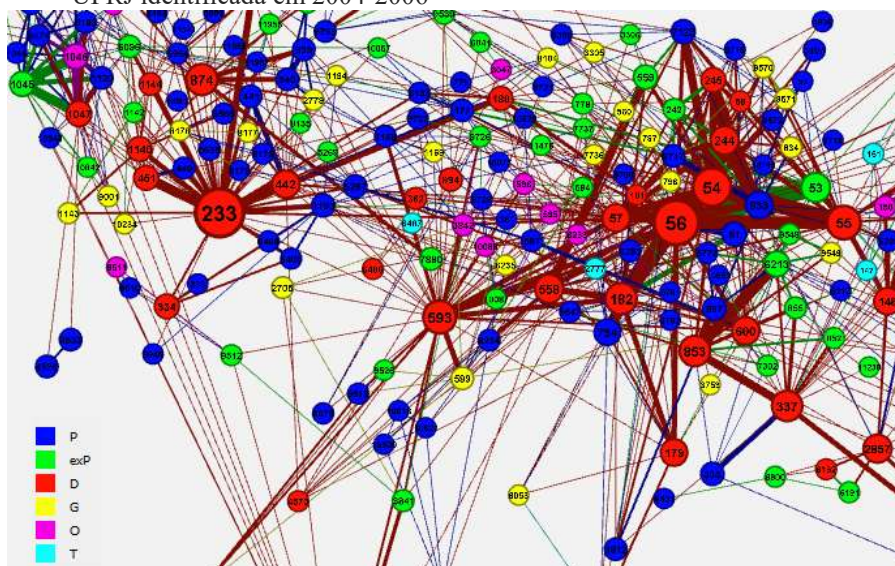
* Tipos de Vínculo: D (Docente); exP (Egresso Discente de Pós-graduação); P (Discente de Pós-graduação).

Observa-se que dos 20 autores identificados, 17 são docentes, dois são egressos da pós-graduação e um é aluno de pós-graduação.

No tocante às grandes áreas dos autores, verifica-se que 17 deles atuam nas Ciências Biológicas, e dois nas Ciências da Saúde. Apenas um autor atua em duas áreas do conhecimento: Engenharias e Ciências Biológicas.

Uma visão detalhada da rede em torno do grupo de autores mais centrais revela algumas informações importantes (Figura 6). Podemos notar dois grandes polos de colaboração. Um envolvendo os docentes Id = 56 e 54, os quais tem fortes ligações com outros docentes e, em seu entorno, alunos e egressos da pós-graduação.

Figura 6 - Grafo dos autores mais centrais da maior comunidade de autores de publicações da UFRJ identificada em 2004-2006



Fonte: Elaborado pelo autor (2018).

Por outro lado, temos o docente Id= 233, que concentra sua colaboração com alunos de pós-graduação.

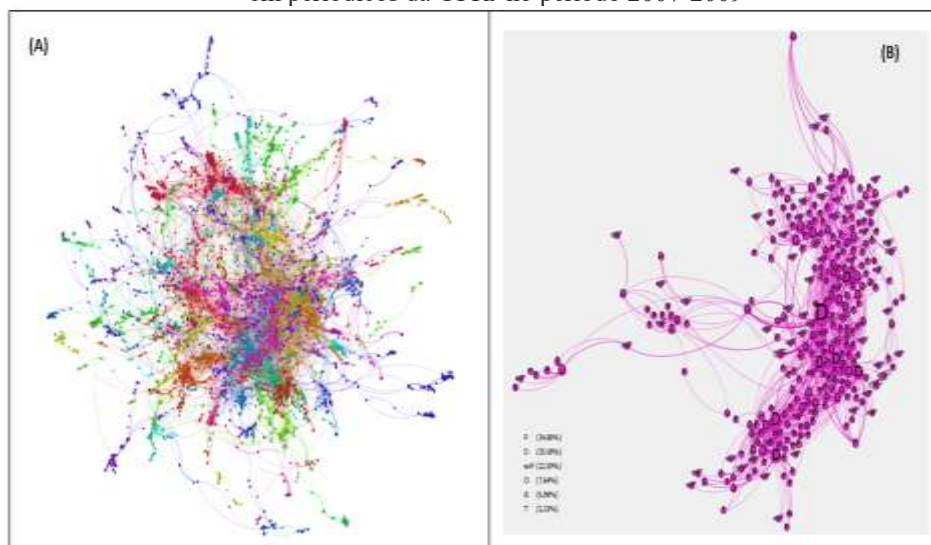
3.3 Comunidades de autores no triênio 2007-2009

No período 2007-2009, a rede de autores da UFRJ é composta por 7.484 autores (total de autores em artigos com colaboração).

A identificação de agrupamentos dessa população, aplicando o algoritmo Blondel et al. (2008), identificou um total de 50 comunidades, envolvendo 5.268 autores, os quais mantiveram entre si 17.548 ligações, representadas na rede por arestas (Figura 7-A). O grau médio foi de 7,48 e a transitividade (coeficiente de *clustering*) de 0,78. Há, portanto, relativo aumento na formação de agrupamentos em relação ao período anterior.

A maior comunidade de autores identificada no período de 2007-2009 é constituída por 301 autores (o que representa 5,71% do total de autores das 50 comunidades), número similar ao triênio anterior. Esses autores estabeleceram 1.133 relacionamentos, arestas (Figura 7-B). O grau médio é 7,52, isto é, cada autor colabora, em média, com mais de sete pesquisadores, e o coeficiente de *clustering* é de 0,79.

Figura 7 - (A) Grafo das comunidades de autores da UFRJ de publicações em periódicos identificadas no período 2007-2009. (B) Grafo da maior comunidade de autores de artigos em periódicos da UFRJ no período 2007-2009



Fonte: Elaborado pelo autor (2018)

A classificação de autores dessa comunidade por tipo de vínculo mostra que a maioria é composta por alunos de pós-graduação - P (34,88%), seguida por docentes - D (25,58%) e egressos da pós-graduação - exP (22,59%). Autores com outros tipos de vínculo (O) representam 7,64%, alunos de graduação (G) 6,98% e técnicos (T) 2,33%.

Os 20 autores com as maiores medidas de centralidade de grau são mostrados na Tabela 3. Podemos notar que 18 dos autores são docentes, um é aluno de pós-graduação e o outro é egresso da pós-graduação.

Em relação às áreas dos autores, verificamos que todos atuam na área de Ciências Biológicas, sendo que um deles também participa das Ciências Humanas.

Tabela 3 - Vínculo e grande área dos autores com maiores medidas de centralidade de grau da maior comunidade identificada das publicações da UFRJ em 2007-2009

Id	Vínculo*	Grau	Áreas de Atuação
290	D	54	Ciências Biológicas
194	D	41	Ciências Biológicas
1095	D	38	Ciências Biológicas
1365	D	34	Ciências Biológicas
1368	D	33	Ciências Biológicas
193	exP	31	Ciências Biológicas

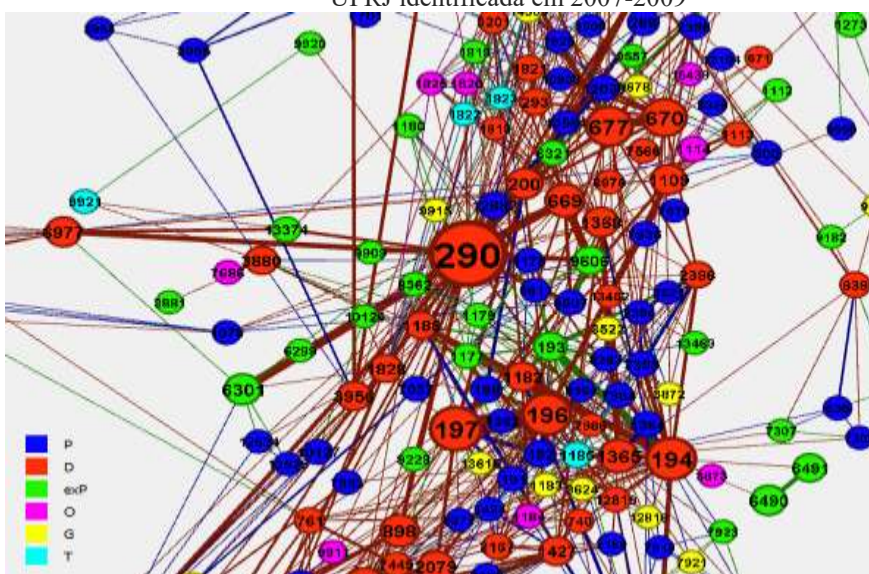
1828	D	29	Ciências Biológicas
2078	D	29	Ciências Biológicas
677	D	27	Ciências Biológicas
19	D	26	Ciências Biológicas
293	D	26	Ciências Biológicas
1427	D	26	Ciências Biológicas
294	D	25	Ciências Biológicas
1829	D	25	Ciências Biológicas, Ciências Humanas
3957	D	25	Ciências Biológicas
602	D	24	Ciências Biológicas
1827	D	24	Ciências Biológicas
670	D	23	Ciências Biológicas
1826	P	23	Ciências Biológicas
669	D	22	Ciências Biológicas

Fonte: Elaborado pelo autor (2018).

* Tipos de Vínculo: D (Docente); exP (Egresso Discente de Pós-graduação);
P (Discente de Pós-graduação).

Uma visualização detalhada da rede em torno do grupo de autores mais centrais (Figura 8) mostra maior concentração de colaboração entre docentes. Chama a atenção, por outro lado, a significativa participação, além dos alunos de pós-graduação, dos egressos da pós-graduação em torno dos autores mais representativos em termos de grau (Id = 290 e 194).

Figura 8 - Grafo dos autores mais centrais da maior comunidade de autores de publicações da UFRJ identificada em 2007-2009



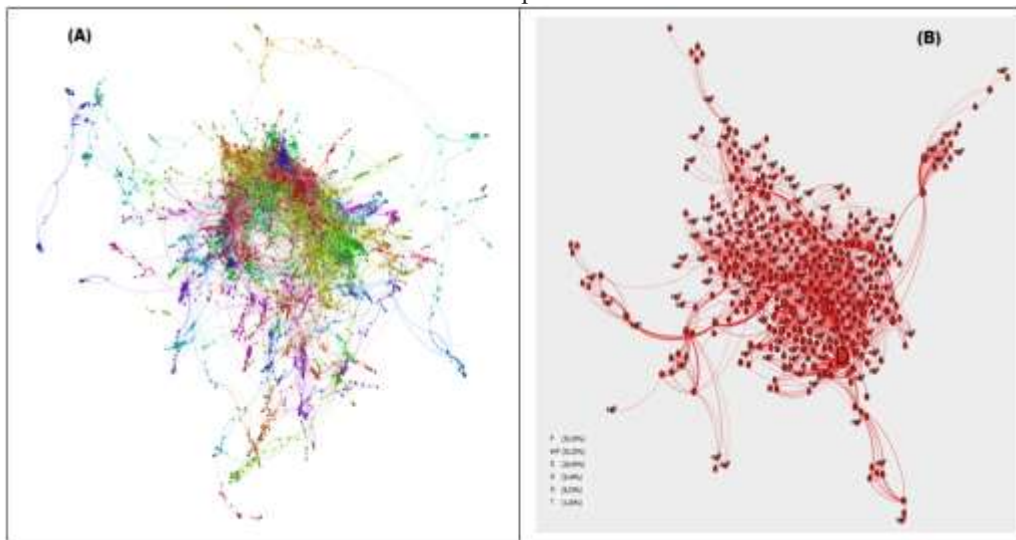
Fonte: Elaborado pelo autor (2018).

3.4 Comunidades de autores no triênio 2010-2012

A rede de autores da UFRJ, no período 2010-2012, é composta por 7.503 autores (total de autores em artigos com colaboração). A identificação de agrupamentos dessa população, aplicando o algoritmo Blondel et al. (2008), apontou um total de 42 comunidades, envolvendo 5.169 autores, os quais mantiveram entre si 19.328 conexões, representadas na rede por arestas (Figura 9-A). O grau médio foi de 7,47 e a transitividade (coeficiente de *clustering*) de 0,78.

Note-se que o número de autores e comunidades diminuiu em relação ao período anterior, 2007-2009, entretanto, o número de conexões aumentou. Isto pode indicar que o aumento da colaboração ocorreu pelo incremento dos agrupamentos já existentes em períodos anteriores.

Figura 9 - (A) Grafo que mostra as 40 comunidades de autores da UFRJ identificadas no período 2010-2012. (B) Grafo que destaca a maior comunidade de autores da UFRJ identificada no período 2010-2012



Fonte: Elaborado pelo autor (2018).

A maior comunidade de autores identificada no período de 2010-2012 é constituída por 396 autores (o que representa 7,66% do total de autores das 40 comunidades), 95 autores a mais que no triênio anterior. Esses autores estabeleceram 1.607 relacionamentos, arestas (Figura 9-B). O grau médio é

8,12, isto é, cada autor colabora, em média, com pouco mais de oito pesquisadores, e o coeficiente de *clustering* é de 0,81.

Ao verificar os autores dessa comunidade de acordo com seus respectivos tipos de vínculo, vemos que os grupos mais representativos são os alunos de pós-graduação - P (35,35%), seguido pelos egressos da pós-graduação - exP (25,25%) e docentes - D (20,45%). Os estudantes de graduação (G) representam 9,34%, os autores com outros tipos de vínculo (O) 8,33% e os técnicos (T) 1,26%. Desse total, identificamos os 20 autores que apresentam as maiores medidas de centralidade (Tabela 4).

Tabela 4 - Vínculo e grande área dos autores com maiores medidas de centralidade de grau da maior comunidade identificada das publicações da UFRJ em 2010-2012

Id	Vínculo*	Grau	Áreas de Atuação
711	D	80	Ciências Biológicas
536	D	47	Ciências Biológicas
2712	D	44	Ciências Biológicas
180	D	43	Ciências Biológicas
179	D	37	Ciências Biológicas
678	D	37	Ciências Biológicas
350	D	36	Ciências Biológicas
9049	D	36	Ciências Biológicas
167	D	34	Ciências Biológicas
168	D	34	Ciências Biológicas
334	D	33	Ciências Biológicas
177	P	32	Ciências Biológicas
1436	D	27	Ciências Biológicas
1623	D	27	Ciências Biológicas
563	D	25	Ciências Biológicas
564	D	25	Ciências Biológicas
1398	D	25	Ciências Biológicas, Ciências Exatas e da Terra
1714	D	25	Ciências Biológicas, Ciências Humanas
1608	D	23	Ciências Biológicas
1609	D	23	Ciências Biológicas, Ciências da Saúde

Fonte: Elaborado pelo autor (2018).

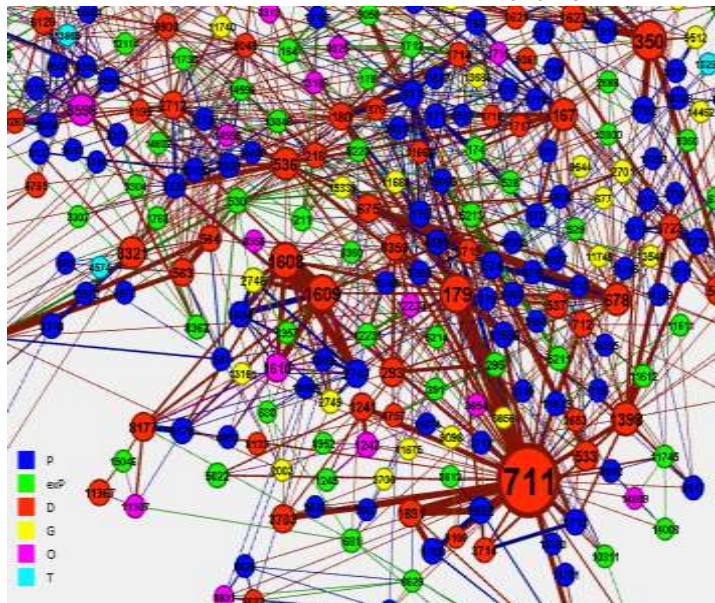
* Tipos de Vínculo: D (Docente); P (Discente de Pós-graduação)

Verifica-se que 19 deles são docentes e um aluno de pós-graduação. Em relação às áreas dos autores, constatamos que todos atuam na área de Ciências Biológicas, sendo que três deles atuam também nas seguintes áreas: Ciências da Saúde; Ciências Exatas e da Terra e; Ciências Humanas.

Uma visão detalhada da rede em torno do grupo de autores mais centrais revela algumas informações interessantes (Figura 10). É possível notar forte interação entre os docentes identificados com os números 711 e 179. O primeiro

mantendo, em seu entorno, colaborações com outros docentes e, o último, colaborando mais intensamente com alunos de pós-graduação.

Figura 10 - Grafo dos autores mais centrais da maior comunidade de autores de publicações da UFRJ identificada em 2010-2012



Fonte: Elaborado pelo autor (2018).

4 Discussão e conclusões

A maior parte das redes do mundo real caracteriza-se por conter comunidades muito entrelaçadas e ocultas, nas quais seus membros formam subgrupos, configurando uma imbricada rede entre todos eles. A identificação de comunidades possibilita definir e melhor conhecer os subgrupos em função da estrutura de escolha de seus enlaces, isto é, com base na premissa de que os membros de cada subgrupo tendem a escolher os mesmos colaboradores e serem escolhidos pelos mesmos autores. Dessa forma, a permanência em um subgrupo estabelece-se em função das similaridades de escolha, dadas e recebidas, por cada autor.

Nesse sentido, as métricas de grau e de transitividade mostraram que, ao longo do período analisado, novas colaborações entre autores da UFRJ foram estabelecidas, configurando novos agrupamentos/comunidades, tornando, dessa maneira, maior a transitividade da rede. Em outras palavras, ao longo dos triênios, a tendência dos autores da UFRJ se agruparem aumenta.

Assim, no período de 2001-2003, se identificou um total de 40 comunidades de autores, envolvendo 3.736 autores. Em média, esses autores participaram de publicações com pouco mais de seis autores (o grau médio foi de 6,05), e a transitividade foi de 0,76. Ao analisar a maior comunidade identificada nesse período, se verificou que ela é composta por 236 autores, e que cada autor colabora, em média, com cinco pesquisadores. O coeficiente de transitividade apresentado foi de 0,79. No período 2004-2006, houve um aumento do nível de colaboração e expansão do número de agrupamentos/comunidades ($n = 45$) em relação ao período anterior. No triênio seguinte (2007-2009), também se observou aumento do nível de colaboração e formação de agrupamentos/comunidades ($n = 50$) em relação ao período anterior.

No último triênio, por outro lado, se identificou um total de 42 comunidades, envolvendo 5.169 autores (o grau médio foi de 7,47, e a transitividade foi de 0,78), ou seja, diminuiu o número de autores e de comunidades em relação ao período anterior. Isto indica que houve uma intensificação das interações nos agrupamentos já existentes em períodos anteriores. Ao analisar a maior comunidade identificada nesse período (2010-2012), verificou-se que ela é composta por 396 autores, 95 autores a mais que no triênio anterior, e que cada autor colabora, em média, com pouco mais de oito pesquisadores (grau médio 8,12). A transitividade apresentada foi de 0,81.

Este quadro constitui traço característico da ciência contemporânea, sendo cada vez mais uma atividade coletiva, multidisciplinar e na qual participam diversos atores. A complexidade cada vez maior dos problemas de pesquisa supõe que pesquisadores isolados são capazes de tratar uma quantidade limitada de informações. Assim sendo, pesquisadores procuram se agrupar e interagir através de colégios invisíveis (SOLLA PRICE, 1965; CALLON et al., 1995).

Observar, portanto, o desempenho científico do corpo social de uma instituição é de particular interesse para obter informações sobre o grau de eficiência e envolvimento dos diferentes recursos humanos na ciência. Muitos estudos têm buscado esse relacionamento por meio de diferentes abordagens. Larivière (2012), por exemplo, investigou o desempenho de todos os alunos de

doutorado matriculados em cursos de pós-graduação em universidades de Quebec, que, segundo o autor, é o primeiro estudo em grande escala com esse foco. Entre os resultados, o autor destaca a alta contribuição de estudantes de doutorado: participando da autoria de cerca de 30% dos artigos de toda a província.

Em nosso caso, envolvendo a UFRJ, o foco não foi apenas no desempenho dos alunos de doutorado, mas em todos os tipos de relações institucionais. Dessa forma, ao observar a maior comunidade de autores identificada em 2001-2003, se verificou que a maioria de autores é composta por alunos de pós-graduação (38,02%), seguida por docentes (25,48%) e alunos de graduação (12,93%). Alunos egressos da pós-graduação representam 9,89% e os autores com outros tipos de vínculo 8,75%. Já os técnicos representam 4,94%.

No último triênio observado (2010-2012), a classificação de autores da maior comunidade identificada nesse período mostrou que a maioria é composta por alunos de pós-graduação (35,35%), seguido pelos ex-alunos de pós-graduação (25,25%) e docentes (20,45%). Os estudantes de graduação representam 9,34%, os autores com outros tipos de vínculo 8,33% e os técnicos 1,26%.

Esta análise só foi possível graças à principal fonte de informação deste trabalho, o SIGMA.UFRJ. Trata-se, portanto, de um estudo singular, de ampla abrangência, uma vez que envolve todo o corpo social de uma das principais instituições de ensino e pesquisa do Brasil.

Por fim, cabe ressaltar que uma das contribuições destacáveis deste estudo é a particular visão que proporciona sobre a produção científica da UFRJ, já que, a diferença de outros estudos bibliométricos, a abordagem conduzida não se restringe às limitações de realizar estudos deste tipo utilizando as tradicionais bases de dados internacionais – *Web of Science* (WoS) ou *Scopus*. Bases como essas não permitem observar determinadas especificações da produção científica como, por exemplo, o tipo de vínculo institucional dos autores como aqui mostrado.

Financiamento

Este trabalho contou com financiamento da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Referências

BARABÁSI, A.; JEONG, H.; NÉDA, Z.; RAVASZ, E.; SCHUBERT, A.; VICSEK, T. Evolution of the social network of scientific collaborations. **Physica A**, v. 311, n. 3, p. 590-614, 2002.

BARBOSA, Diego Andrés de Barros Lima.; AVELINO, Leonardo Borges; SOUZA, Rarylson Freitas de.; OLIVEIRA, Camila Cristina Gomes Ferreira de; JUSTEL, Claudia. Medidas de centralidade e detecção de comunidades em rede de co-autoria. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 43., 2011, Ubatuba. **Anais...** Ubatuba, 2011. p. 1-10.

BLONDEL, Vincent D.; GUILLAUME, Jean-Loup; LAMBIOTTE, Renaud; LEFEBVRE, Etienne. Fast unfolding of communities in large networks. **J. Stat. Mech**, p.100-108, 2008.

CALLON, Michel; COURTIAL, Jean-Pierre; PENAN, Hérve. **Cienciometría. El estudio cuantitativo de la actividad científica: de la bibliometría a la vigilancia tecnológica**. Gijón: Trea, 1995.

CANCHUMANI, Roberto Mario Lóvon. Redes de Interação Científica: identificação de comunidades de autores na UFRJ. In: ENCONTRO BRASILEIRO DE BIBLIOMETRIA E CIENTOMETRIA, 6., 2018, Rio de Janeiro. **Anais....** Rio de Janeiro: UFRJ, 2018. p. 333-343.

CANCHUMANI, Roberto Mario Lóvon; LETA, Jacqueline; FIGUEIREDO, Antonio MacDowell de. Domínios Científicos: mapeamento de áreas do conhecimento da Universidade Federal do Rio de Janeiro. **Inf. & Soc: Est**, v. 27, n. 2, p.199-218, 2017.

FORTUNATO, Santo. Community detection in graphs. **Physics Reports**, v. 486, p. 75-174, 2010.

LARIVIERE, Vincent. On the shoulders of students? The contribution of PhD students to the advancement of knowledge. **Scientometrics**, v. 90, n. 2, p. 463-481, 2012.

PALLA, Gergely; DERÉNYI, Imre; FARKAS, Illes; VICSEK, Tamas. Uncovering the overlapping community structure of complex networks in nature and society. **Nature**, 435, p. 814-818, 2005.

RIVELLINI, Giulia; RIZZI, Ester; ZACCARIN, Susanna. The science network in Italian population research: an analysis according to the social network perspective. *Scientometrics*, v. 67, n. 3, p. 407-418, 2006.

PERIANES-RODRIGUEZ, Antonio; OLMEDA-GÓMEZ, Carlos; MOYA-ANEGON, Félix. Detecting, identifying and visualizing research groups in co-authorship networks. *Scientometrics*, v. 82, n. 2, p. 307-319, 2010.

RODRIGUEZ, Josep A. **Análisis estructural y de redes**. Madrid: Centro de Investigaciones Sociológicas, 1995.

SOLLA PRICE, Derek J. Networks of scientific papers. *Science*, v. 149, n. 3.683, p. 510-515, 1965.

Scientific Interaction Networks: mapping of author communities

Abstract: It provides the results of a proposal of identification of authors communities realized at the Federal University of Rio de Janeiro - UFRJ. The objective was to identify the clustering tendency of authors of the institution over time, as well as to observe more specific characteristics of the main communities identified. The study investigated 44,233 records retrieved from the institutional database, SIGMA.UFRJ, articles published in journals from 2001 to 2012. Data were organized and analyzed using *Excel* and *Gephi* software. The metrics of degree and transitivity/modularity showed that, over the study period, new collaborations among authors of UFRJ were established, forming new clusters/communities of authors. The largest authors' clusters/communities were detailed and revealed their main specifications

Keywords: Authors Communities. Collaboration Networks. Co-authorship.

Recebido: 23/09/2017

Aceito: 03/12/2018

