

RICARDO LOPES DE ANDRADE

**A Influência das Redes de Coautoria na Performance dos
Bolsistas de Produtividade e nos Programas de Pós-
Graduação em Engenharia de Produção**

Recife
2016

RICARDO LOPES DE ANDRADE

A Influência das Redes de Coautoria na Performance dos Bolsistas de Produtividade e nos Programas de Pós-Graduação em Engenharia de Produção

Dissertação de Mestrado apresentada à UFPE para a obtenção de grau de Mestre como parte das exigências do Programa de Pós-Graduação em Engenharia de Produção (Área de Concentração: Pesquisa Operacional).

Orientador: Leandro Chaves Rêgo, PhD

Recife
2016

Catálogo na fonte
Bibliotecária Valdicea Alves, CRB-4 / 1260

A543i Andrade, Ricardo Lopes de.
A Influência das redes de coautoria na performance dos bolsistas de produtividade e nos programas de pós-graduação em engenharia de produção. / Ricardo Lopes de Andrade. - Recife: O Autor, 2016.
144 folhas, Il., Abr., Tabs.

Orientador: Prof^o. PhD. Leandro Chaves Rêgo.

Dissertação (Mestrado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Engenharia de Produção, 2016.
Inclui: Referências.

1. Engenharia de Produção. 2. Rede de coautorias. 3. Redes Ponderadas. 4. Análise de redes sociais. 5. Nível de produtividade. 6. Programas de pós-graduação. I. Rêgo. Leandro Chaves (Orientador). II. Título.

UFPE

658.5 CDD (22. ed.)

BCTG/2016-90

UNIVERSIDADE FEDERAL DE PERNAMBUCO

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

**PARECER DA COMISSÃO EXAMINADORA
DE DEFESA DE DISSERTAÇÃO DE MESTRADO DE**

RICARDO LOPES DE ANDRADE

**A Influência das Redes de Coautoria na Performance dos Bolsistas
de Produtividade e nos Programas de Pós-Graduação em
Engenharia de Produção**

ÁREA DE CONCENTRAÇÃO: Pesquisa Operacional

A comissão examinadora composta pelos professores abaixo, sob a presidência do primeiro, considera o candidato **RICARDO LOPES DE ANDRADE, APROVADO .**

Recife, 26 de fevereiro de 2016.

Prof. LEANDRO CHAVES RÊGO, PhD (UFPE)

Prof. CAROLINE MARIA DE MIRANDA MOTA, Doutora (UFPE)

Prof. RAYDONAL OSPINA MARTÍNEZ, Doutor (UFPE)

AGRADECIMENTOS

A DEUS, pela coragem, saúde e sabedoria para superar as dificuldades, enxergar com otimismo e vencer.

Ao meu pai João e a minha mãe Edilva que amo tanto e sempre batalharam e apoiaram minhas decisões. Ao meu irmão Rafael, amigo, companheiro e sonhador. Presente em todas as minhas conquistas. Amo-te muito.

Ao professor Leandro, orientador deste trabalho, por todos os ensinamentos, dedicação, confiança no meu trabalho e pelos desafios dados a mim.

Aos meus amigos de caminhada Lucas, Guilherme, Henrique e Túlio. Muito feliz por ter conhecidos vocês e pelos momentos de estudos, dificuldades e entretenimentos que passámos juntos.

A banca examinadora, por contribuírem de forma significativa para as melhorias deste trabalho, em especial o professor Raydonal.

Ao professor Jesús Mena-Chalco pelo auxílio no scriptLattes, ferramenta essencial para obtenção dos dados deste trabalho.

E ao CNPq, pelo auxílio financeiro concedido durante o período de mestrado.

RESUMO

O estudo, por meio da análise de redes sociais (ARS), das redes de coautoria permite identificar e analisar porque os pesquisadores colaboram e quais as consequências dessas colaborações. Pesquisas apontam a melhoria da qualidade dos artigos e o aumento do número de publicações como benefícios da colaboração. É crescente o número de estudos que buscam analisar as redes de coautorias e os impactos nos desempenhos dos autores, medidos, por exemplo, por meio dos índices h e g. Outros trabalhos também analisam o desempenho de grupos acadêmicos com base na estrutura e dinâmica de relacionamento entre os pesquisadores. O desempenho dos pesquisadores e dos grupos acadêmicos (programas de pós-graduação) também são avaliados por instituições de fomento à pesquisa como a CAPES e o CNPq. A CAPES avalia os programas de pós-graduação para assegurar e manter a qualidade dos programas. O CNPq financia projetos de pesquisa e concede bolsas de produtividade a pesquisadores de diferentes áreas do conhecimento. Ambas têm como indicadores, entre outros, as produções bibliográficas dos pesquisadores. Esta dissertação buscou analisar como a estrutura social, por meio das métricas de ARS, influenciou o nível da bolsa de produtividade dos pesquisadores bolsistas de produtividade do CNPq na área de engenharia de produção e os conceitos dos programas de pós-graduação em engenharia de produção avaliados pelo CAPES. Foi desenvolvido também um método que integrou a importância do nó na estrutura da rede. Com as informações bibliométricas dos pesquisadores bolsistas foi criada uma rede de coautoria e calculadas as métricas de ARS de três maneiras distintas: não-ponderadas, ponderadas com pesos das arestas e ponderadas com pesos das arestas e dos nós. Essas métricas foram correlacionadas (correlação de Spearman) com os níveis de bolsa de produtividade dos pesquisadores e concluiu-se que as métricas ponderadas com os pesos das arestas e dos nós (*E-I index*, centralidade de intermediação, de proximidade e utilidade) e as métricas não-ponderadas (*E-I index*, PageRank, centralidade de grau e de autovetor) são as que obtiveram maiores correlações. Com as informações bibliográficas dos professores permanentes dos programas foram construídas, para cada triênio (2007-2009 e 2010-2012), três conjuntos de redes: redes entre professores de cada programa analisado, rede de coautoria entre os programas, e rede entre os professores de todos os programas. Para o primeiro conjunto de redes foram analisadas as variações das medidas globais entre os triênios e os impactos dessas variações nos conceitos e concluiu-se que: a variação da quantidade de artigos publicados (com ou sem coautoria) impacta nas variações dos conceitos dos programas; as variações do número de arestas, da média das arestas, da densidade, do coeficiente médio, do grau médio e do grau médio ponderado apontaram correlações fortes e significativas com os conceitos dos programas. Para o segundo conjunto de redes, onde cada programa é um nó da rede, foram analisadas as variações das métricas individuais de ARS entre os triênios e as correlações com os conceitos dos programas, nenhuma das métricas calculadas apresentou correlação significativa. O último conjunto de redes, os valores médios, de uma métrica, obtidos por um determinado grupo de professores pertencente a um dado programa foram considerados com o valor médio do programa. As variações das métricas foram correlacionadas com as variações dos conceitos dos programas, concluindo-se que as variações da centralidade de grau e do grau ponderado impactam os conceitos dos programas.

Palavras-chave: Rede de Coautorias. Redes Ponderadas. Análise de Redes Sociais. Nível de Produtividade. Programas de Pós-Graduação. Engenharia de Produção.

ABSTRACT

The study, through social network analysis (SNA), of co-authorship networks allows to identify and to analyze why researchers collaborate and which are the consequences of these collaborations. Researches suggest the improvement on the quality of papers and the increasing number of publications as benefits of collaboration. An increasing number of studies seek to analyze the co-authorship networks and the impact on the performance of the authors, measured, for example, through the h and the g indexes. Other works analyze the performance of academic groups based on the structure and relationship dynamics among researchers. The performance of researchers and academic groups (Graduate Programs) are also evaluated by funding institutions, such as CAPES and CNPq. CAPES evaluates Graduate Programs to ensure and maintain the quality of these programs. CNPq finances research projects and provides productivity fellowships to researchers from different areas of knowledge. Both have as indicators, among others, the bibliography production of the researchers. This dissertation sought to analyze how the social structure, through SNA metrics, influenced the level of productivity fellowship of CNPq's research productivity fellows in the area of Industrial Engineering and the grades of Graduate Programs in Industrial Engineering evaluated by CAPES. A method that integrated the importance of the node in the network structure was developed. With bibliometric information of the fellows, it was created a network of co-authorship and calculated the SNA metrics in three distinct ways: unweighted, with weights of edges and with weights of both edges and nodes. These metrics were correlated (Spearman correlation) with the researchers' productivity fellowship levels and it was found that the metrics with weights of edges and nodes (E-I index, intermediary centrality, closeness and utility) and unweighted metrics (E-I index, PageRank, degree centrality and eigenvector) are the ones that had higher correlations. With the bibliographic information of the permanent professors of the Graduate Programs, three sets of networks were built for each three-year period (2007-2009 and 2010-2012): networks among professors in each program analyzed, co-authoring network among programs and network among professors of all programs. For the first set of networks, the variations of overall measures between the studied periods and the effects of these variations on the grades have been analyzed. It was found that: the variation of the amount of published articles (with or without co-authorship) impact on variations of the programs' grades; variations in the number of edges, the mean edges, the density, the average coefficient, average weighted degree and average degree indicated strong and significant correlation with programs' grades. For the second set of networks, where each program is a network node, variations of the individual metrics of SNA within the study periods and correlations with the grade of the programs were analyzed, none of the computed metrics presented significant correlation. For the last set of networks, the average values of a metric, obtained for a certain group of professors that belong to a certain program were considered to be the mean value of the program. Variations of the metrics were correlated with variations on the grades of the programs, so that the variations of the degree centrality and the weighted degree influence the grade of programs.

Keywords: Co-authorship Network. Weighted Networks. Social Network Analysis. Productivity level. Graduate Programs. Industrial Engineering.

LISTA DE FIGURAS

Figura 2.1 – Desenho do problema das sete pontes.....	20
Figura 2.2 – Ilustração de exemplos de grafos	21
Figura 2.3 – Exemplo de grafo bipartido	24
Figura 4.1 – Exemplo de aplicação do modelo	38
Figura 5.1 - Rede de coautoria entre os bolsistas	48
Figura 5.2 - Gráfico boxplot para o E-I index	49
Figura 5.3 – Gráfico boxplot para o E-I index_W.....	50
Figura 5.4 – Gráfico boxplot para o E-I index_Z.....	50
Figura 5.5 – Relação entre E-I Index e E-I index_W	52
Figura 5.6 – Relação entre E-I index e o E-I index_Z.....	52
Figura 5.7 – Relação entre o E-I index_W e o E-I index_Z.....	53
Figura 5.8 – Gráfico boxplot para CGNP.....	54
Figura 5.9 – Gráfico boxplot para CGPW	55
Figura 5.10 – Gráfico boxplot para CGPZ	55
Figura 5.11 – Relação entre a CGNP e a CGPW	56
Figura 5.12 – Relação entre a CGNP e a CGPZ.....	57
Figura 5.13 – Relação entre a CGPW e a CGPZ.....	57
Figura 5.14 – Gráfico boxplot para a TS.....	59
Figura 5.15 – Gráfico boxplot para a TSZ	59
Figura 5.16 – Relação entre a TS e a TSZ.....	60
Figura 5.17 – Gráfico boxplot para CPNP	62
Figura 5.18 – Gráfico boxplot para CPPW	63
Figura 5.19 – Gráfico boxplot para CPPZ.....	63
Figura 5.20 – Relação entre a CPNP e a CPPW.....	64
Figura 5.21 – Relação entre a CPNP e a CPPZ.....	65
Figura 5.22 – Relação entre a CPPW e a CPPZ.....	65
Figura 5.23 – Gráfico boxplot para a CIMP	67
Figura 5.24 – Gráfico boxplot para a CIPW.....	68
Figura 5.25 – Gráfico boxplot para a CIPZ.....	68
Figura 5.26 – Relação entre a CIMP e a CIPW	69
Figura 5.27 – Relação entre a CIMP e a CIPZ.....	70
Figura 5.28 – Relação entre a CIPW e a CIPZ.....	70
Figura 5.29 – Gráfico boxplot da CANP	72
Figura 5.30 – Gráfico boxplot da CAPW	72
Figura 5.31 – Gráfico boxplot da CAPZ	73
Figura 5.32 – Relação entre a CANP e a CAPW	74
Figura 5.33 – Relação entre a CIMP e a CIPZ.....	74
Figura 5.34 – Relação entre a CAPW e a CAPZ.....	75
Figura 5.35 – Gráfico boxplot do PRNP	77
Figura 5.36 – Gráfico boxplot do PRPW	77
Figura 5.37 – Gráfico boxplot do PRPZ.....	78
Figura 5.38 – Relação entre o PRNP e o PRPW	79
Figura 5.39 – Relação entre o PRNP e o PRPZ	79
Figura 5.40 – Relação entre o PRPW e o PRPZ.....	80
Figura 5.41 – Gráfico boxplot do CLNP	81
Figura 5.42 – Gráfico boxplot do CLPW	82

Figura 5.43 – Gráfico boxplot do CLPZ	82
Figura 5.44 – Relação entre o CLNP e o CLPW	83
Figura 5.45 – Relação entre o CLNP e o CLPZ	84
Figura 5.46 – Relação entre o CLPW e o CLPZ	84
Figura 5.47 – Gráfico boxplot da ENP	86
Figura 5.48 – Gráfico boxplot da EPW	86
Figura 5.49 – Gráfico boxplot da EPZ	87
Figura 5.50 – Relação entre a ENP e a EPW	88
Figura 5.51 – Relação entre a ENP e a EPZ	88
Figura 5.52 – Relação entre a EPW e a EPZ	89
Figura 5.53 – Gráfico boxplot da UNP	90
Figura 5.54 – Gráfico boxplot da UPW	91
Figura 5.55 – Gráfico boxplot da UPZ	91
Figura 5.56 – Relação entre a UNP e a PW	92
Figura 5.57 – Relação entre a UNP e a UPZ	93
Figura 5.58 – Relação entre a UPW e a UPZ	93
Figura 6.1 – Evolução das redes que diminuíram o conceito	105
Figura 6.2 – Evolução das redes que mantiveram o conceito	107
Figura 6.3 - Evolução das redes que aumentaram o conceito	110
Figura 6.4 – Rede de coautoria entre programas no triênio 2007-2009	118
Figura 6.5 - Rede de coautoria entre programas no triênio 2010-2012	120
Figura 6.6 - Rede de coautoria entre professores permanentes dos programas no triênio 2007-2009	126
Figura 6.7 - Rede de coautoria entre professores permanentes dos programas no triênio 2010-2012	128

LISTA DE TABELAS

Tabela 4.1 – Posições dos vértices nas três métricas.....	39
Tabela 4.2 – Classificação dos vértices	39
Tabela 5.1 – Critérios específicos do CA-EP	44
Tabela 5.2 - Descrição geral do nível macro da rede	47
Tabela 5.3 – Correlações entre as métricas de E-I index.....	51
Tabela 5.4 – Os 10 mais bem posicionados na centralidade de grau	54
Tabela 5.5 – Correlações entre as métricas de centralidade de grau	56
Tabela 5.6 – Correlações das métricas de centralidades de grau com os níveis das bolsas	58
Tabela 5.7 – Os 10 mais bem posicionados na média das ligações fortes	58
Tabela 5.8 – Correlações entre médias de ligações fortes	60
Tabela 5.9 – Correlações das métricas média das ligações fortes com os níveis das bolsas....	61
Tabela 5.10 – Os 10 mais bem posicionados na centralidade de proximidade	61
Tabela 5.11 – Correlações entre as métricas de centralidade de proximidade	64
Tabela 5.12 – Correlações das métricas de centralidade de proximidade com os níveis das bolsas	66
Tabela 5.13 – Os 10 mais bem posicionados na centralidade de intermediação.....	66
Tabela 5.14 – Correlações entre as métricas de centralidade de intermediação.....	69
Tabela 5.15 – Correlações das métricas de centralidade de intermediação com os níveis das bolsas	71
Tabela 5.16 – Os 10 mais bem posicionados na centralidade de autovetor	71
Tabela 5.17 – Correlações entre as métricas de centralidade de autovetor	73
Tabela 5.18 – Correlações das métricas de centralidade de autovetor com os níveis das bolsas	75
Tabela 5.19 – Os 10 mais bem posicionados no PageRank	76
Tabela 5.20 – Correlações entre os PageRank	78
Tabela 5.21 – Correlações dos PageRank com os níveis das bolsas	80
Tabela 5.22 – Os 10 mais bem posicionados no Cluster	81
Tabela 5.23 – Correlações entre os Coeficientes de Cluster	83
Tabela 5.24 – Correlações das métricas de Coeficientes de Cluster com os níveis das bolsas	85
Tabela 5.25 – Os 10 mais bem posicionados na excentricidade	85
Tabela 5.26 – Correlações entre as Excentricidades	87
Tabela 5.27 – Correlações das métricas de excentricidades com os níveis das bolsas	89
Tabela 5.28 – Os 10 mais bem posicionados na Utilidade.....	90
Tabela 5.29 – Correlações entre as utilidades	92
Tabela 5.30 – Correlações d os níveis das bolsas com a utilidade	94
Tabela 6.1– Conceitos dos PPGEP.....	100
Tabela 6.2 – Distribuições dos conceitos dos programas	101
Tabela 6.3 – Dados de publicações	102
Tabela 6.4 – Relação entre publicações e número de professores.....	102
Tabela 6.5 – Correlações entre Conceito e publicações	103
Tabela 6.6 – Dados sobre nós e arestas	110
Tabela 6.7 – Dados sobre densidade, cluster e componente gigante.....	111
Tabela 6.8 – Grau médio, proximidade média e intermediação média dos programas.....	112
Tabela 6.9 – Média da centralidade de autovetor, do PageRank e da utilidade	114
Tabela 6.10 – Correlações entre as médias das medidas.....	115
Tabela 6.11 – Correlações entre as métricas globais de rede e os conceitos dos programas .	116
Tabela 6.12 – Correlações das métricas ponderadas por W e cos conceitos dos programas .	117

Tabela 6.13 – Correlações das métricas ponderadas por Z com os conceitos dos programas	117
Tabela 6.14 – Descrição geral da rede de coautoria entre programas no triênio 2007-2009 .	119
Tabela 6.15 – Descrição geral da rede de coautoria entre programas no triênio 2010-2012 .	119
Tabela 6.16 – Centralidade de grau e grau ponderado	121
Tabela 6.17 – Centralidade de proximidade e centralidade de intermediação	122
Tabela 6.18 – Centralidade de autovetor e PageRank	123
Tabela 6.19 – Excentricidade, cluster e utilidade	124
Tabela 6.20 – Correlações com os conceitos dos programas	124
Tabela 6.21 – Correlações das métricas ponderadas por W com os conceitos dos programas	125
Tabela 6.22 – Correlações das métricas ponderadas por Z com os conceitos dos programas	125
Tabela 6.23 – Descrição geral da rede de coautoria entre professores permanentes dos programas (2007-2009)	127
Tabela 6.24 – Descrição geral da rede de coautoria entre professores permanentes dos programas (2010-2012)	127
Tabela 6.25 – Média da centralidade de grau e grau ponderado	129
Tabela 6.26 – Média da centralidade de proximidade e de intermediação.....	130
Tabela 6.27 – Média do coef. de cluster, do PageRank e centralidade de autovetor	131
Tabela 6.28 – Utilidade e E-I Index	132
Tabela 6.29 – Correlações das métricas não-ponderas com os conceitos dos programas.....	133
Tabela 6.30 – Correlações das métricas ponderadas por W com os conceitos dos programas	134
Tabela 6.31 – Correlações ponderadas por Z com os conceitos dos programas	135

LISTA DE ABREVIATURAS

ARS – Análise de Redes Sociais

CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CA-EP – Comitê Assessor de Engenharia de Produção e de Transporte do CNPq

CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico

CTC – Conselho Técnico-Científico

PQ – Bolsa de Produtividade em Pesquisa

SNPG – Sistema Nacional de Programas de Pós-Graduação

CGNP – Centralidade de Grau Não-Ponderada

CGPW – Centralidade de Grau Ponderada com pesos das arestas

CGPZ – Centralidade de Grau Ponderada com pesos das arestas e dos nós

MLFW – média das ligações fortes com os pesos das arestas

MLFZ – média das ligações fortes com os pesos das arestas e dos nós

CPNP – Centralidade de Proximidade

CPPW – Centralidade de Proximidade com pesos das arestas

CPPZ – Centralidade de Proximidade com pesos das arestas e dos nós

CINP – Centralidade de Intermediação Não-Ponderada

CIPW – Centralidade de Intermediação com pesos das arestas

CIPZ – Centralidade de Intermediação com pesos das arestas e dos nós

CANP – Centralidade de Automação

CAPW – Centralidade de Automação com pesos das arestas

CAPZ – Centralidade de Automação com pesos das arestas e dos nós

PRNP – PageRank Não-Ponderado

PRPW – PageRank Ponderado com pesos das arestas

PRPZ – PageRank Ponderado com pesos das arestas e dos nós

CLNP – Coeficiente de Cluster Não-Ponderado

CLPW – Coeficiente de Cluster com pesos das arestas

CLPZ – Coeficiente de Cluster com pesos das arestas e dos nós

ENP – Excentricidade Não-Ponderada

EPW – Excentricidade Ponderada com pesos das arestas

EPZ – Excentricidade Ponderada com pesos das arestas e dos nós

UNP – Utilidade Não-Ponderada

UPW – Utilidade Ponderada com pesos das arestas

UPZ – Utilidade Ponderada com pesos das arestas e dos nós

SUMÁRIO

1	INTRODUÇÃO	14
1.1	JUSTIFICATIVA	15
1.2	OBJETIVO GERAL	17
1.3	OBJETIVO ESPECÍFICO	17
1.4	ESTRUTURA DO TRABALHO	18
2	REFERENCIAL TEÓRICO	19
2.1	REDES DE COAUTORES.....	19
2.2	CARACTERIZAÇÃO DE REDES DE COAUTORIA	20
2.2.1	Grafos.....	20
2.2.2	Métricas	24
3	METODOLOGIA	32
3.1	OBTENÇÃO DE DADOS.....	32
3.2	ANÁLISE DE REDES SOCIAIS	33
4	MÉTODO PROPOSTO	37
5	REDE DE COAUTORIA ENTRE OS BOLSISTAS DE PRODUTIVIDADE	41
5.1	INTRODUÇÃO	41
5.2	SISTEMA DE AVALIAÇÃO DA PRODUTIVIDADE EM PESQUISA	43
5.3	METODOLOGIA	45
5.4	CONHECENDO A REDE DE COAUTORIA ENTRE OS BOLSISTAS DE PRODUTIVIDADE.....	46
5.5	IMPACTO DAS MÉTRICAS DE ARS NO NÍVEL DE PRODUTIVIDADE.....	48
5.5.1	<i>E-I Index</i>	48
5.5.2	Centralidade de Grau	53
5.5.3	Média das Ligações Fortes.....	58
5.5.4	Centralidade de Proximidade.....	61
5.5.5	Centralidade de Intermediação	66
5.5.6	Centralidade de Autovetor	71
5.5.7	PageRank	76
5.5.8	Coeficiente de <i>Cluster</i>	80
5.5.9	Excentricidade	85
5.5.10	Utilidade.....	89
5.6	CONCLUSÕES	94
6	REDES DE COAUTORIA ENTRE OS PROFESSORES PERMANENTES	96
6.1	INTRODUÇÃO	96
6.2	SISTEMA DE AVALIAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO	97
6.3	METODOLOGIA	98
6.4	REDES DE COAUTORIAS.....	99

6.4.1	Redes de Coautorias dos Programas de Pós-Graduação em Engenharia de Produção	101
6.4.2	Redes de Coautorias entre os Programas de Pós-Graduação em Engenharia de Produção ...	118
6.4.3	Redes de Coautorias entre os Professores Permanentes dos PPGEPP.....	126
6.5	CONCLUSÕES	135
7	CONCLUSÕES E TRABALHOS FUTUROS	137
	REFERÊNCIAS	140

1 INTRODUÇÃO

Uma rede social, para Newman (2001a), é composta por um conjunto de pessoas ou grupos, também chamados de atores, onde cada um dos quais tem algum tipo de ligação (vínculo) para alguns ou todos os outros. Tanto os atores quanto os vínculos podem ser constituídos de diferentes formas, dependendo das questões de interesses. Um ator pode ser uma única pessoa, um grupo, ou uma empresa. O vínculo pode ser uma amizade entre duas pessoas, uma colaboração ou membro comum entre duas equipes, ou uma relação de negócios entre empresas (NEWMAN, 2001b). Uma rede de colaboração científica é um tipo de rede social no qual os atores são os pesquisadores e os vínculos podem ser manifestados pelas diversas formas de colaboração existentes no contexto científico. A coautoria, elaboração de uma mesma publicação por dois ou mais autores, é uma forma de colaboração.

Com o desenvolvimento da ciência moderna, mais e mais pesquisas são concluídas pela colaboração de vários pesquisadores, promovendo a eficiência e a fertilização de novas ideias (Liu *et al.*, 2014). Um dos motivos para esse aumento, segundo Wagner & Leydesdorff (2003), é a grande disponibilidade de tecnologia de informação para os cientistas. Já para Gibbons *et al.* (1994) a colaboração científica é engrenada pelo aumento da complexidade do conhecimento e pela demanda de competência mais especializada e interdisciplinar nas pesquisas.

Paralelo ao aumento da colaboração entre pesquisadores, tem havido também, muitas pesquisas de análise de redes de coautoria que identificam os autores influentes nas mesmas (KEMPE&KLEINBERG, 2005). Para Barabasi *et al.* (2002) estudar redes de coautoria é interessante porque permite determinar como o campo de pesquisa está evoluindo e fazer previsões sobre a direção desse campo e onde os avanços terão maiores probabilidades de ocorrerem. O mesmo é relatado por Newman (2004), onde segundo ele a estrutura dessas redes revela muitas características e padrões interessantes das comunidades acadêmicas. O grau de transitividade ou coeficiente de *cluster* de uma rede de coautoria, por exemplo, informa que existe uma maior probabilidade de que dois cientistas tenham um artigo em coautoria se eles tiverem colaborado com um ou mais cientistas em comum.

A análise dos padrões, das interações e implicações das redes é facilitada pela utilização da Teoria dos Grafos, por meio de um conjunto de algoritmos matemáticos (WASSERMAN & FAUST, 1994). Tal abordagem é conhecida como Análise de Rede Social (ARS) e será abordada neste trabalho como metodologia que explica os fenômenos indagados.

1.1 JUSTIFICATIVA

Uma grande parte das publicações bibliográficas realizadas no Brasil é proveniente dos programas de pós-graduação. Esses, por sua vez, recebem recursos de instituições de fomento como a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) que utilizam, dentre outras ferramentas, os dados bibliométricos dos pesquisadores para avaliar e distribuir suas verbas para os programas de pós-graduação e para os pesquisadores.

A avaliação dos programas de pós-graduação é uma atividade essencial que visa assegurar e manter a qualidade dos cursos de Mestrado e Doutorado no país (CAPES, 2014). Entre os critérios de avaliação, a Produção Intelectual (principalmente as publicações qualificadas do Programa por docentes permanentes) é a mais relevante na definição dos conceitos dos programas.

O CNPq inclui-se, entre as formas de financiamento, a Bolsa de Produtividade em Pesquisa (PQ), que é atribuída a pesquisadores de todas as áreas. E a escolha do pesquisador a ser contemplado pela bolsa é baseada não só na qualidade de um projeto submetido, mas, principalmente, na “qualidade” do pesquisador (WAINER & VIEIRA, 2013).

O desempenho ou a “qualidade” do pesquisador pode ser avaliada por meio da quantificação das atividades de publicação. Para Abbasi *et al.* (2011), o pesquisador terá uma alta apreciação se os trabalhos publicados por ele são citados por outros. Desta forma, o número de citações qualifica as publicações (LEHMANN *et al.*, 2006).

Para auferir o impacto das publicações vários índices vêm sendo propostos. O índice h, índice mais conhecido, foi proposto por Hirsch (2005) como uma medida que combina de uma forma simples, a quantidade de publicações e o impacto das publicações. O índice h, apesar de algumas deficiências – por exemplo, não considera uma quantidade de citações mínima para se obter um valor específico, como relatado por Egghe (2006) – é o índice geralmente usado por muitos bancos de dados acadêmicos (por exemplo, *Web of Science* e *Scopus*) para medir o desempenho dos pesquisadores.

No entanto, as diferentes medidas de desempenho científico: (como, número de publicações, número de citações, entre outros) e os vários índices disponíveis na literatura observam somente o passado dos pesquisadores, o que pode levar a uma percepção viesada do resultado futuro da pesquisa, por parte dos avaliadores. O meio onde os pesquisadores estão inseridos não está ao alcance dos indicadores. Dessa forma, não são percebidos os padrões

relacionais desse meio, as estratégias de sociabilidade a partir do modo como se mobilizam e são mobilizados, assim como a dependência entre esses padrões relacionais e atributos e comportamentos individuais (MARTINS *et al.*, 2013). Por isso, torna-se necessário agregar mais fatores no processo de avaliação de desempenho que possam preencher lacunas deixadas pelos indicadores.

Entretanto, quais outros fatores poderiam ser incorporados ao processo de avaliação de desempenho? Vários estudos, como os de Eaton *et al.* (1999) e Lee & Bozeman (2005), apontam uma forte relação entre produtividade e colaboração. Outra pesquisa, a de Hart (2000), mostra que a colaboração leva à “melhoria da qualidade das publicações”. Logo, um dos fatores a serem considerados como forma de elucidar os resultados de produtividade científica dos pesquisadores está relacionado ao modo como eles colaboram entre si.

Neste sentido, é possível analisar a influência da estrutura social na conduta e na performance dos pesquisadores por meio da abordagem de análise de redes sociais. Por exemplo, os trabalhos realizados por Abbasi & Altman (2011) e Abbasi *et al.* (2011) estudaram a correlação existente entre métricas de centralidade de ARS, obtidas por meio de uma rede de coautoria, e os índices de desempenho (índice h e índice g, respectivamente) dos pesquisadores participantes dessa rede. Os resultados encontrados nesses trabalhos apontaram uma correlação positiva entre algumas métricas de centralidade de ARS e os índices de desempenho dos pesquisadores analisados.

De outro modo, LIU *et al.* (2014) observaram que os métodos baseados em redes se concentram apenas na estrutura topológica, ou seja, no conjunto de relações que os indivíduos estabelecem por meio das suas interações uns com os outros, ignorando as características ou atributos dos nós. Portanto, para suprir essa lacuna, esses autores desenvolveram um novo método para construção de redes de coautoria, combinando o topológico com as informações dos autores (impacto das publicações) para avaliar a influência dos mesmos.

Geralmente, esses trabalhos tomam pesquisadores de determinado sistema social (grupos) para formação da rede social, como objeto de estudo. Tais sistemas sociais podem ser: campos científicos, instituições de ensino e pesquisa, ou formados por pesquisadores que publicam em uma determinada revista, entres outros, e alguns trabalhos buscam analisar o desempenho desses grupos com base na estrutura da rede. Os resultados do trabalho de caracterização dos programas de pós-graduação em Ciência da Computação no Brasil e as relações entre eles, foram analisadas por Digiampietri *et al.* (2014), os autores apontam que os programas com

maiores densidades nas redes de coautoria também possuem maiores índices produtividade em pesquisa, entre 2004 e 2009. Martins *et al.* (2014) analisou a influência da rede de coautoria na performance dos programas de pós-graduação no sistema de avaliação da CAPES, concluindo que há uma relação positiva entre a formação de grupos de coautoria e a pontuação obtida pelos programas estudados. Amorim (2014) abordou no mesmo trabalho duas análises das redes de colaboração, uma entre bolsistas de produtividade e a outra entre programas de doutorado, ambas na área de estatística no Brasil. A primeira revelou, entre outros resultados, que os bolsistas mais produtivos também são os mais centrais na rede e que as métricas de centralidade de grau e de proximidade tiveram maiores impactos sobre o número de artigos publicados por um bolsista. A segunda mostrou que de forma concomitante a evolução das redes dos programas, entre 2004 a 2012, também evoluíram nas medidas das redes, bem como no número de colaborações.

1.2 OBJETIVO GERAL

Com base na revisão apresentada, o objetivo desta dissertação é caracterizar as redes de coautoria dos pesquisadores da área de Engenharia de Produção no Brasil, tendo como base de dados a produção bibliográfica dos bolsistas de produtividade do CNPq e dos professores permanentes dos programas de pós-graduação (mestrado acadêmico e doutorado) em Engenharia de Produção.

1.3 OBJETIVO ESPECÍFICO

Os objetivos específicos são:

- Desenvolver um novo método para construção de rede social que combine estrutura topológica e informações características (ou atributos individuais) dos atores dessa rede;
- Construir uma rede de coautoria entre os pesquisadores bolsistas de produtividade do CNPq da área de Engenharia de Produção. E analisar o impacto das métricas de análise de redes sociais no nível de produtividade. Estas métricas são calculadas de três maneiras distintas: não-ponderadas, ponderadas só com peso das arestas e ponderadas com pesos das arestas e dos nós;
- Analisar se existe uma relação entre a evolução da performance dos programas de pós-graduação em Engenharia de Produção e a evolução das redes de coautoria entre

os docentes dos programas, das redes de coautoria entre esses programas de pós-graduação e das redes de coautorias entre todos os professores permanentes.

1.4 ESTRUTURA DO TRABALHO

O trabalho apresenta a seguinte estrutura:

- O Capítulo 1 abrange a descrição do tema e o contexto da pesquisa, a justificativa e os objetivos;
- O Capítulo 2 inclui o referencial teórico, abordando rede de coautoria e caracterização de redes;
- O Capítulo 3 traz a metodologia, descrevendo como os dados foram obtidos e tratados para alcançar os objetivos pretendidos.
- No Capítulo 4 um novo método que integra as informações características dos atores à estrutura da rede é desenvolvido;
- Os resultados empíricos são apresentados em dois capítulos: o Capítulo 5 analisa o impacto das métricas de ARS no nível de produtividade dos bolsistas; o Capítulo 6 analisa a evolução dos conceitos dos programas de pós-graduação em engenharia de produção e a relação com a evolução das redes de coautoria;
- O trabalho é finalizado com o Capítulo 7 que apresenta as principais conclusões frente aos objetivos pretendidos. Também trata algumas sugestões para trabalhos futuros.

2 REFERENCIAL TEÓRICO

Este capítulo apresenta uma pequena abordagem sobre redes de coautoria relatando as principais características e alguns estudos sobre o tema. Em seguida, como a rede de coautoria é caracterizada como um grafo, são levantados alguns conceitos e métricas em grafos.

2.1 REDES DE COAUTORES

A coautoria, elaboração de um mesmo artigo por dois ou mais autores, é uma forma de colaboração que implica numa relação temporal e acadêmica, onde autores compartilham ideias e recursos. Ela é, por exemplo, diferente das citações que podem ocorrer sem que os autores se conheçam ou tenham se comunicado, e se estendem ao longo do tempo.

Uma das mais famosas redes de coautoria é a do matemático Paul Erdos, o qual tem mais de 500 coautores e mais de 1400 trabalhos publicados. O papel de Erdos como colaborador foi tão significativo no campo da matemática que o número Erdos é definido para medir a proximidade com Erdos por meio da rede de coautoria (LIU *et al.*, 2014). Quem já publicou com Erdos tem um número Erdos de 1, aqueles que têm uma publicação com um coautor de Erdos têm um número Erdos de 2, e assim por diante (NEWMAN, 2001c).

Para Kumar (2015), os estudos sobre coautoria ganharam novos interesses após Newman (2001a, b, c; 2004) ter usado métodos de análise de redes sociais para investigar as características e padrões interessantes das comunidades acadêmicas. Kempe & Kleinberg (2005) também relatam o surgimento de muitas pesquisas de análise de rede de coautoria e na identificação dos autores influentes na mesma. Observa-se também que os trabalhos de Huang *et al.* (2013) e de LIU *et al.* (2014) buscaram estudar a rede de coautoria na possibilidade de avaliar o *status* do autor em um determinado campo, e assim, melhorar as relações para chegar mais perto do núcleo da comunidade por meio da identificação dos pesquisadores mais influentes.

Diferentes motivos podem levar dois ou mais pesquisadores a produzir um artigo conjuntamente. Para Gibbons *et al.* (1994), a colaboração científica é engrenada pelo aumento da complexidade do conhecimento e pela demanda de competência mais especializadas e interdisciplinares nas pesquisas. A pesquisa de Hart (2000), dos autores da literatura de biblioteconomia acadêmica, mostrou que os autores consideram "a melhoria da qualidade do artigo" como um benefício da coautoria. Outra motivação que levam os pesquisadores a colaborar é a pressão que sentem de instituições de fomento para aumentar a produção de

publicações, assim eles veem a coautoria como um meio para publicar mais artigos. Destaca-se que uma forte relação entre produtividade e colaboração foi encontrada nos estudos de Eaton *et al.* (1999).

2.2 CARACTERIZAÇÃO DE REDES DE COAUTORIA

Redes de coautoria são um tipo de rede social e, portanto, têm a análise baseada na premissa de que as relações entre atores sociais podem ser descritas mediante um grafo. Essa representação apresenta a vantagem, segundo Wasserman & Faust (1994), de permitir a utilização da Teoria dos Grafos para analisar as interações sociais, os padrões e as implicações.

2.2.1 Grafos

A solução proposta pelo matemático suíço Leonhard Euler, para o problema das sete pontes de Königsberg (região hoje conhecida por Kaliningrado, localizada em uma pequena porção da Rússia, entre a Polônia e a Lituânia), foi o primeiro registro conhecido relacionado com o que hoje em dia se chama de teoria dos grafos. A cidade de Königsberg possuía um rio, Pregel, com duas ilhas conectadas por uma ponte. As duas ilhas se ligavam ainda às margens por mais seis pontes ao todo, como aparece no desenho da época, Figura 2.1. O problema consistia em encontrar o percurso para um passeio que partisse de uma das margens e, atravessando uma única vez cada uma das sete pontes, retomasse a margem de partida.

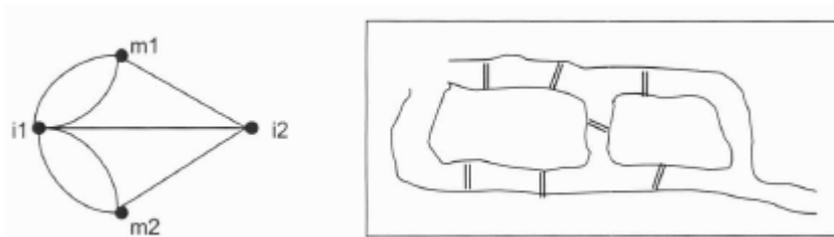


Figura 2.1 – Desenho do problema das sete pontes
Fonte: Netto & Jurkiewicz (2009)

Euler publicou um artigo em 1736, *Solutio prolemactis ad geometriam situs pertinentis*, em que provou não ser possível percorrer toda cidade de Königsberg atravessando cada ponte uma única vez. Usando, para solução de tal problema, o que hoje é conhecido como grafo, ao representar cada ponto da cidade onde se desejava chegar por um ponto e as pontes representadas por linhas, ligando os pontos. Como pode-se ver na Figura 2.1.

O problema para Euler não passou de um desafio intelectual, e não procurou desenvolvê-lo, ou achar outras aplicações. Portanto, a teoria dos grafos não foi muito estudada durante um bom tempo. Somente em 1936 com a publicação do primeiro livro sobre teoria dos grafos (*Theorie der Endlichen und Unendlichen Graphen*) de Denes König que os estudos sobre o tema receberam mais atenção.

Para formalizar a definição de grafo, são descritos abaixo alguns conceitos básicos, seguindo as ideias de Furtado (1973), Feofiloff *et al.* (2011) e Amorim (2014). A Figura 2.2 ilustra exemplos dos conceitos descritos.

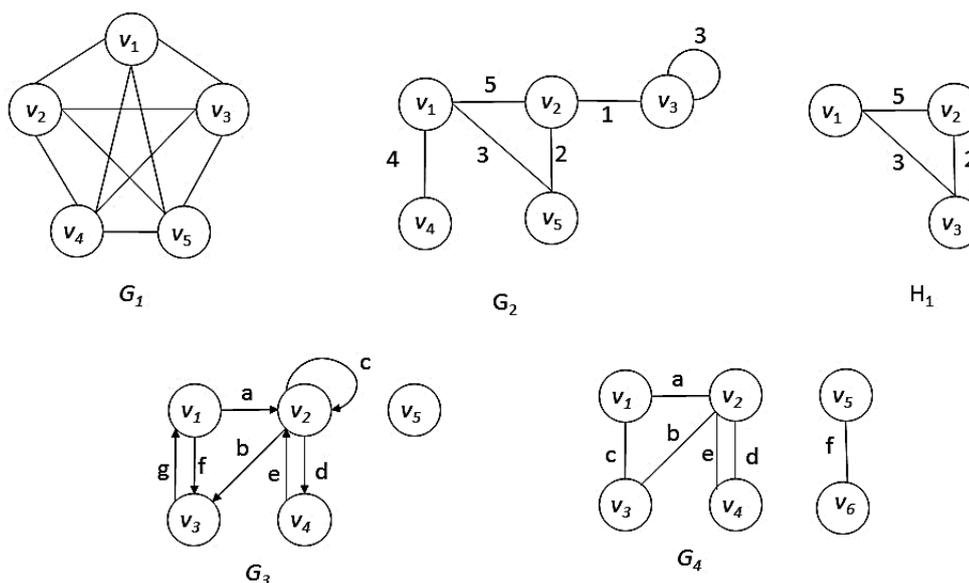


Figura 2.2 – Ilustração de exemplos de grafos
 Fonte: adaptado de Amorim (2014)

Matematicamente, tem-se que um **grafo** G é uma estrutura constituída por um conjunto finito e não vazio de **nós** (vértices), $V(G) = \{v_1, v_2 \dots v_n\}$, e um conjunto finito de **arestas** $E(G)$, formado por pares de elementos (não necessariamente ordenados ou distintos) de $V(G)$.

Ordem ou **cardinalidade** de um grafo é a quantidade de elementos do conjunto dos vértices $V(G)$, ou seja, $\#V(G)$. Na Figura 2.2, o grafo G_2 tem uma ordem igual a 5, enquanto H_1 tem ordem 3. O conjunto de arestas $E(G)$ conduz a uma relação binária em $V(G)$ que é chamada de **relação de adjacência** dos vértices de G . Uma rede com n vértices é representada por uma **matriz de adjacência** $A(G)$ com $n \times n$ elementos,

$$a(v_i, v_j) = \begin{cases} 1 & \text{se } v_i \text{ e } v_j \text{ estão conectados,} \\ 0 & \text{caso contrário.} \end{cases}$$

Uma aresta **incidente** liga os vértices v_i e v_j de um grafo, então é dito que esses vértices são **vizinhos** ou **adjacentes**. A **vizinhança** do vértice v_i é o conjunto formado por todos os vértices adjacentes a ele. Por outro lado, se uma aresta apresenta apenas uma extremidade, então é chamada de **Laço**. Em G_2 , tem-se um laço em v_3 . Se todos os pares de $E(G)$ são de elementos distintos, diz-se que o grafo não possui laços. Para um grafo direcionado (cuja descrição será dada mais adiante) existe uma distinção entre aresta exteriormente incidente (aresta que sai do vértice); e aresta interiormente incidente (aresta que entra no vértice). Um vértice que não possui aresta incidente é dito isolado. Em G_3 tem-se que b é uma aresta exteriormente incidente a v_2 e interiormente incidente a v_3 , enquanto v_5 não possui aresta incidente, sendo um vértice isolado.

Se em um grafo cada aresta tem um valor ou peso associado, esse grafo é chamado de **ponderado**, podendo ser representado por uma matriz de adjacência ponderada $W(G)$ cujas entradas não são simplesmente 0 ou 1, mas são iguais aos pesos das arestas (NEWMAN, 2004):

$$w(v_i, v_j) = (\text{peso da ligação ente } v_i \text{ e } v_j)$$

Por exemplo, o grafo H_1 pode ser representado por uma matriz de adjacência ponderada:

$$\begin{matrix} & v_1 & v_2 & v_3 \\ \begin{pmatrix} 0 & 5 & 3 \\ 5 & 0 & 2 \\ 3 & 2 & 0 \end{pmatrix} & v_1 \\ & v_2 \\ & v_3 \end{matrix}$$

Nota-se que nesse exemplo os pesos são números inteiros, porém cada elemento pode assumir qualquer valor real, inclusive negativo, segundo Newman (2004).

Arestas incidentes ao mesmo par de vértices são chamadas de **arestas múltiplas ou arestas paralelas**. A matriz de adjacência de um grafo que apresentam arestas múltiplas é similar à de um grafo ponderado, nesse caso o número de arestas múltiplas representa o peso da ligação entre esses vértices. Em G_4 , as arestas e , d são arestas múltiplas, no entanto a , b e c não são. Por outro lado, quando as arestas estão ligadas a um único vértice elas são chamadas de **arestas adjacentes**. Em G_4 , a e c são arestas adjacentes, enquanto as arestas c e d não são.

Um **Caminho** é uma sucessão de vértices e arestas, onde cada aresta liga o vértice que a precede ao vértice que a segue, não repetindo arestas. O comprimento de um caminho entre dois determinados vértices será dado pelo número de arestas presentes no caminho. No grafo G_2 , $v_4 \rightarrow v_1 \rightarrow v_5 \rightarrow v_2$ é um caminho que vai do vértice v_4 ao v_2 , sendo o comprimento desse caminho, que contém v_1 e v_5 , igual a 3.

Um **caminho geodésico**, ou caminho mais curto, é o caminho mais curto entre dois vértices. O comprimento do caminho geodésico, $d(v_i, v_j)$, também chamado de **distância geodésica** ou distância mais curta, é, portanto, a mais curta distância na rede, entre esses dois vértices. Dado um caminho $c = (v_0, v_1, v_2, \dots, v_k)$ entre os vértices v_i e v_j , o comprimento desse caminho é dado por $d_c = k$. Seja $C(v_i, v_j)$, o conjunto de todos os caminhos entre os vértices v_i e v_j , então a distância geodésica é definida por:

$$d(v_i, v_j) = \min\{d_c : c \in C(v_i, v_j)\}. \quad (2.1)$$

Em caso de redes ponderadas, o comprimento de um caminho $c = (v_0, v_1, v_2, \dots, v_k)$ entre os vértices v_i e v_j , é dado por:

$$d_c^w = \left(\frac{1}{w(v_0, v_1)} + \frac{1}{w(v_1, v_2)} + \dots + \frac{1}{w(v_{k-1}, v_k)} \right),$$

onde $w(v_j, v_k)$ é o peso da matriz adjacente ponderada. E a distância geodésica ponderada é dada por:

$$d^w(v_1, v_2) = \min\{d_c^w : c \in C(v_i, v_j)\}. \quad (2.2)$$

Diferentes tipos de grafos, de acordo com os tipos de ligações, podem existir entre os vértices. Um **grafo simples**, como em H_1 , é um grafo que não possui laços, nem arestas múltiplas. Um grafo cujas arestas possuem uma direção é chamado de **direcionado ou dígrafo**, e as arestas são chamadas de arcos. Se as arestas não possuem direção, o grafo é dito **não direcionado**. O grafo G_1 é não direcionado, enquanto G_3 é um dígrafo.

Um grafo é **completo** quando todos os pares de vértices distintos são adjacentes. Para grafos direcionados, basta que a ligação entre cada par de vértices ocorra em pelo menos um sentido. Um grafo onde todos os vértices têm o mesmo grau (número de arestas) é chamado de **regular**. G_1 é um grafo completo e regular.

Um grafo é **bipartido** se os vértices são particionados em dois subconjuntos V_1 e V_2 não vazios, disjuntos e independentes e que nenhuma aresta seja incidente a dois vértices do mesmo subconjunto, ou seja, dois vértices do mesmo subconjunto não são adjacentes. Na Figura 2.3, G_5 é um grafo bipartido. Grafo **conexo** é aquele que entre qualquer par de vértices existe sempre um caminho que os une, caso contrário, o grafo é **desconexo**. Os grafos G_1, G_2 e H_1 são conexos, enquanto G_3 e G_4 são desconexos.

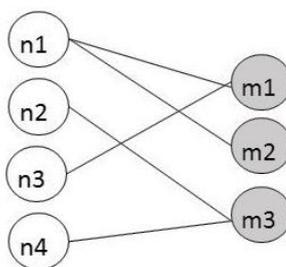


Figura 2.3 – Exemplo de grafo bipartido (G_5)
 Fonte: adaptado de Amorim (2014)

Um **Subgrafo** de um grafo G é qualquer grafo H tal que $V(H) \subseteq V(G)$ e $A(H) \subseteq A(G)$. Em outras palavras um subgrafo de G é aquele cujo conjunto de vértices e o conjunto de arestas são subconjuntos do conjunto de vértices e arestas de G , respectivamente. Um subgrafo H de G é **próprio** se $V(H) \neq V(G)$ ou $A(H) \neq A(G)$. Um subgrafo completo é dito ser um **clique**. H_1 , na Figura 2.2, é um subgrafo completo de G_2 .

Um subgrafo conexo H de um grafo G é **maximal** se H não é subgrafo próprio de nenhum subgrafo conexo de G . Dado um grafo desconexo, uma **componente conexa** é um subgrafo conexo e maximal, ou seja, não existe caminho entre um nó pertencente ao subgrafo e outro nó não pertencente ao subgrafo. O subgrafo H_1 é uma componente conexa de G_2 . A **componente gigante** de um grafo G é a componente conexa que contém o maior número de nós.

2.2.2 Métricas

As métricas podem ser divididas em **globais**, que descrevem as características sobre todo o grafo e **individuais**, que estão relacionadas com a análise das propriedades individuais de atores (nós ou vértices) da rede.

O **número de arestas** como o próprio nome diz, refere-se à cardinalidade do conjunto das arestas ($E(G)$), denotada por $\#E(G)$. Em uma rede de coautoria o número de arestas reflete o número de ligações entre os pesquisadores e é comumente utilizada para avaliar a evolução temporal do número de coautoria em redes obtidas em diferentes períodos de tempo (MENA-CHALCO *et al.*, 2012).

A **densidade** mede o quão perto o grafo está de ser completo. Ou seja, a relação entre o total de ligações existentes no grafo e o total de ligações caso todos os pontos fossem adjacentes uns aos outros. Para um grafo não direcionado com n nós e sem laços, a densidade é definida como:

$$Dens(G) = \frac{2 \times (\#E(G))}{n \times (n - 1)} \quad (2.3)$$

A maior distância geodésica entre qualquer par de vértices é chamada de **diâmetro** de um grafo, podendo variar de um mínimo de 1, se o grafo for completo, a um máximo de $n - 1$, onde n é a ordem do grafo. Seja $d(v_i, v_j)$ o comprimento do menor caminho que liga o vértice v_i ao vértice v_j , então o diâmetro do grafo conexo G será dado por:

$$Dim(G) = \max_{\{v_i, v_j \in V(G)\}} d(v_i, v_j) \quad (2.4)$$

O diâmetro, em grafos desconexos, é dado pelo maior valor encontrado entre os diâmetros das componentes conexas. Em caso de redes ponderadas, é utilizada a distância geodésica ponderada, $d^w(v_1, v_2)$.

Conhecida também como componente gigante, o **tamanho da maior componente conexa**, refere-se à cardinalidade da componente conexa com maior número de nós. Por meio da componente gigante em uma rede de coautoria é permitido investigar o comportamento do maior grupo colaborativo e identificar a quantidade de autores que participa desse grupo. É possível que na rede de coautoria, assim como em qualquer outra, existam mais de uma componente conexa, evidenciando a presença de grupos que publicam isoladamente (MENA-CHALCO *et al.*, 2012).

Dado um vértice v_i , a **excentricidade ou afastamento**, $e(v_i)$, é a distância máxima a partir dele para qualquer outro vértice do grafo. O relacionamento de um vértice com outros vértices é melhor quanto menor for a excentricidade. A excentricidade de v_i será dada por:

$$e(v_i) = \max_{v_j \in V(G)} d(v_i, v_j) \quad (2.5)$$

O diâmetro, como definido acima, é igual à excentricidade máxima, enquanto a excentricidade mínima é chamada de **raio**. O centro de um grafo G é formado pelos vértices com excentricidade mínima, enquanto os vértices com excentricidade máxima estão localizados na periferia. No caso de redes que apresentam mais de uma componente, a excentricidade de cada vértice é calculada levando em consideração a componente à qual pertence. Em caso de redes ponderadas, a excentricidade também é calculada utilizando a distância geodésica ponderada, $d^w(v_i, v_j)$.

O **grau do vértice** é o número de arestas incidente a um vértice. O grau do vértice, em grafos direcionados, é a soma das arestas exteriormente incidentes e interiormente incidentes a um vértice (MENA-CHALCO *et al.* 2012).

Como o grau de um vértice é calculado em termos do número de vértices adjacentes, o grau pode ser considerado como uma métrica de centralidade local. Portanto, a **centralidade de grau** de um vértice v_i em um grafo G não ponderado é:

$$C_d(v_i) = \sum_{j=1}^n a(v_i, v_j) \quad (2.6)$$

Caso a rede seja ponderada, a centralidade de grau é definida por:

$$C_d^w(v_i) = \sum_{j=1}^n w(v_i, v_j), \quad (2.7)$$

onde $a(v_i, v_j)$ e $w(v_i, v_j)$ são elementos da matriz de adjacências $A(G)$ e $W(G)$, respectivamente.

Para uma métrica mais genérica, afim de comparar a centralidade de grau em redes com diferentes tamanhos, foi proposta por Freeman (1979) uma métrica relativa. Esta métrica normaliza o número real de ligações pelo número máximo de ligações que se poderia ter, $n - 1$. Assim a **centralidade relativa de grau**, em uma rede não ponderada, de um vértice v_i é dada por:

$$C'_d(v_i) = \frac{C_d(v_i)}{(n - 1)} \quad (2.8)$$

A centralidade de grau é a forma mais simples e fácil de medir a influência de um nó (ABBASI *et al.*, 2012; LIU *et al.*, 2005). Em uma rede de coautoria, essa métrica identifica os autores mais ativos e populares (ABBASI *et al.*, 2011; ANASTASIOS *et al.*, 2012; FREEMAN 1979).

Outra métrica para analisar um nó na rede partiu da teoria da “força das ligações fortes” de Krackhardt (1992). Para o cálculo da força da ligação de um vértice divide-se o grau ponderado, $C_d^w(v_i)$, pela centralidade de grau, $C_d(v_i)$.

$$TS(i) = \frac{C_d^w(v_i)}{C_d(v_i)} \quad (2.9)$$

Sendo assim, $TS(v_i)$ representa o **peso médio** das ligações de um nó. A teoria da “força das ligações fortes” está em contraposição à teoria da “força das ligações fracas”, de Granovetter (1973). As ligações fracas proporcionam aos indivíduos da rede acesso à informação e recursos além daqueles disponíveis em seus próprios círculos sociais (vizinhos), ou seja, funcionam como uma ponte; porém as ligações fortes conduzem ao recebimento de conhecimentos úteis e são geralmente mais acessíveis.

A distância média de um determinado nó inicial para todos os demais vértices da rede é calculada pela **centralidade de proximidade**. Seja $d(v_i, v_j)$ a distância geodésica entre os vértices v_i e v_j , como definido na Equação (2.1). Define-se a centralidade de proximidade do vértice v_i , $C_c(v_i)$, como:

$$C_c(v_i) = \frac{1}{\sum_j d(v_i, v_j)} \quad (2.10)$$

Os vértices mais centrais da rede de acordo com essa métrica são aqueles que possuem uma distância menor dos outros vértices. Para redes ponderadas, $d^w(v_i, v_j)$ segue a equação (2.2), e a centralidade de proximidade ponderada é dada por:

$$C_c^w(v_i) = \frac{1}{\sum_j d^w(v_i, v_j)} \quad (2.11)$$

A métrica pode ser normalizada utilizando a maior distância possível entre quaisquer dois nós de uma rede de n nós. Este valor é $n - 1$, então, a **centralidade relativa de proximidade** do vértice v_i , em uma rede não ponderada, é dada por:

$$C'_c(v_i) = C_c(v_i)(n - 1) \quad (2.12)$$

Um nó que está, em média, numa posição mais próxima dos outros nós pode obter informações de maneira mais eficiente, ou seja, a métrica de proximidade está relacionada a independência e eficiência na comunicação com outros nós (FREEMAN, 1979). Neste sentido, na rede de coautoria, um autor com uma centralidade de proximidade alta pode indicar uma maior possibilidade de estabelecer parcerias de publicação (SOUZA *et al.*, 2012).

A **centralidade de intermediação** do vértice v_i é a soma para todos os pares de nós diferentes v_j da razão entre o número de caminhos mais curtos entre o dado par de nós que passam por v_i , dividido pelo número total de caminhos mais curtos entre o dado par de nós. A centralidade de intermediação $C_b(v_i)$ do vértice v_i é dada por:

$$C_b(v_i) = \sum_{j,k} \frac{g(v_j, v_i, v_k)}{g(v_j, v_k)}, j \neq k \neq i, \quad (2.13)$$

onde $g(v_j, v_k)$ é o número de caminhos mais curtos do nó v_j para o nó v_k e $g(v_j, v_i, v_k)$ é o número de caminhos mais curtos do nó v_j para o nó v_k passando por v_i .

Em uma rede ponderada a centralidade de intermediação é dada por:

$$C_b^w(v_i) = \sum_{j,k} \frac{g^w(v_j, v_i, v_k)}{g^w(v_j, v_k)}, \quad (2.14)$$

em que $g^w(v_j, v_k)$ é o número de caminhos mais curtos do nó v_j para o nó v_k e $g^w(v_j, v_i, v_k)$ é o número de caminhos mais curtos do nó v_j para o nó v_k passando por v_i , considerando a distância ponderada, $d^w(v_i, v_j)$.

Da mesma forma que a centralidade de grau, existe a necessidade de normalizar essa métrica. A centralidade de intermediação do vértice v_i é normalizada pelo número máximo possível de caminhos mais curtos, excluindo o vértice v_i . Dado uma rede não direcionada, o máximo é: $[(n-1)(n-2)]/2 = (n^2 - 3n + 2)/2$. Assim a **centralidade relativa de intermediação**, em uma rede não ponderada, é dada por:

$$C'_b(v_i) = \frac{2 \times C_b(v_i)}{(n^2 - 3n + 2)} \quad (2.15)$$

A intermediação é um indicador do potencial de um nó de desempenhar um papel de “intermediador” ou “porteiro”, (FREEMAN, 1979; ABBASI *et al.*, 2012), podendo controlar com maior frequência o fluxo de informação na rede. Neste sentido, 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 autor (SOUZA, 2012).

Uma métrica de importância do nó na rede baseada nas conexões, a **centralidade de autovetor** é baseado na ideia que um determinado nó terá centralidade alta se estiver conectado a nós com posições centrais na rede (BONACICH, 1987), ou seja, a centralidade de um nó não depende apenas do número de nós adjacente mais também da centralidade desses nós. Seja λ uma constante, $n = \# V(G)$ e $a(v_i, v_j)$ elementos da matriz de adjacência associada ao i -ésimo vértice da rede como definida anteriormente, então a centralidade de autovetor $C_e(v_i)$, será dada por:

$$C_e(v_i) = \frac{1}{\lambda} \sum_{j=1}^n a(v_i, v_j) C_e(v_j) \quad (2.16)$$

Usando a notação vetorial, seja $X = (C_e(1), C_e(2) \dots C_e(n))$ o vetor de centralidade, pode-se reescrever a equação acima como, $\lambda X = AX$.

Ao assumir que a centralidade são valores não negativos (usando o teorema de Perron-Frobenius) pode-se mostrar que λ é o maior autovalor da matriz de adjacência, sendo X o autovetor correspondente (JACKSON, 2008).

Para o caso de redes ponderadas, a centralidade de autovetor é definida por:

$$C_e^w(v_i) = \frac{1}{\lambda} \sum_{j=1}^n w(v_i, v_j) C_e^w(v_j) \quad (2.17)$$

Para normalizar a centralidade de autovetor é usada a norma Euclidiana (ABBASI *et al.*, 2011). A **centralidade relativa de autovetor**, em redes não ponderadas, é dada por:

$$C'_e(v_i) = \frac{C_e(v_i)}{\sqrt{\sum_{j=1}^n C_e(v_j)^2}} \quad (2.18)$$

O **Coefficiente de cluster (agrupamento)** indica como os nós estão inseridos na vizinhança e em conjunto com o valor médio de caminho mais curto, pode identificar um efeito “*small-world*” (redes com coeficiente de cluster grande e distância relativamente curta entre os nós). O coeficiente de agrupamento de um vértice v_i é a razão entre o número de triângulos que contém o vértice v_i e o número de arestas possíveis entre os vizinhos. Sendo $NT(v_i)$ o número de triângulos (consiste em três nós conectados por três ligações) contendo o nó v_i então tem-se o coeficiente de *cluster* como:

$$CL(v_i) = \frac{NT(v_i)}{\binom{C_d(v_i)}{2}} \quad (2.19)$$

O coeficiente de cluster ponderado foi proposto por Onnela *et al.* (2005) *apud* Saramäki *et al.* (2007), e é dado por:

$$CL^w(v_i) = \frac{1}{C_d(v_i)(C_d(v_i) - 1)} \sum_{j,k} (\widehat{w}(v_i, v_j) \widehat{w}(v_i, v_k) \widehat{w}(v_j, v_k))^{1/3}, \quad (2.20)$$

onde os pesos das arestas são normalizados pelo peso máximo da rede, $\widehat{w}(v_i, v_j) = w(v_i, v_j) / \max_{i,j \in V(G)} (w(v_i, v_j))$ e a contribuição de cada triângulo depende de todos os pesos das arestas.

O coeficiente médio de agrupamento é o valor médio dos coeficientes individuais, e é dado por:

$$CL(G) = \frac{1}{n} \sum_i CL(v_i) \quad (2.21)$$

O coeficiente de *cluster* $CL(G)$ para a rede de coautoria refere-se à probabilidade de que quaisquer dois colaboradores de um pesquisador tenham colaborado entre si (ONEL *et al.*, 2011). No caso individual, o coeficiente de agrupamento de um determinado autor indica o quanto os colaboradores dele estão colaborando entre si.

O **PageRank** é um método de ranking das páginas web, medindo de maneira efetiva o interesse dos navegadores e a atenção dedicada por eles. PageRank trabalha contando o número e a qualidade dos links de uma página, afim de determinar uma estimativa aproximada de quão influente uma página web é (LIU *et al.*, 2014). Seja T uma página web. O T_i representa uma das páginas da web. Assim, o PageRank é definido como função:

$$PR(T_A) = (1 - \delta) + \delta \left(\frac{PR(T_1)}{C(T_1)} + \dots + \frac{PR(T_n)}{C(T_n)} \right), \quad (2.22)$$

em que $PR(T_A)$ é o PageRank da página T_A , $PR(T_i)$ é o PageRank das páginas T_i que se conectam a página T_A , $C(T_i)$ é o número de links de saída na página T_i e δ é um fator de amortecimento (probabilidade de que qualquer uma das páginas seja selecionada), que pode estar situado entre 0 e 1.

No estudo de Santos (2014) sobre redes de coautoria, foi proposta uma métrica para avaliar o benefício ou **utilidade** para um determinado autor de pertencer a uma certa estrutura de rede. De acordo com essa métrica, considera-se que cada autor possui uma quantidade finita de tempo para se dedicar a colaborações científicas, e que cada autor recebe uma utilidade proveniente de um autor adjacente que é igual à proporção de trabalhos que o coautor tem com ele, mais a formação de uma sinergia, que é dada pelo produto da dedicação de cada autor a colaboração. De maneira formal, a utilidade $U^w(v_i)$ de um dado autor v_i em um determinado grafo G é dada por:

$$U^w(v_i) = \sum_j \left(\frac{w(v_i, v_j)}{C_d^w(v_i)} + \frac{w(v_i, v_j)}{C_d^w(v_j)} + \frac{w(v_i, v_j)^2}{C_d^w(v_i)C_d^w(v_j)} \right), \quad (2.23)$$

onde $w(v_i, v_j)$ corresponde ao número total de trabalhos entre os autores v_i e v_j e $C_d^w(v_i)$ e $C_d^w(v_j)$ os graus ponderados desses autores, respectivamente.

A utilidade desenvolvida por Santos (2014) foi baseada no modelo original de utilidade de Jackson & Wolinsky (1996). Esse modelo considera apenas se o autor está ou não ligado a outro autor, desconsiderando a quantidade de trabalhos desenvolvidos em conjunto. Desta forma a utilidade de um dado autor v_i em um determinado grafo G é dado por:

$$U(v_i) = \sum_j \left(\frac{1}{C_d(v_i)} + \frac{1}{C_d(v_j)} + \frac{1}{C_d(v_i)C_d(v_j)} \right), \quad (2.24)$$

em que $C_d(v_i)$ e $C_d(v_j)$ são as centralidades de grau dos vértices v_i e v_j , respectivamente.

Uma métrica que analisa o **grau de externalidade** e **internalidade** das relações (heterofilia e homofilia, respectivamente) em uma rede em que os atores são rotulados ou

particionados por uma ou várias de suas características. Por exemplo, em uma rede de coautoria pode-se particionar os autores pelas instituições a que pertencem, pelas unidades federativas ou regiões onde estão localizadas tais instituições, pelas áreas de atuações, entre outros. Ligações externas são relações entre atores de diferentes células da partição, enquanto ligações internas são aquelas entre atores membros da mesma célula da partição.

Para esse fim, Krackhardt & Stern, (1988) propuseram uma métrica chamada *E-I index* que avalia as tendências das ligações entre membros das células da partição, comparando os números de ligações dentro e fora das células da partição (Hanneman & Riddle, 2005).

$$E - I \text{ index} = \frac{EL - EI}{EL + EI} \quad (2.25)$$

em que *EL* é o número de relações externas e *EI* o número de relações internas.

O *E-I index* possui valores que variam de -1 a +1. Valores próximos de +1 indicam maior tendência de relacionamento entre atores de células diferentes da partição (heterofilia), enquanto valores mais próximos de -1 revelam propensão dos atores a se relacionarem internamente à célula da partição de que fazem parte (homofilia). Se as ligações estão divididas igualmente, o *E-I index* é igual à zero.

Para uma rede ponderada o *E-I index* pode ser calculado usando o peso das arestas, dessa forma *EL* é a soma dos pesos das arestas que ligam diferentes células da partição e *EI* é a soma dos pesos das arestas que ligam atores de uma mesma célula da partição.

3 METODOLOGIA

Este capítulo aborda a metodologia adotada para obtenção dos dados bibliográficos e estabelecimento das relações de coautoria. Comenta sobre a análise de rede social como um método de estudo da estrutura de rede social e alguns softwares de análise e visualização de redes. Apresenta também o teste de correlação de Spearman, teste adotado nas análises dos dados.

3.1 OBTENÇÃO DE DADOS

Os dados acadêmicos apresentados neste trabalho foram obtidos a partir da plataforma Lattes, a qual retrata a experiência do CNPq na integração de bases de dados de currículos. Considerado um padrão nacional de avaliação, o chamado "Currículo Lattes" apresenta um histórico das atividades científicas, acadêmicas e profissionais dos pesquisadores cadastrados na Plataforma Lattes (AMORIN, 2003) (PEREZ-CERVANTES *et al.*, 2012) (MENA-CHALCO & CESAR-JR, 2011).

O currículo Lattes, hoje adotado pela maioria das instituições de fomento, universidades e institutos de pesquisa do País, constitui-se cada vez mais abrangente e confiável (CNPq, 2015). Torna-se uma fonte extremamente rica em informações para pesquisadores interessados em investigar e compreender o comportamento de diversos grupos de pesquisa de grande porte, sejam esses relacionados a áreas do conhecimento ou grandes áreas do conhecimento, e até o banco de dados inteiro dos currículos Lattes (MENA-CHALCO *et al.*, 2012).

Para analisar as produções bibliográficas de grandes grupos de pesquisadores, como os apresentados neste trabalho requereria um esforço exaustivo e suscetível a falhas, caso fosse realizado de forma manual. Nesse caso, foi utilizado o scriptLattes, ferramenta de código livre projetada para a extração e compilação automática de produções bibliográficas, técnicas e artísticas, entre outras. Além disso, esta ferramenta possibilita a geração de grafo de colaboração e de mapa de geolocalização de um conjunto de pesquisadores cadastrados na plataforma Lattes (MENA-CHALCO & CESAR-JR, 2009).

Os currículos Lattes de um conjunto de pesquisadores são coletados automaticamente em formato HTML pelo scriptLattes, que compila as listas de produções, tratando apropriadamente as produções duplicadas e similares. São geradas páginas HTML com listas de produções e orientações separadas por tipo e colocadas em ordem cronológica invertida. Para isso, a ferramenta utiliza como entrada um arquivo em formato de texto, onde cada membro de um

grupo está representado por uma linha do arquivo de texto, contendo obrigatoriamente o ID Lattes do membro (código de 16 dígitos que o CNPq utiliza como identificador de cada currículo Lattes), seguido, de maneira opcional, pelo nome do membro, o período de permanência no grupo, ou seja, os anos em que cada membro permaneceu no grupo, e um rótulo ou etiqueta que identifica uma característica ou informação do membro. Abaixo um exemplo de como cada linha deve ser preenchida:

7258792374108591, Ricardo Lopes de Andrade, 2010-2011, Aluno

O scriptLattes utiliza também como entrada um arquivo de configurações, onde é possível limitar o tipo de atividade acadêmica a ser extraída e o intervalo de tempo desta.

Por meio de um *Paser* (programa baseado em análise textual, que permite extrair regiões ou trechos específicos de texto) são automaticamente extraídas as informações de cada currículo Lattes, indicado no arquivo de entrada para a ferramenta. As listas completas de produções acadêmicas são limitadas pelos anos indicados no tempo de permanência de cada membro do grupo, e os tipos de produções científicas consideradas pelo scriptLattes são os registrados na plataforma Lattes.

Com o intuito de evitar que uma mesma publicação seja considerada mais de uma vez, a ferramenta mantém um módulo de tratamento de redundâncias que permite a detecção de produções acadêmicas iguais ou similares, e é realizada por meio da comparação dois a dois de todas as produções do conjunto de dados separados por ano e tipo de produção. A comparação de duas produções quaisquer é realizada por meio de um casamento aproximado entre os títulos associados a cada cadastro e são consideradas iguais se a porcentagem de similaridade entre os títulos for maior a uma determinada porcentagem. Tal similaridade baseia-se na distância proposta por Levenshtein, que é obtida pelo número mínimo de operações (inserções, eliminações ou substituições de caracteres). Neste trabalho, dois títulos foram considerados equivalentes se ambos são pelo menos 80% similares.

As produções duplicadas são usadas para detectar colaboração entre os membros do grupo: dois ou mais membros são considerados como colaboradores se existe uma produção comum entre eles.

3.2 ANÁLISE DE REDES SOCIAIS

A análise de redes sociais (ARS) tem uma história que remota, pelo menos, meio século e produziu muitos resultados a respeito da influência social, grupos sociais, a desigualdade, a

propagação de doenças, a comunicação de informações, e, na verdade quase todos os tópicos que têm interessado a sociologia do século XX (NEWMAN, 2001c).

A análise de redes constitui um novo paradigma na pesquisa sobre a estrutura social, segundo Marteleto (2001, 2007), e é o método para realizar uma análise estrutural cuja finalidade é mostrar que a forma da rede é explicativa dos fenômenos analisados. Essa análise baseia-se na premissa de que as relações entre os atores sociais podem ser descritas por um grafo. Os nós do grafo representam os atores sociais e as arestas conectam pares de nós e, portanto, representam as interações sociais. Tal representação permite aos pesquisadores a análise do que seria considerado um problema inerentemente evasivo e pouco compreendido: o emaranhado de interações sociais (WASSERMAN & FAUST, 1994). Ainda segundo Wasserman & Faust (1994), um conjunto desenvolvido de algoritmos matemáticos são aplicados para a análise e visualização de redes.

Recursos computacionais, para a análise das conexões existentes entre os atores têm auxiliado tanto nos cálculos de métricas de redes, quanto na visualização das mesmas. Alguns programas utilizados para ARS são: UCINET, VISONE, MultiNet, NetDraw, Netminer e SocNetV (AZEVEDO & RODRIGUEZ, 2010), Gephi (BASTIAN *et al.*, 2009) e NetworkX (ALBERT & BARABÁSI, 2002). Neste trabalho serão utilizados o Gephi, o UNICET e o NetworkX. Também foi utilizado o software R para o cálculo da utilidade (SANTOS, 2014) de cada bolsista, usando o código desenvolvido por Amorim (2014). R é uma linguagem e ambiente de desenvolvimento integrado, para cálculos estatísticos e gráficos.

2.2.1 Gephi

O Gephi é um software de código aberto, usado para exploração e compressão de redes. Inicialmente, foi desenvolvido por estudantes da *University of Technology of Compiègne* (UTC) na França. Os módulos desenvolvidos podem importar, visualizar, tomar características espaciais, filtrar, manipular e exportar todos os tipos de redes. O software oferece um conjunto de métricas para análise de redes divididas em três níveis: visão geral da rede, visão geral dos nós e visão geral das arestas, sendo possível calcular as métricas mais comuns de análise de redes. Tais como, grau, diâmetro, densidade, coeficiente de *cluster*, modularidade, componentes conectados, excentricidade e métricas de centralidade.

2.2.2 NetworkX

O NetworkX é um pacote de software, de código aberto, na linguagem Python para a criação, manipulação e estudo da estrutura, dinâmica e funções de redes complexas. O

NetworkX destina-se a fornecer ferramentas para o estudo da estrutura e dinâmica das redes sociais, biológicas e de infraestrutura. Possui uma interface de programação padrão e uma implementação gráfica adequada para muitas aplicações, num ambiente de desenvolvimento rápido para colaboração e projetos multidisciplinares.

2.3.3 UNICET

Desenvolvidos por Steve Borgatti, Everett e Martin Freeman Lin (2002), o UNICET é um software de análise de redes sociais distribuído pela Analytic Technologies. E trabalha em conjunto com outro programa chamado NETDRAW, que desenha e visualiza diagramas de redes sociais. O UCINET é um pacote completo para a análise de rede social, podendo ler e gravar uma infinidade de arquivos de texto formatado de forma diferente, bem como arquivos de Excel.

3.3 METODOLOGIA PARA ANÁLISE DE CORRELAÇÃO

A análise de correlação é usada para descrever a força e a direção da relação linear entre duas variáveis. Para analisar as correlações entre as métricas de ARS com o nível de produtividade dos bolsistas PQ e com os conceitos dos programas de pós-graduação foi utilizada a Correlação de Postos de Spearman.

A Correlação de Postos de Spearman é um teste estatístico não-paramétrico proposto por Spearman (1904) como uma medida da força de associação entre duas variáveis (Hauke e Kossowski, 2011). A correlação de Spearman é usada quando a distribuição dos dados faz com que a correlação de Pearson seja inadequada. Ela avalia o quão bem uma função monotônica arbitrária pode descrever a relação entre duas variáveis, sem fazer quaisquer suposições sobre a frequência de distribuição das variáveis. Ao contrário da correlação de Pearson, não requer que as variáveis sejam medidas em escalas de intervalo, podendo ser usado para as variáveis medidas no nível ordinal (Hauke e Kossowski, 2011).

O coeficiente de correlação de Spearman, designado por “rho” e representado por ρ_s , só pode assumir valores de -1 a +1, em que o sinal (-, +) indica se existe uma correlação positiva (quando uma variável aumenta, também aumenta a outra) ou uma correlação negativa (quando uma variável aumenta, a outra diminui). O valor absoluto (ignorando o sinal) proporciona uma indicação da força da relação: uma correlação de 0 indica nenhuma relação, uma correlação de +1 indica uma perfeita correlação positiva, e um valor de -1 indica uma perfeita correlação

negativa. Para interpretar os valores (entre 0 e 1) dos coeficientes de correlação, Cohen (1988) estabeleceu as seguintes classificações:

- correlação fraca: $\rho_s = 0,10 - 0,29$;
- correlação média: $\rho_s = 0,30 - 0,49$;
- correlação forte: $\rho_s = 0,50 - 1,0$.

No próximo capítulo será desenvolvido um método que integra a importância do nó na estrutura topológica da rede.

4 MÉTODO PROPOSTO

O foco principal do estudo de redes sociais é analisar como as relações que os indivíduos estabelecem entre si (a estrutura social) influencia o comportamento e a performance deles (LAZENGA *et al.*, 2008), não se preocupando com os atributos individuais dos mesmos (MARTELETO, 2001), ou seja, os métodos baseados em rede se concentram apenas na estrutura topológica, ignorando as características dos nós (LIU *et al.*, 2014). No entanto essas características podem influenciar e sofrer influência da estrutura social.

LIU *et al.* (2014) em um estudo de redes de coautoria, desenvolveu um método que combinou a importância dos nós com a estrutura topológica da rede. Dessa forma, a posição ou o *status* de um nó na rede não depende somente da estrutura, mas também da importância individual de cada nó. A maneira utilizada pelo autor para inserir a importância dos nós foi modificar o peso das arestas por meio da Equação (4.1), que tem semelhança com a lei da gravidade.

$$S(v_i, v_j) = \frac{k \times s(v_i) \times s(v_j)}{D^2(v_i, v_j)}, \quad (4.1)$$

em que, $S(v_i, v_j)$ é o peso da aresta que liga os vértices v_i e v_j , considerando os pesos, $s(v_i)$ e $s(v_j)$ dos vértices v_i e v_j , respectivamente. k é uma constante. Os pesos dos vértices foram definidos pelo número total de citações ponderadas, por meio de uma equação desenvolvida pelos autores, que considera o tempo da publicação. A Equação (4.2) denota a distância efetiva, $D(v_i, v_j)$, entre dois vértices, v_i e v_j . Enquanto $w(v_i, v_j)$ é o peso da aresta sem considerar o peso do nó e $C_d^w(v_i)$ o número total de ligações do vértice v_i .

$$D(v_i, v_j) = 1 - \log(w(v_i, v_j)/C_d^w(v_i)). \quad (4.2)$$

Como $D(v_i, v_j)$ e $D(v_j, v_i)$ podem ser diferentes, então podem existir também dois pesos diferentes entre os vértices v_i e v_j , tratando-se, portanto, de um grafo ponderado direcionado.

Neste sentido, buscando também incorporar a estrutura topológica com as informações características dos nós em uma rede social qualquer, foi desenvolvido um novo método de ponderação, definida por:

$$Z(v_i, v_j) = w(v_i, v_j) \times \left(\frac{s(v_i) + s(v_j)}{2} \right), \quad (4.3)$$

onde $Z(v_i, v_j)$ é igual ao peso da aresta $w(v_i, v_j)$, entre os vértices v_i e v_j , combinado com os pesos desses vértices $s(v_i)$ e $s(v_j)$. Os pesos dos vértices são características mensuráveis, associadas ao tipo de relação que os conectam.

Como apresentado no Capítulo 2 uma rede é definida como um conjunto de nós, $V(G)$, um conjunto de arestas, $E(G)$, e uma matriz de adjacência, $W(G)$, que é um conjunto de pesos $w(v_i, v_j)$ associados a cada aresta que conecta um par de vértices, $v_i e v_j$. Com a incorporação dos pesos dos nós na rede, $Z(G)$ passa ser a nova matriz de adjacência e $Z(v_i, v_j)$ o conjunto de pesos.

O peso dos vértices é um número real positivo. Quando não são considerados, assume-se que eles são todos iguais a um, e, portanto, o peso $Z(v_i, v_j)$ é igual a $w(v_i, v_j)$. Observa-se que, diferente do modelo posposto por LIU *et al.* (2014), cada aresta terá somente um único peso, ou seja, $Z(v_i, v_j) = Z(v_j, v_i)$. Logo ao ser aplicada em uma rede não-direcionada, permaneceria não-direcionada.

Para calcular as métricas ponderadas com os pesos das arestas e dos nós, substitui-se nas métricas ponderadas, apresentadas no Capítulo 2, o peso das arestas $w(v_i, v_j)$ pelo peso definido na equação 4.3.

A Figura 4.1 é utilizada para comparar as posições dos vértices em relação as métricas de centralidade de grau, centralidade de proximidade e centralidade de intermediação sem considerar os pesos das arestas (NP), considerando os pesos das arestas (W) e pelo novo método de ponderação proposto (Z).

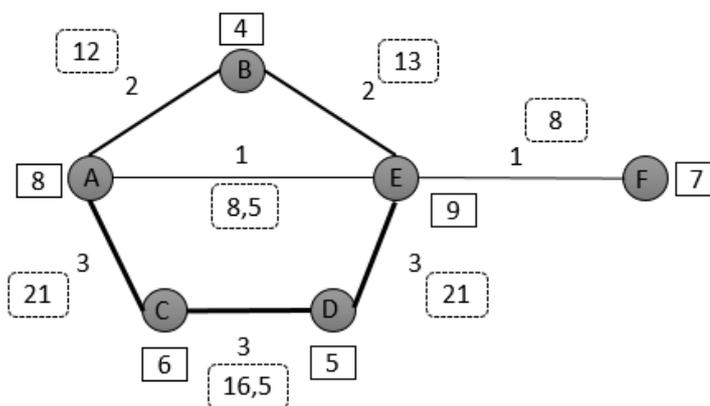


Figura 4.1 – Exemplo de aplicação do modelo

Tabela 4.1 – Posições dos vértices nas três métricas

Vértices	Grau			Proximidade			Intermediação		
	NP	W	Z	NP	W	Z	NP	W	Z
A	3	6	41,50	0,71	1,12	0,06	0,20	0,10	0,30
B	2	4	25,00	0,63	1,20	0,04	0,00	0,00	0,00
C	2	6	37,50	0,56	1,31	0,04	0,05	0,30	0,00
D	2	6	37,50	0,63	1,44	0,04	0,10	0,40	0,00
E	4	7	50,50	0,83	1,44	0,06	0,55	0,50	0,60
F	1	1	8,00	0,50	0,67	0,04	0,00	0,00	0,00

A Tabela 4.1 apresenta o resultado. Os valores que estão nos retângulos ao lado dos vértices são os pesos, já nos retângulos tracejados são os pesos das arestas calculados pelo modelo proposto. A Tabela 4.2 classifica os vértices de acordo com os resultados apresentados na Tabela 4.1.

Tabela 4.2 – Classificação dos vértices

Ranking	Grau			Proximidade			Intermediação			Importância do vértice
	NP	W	Z	NP	W	Z	NP	W	Z	
1	E	E	E	E	E	E	E	E	E	E
2	A	A	A	A	D	A	A	D	A	A
3	B	C	C	D	C	F	D	C	D	F
4	C	D	D	B	B	B	C	A	C	C
5	D	B	B	C	A	D	F	F	F	D
6	F	F	F	F	F	C	B	B	B	B

A centralidade de grau não-ponderada de um nó valoriza o número de ligações. A centralidade de grau ponderada, com os pesos das arestas, valoriza as frequências das ligações. A centralidade de grau ponderada, com os pesos das arestas e dos nós, valoriza tanto a frequência das ligações quanto a importância dos nós.

Observando a classificação na Tabela 4.2, o vértice E é o mais central independente da métrica e dos pesos das ligações. Os vértices B, C e D apresentam a mesma centralidade de grau, no entanto C apresenta maior grau ponderado com e sem a importância dos vértices, levando C a uma posição mais alta.

Na centralidade de proximidade não-ponderada as distâncias são formadas pelas somas das ligações sem considerar a frequência das mesmas. Por outro lado, na centralidade de proximidade ponderada, com os pesos das arestas, as frequências das ligações encurtam os caminhos. Portanto, caminhos antes não considerados, por serem mais longos, podem, se apresentarem ligações frequentes, serem preferíveis a caminhos com ligações menos

frequentes. Já a centralidade de proximidade ponderada, com os pesos das arestas e dos nós, considera a importância dos nós que estão nos caminhos entre dois nós.

Os vértices B e D apresentam a mesma centralidade de proximidade, mas ao considerar os pesos das arestas o vértice D obteve uma classificação melhor, ambos voltam a ter a mesma proximidade quando é considerada a importância dos vértices.

Na centralidade de intermediação, o vértice C subiu uma posição quando o peso da aresta foi considerado, porém, ao considerar os pesos dos vértices, perdeu a centralidade de intermediação, ou seja, ele não pertence mais ao caminho mais curto que liga dois nós. O contrário ocorreu com A que caiu uma posição quando considerou os pesos das arestas e voltou a subir quando foi considerado a importância dos vértices.

O peso dado ao vértice deve estar associado ao tipo de rede. Em uma rede de coautoria a importância individual do vértice pode estar relacionado ao número de publicações, número de citações ou índices de impacto. Em estudos de redes que analisam as formações de casais, Kennedy *et al.* (2015), os pesos dos vértices podem estar ligados, por exemplo, a fatores econômicos culturais. Redes que analisam as propagações de doenças os pesos dos vértices podem estar relacionados a fatores de riscos. Para redes que analisam a difusão da inovação, Leyden *et al.* (2014), os pesos dos vértices podem estar ligados ao grau de inovação, ao perfil empreendedor, poder de capital, acesso a informação, entre outras características.

O método proposto definido neste capítulo será abordado nos capítulos seguintes, na rede de coautoria entre os bolsistas de produtividade do CNPq da área de engenharia de produção e nas redes formadas pelos professores permanentes dos programas de pós-graduação em engenharia de produção. A importância dos vértices, nesse caso dos bolsistas e dos professores, será indicada pelo índice h da *Scopus*.

5 REDE DE COAUTORIA ENTRE OS BOLSISTAS DE PRODUTIVIDADE

5.1 INTRODUÇÃO

A coautoria, elaboração de uma mesma publicação por dois ou mais autores, é uma forma de colaboração. Para Hudson (1996), a coautoria é a manifestação mais formal da colaboração intelectual em pesquisa científica, e o maior ganho da colaboração é permitir uma divisão de trabalho eficiente, por intermédio da complementariedade de habilidades ou sinergia (criação conjunta de novas ideias que não seriam alcançadas de maneira individual). A consequência dessa divisão de trabalho eficiente é uma produção científica de maior qualidade ou quantidade. Resultados esses que Barnett *et al.* (1988) já haviam relatados como justificativas que levam pesquisadores a colaborarem entre si. Outros trabalhos, por exemplo Eaton *et al.* (1999) e Lee & Bozeman (2005), também apontam a produtividade como consequência da colaboração. Hart (2000) mostra que a colaboração leva à melhoria da qualidade das publicações.

A qualidade das publicações pode ser determinada, entre outros fatores, por medidas baseadas em citações. Observa-se um crescente interesse pelo estudo da análise da influência da estrutura social na atuação e no desempenho dos pesquisadores, por meio da abordagem de análise de redes sociais. Muitos desses estudos buscam correlacionar as principais métricas de centralidades com medidas baseadas nos números de citações.

O trabalho de Yan & Ding (2009) correlacionou quatro métricas de centralidades (grau, proximidade, intermediação e PageRank) com o número de citações das publicações dos autores de uma rede de coautoria. Essas métricas tiveram correlações significativas com as contagens de citações, principalmente a centralidade de intermediação.

No estudo de Abbasi & Altmann (2011) foram utilizadas as métricas de centralidades normalizadas de grau, de intermediação, de proximidade, assim como a centralidade de grau ponderada e a eficiência¹, para correlacionar com o índice h. O resultado mostrou que o índice h dos pesquisadores tinha correlações significativas positivas somente com a centralidade de grau e a eficiência. No mesmo ano Abbasi *et al.* (2011) publicaram um estudo que analisou a influência de seis métricas de ARS: centralidades de grau, de proximidade, de intermediação, de autovetor (todas essas normalizadas), média das forças das ligações e eficiência; sobre o

¹Razão entre o número total de grupos distintos, com nós diretamente conectados, ligados por um único nó central e grau desse nó.

índice g^2 . Os autores chegaram à conclusão que somente a centralidade normalizada de grau, a eficiência e a média das forças das ligações tinham influências significativas sobre o índice g .

Outro trabalho que também correlacionou as métricas de análise de redes sociais com o índice h foi publicado por Wanderley *et al.* (2014). Os autores criaram uma rede de coautoria entre pesquisadores de Ciência da Computação e calcularam as métricas de centralidades normalizadas de grau, de proximidade e de intermediação, grau ponderado e autoridade³. Somente a centralidade de intermediação e o grau ponderado tiveram correlações positivas significativas. A autoridade também apresentou correlação significativa, porém negativa.

O trabalho de Amorim (2014), em uma rede de coautoria entre bolsistas de produtividade do CNPq na área de Estatística, revelou que os bolsistas mais produtivos também são os mais centrais na rede e que as métricas de centralidades de grau e de proximidade tiveram maiores impactos sobre o número de artigos publicados por um bolsista. Outros trabalhos que analisam os impactos das redes de coautoria na performance dos pesquisadores foram apresentados em Andrade e Rêgo (2015a, 2015b) e são versões preliminares dos resultados deste capítulo.

Nos trabalhos acima citados não foram utilizadas métricas ponderadas, ou seja, métricas calculadas considerando a frequência de colaboração, com exceção do grau ponderado. Foram realizadas buscas por trabalhos com essa abordagem, mas poucos foram obtidos. O trabalho de LIU *et al.* (2015) propôs um método que insere a importância (baseada em citações) dos pesquisadores na estrutura de rede de coautoria e define o peso das arestas. Esse peso é usado no cálculo do PageRank, aplicado na rede de Erdos, para identificar os autores mais influentes.

O objetivo deste capítulo é construir uma rede de coautoria entre os pesquisadores bolsistas de produtividade do CNPq da área de Engenharia de Produção e analisar o impacto das métricas de análise de redes sociais no nível de produtividade. Essas métricas são calculadas de três maneiras distintas: não-ponderadas, ponderadas só com peso das arestas e ponderadas com pesos das arestas e dos nós (conforme o método proposto no Capítulo 4). Os pesos dos nós são os índices h dos pesquisadores e o impacto das métricas foi mensurado por meio da correlação de Spearman.

²Segundo Egghe (2006), o índice g é definido: “Dado um conjunto de artigos ordenados de forma decrescente com relação ao número de citações, o índice g é o maior valor de g em que os primeiros g artigos recebem juntos ao menos g^2 citações”

³Calculado somando a quantidade de *hubs* (nós com muitas ligações) com o qual um determinado nó está conectado.

A estrutura deste capítulo está dividida em seis subcapítulos. O primeiro é a introdução, onde foi apresentada uma revisão de trabalhos que analisaram a influência das métricas de ARS, em redes de coautoria, na performance dos pesquisadores. O Subcapítulo 5.2 aborda os critérios de atribuição de bolsas de produtividades entre os pesquisadores da área de Engenharia de Produção. O Subcapítulo 5.3 descreve a metodologia utilizada para criação da rede de coautoria. No Subcapítulo 5.4 é apresentada as métricas globais da rede de coautoria entre os bolsistas de produtividade. O impacto das métricas individuais de ARS sobre o nível de produtividade é apresentada no Subcapítulo 5.5. Por fim, no Subcapítulo 5.6, apresenta-se uma conclusão que encerra este capítulo.

5.2 SISTEMA DE AVALIAÇÃO DA PRODUTIVIDADE EM PESQUISA

No Brasil o CNPq é um dos principais órgãos de financiamento da pesquisa científica. Entre as formas de financiamento do CNPq inclui-se a Bolsa de Produtividade em Pesquisa (PQ), que é atribuída a pesquisadores de todas as áreas, baseado não só na qualidade de um projeto submetido, mas principalmente na “qualidade” do pesquisador (WAINER & VIEIRA, 2013). Segundo o CNPq (2015) a PQ é organizada em níveis, em ordem crescente: 2, 1D, 1C, 1B, 1A, sendo que os últimos quatro níveis são coletivamente chamados de “níveis 1”. Para o pesquisador se enquadrar na categoria 1, exige-se que ele tenha, no mínimo, 8 (oito) anos de doutorado por ocasião da implementação da bolsa. O enquadramento do pesquisador na categoria 2 exige que ele tenha, no mínimo, 3 (três) anos de doutorado por ocasião da implementação da bolsa. O desempenho do pesquisador é avaliado por meio de indicadores referentes ao quinquênio anterior, no caso da categoria 2, e do decênio anterior, no caso da categoria 1.

O Comitê Assessor de Engenharia de Produção e de Transporte do CNPq (CA-EP) considera importante, para ingressar e permanecer no sistema, que o pesquisador tenha uma clara participação em atividades integradas de pesquisa, ensino e extensão, que se some a uma produção científica relevante, caracterizada por regularidade e qualidade tanto em nível científico como dos meios utilizados para a sua divulgação. No Curriculum Vitae Lattes do proponente a bolsa, todos os aspectos acima devem ser evidenciados. Também é necessário um projeto de pesquisa consistente e bem elaborado. Junto ao projeto de pesquisa, o proponente deverá submeter uma súmula curricular, com no máximo 4 páginas, em que ele deverá relacionar as suas principais publicações (no máximo, dos últimos 5 anos) relacionadas ao projeto, comentando como cada uma delas se relaciona ao referido projeto. Além disso, para

cada artigo relacionado em que o proponente não é o único autor, ele deverá comentar qual foi a sua participação/contribuição no referido artigo.

Tabela 5.1 – Critérios específicos do CA-EP

Categoria	Nível	É desejável que:
Pesquisador 1	A	possua produção científica relevante ao longo de sua carreira e nos últimos anos, com publicações em periódicos internacionais de ampla circulação e rigorosa arbitragem na área de Engenharia de Produção e de Transportes; tenha regularidade na divulgação da produção científica em periódicos e congressos nacionais e internacionais de bom nível científico; tenha orientado vários mestres, doutores e pesquisadores; tenha indiscutível liderança nacional e reconhecimento internacional, tendo contribuído para formar grupos de competência com reconhecimento nacional e internacional e para desenvolver sua área no país.
	B	possua produção científica relevante ao longo de sua carreira e nos últimos anos, com publicações em periódicos internacionais de ampla circulação e rigorosa arbitragem na área de Engenharia de Produção e de Transportes; tenha regularidade na divulgação da produção científica em periódicos nacionais e congressos nacionais e internacionais de elevado nível científico, particularmente nos eventos da área de Engenharia de Produção e de Transportes; tenha orientado vários mestres e doutores, quando vinculado a instituição de ensino de pós-graduação; tenha comprovada independência e liderança na sua área de atuação, tendo contribuído para formar grupos de pesquisa com reconhecimento nacional e internacional.
	C	possua produção científica relevante ao longo de sua carreira e nos últimos anos, com publicações em periódicos internacionais de ampla circulação e rigorosa arbitragem na área de Engenharia de Produção e de Transportes; tenha regularidade na divulgação da produção científica em periódicos nacionais e congressos nacionais e internacionais de elevado nível científico, particularmente nos eventos da área de Engenharia de Produção e de Transportes; tenha orientado vários mestres e doutores, quando vinculado a instituição de ensino de pós-graduação; tenha comprovada independência e liderança na sua área de atuação, tendo contribuído para formar pesquisadores.
	D	possua produção científica relevante nos últimos dez anos, com publicações em periódicos internacionais de ampla circulação e rigorosa arbitragem na área de Engenharia de Produção e de Transportes; tenha regularidade na divulgação da produção científica em periódicos nacionais e congressos nacionais e internacionais de elevado nível científico, particularmente nos eventos da área de Engenharia de Produção e de Transportes ; tenha orientado mestres e doutores, quando vinculado a instituição de ensino de pós-graduação; tenha independência intelectual e reconhecimento na sua área de atuação.
Pesquisador 2	-	possua produção científica relevante com pelo menos uma publicação em periódicos de ampla circulação e rigorosa arbitragem na área de Engenharia de Produção e de Transportes; tenha regularidade na divulgação da produção científica em congressos nacionais e internacionais de bom nível científico, particularmente nos eventos da área de Engenharia de Produção e de Transportes; tenha envolvimento no processo de formação de mestres, quando vinculado a instituição de ensino de pós-graduação.

Fonte: Site do CNPq em Critérios de Julgamento – CA-EP

O CA-EP ressalta que a avaliação não se baseia somente em elementos estritamente numéricos, na medida em que buscasse identificar o perfil do pesquisador, avaliando a

produtividade, qualidade e o equilíbrio de suas atividades. E que os critérios definidos na Tabela 5.1 são os desejáveis, porém o rigor de sua aplicação dependerá da demanda de cada avaliação.

São definidos também critérios para a progressão ou regressão do pesquisador e obedecem às recomendações da Diretoria do CNPq, sendo rigorosamente aplicados:

- para que o pesquisador progrida da categoria 2 para 1D, categoria 1 nível D, o mesmo deve apresentar uma crescente autonomia e produção científica; pressupõe uma produção científica internacional relevante;
- a partir da categoria 1C, a progressão do pesquisador estará associada a uma produção científica independente e regular e uma demonstração de liderança e reconhecimento “inter-pares”;
- a renovação não ocorrerá, caso não haja desempenho que justifique sua permanência no sistema no nível atual ou no nível imediatamente abaixo.

5.3 METODOLOGIA

Para construção da rede de coautoria entre os bolsistas de produtividade do CNPq da área de Engenharia de Produção, foi considerado como única fonte de dados as listas de artigos publicados em periódicos e aqueles aceitos para publicações entre 2005 e 2014. As seguintes etapas foram adotadas:

- Etapa 1 – identificação dos bolsistas de produtividade da área de engenharia de produção do CNPq e o nível da bolsa;
- Etapa 2 – identificação do índice h de cada bolsista de produtividade;
- Etapa 3 – identificação do currículo Lattes de cada bolsista de produtividade;
- Etapa 4 – extração das publicações de cada bolsista de produtividade;
- Etapa 5 – identificação das publicações realizadas em coautoria;
- Etapa 6 – produção da rede de coautoria;
- Etapa 7 – cálculo das métricas de ARS.

A identificação dos bolsistas de produtividade da área de engenharia de produção, Etapa 1, foi obtida no site do CNPq, onde, na busca, utilizou-se os seguintes filtros: grande área – engenharias; área – engenharia de produção; modalidade – produtividade em pesquisa e; níveis – todos. No total foram apontados, em 2 de março de 2015, 145 pesquisadores bolsistas de produtividade, sendo esses utilizados na construção da rede.

O índice h dos bolsistas, Etapa 2, foi obtido na aba “Indicadores da Produção”, no site do CNPq, quando se utiliza a ferramenta de busca por currículos Lattes e clica sobre o nome do pesquisador. Nessa aba, está disponível o índice h calculado pela *Web of Science* e *Scopus*. Foi considerado o índice h da *Scopus*, apesar que em ambos são usados no cálculo, somente, os artigos publicados em periódicos e registrados no currículo Lattes, os quais foram utilizados para a formação da rede de coautoria.

A Etapa 3, identificação do currículo Lattes, foi realizada em paralelo com a Etapa 1, porque ao identificar os bolsistas de produtividade, também se identificava os IDs Lattes. Na Etapa 4 e na Etapa 5 foi utilizado o scriptLattes, cuja descrição se encontra no Subcapítulo 2.1. No arquivo texto, fonte de entrada para a ferramenta, incluía em cada linha nesta ordem e separado por vírgula: o ID; o nome do bolsista; espaço, pois foi preferível por essa informação no arquivo de configurações, por ser comum a todos os bolsistas (2005 – 2014); e o nível de produtividade como rótulo.

Na Etapa 6 foi construída uma rede com as relações de coautoria encontradas na Etapa 5. O cálculo das métricas dessa rede, Etapa 7, ocorreu pelo software NetworkX e de três formas: não-ponderadas; ponderadas só com peso das arestas e; ponderadas com pesos das arestas e dos nós, definidos pelo método proposto no Capítulo 4. O NetworkX foi utilizado no cálculo das métricas, por ser possível programar a inserção ou não dos pesos.

5.4 CONHECENDO A REDE DE COAUTORIA ENTRE OS BOLSISTAS DE PRODUTIVIDADE

A rede de coautoria entre os bolsistas de produtividade do CNPq foi construída utilizando os dados bibliométricos entre o período de 2005 a 2014. Foram analisados nesse período 3.796 artigos completos publicados em periódicos e 89 aceitos para publicação totalizando 3.885 artigos. Distribuídos entre os 145 bolsistas de produtividade, uma média de 26,79 artigos para cada bolsista. Desses 10,26 foram realizados em coautoria. A Tabela 5.2 apresenta uma descrição geral do nível macro da rede.

A rede está dividida em 33 componentes, sendo a componente gigante composta por 92 vértices, representando aproximadamente 63,45% dos nós da rede; a segunda possui 8 vértices (5,52%) e 21 pesquisadores possuem componente tamanho 1, ou seja, cerca de 14,48% dos bolsistas não possuem colaboradores na rede.

Ao analisar a densidade vemos que a rede é dispersa, pois somente possui densidade igual a 0,015, ou seja, apenas estão ocorrendo 1,5% das ligações possíveis na rede. A densidade é

considerada muito baixa se comparada com a densidade máxima que é igual a 1. Porém, tal densidade pode ser justificada pelo fato de a rede ser formada somente por um grupo restrito da produção dos pesquisadores (apenas artigos publicados em periódicos e artigos aceitos para publicações, entre 2005 a 2014) e que são bolsistas de produtividade da área de engenharia de produção, desconsiderando a colaboração com outros pesquisadores.

Tabela 5.2- Descrição geral do nível macro da rede

Número de autores:	145
Número de artigos:	3.885
Artigos/autores	26,79
Autores/artigos	0,037
Número de arestas:	161
Número de componentes:	33
Números de autores na componente principal (%):	63,45
Coefficiente de <i>Cluster</i> médio:	0,293
Densidade:	0,015
Diâmetro*:	13
Raio*:	7
Distância média*:	6,00
Número de caminhos mais curtos*:	8.464

*Relativo a Componente Principal

O diâmetro da rede é igual a 13 e o raio 0, representando a excentricidade máxima e a mínima, respectivamente, sendo o raio da componente gigante igual a 7. O coeficiente de *cluster* médio é igual a 0,293, sabendo que esse coeficiente pode variar de 0 a 1, temos então que pouco menos de um terço das possíveis coautorias entre os coautores de um dado autor estão presentes na rede.

A distância média de um caminho entre um par de vértices é aproximadamente 6,00. Esse valor refere-se à componente gigante, e significa que, em média, 6,00 ligações separam dois pesquisadores quaisquer na componente. O número de caminhos mais curtos é de 8.464.

A Figura 5.1 ilustra a rede de coautoria dos bolsistas gerada pelo software Gephi, em que a espessura das arestas é proporcional ao peso (total de artigos em coautoria) e o diâmetro do nó é proporcional ao grau.

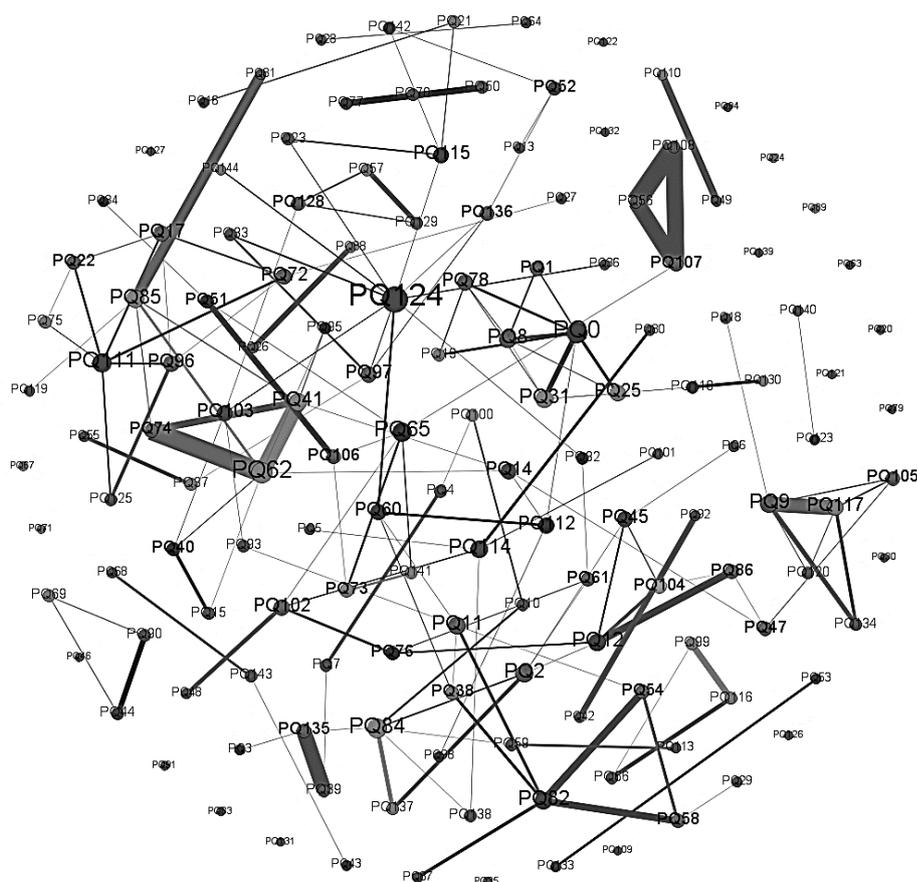


Figura 5.1 - Rede de coautoria entre os bolsistas

5.5 IMPACTO DAS MÉTRICAS DE ARS NO NÍVEL DE PRODUTIVIDADE

Com a rede descrita no subcapítulo anterior foram calculadas as métricas de ARS de três maneiras distintas: não-ponderadas (NP), ponderadas só com os pesos das arestas (W) e ponderadas com os pesos das arestas e dos nós (Z), método proposto. Os resultados são apresentados a seguir.

5.5.1 *E-I Index*

Para analisar o grau de externalidade e internalidade dos pesquisadores foi utilizada a métrica *E-I index* e os pesquisadores foram rotulados pelos níveis das bolsas. A Figura 5.2, 5.3 e 5.4 apresentam as distribuições do *E-I index*, do *E-I index_W* (com peso das arestas) e do *E-I index_Z* (com peso das arestas e dos nós), respectivamente, nos diferentes níveis de bolsas. Na Figura 5.2, as maiores variações do *E-I index* são apresentadas pelos níveis 2 e 1D. Os níveis 1C, 1B e 1A apresentam a maior mediana. O nível 1C apresenta a menor variação quase todos

os pesquisadores desse nível têm *E-I index* igual a 1, significando que as relações são estritamente externas. Somente alguns pesquisadores dos níveis 2 e 1D possuem quantidades de relações internas, ou seja, entre pesquisadores de um mesmo nível de bolsa, superiores as relações externas.

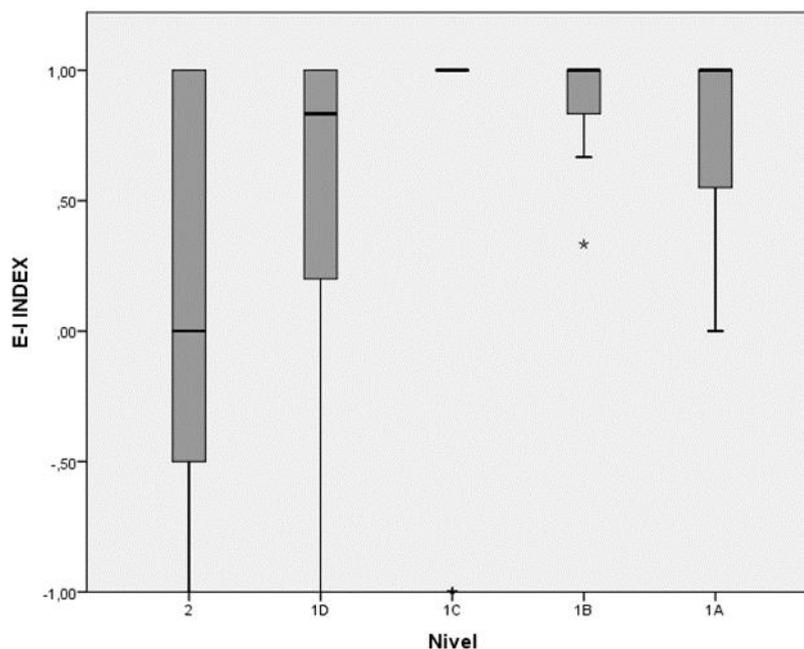


Figura 5.2 - Gráfico boxplot para o *E-I index*

Nas Figuras 5.3 e 5.4 observa-se que 50% dos pesquisadores dos níveis 1D, 1B e 1A apresentam relações totalmente externas. Os pesquisadores nível 1C não se relacionam entre si.

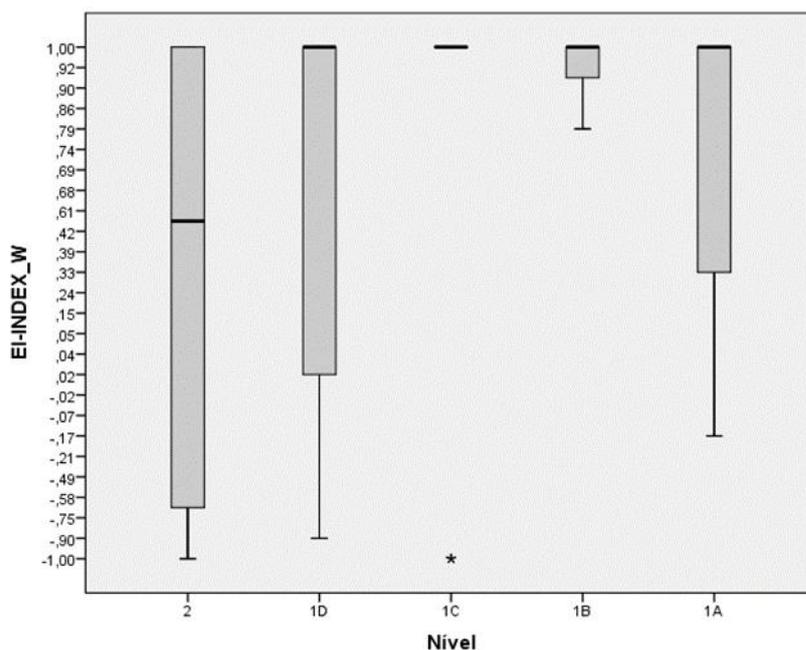


Figura 5.3 – Gráfico boxplot para o E-I index_W

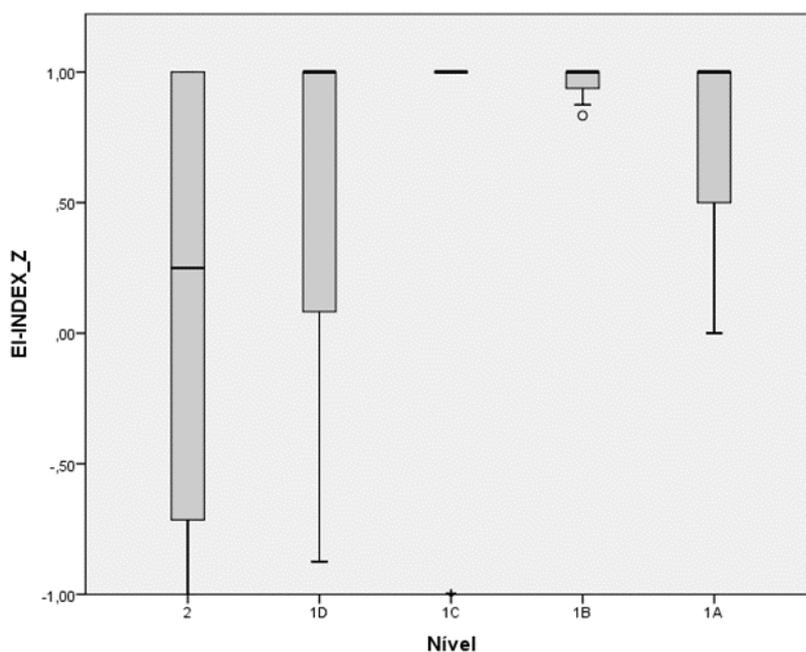


Figura 5.4 – Gráfico boxplot para o E-I index_Z

Existe uma moderada e significativa correlação entre o *E-I index* com o nível da bolsa, 0,424 (ao nível de significância de 0,01). Desta forma, pesquisadores que estabelecem relações com pesquisadores de níveis diferentes tendem a ter maiores níveis de bolsas. Se as relações levam em consideração os pesos das arestas (*E-I index_W*), a correlação com o nível da bolsa tem uma pequena queda, sendo igual a 0,354 (ao nível de significância de 0,01). Considerando

os pesos das arestas e dos nós (*E-I index_Z*), a correlação é igual a 0,390 (ao nível de significância de 0,01).

No entanto, esse resultado é tendencioso quando são observadas as distribuições dos níveis das bolsas. O cálculo do *E-I index* foi obtido somente para pesquisadores que possuem no mínimo uma relação de coautoria. Como apresentado anteriormente, foram 21 pesquisadores isolados na rede (17 possuem nível 2 e 4 possuem nível 1D). Portanto 124 pesquisadores possuíam relações de coautorias: 69 (55,6%) possuem nível 2, 24 (19,4%) possuem nível 1D, 12 (9,7%) possuem nível 1C, 7 (5,6%) possuem nível 1B e 12 (9,7%) possuem nível 1A. Desta forma, assumindo que a formação de coautorias é aleatória, existe uma maior probabilidade de os pesquisadores de níveis mais baixos se relacionarem com pesquisadores do mesmo nível e de pesquisadores de níveis mais altos se relacionarem com pesquisadores de níveis diferentes.

A Tabela 5.3 mostra a correlação dessas três métricas de *E-I index*. Nota-se que elas são altamente correlacionadas, principalmente a *E-I index_W* e o *E-I index_Z*.

Tabela 5.3 – Correlações entre as métricas de *E-I index*

	Correlações		
	<i>E-I index</i>	<i>E-I index_W</i>	<i>E-I Index_Z</i>
<i>E-I index</i>	1		
<i>E-I index_W</i>	0,893**	1	
<i>E-I index_Z</i>	0,901**	0,982**	1

** A correlação é significativa no nível 0,01 (teste bilateral).

Para visualizar as dispersões e as correlações entre as métricas de *E-I index* são apresentadas as Figuras 5.5, 5.6 e 5.7. Nas três figuras é possível notar que os pontos estão mais dispersos abaixo e a esquerda das medianas, representadas pela linha horizontal e vertical⁴, respectivamente, e aglutinados acima e a direita dessas medianas.

⁴ Em todos os gráficos do Capítulo 5 que relacionam duas métricas, as linhas horizontais e verticais representam as medianas das métricas que estão sendo relacionadas.

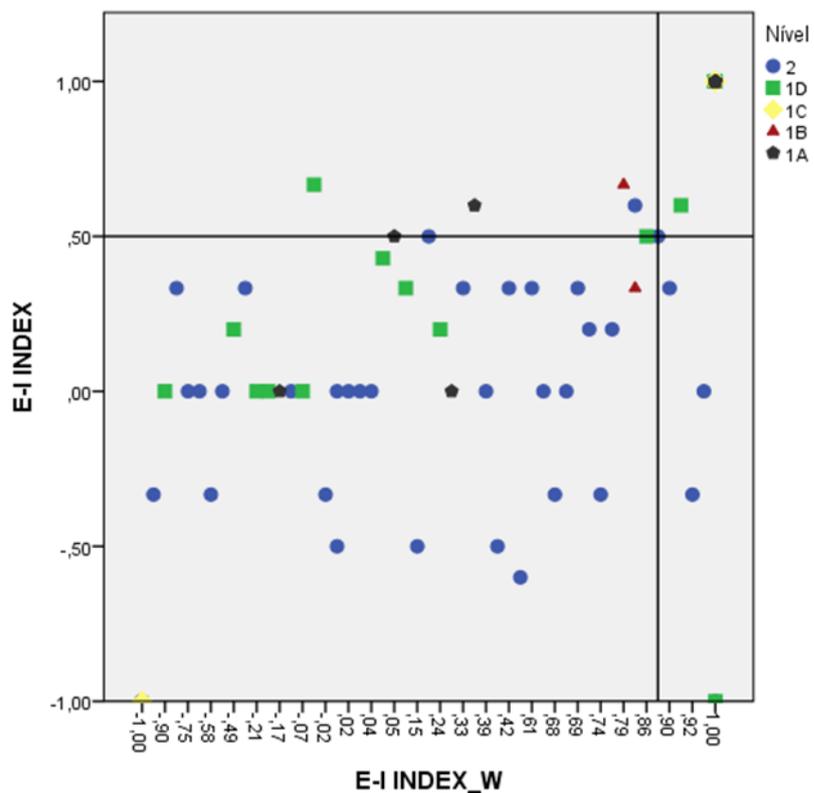


Figura 5.5 – Relação entre E-I Index e E-I index_W

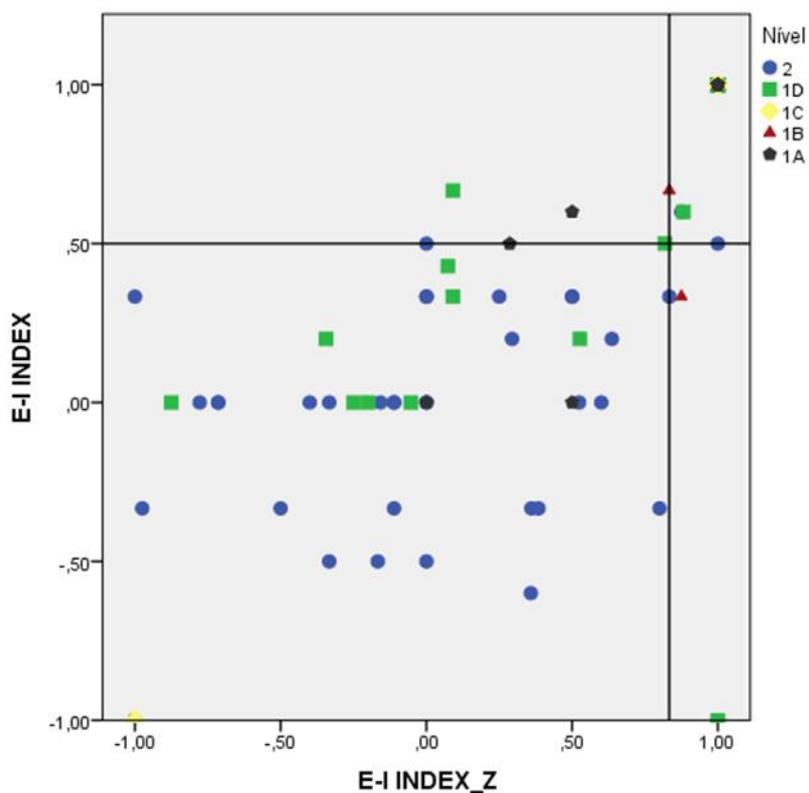


Figura 5.6 – Relação entre E-I index e o E-I index_Z

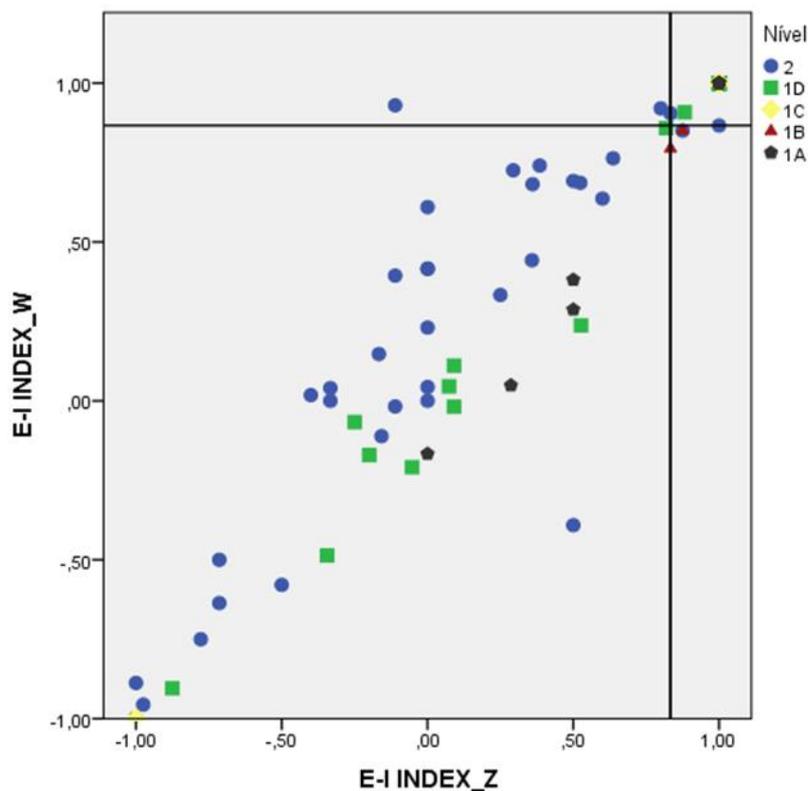


Figura 5.7 – Relação entre o E-I index_W e o E-I index_Z

5.5.2 Centralidade de Grau

A Tabela 5.4 classifica os 10 pesquisadores com maiores centralidades de grau, calculadas de três maneiras distintas: Centralidade de Grau Não-Ponderada (CGNP); Centralidade de Grau Ponderada (CGPW), só com pesos das arestas; e Centralidade de Grau Ponderada (CGPZ), com pesos das arestas e dos nós. Por esta tabela é possível observar como o número de ligações, as frequências das ligações e a combinação das frequências das ligações com os pesos dos nós alteram as posições dos pesquisadores. Por exemplo, o pesquisador PQ124 figura na primeira posição na CGNP, porém ao considerar o peso das ligações e a importância dos nós esse pesquisador não figura entre as dez primeiras posições. Nesse caso, o pesquisador PQ124 tem o maior número de coautores, mas com baixas frequências de colaboração quando comparado, por exemplo, com o pesquisador PQ62 que estava em segundo na CGNP e assume a primeira posição na CGPW. Esse pesquisador assume também a primeira posição quando são considerados os pesos dos nós.

Quanto ao nível de produtividade os pesquisadores de níveis mais baixos (2 e 1D) são os mais presentes nas 10 primeiras posições, principalmente na CGPW e na CGPZ.

Tabela 5.4– Os 10 mais bem posicionados na centralidade de grau

Posição	CGNP			CGPW			CGPZ		
1	PQ124	(10)	1A	PQ62	(95)	1D	PQ62	(1247)	1D
2	PQ62	(7)	1D	PQ108	(87)	2	PQ74	(764)	2
3	PQ0	(7)	1A	PQ56	(86)	2	PQ41	(740)	1D
4	PQ111	(6)	1A	PQ107	(80)	2	PQ82	(490)	1A
5	PQ84	(6)	1D	PQ74	(65)	2	PQ0	(461)	1A
6	PQ65	(6)	1B	PQ41	(61)	1D	PQ12	(396)	1A
7	PQ41	(5)	1D	PQ82	(57)	1A	PQ108	(368)	2
8	PQ82	(5)	1A	PQ9	(53)	2	PQ56	(364)	2
9	PQ9	(5)	2	PQ117	(50)	1D	PQ107	(362)	2
10	PQ85	(5)	1D	PQ0	(43)	1A	PQ9	(323)	2

Para avaliar e comparar as distribuições das centralidades de grau entre os níveis das bolsas são apresentados os gráficos boxplots nas Figuras 5.8, 5.9 e 5.10. A Figura 5.8 aponta o nível 1A como o de maior variabilidade de CGNP e 1B como o de menor variabilidade. O nível 1A apresenta também a maior CGNP em termos da mediana, enquanto os níveis 2, 1D e 1C as menores. O comportamento apresentado na Figura 5.8 também se repete nas Figuras 5.9 e 5.10, entretanto, em termos da mediana o nível 1C, na Figura 5.9, e o nível 2, na Figura 5.10, são as menores. Também, nessas figuras, há um maior número de *outliers* é observado nos níveis mais inferiores, revelando que a alta CGPW e CGPZ obtidas por alguns pesquisadores são atípicas (raros) para esses níveis de bolsas.

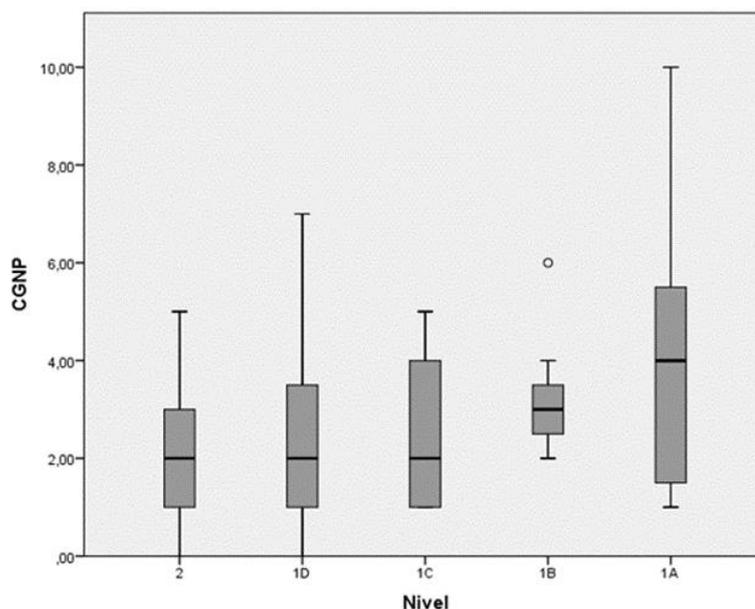


Figura 5.8 – Gráfico boxplot para CGNP

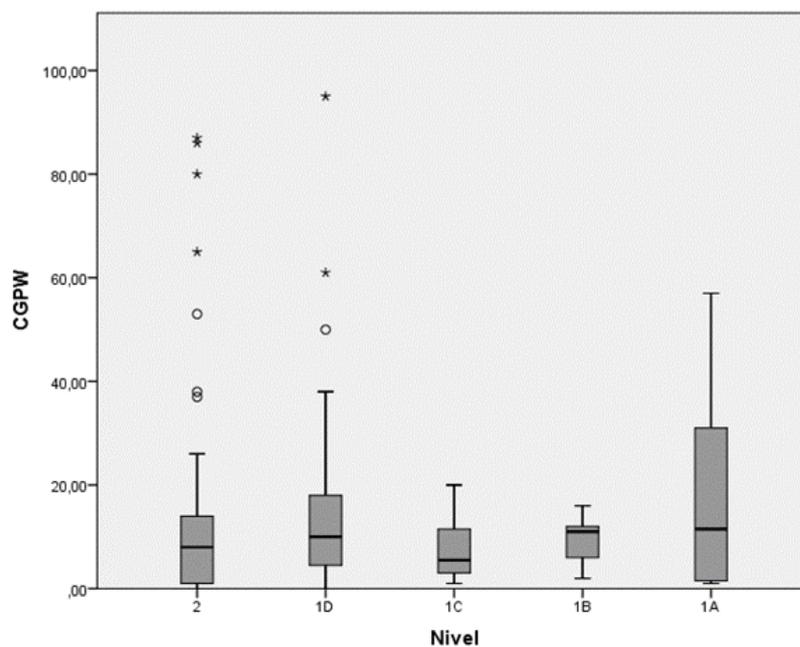


Figura 5.9 – Gráfico boxplot para CGPW

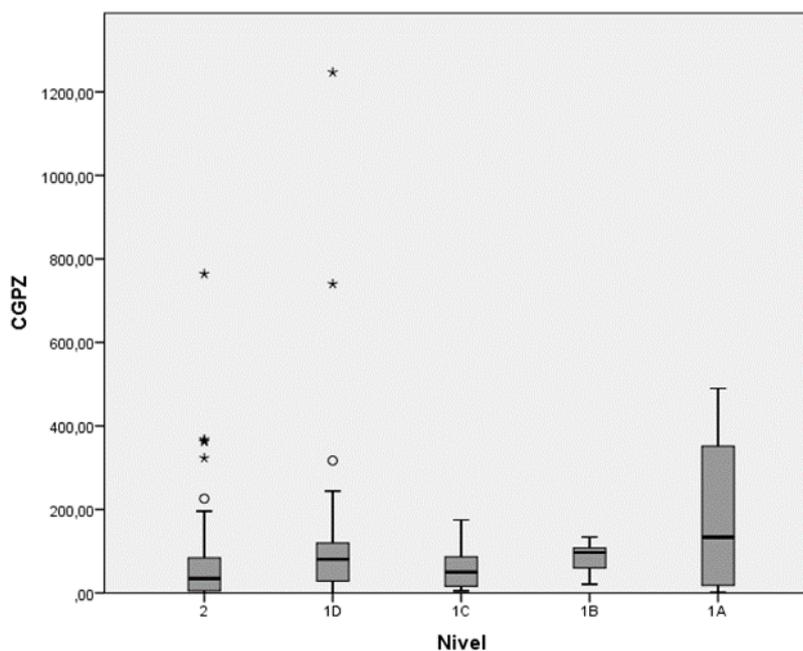


Figura 5.10 – Gráfico boxplot para CGPZ

A Tabela 5.5 mostra a correlação dessas três métricas de centralidade de grau. Nota-se que elas são altamente correlacionadas, principalmente a CGPW e a CGPZ. Desta forma, a inserção dos pesos dos nós na estrutura da rede não alterou a natureza funcional da CGNP, apesar de alterar as posições dos pesquisadores mais centrais.

Tabela 5.5 – Correlações entre as métricas de centralidade de grau

	Correlações		
	CGNP	CGPW	CGPZ
CGNP	1		
CGPW	0,717**	1	
CGPZ	0,757**	0,952**	1

** . A correlação é significativa no nível 0,01 (teste bilateral).

Nas Figuras 5.11, 5.12 e 5.13 é possível visualizar as dispersões e as correlações entre as métricas de centralidades de grau. A relação entre a CGNP e a CGPW é apresentada na Figura 5.11, onde as linhas horizontal e vertical indicam as medianas da CGNP e da CGPW, respectivamente. E aponta que a CGNP e a CGPW estão aglutinadas abaixo da mediana e esquerda da mediana, respectivamente. Acima e a direita dessas medianas os pontos estão mais dispersos. Nota-se que os pesquisadores não conseguem ter simultaneamente o máximo CGNP e o máximo CGPW. Relação similar é apresentada entre a CGNP e a CGPZ, na Figura 5.12. A relação entre a CGPW e a CGPZ é representada na Figura 5.13, a qual é mais linear que as anteriores, para todos os níveis dos pesquisadores.

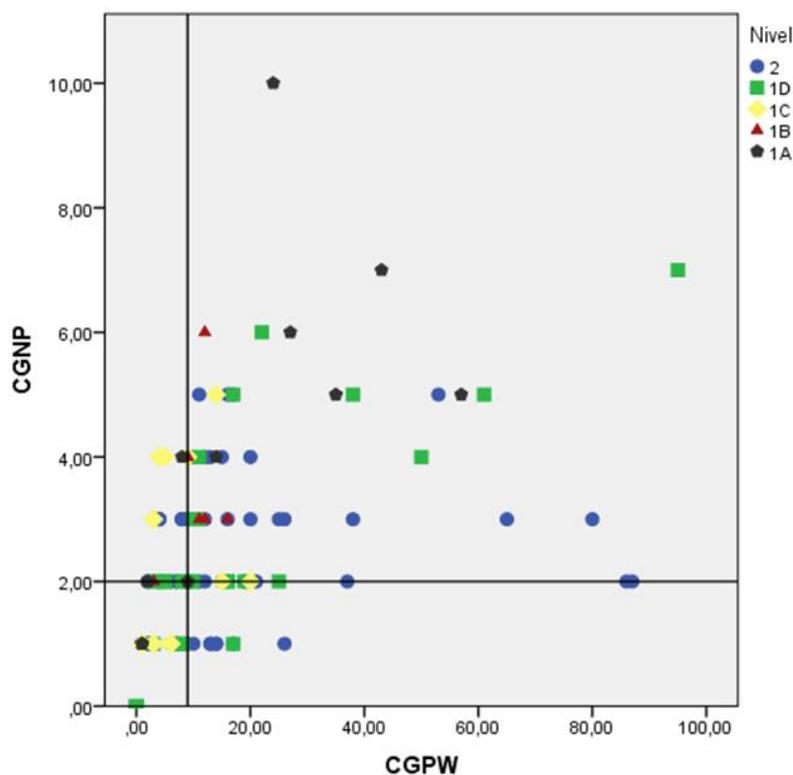


Figura 5.11 – Relação entre a CGNP e a CGPW

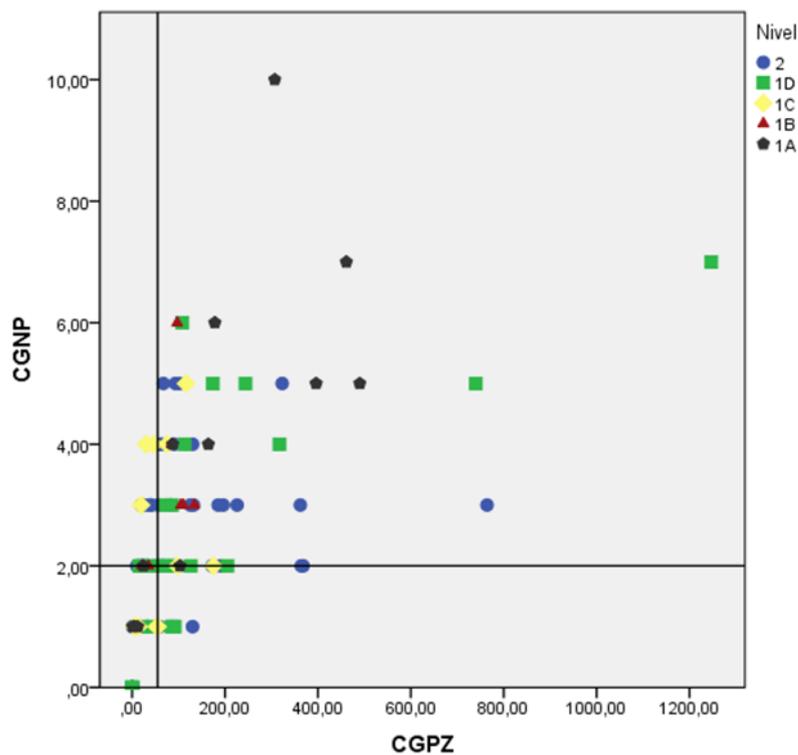


Figura 5.12 – Relação entre a CGNP e a CGPZ

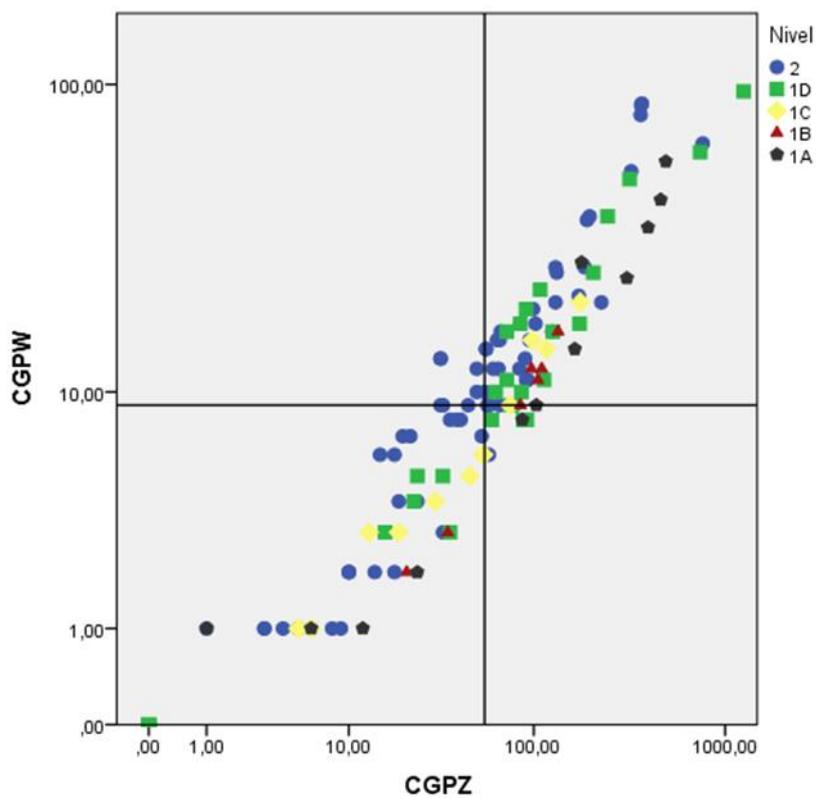


Figura 5.13 – Relação entre a CGPW e a CGPZ

A Tabela 5.6 apresenta as correlações das três métricas de centralidades de grau com o nível de produtividade.

Tabela 5.6 – Correlações das métricas de centralidades de grau com os níveis das bolsas

	Correlações		
	CGNP	CGPW	CGPZ
Nível de produtividade	0,287**	0,111	0,261**

** . A correlação é significativa no nível 0,01 (teste bilateral).

Observa-se que, apenas, a CGPW não tem correlação significativa com o nível de produtividade. A centralidade de grau que apresentou maior correlação com o nível de produtividade foi a que considera somente o número de ligações. Desta forma, cooperar com mais autores ou cooperar com frequência com autores com maiores desempenhos (índice h) impacta o nível de produtividade positivamente.

5.5.3 Média das Ligações Fortes

Os resultados dos 10 autores mais bem posicionados, de acordo com a média das ligações fortes, são apresentados na Tabela 5.7. A média das ligações fortes ponderada por W (MLFW), é o resultado da razão entre CGPW e CGNP e a média das ligações fortes ponderada por Z (MLFZ), é a razão entre CGPZ e CGNP. Observe que a participação dos pesquisadores nível 2 predomina entre as 10 primeiras posições na MLFW, ao considerar a importância dos nós essa participação diminui, mas ainda é alta. Isto indica que bolsistas com nível menor tendem a concentrar seus trabalhos com alguns outros bolsistas, enquanto bolsistas com maior nível tendem a diversificar mais suas colaborações.

Tabela 5.7 - Os 10 mais bem posicionados na média das ligações fortes

Posição	MLFW			MLFZ		
1	PQ108	(43,50)	2	PQ74	(254,67)	2
2	PQ56	(43,00)	2	PQ108	(184,00)	2
3	PQ107	(26,67)	2	PQ56	(182,00)	2
4	PQ81	(26,00)	2	PQ62	(178,14)	1D
5	PQ74	(21,67)	2	PQ41	(148,00)	1D
6	PQ39	(18,50)	2	PQ81	(130,00)	2
7	PQ49	(17,00)	2	PQ107	(120,67)	2
8	PQ110	(17,00)	1D	PQ116	(102,50)	1D
9	PQ42	(14,00)	2	PQ82	(98,00)	1A
10	PQ92	(14,00)	2	PQ39	(95,00)	2

Para avaliar e comparar as distribuições das médias das ligações fortes entre os níveis das bolsas são apresentados os gráficos boxplots nas Figuras 5.14 e 5.15.

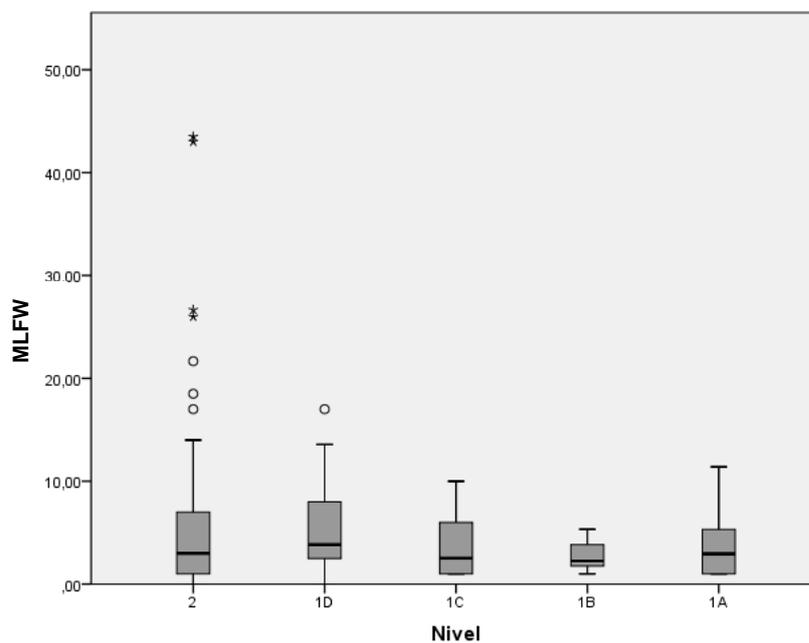


Figura 5.14 – Gráfico boxplot para a MLFW

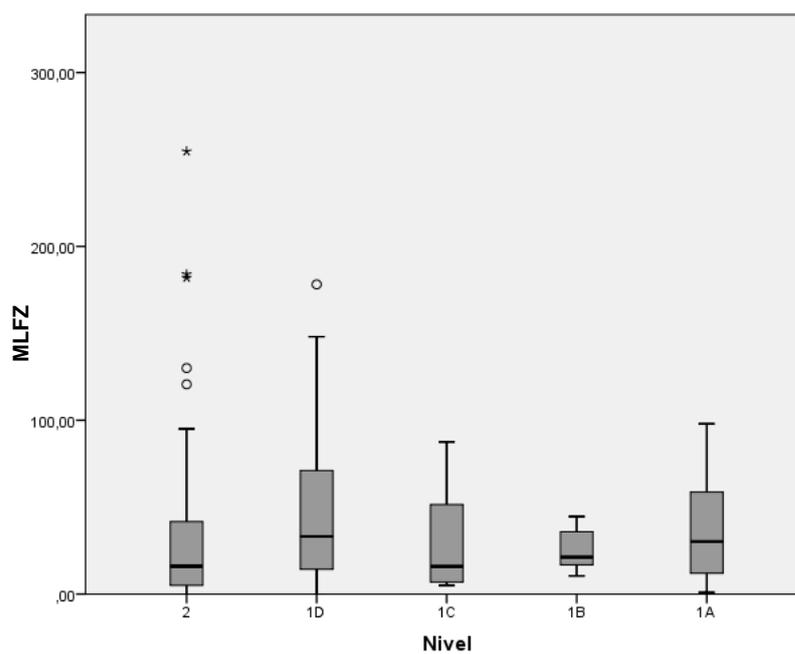


Figura 5.15 – Gráfico boxplot para a MLFZ

A Figura 5.14 aponta o nível 2 como o de maior variabilidade da MLFW e possui a menor MLFW. 1B apresenta a menor variabilidade. Na Figura 5.15, o nível 1D aponta a maior

variabilidade de MLFZ, enquanto a menor variabilidade é apresentada pelo nível 1B. Em ambas as figuras, os níveis 2 e 1D apresentam valores *outliers*.

A Tabela 5.8 mostra as correlações dessas duas médias das ligações fortes. Nota-se que elas apresentam correlações altas. No entanto 40% dos pesquisadores entre as 10 primeiras posições na MLFW não figuram neste *ranking* na MLFZ.

Tabela 5.8 - Correlações entre médias de ligações fortes

	Correlações	
	MLFW	MLFZ
MLFW	1	
MLFZ	0,944**	1

** . A correlação é significativa no nível 0,01 (teste bilateral)

Na Figura 5.16 é possível visualizar a dispersão e a correlação entre as métricas de média das ligações fortes. As linhas horizontal e vertical indicam as medianas da MLFW e da MLFZ, respectivamente. É possível observar que os pesquisadores níveis 2 e 1D apresentam simultaneamente as mais altas MLFW e MLFZ.

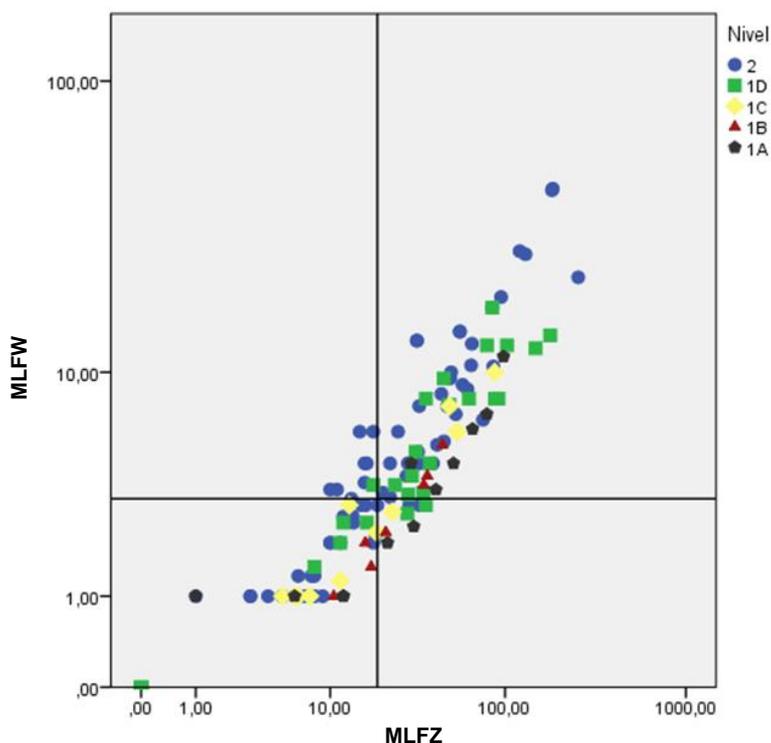


Figura 5.16 – Relação entre a MLFW e a MLFZ

A Tabela 5.9 apresenta as correlações das duas métricas de média das ligações fortes com o nível de produtividade. A média das forças das ligações pouco impacta o nível de produtividade, quando considerada a importância dos nós e, o impacto é quase zero e não significativo no nível 0,05, quando não é considerada. Desta forma, a média das forças das ligações pouco contribui para o nível de produtividade.

Tabela 5.9 - Correlações das métricas média das ligações fortes com os níveis das bolsas

	Correlações	
	MLFW	MLFZ
Nível de produtividade	0,011	0,172*

*. A correlação é significativa no nível 0,05 (teste bilateral)

5.5.4 Centralidade de Proximidade

A Tabela 5.10 apresenta os resultados dos 10 autores mais bem posicionados, de acordo com a centralidade de proximidade: não-ponderada (CPNP); ponderada com os pesos das arestas (CPPW) e; ponderada com os pesos das arestas e dos nós (CPPZ).

Tabela 5.10 - Os 10 mais bem posicionados na centralidade de proximidade

Posição	CPNP			CPPW			CPPZ		
1	PQ124	(0,1834)	1A	PQ124	(0,3010)	1A	PQ124	(2,2633)	1A
2	PQ32	(0,1740)	1B	PQ60	(0,2915)	1A	PQ60	(2,2266)	1A
3	PQ60	(0,1713)	1A	PQ112	(0,2838)	1B	PQ32	(2,1843)	1B
4	PQ61	(0,1704)	2	PQ32	(0,2808)	1B	PQ112	(2,1621)	1B
5	PQ14	(0,1650)	1C	PQ33	(0,2769)	2	PQ33	(2,1269)	2
6	PQ47	(0,1618)	2	PQ14	(0,2763)	1C	PQ61	(2,1004)	2
7	PQ103	(0,1618)	1C	PQ61	(0,2739)	2	PQ103	(2,0902)	1C
8	PQ76	(0,1595)	1B	PQ47	(0,2681)	2	PQ144	(2,0877)	1D
9	PQ112	(0,1595)	1B	PQ76	(0,2678)	1B	PQ36	(2,0732)	2
10	PQ2	(0,1592)	2	PQ36	(0,2654)	2	PQ76	(2,0520)	1B

Para exemplificar a mudança nas posições dos nós nas três métricas de centralidade de proximidade observe os pesquisadores PQ32 e PQ60. O pesquisador PQ32 é o segundo mais próximo dos outros nós, pela CPNP, neste caso, a soma das distâncias entre ele e os outros nós é menor que a soma das distâncias PQ60 aos demais nós. No entanto pela CPPW os caminhos que ligam o pesquisador PQ60 aos demais nós são formadas por ligações mais frequentes que aquelas que formam os caminhos que ligam o pesquisador PQ32 aos outros nós. Como as frequências das ligações encurtam os caminhos, o pesquisador PQ60 obteve uma melhor

posição em relação à CPPW. Esse pesquisador se manteve também na segunda posição na CPPZ.

Quanto ao nível de produtividade, pesquisadores com níveis mais altos predominam entre as 10 posições nas três métricas de centralidade de proximidade. Os pesquisadores nível 2 ocupam em média três posições nessa tabela.

Para avaliar e comparar as distribuições das centralidades de proximidade entre os níveis das bolsas são apresentados os gráficos boxplots nas Figuras 5.17, 5.18 e 5.19. Nos dois primeiros gráficos a maior variabilidade é apontada pelo nível 1C e, a maior mediana e a menor variação são apresentadas pelo nível 1B. No terceiro gráfico o nível 2 apresenta a maior variabilidade.

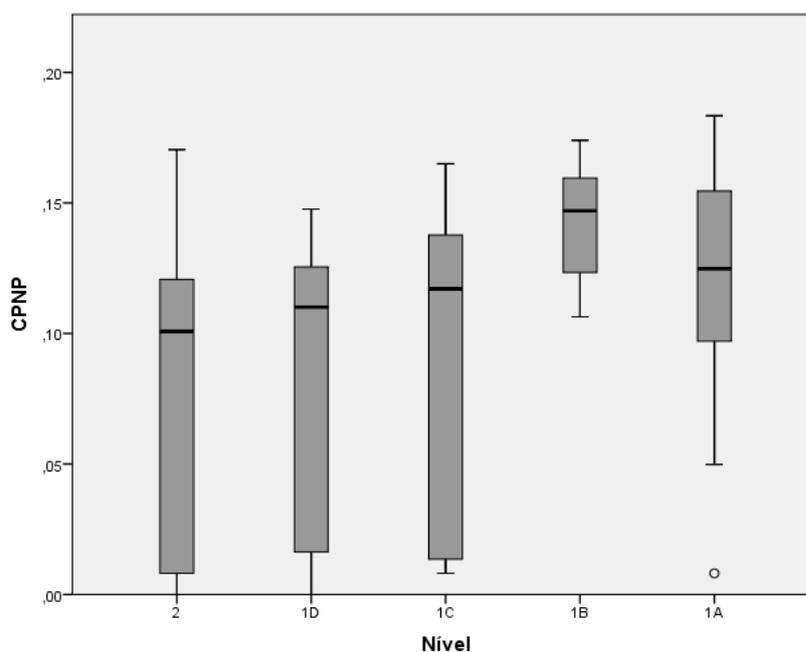


Figura 5.17 – Gráfico boxplot para CPNP

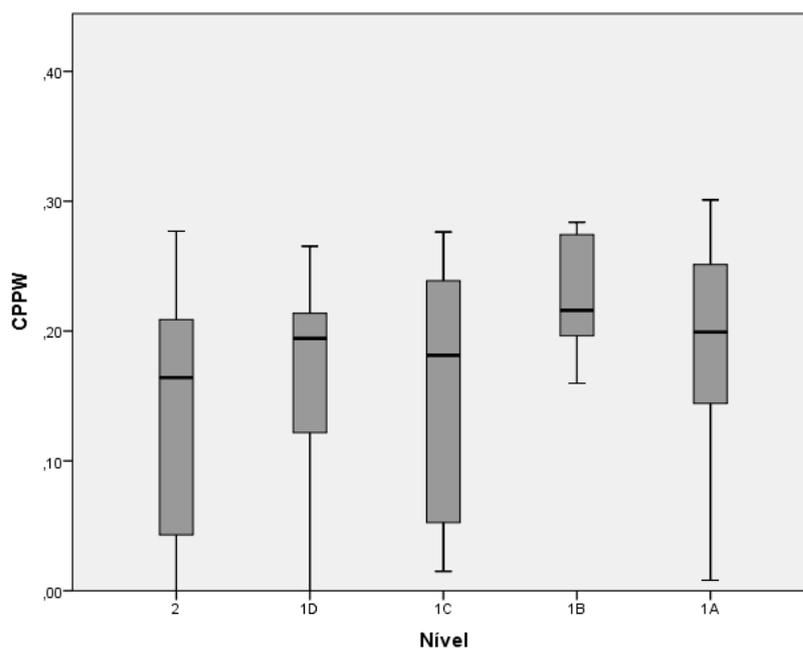


Figura 5.18 – Gráfico boxplot para CPPW

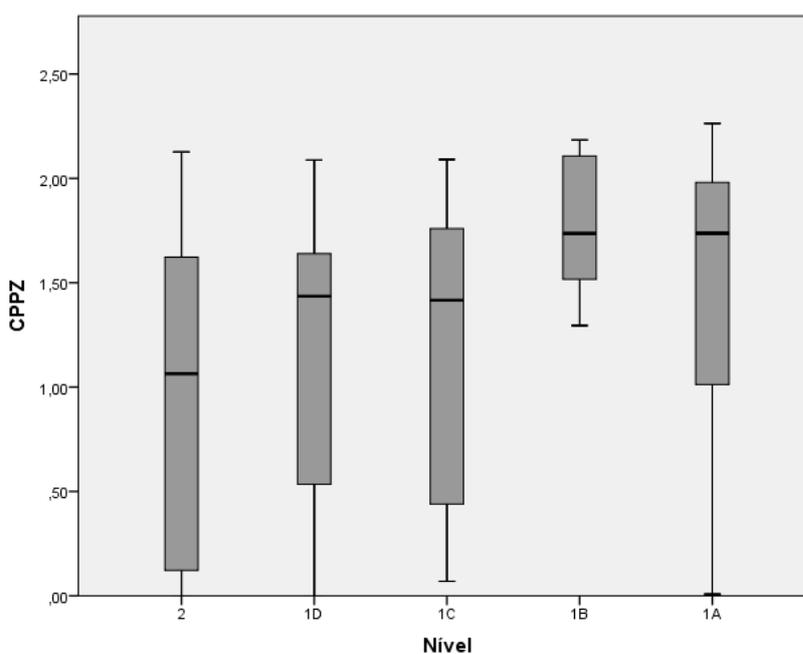


Figura 5.19 – Gráfico boxplot para CPPZ

A Tabela 5.11 mostra as correlações das três métricas de centralidade de proximidade. Nota-se que elas são fortemente correlacionadas, principalmente CPPZ e a CPPW. Desta forma, a inserção do peso dos nós na estrutura da rede não alterou a natureza da centralidade de proximidade calculada pela forma tradicional, mas provocou pequenas alterações nas posições dos pesquisadores.

Tabela 5.11 - Correlações entre as métricas de centralidade de proximidade

	Correlações		
	CPNP	CPPW	CPPZ
CPNP	1		
CPPW	0,940**	1	
CPPZ	0,920**	0,966**	1

**A correlação é significativa no nível 0,01 (teste bilateral).

Nas Figuras 5.20 e 5.21, as dispersões e as correlações entre a CPNP com a CPPW e com CPPZ, respectivamente apresentam comportamento similares, com pontos mais aglutinados e mais visíveis acima da mediana da CPNP e a direita da CPPW e da CPPZ. Também se verifica nesses gráficos a presença de dois grupos, o primeira é formado por nós pertencentes a componente principal o segundo é formado pelas demais componentes e por nós isolados. A relação entre a CPPW e a CPPZ, na Figura 5.22, é a mais linear e não é identificado, de forma aparente, a presença de grupos.

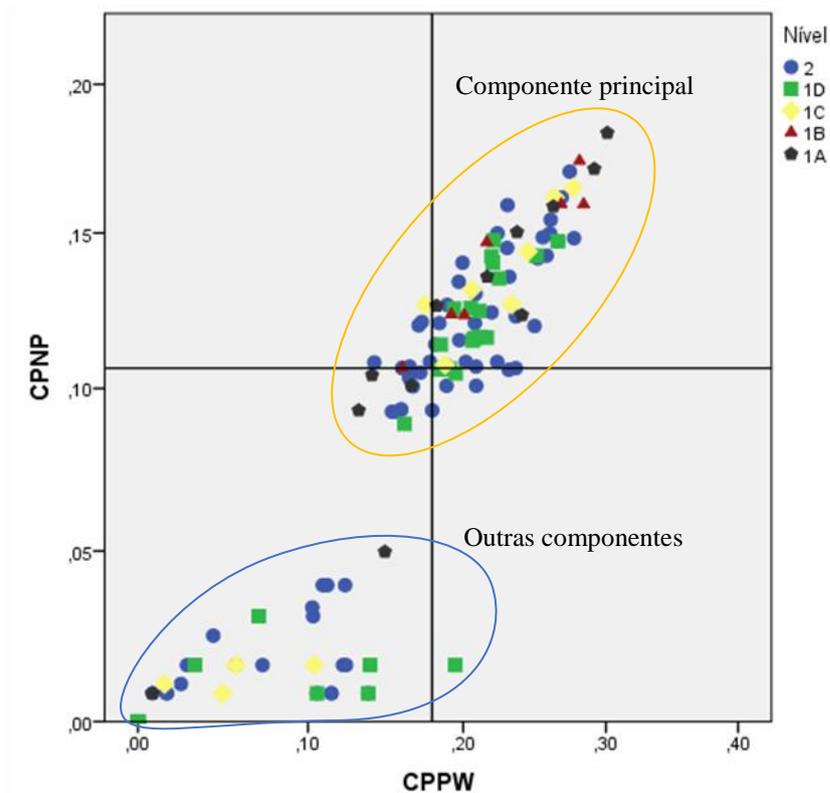


Figura 5.20 – Relação entre a CPNP e a CPPW

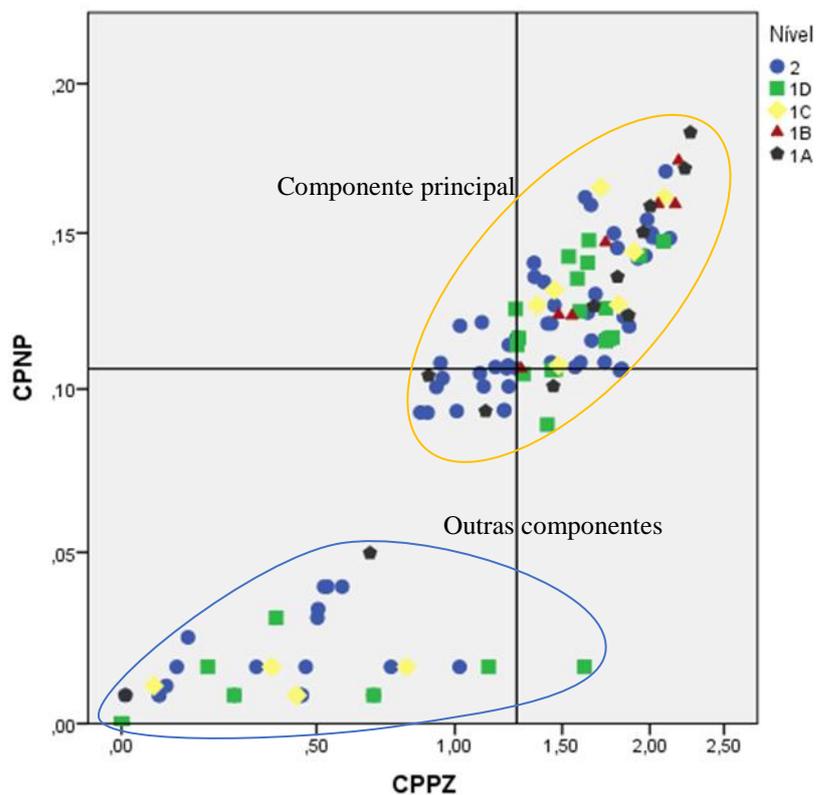


Figura 5.21 – Relação entre a CPNP e a CPPZ

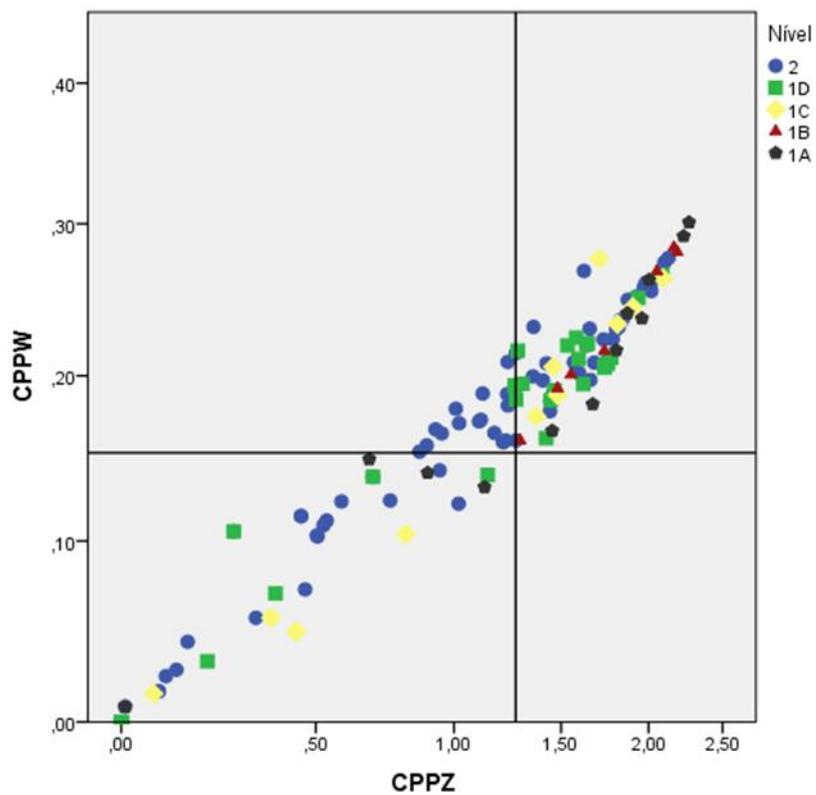


Figura 5.22 – Relação entre a CPPW e a CPPZ

A Tabela 5.12 apresenta as correlações das três métricas de centralidades de proximidade com o nível de produtividade.

Tabela 5.12 - Correlações das métricas de centralidade de proximidade com os níveis das bolsas

	Correlações		
	CPNP	CPPW	CPPZ
Nível de produtividade	0,264**	0,221**	0,284**

** A correlação é significativa no nível 0,01 (teste bilateral).

As três centralidades de proximidade apontaram correlações positivas significativas com o nível de produtividade. A que apresentou maior correlação foi a CPPZ, ou seja, a que considera a importância dos nós. Em seguida a CPNP. Desta forma, por apresentarem maiores possibilidades de estabelecer parcerias de publicações, os pesquisadores com maiores centralidades de proximidade tendem também a apresentar um maior nível de produtividade. Além disso, o pesquisador que está mais próximo de pesquisadores importantes tende a apresentar maior nível de produtividade.

5.5.5 Centralidade de Intermediação

A Tabela 5.13 apresenta os resultados dos 10 autores mais bem posicionados, de acordo com a centralidade de intermediação: não-ponderada (CINP); ponderada pelos pesos das arestas (CIPW) e; ponderada pelos pesos das arestas e dos nós (CIPZ).

Tabela 5.13 - Os 10 mais bem posicionados na centralidade de intermediação

Posição	CINP	CIPW	CIPZ
1	PQ124 (0,2492) 1A	PQ124 (0,2850) 1A	PQ124 (0,3504) 1A
2	PQ61 (0,1572) 2	PQ76 (0,1813) 1B	PQ76 (0,2735) 1B
3	PQ32 (0,1564) 1B	PQ32 (0,1781) 1B	PQ32 (0,2240) 1B
4	PQ112 (0,1453) 1B	PQ61 (0,1781) 2	PQ61 (0,2230) 2
5	PQ60 (0,1433) 1A	PQ60 (0,1766) 1A	PQ60 (0,1943) 1A
6	PQ102 (0,1429) 2	PQ102 (0,1582) 2	PQ12 (0,1843) 1A
7	PQ47 (0,1402) 2	PQ112 (0,1466) 1B	PQ102 (0,1582) 2
8	PQ14 (0,1331) 1C	PQ12 (0,1309) 1A	PQ112 (0,1393) 1B
9	PQ76 (0,1323) 1B	PQ47 (0,1185) 2	PQ103 (0,1298) 1C
10	PQ2 (0,1171) 2	PQ2 (0,1158) 2	PQ2 (0,1158) 2

O pesquisador PQ124 é o mais central em relação às três medidas de centralidades de intermediação. Observa-se que a maioria dos pesquisadores classificados entre os 10 na CINP estão classificados também nas outras duas centralidades de intermediação, com pequenas

variações nas posições. Desta forma, a utilização do peso das arestas e do método proposto não alteram muito as posições dos pesquisadores na rede.

Os pesquisadores níveis 1A e 1B são a maioria nessa tabela, os de nível 2 figuram em média em 3 posições em cada métrica.

Nas Figuras 5.23, 5.24 e 5.25 é possível visualizar o centro, a dispersão, o desvio de simetria e a identificação das observações consideradas atípicas. Nesses três gráficos (que estão na escala de potência, para melhor visualização) o nível 1A apresenta a maior variabilidade e o nível 1B a maior mediana. O nível 2 apresenta a menor variação e a maior quantidade de pontos atípicos.

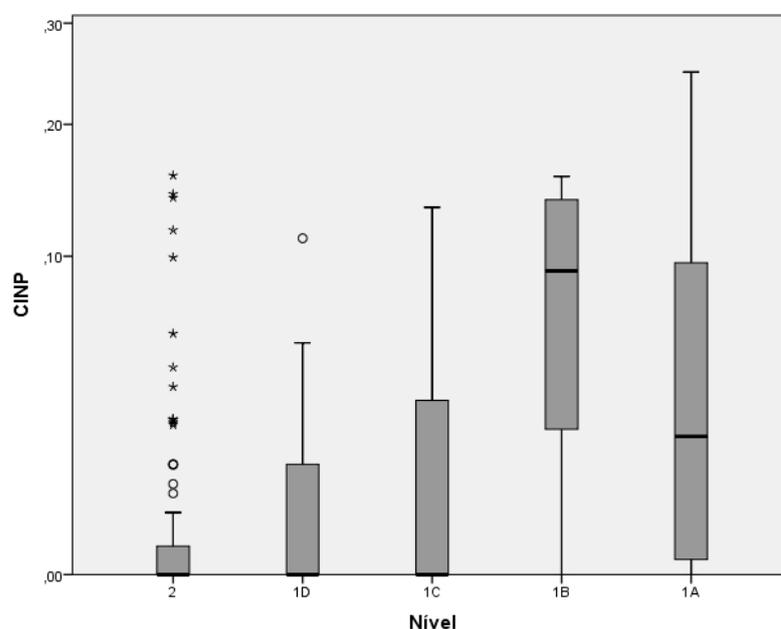


Figura 5.23 – Gráfico boxplot para a CINP

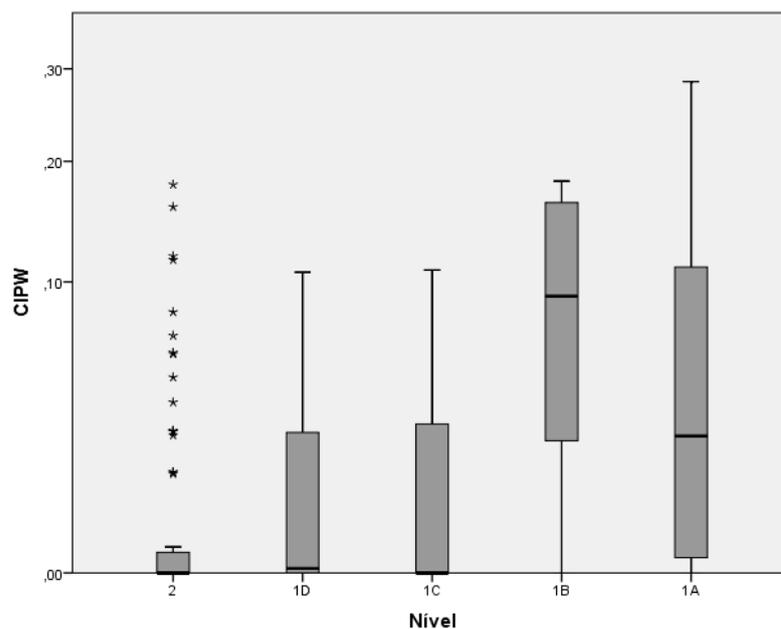


Figura 5.24 – Gráfico boxplot para a CIPW

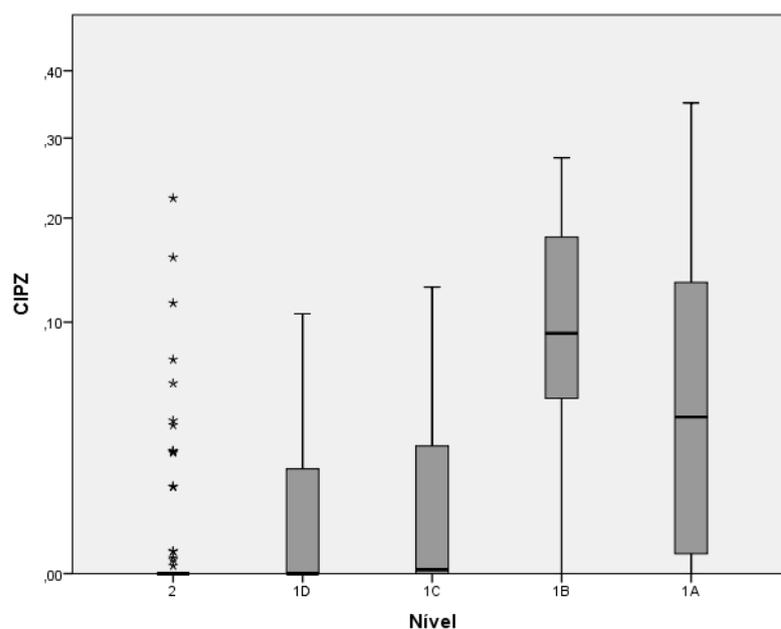


Figura 5.25 – Gráfico boxplot para a CIPZ

A Tabela 5.14 mostra a correlação dessas três métricas de centralidades de intermediação. Nota-se que elas são altamente correlacionadas, principalmente a CIPW e a CIPZ. Desta forma, a inserção da importância dos nós na estrutura da rede não alterou a natureza da centralidade de intermediação calculada pela forma tradicional. Houve pequenas alterações nas posições.

Tabela 5.14- Correlações entre as métricas de centralidade de intermediação

	Correlações		
	CINP	CIPW	CIPZ
CINP	1		
CIPW	0,863**	1	
CIPZ	0,866**	0,937**	1

** . A correlação é significativa no nível 0,01 (teste bilateral).

As Figuras 5.26, 5.27 e 5.28 apresentam gráficos de dispersão da CINP, da CIPW e da CIPZ. Nessas figuras somente é possível visualizar os pontos iguais e superiores a mediana, pois mais da metade dos pesquisadores apresentam valores das métricas iguais ou muito próximos a zero.

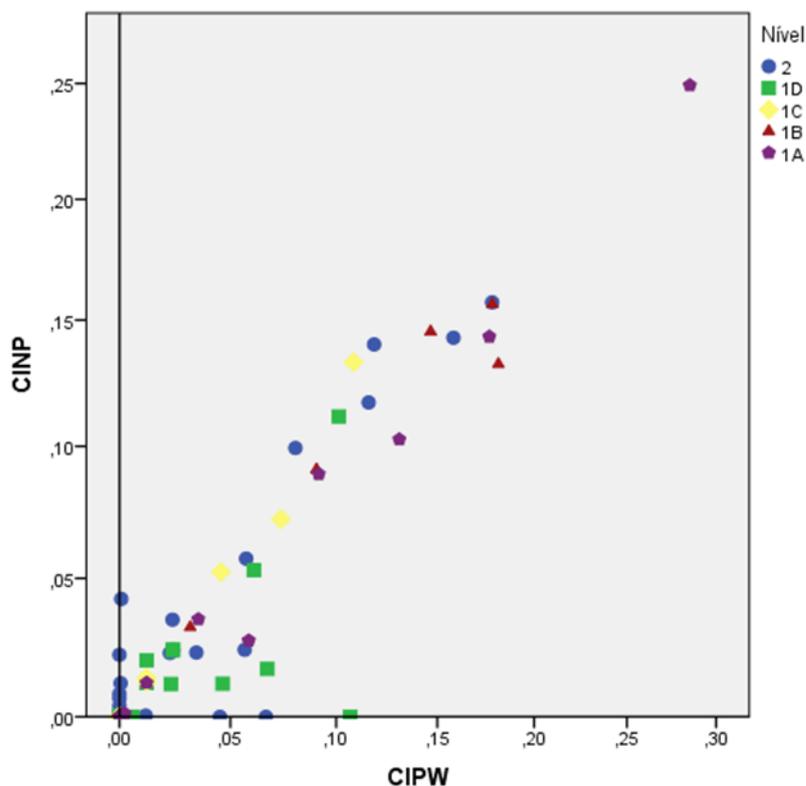


Figura 5.26 – Relação entre a CINP e a CIPW

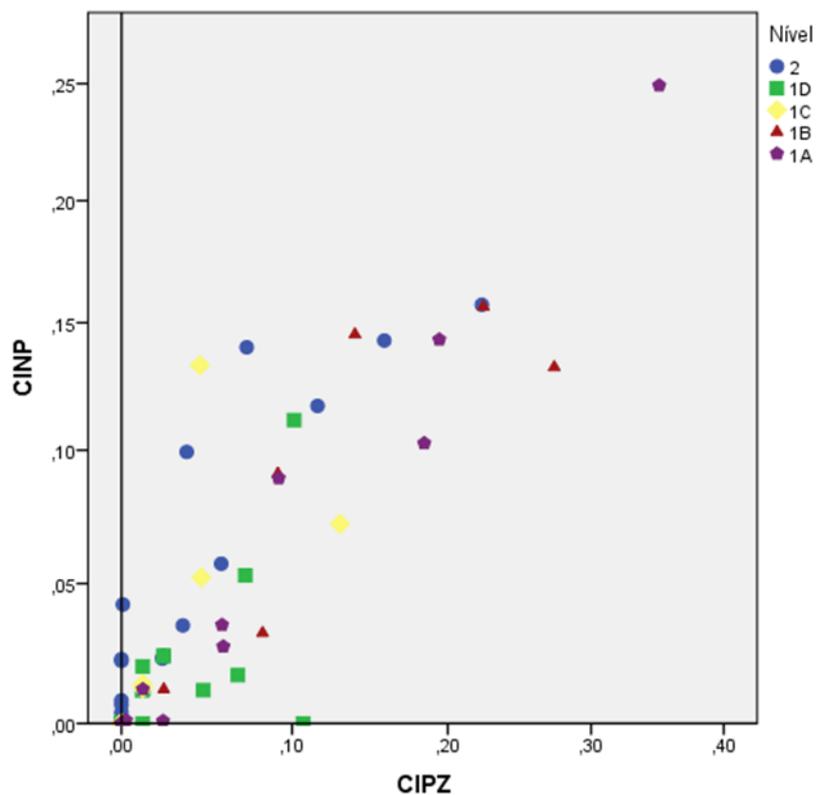


Figura 5.27 – Relação entre a CINP e a CIPZ

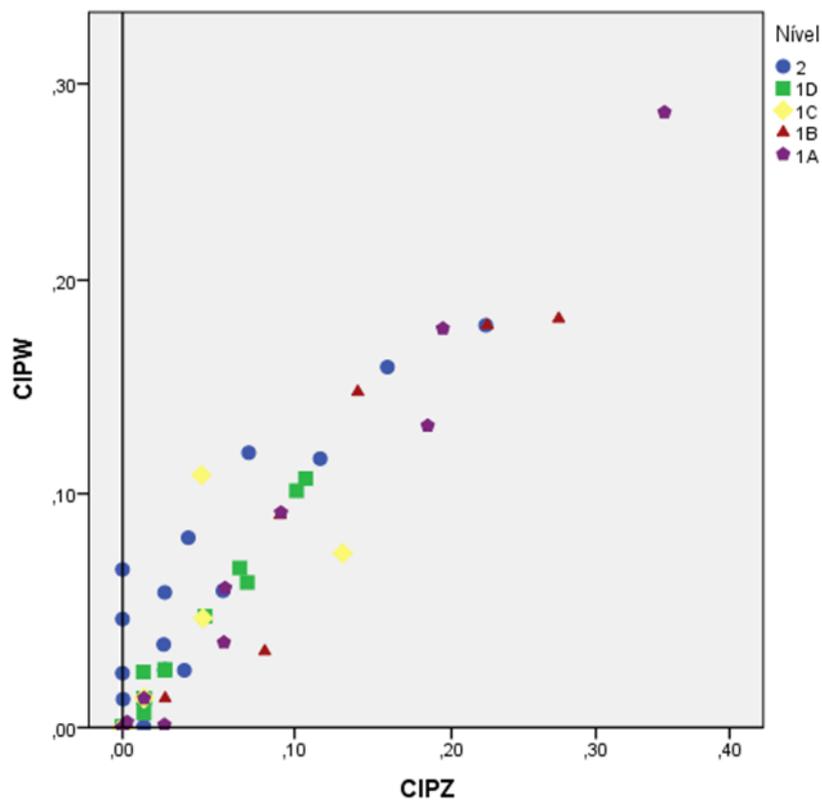


Figura 5.28 – Relação entre a CIPW e a CIPZ

A Tabela 5.15 apresenta as correlações das três métricas de centralidades de intermediação com o nível de produtividade.

Tabela 5.15 - Correlações das métricas de centralidade de intermediação com os níveis das bolsas

	Correlações		
	CINP	CIPW	CIPZ
Nível de produtividade	0,285**	0,354**	0,421**

** . A correlação é significativa no nível 0,01 (teste bilateral)

As três centralidades de intermediação apontaram correlações positivas significativas com o nível de produtividade. A que apresentou maior correlação com o nível de produtividade foi a CIPZ, ou seja, a que considera a importância dos nós. Em seguida a CIPW. Desta forma, os pesquisadores que assumem um papel de “intermediador”, controlando a frequência do fluxo de informação tendem a ter maiores níveis de produtividade, porém, aqueles que intermediam nós em caminhos cujas ligações são mais frequentes ou apresentam nós mais importantes têm níveis de bolsas superiores.

5.5.6 Centralidade de Autovetor

Tabela 5.16 apresenta os 10 autores mais bem posicionados, de acordo com a Centralidade de Autovetor: não-ponderada (CANP); ponderada com os pesos das arestas (CAPW); e ponderada com os pesos das arestas e dos nós (CAPZ).

Tabela 5.16 - Os 10 mais bem posicionados na centralidade de autovetor

Posição	CANP			CAPW			CAPZ		
1	PQ0	(0,4915)	1A	PQ108	(0,5911)	2	PQ62	(0,6693)	1D
2	PQ8	(0,4296)	2	PQ56	(0,5869)	2	PQ74	(0,5408)	2
3	PQ31	(0,3927)	1D	PQ107	(0,5533)	2	PQ41	(0,5022)	1D
4	PQ25	(0,3661)	1D	PQ65	(0,0066)	1B	PQ85	(0,0771)	1D
5	PQ78	(0,3426)	2	PQ73	(0,0004)	1D	PQ95	(0,0226)	1B
6	PQ1	(0,2932)	2	PQ141	(0,0003)	1D	PQ40	(0,0221)	1B
7	PQ19	(0,1900)	1D	PQ62	(0,0002)	1D	PQ14	(0,0129)	1C
8	PQ112	(0,1433)	1B	PQ74	(0,0002)	2	PQ81	(0,0114)	2
9	PQ118	(0,0944)	1A	PQ41	(0,0001)	1D	PQ15	(0,0101)	2
10	PQ14	(0,0554)	1C	PQ106	(0,0001)	2	PQ119	(0,0005)	2

É notório que a composição das 10 primeiras posições nas três métricas de centralidades de autovetor são formadas por pesquisadores diferentes, apenas o pesquisador PQ14 décimo

colocado na CANP figura duas vezes na tabela (posição 7 na CAPZ). Quanto aos níveis de produtividade dos pesquisadores a maioria são de níveis 2 e 1D.

Para avaliar e comparar as distribuições das centralidades de autovetor entre os níveis das bolsas são apresentados os gráficos boxplots nas Figuras 5.29, 5.30 e 5.31. Para obter uma melhor visualização, foram desconsiderados, nestes gráficos os nós isolados (possuem centralidade de autovetor igual a zero, são 21 nós) e a escala utilizada foi a logarítmica.

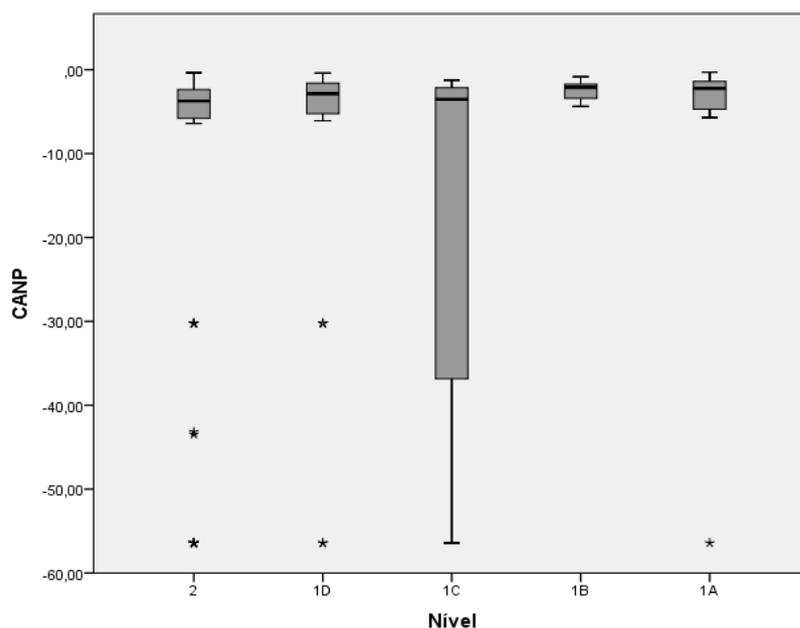


Figura 5.29 – Gráfico boxplot da CANP

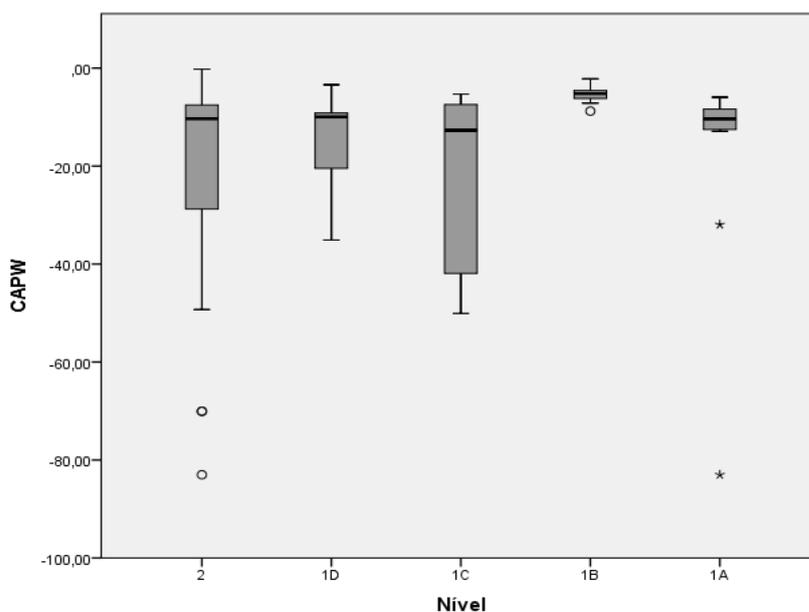


Figura 5.30 – Gráfico boxplot da CAPW

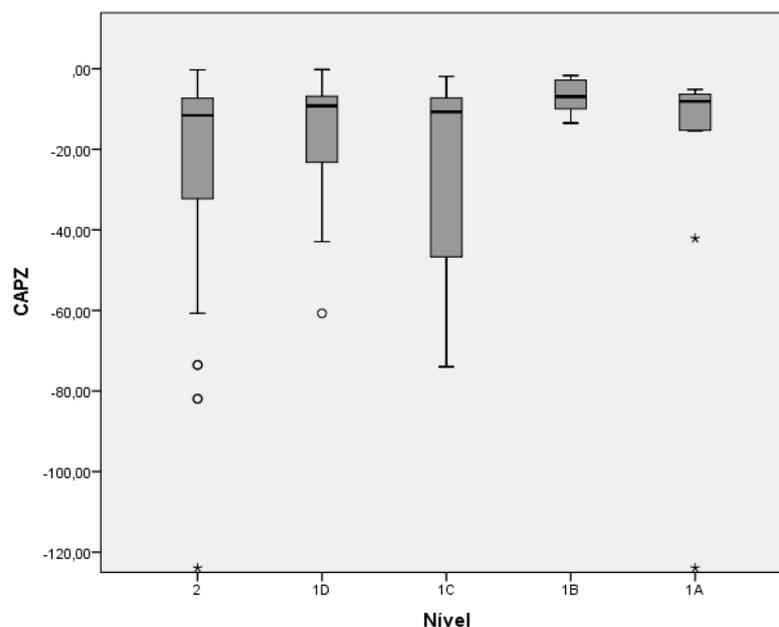


Figura 5.31 – Gráfico boxplot da CAPZ

Nas figuras é possível visualizar que o nível 1C como o de maior variabilidade e que as medianas nos vários níveis de produtividade são bem próximas.

A Tabela 5.17 mostra que as correlações entre essas três métricas são fortes. A maior correlação obtida foi entre a CANP com a CAPZ.

Tabela 5.17 - Correlações entre as métricas de centralidade de autovetor

	Correlações		
	CANP	CAPW	CAPZ
CANP	1		
CAPW	0,609**	1	
CAPZ	0,911**	0,805*	1

** . A correlação é significativa no nível 0,01 (teste bilateral)

As Figuras 5.32, 5.33 e 5.34 apresentam as dispersões da CANP, da CAPW e da CAPZ, também para melhorar a visualização foram excluídos os nós isolados e a escala logarítmica foi utilizada. Observa-se que os pontos estão concentrados ao redor das medianas, onde também se encontram dois grupos mais aparentes, a componente principal e a segunda maior componente.

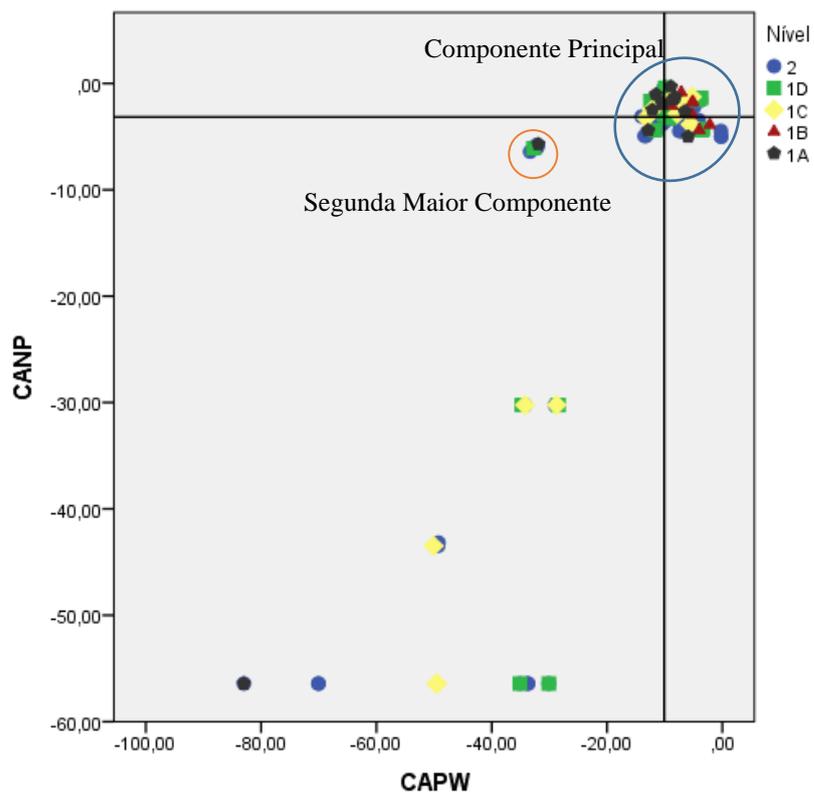


Figura 5.32 – Relação entre a CANP e a CAPW

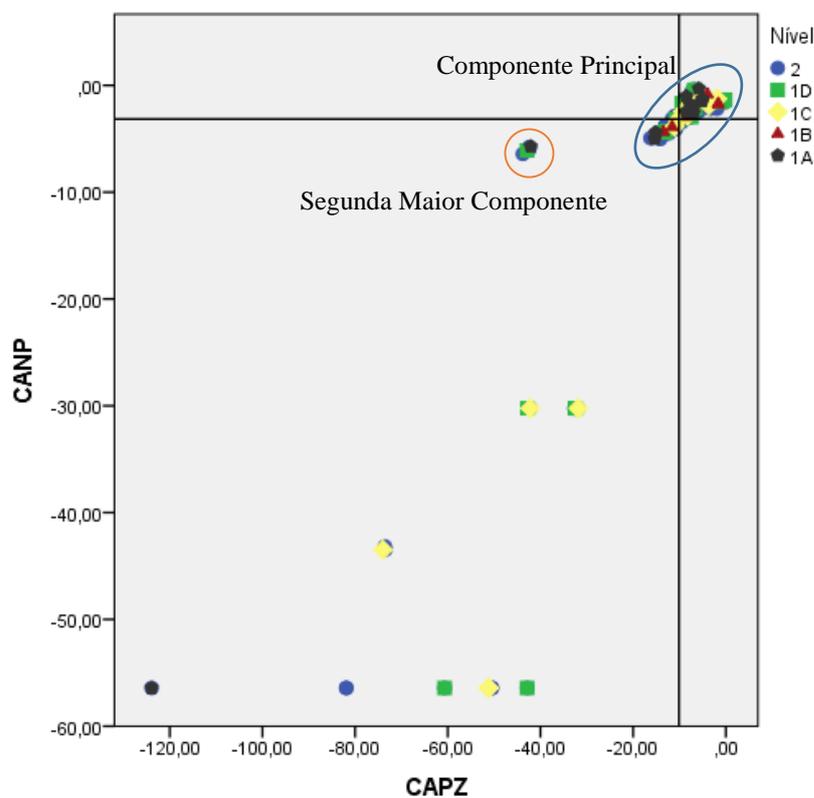


Figura 5.33 – Relação entre a CINP e a CIPZ

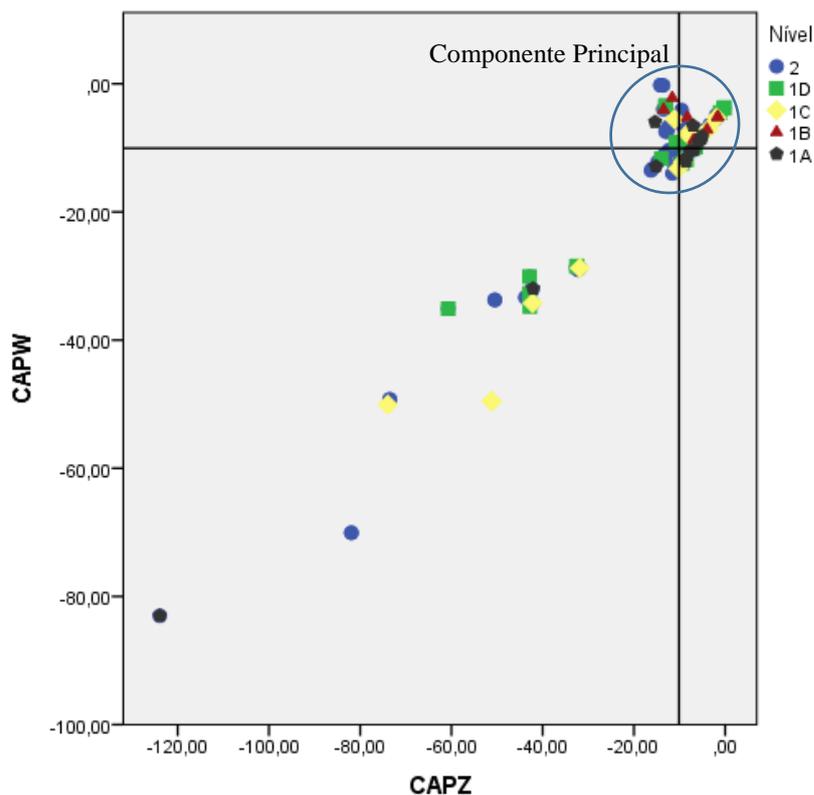


Figura 5.34 – Relação entre a CAPW e a CAPZ

As correlações entre as métricas de centralidades de autovetor e o nível de produtividade são apresentadas na Tabela 5.18.

Tabela 5.18 - Correlações das métricas de centralidade de autovetor com os níveis das bolsas

	Correlações		
	CANP	CAPW	CAPZ
Nível de produtividade	0,244**	0,200*	0,243**

*. A correlação é significativa no nível 0,05 (teste bilateral)

** . A correlação é significativa no nível 0,01 (teste bilateral)

Pela centralidade de autovetor, pesquisadores que estão ligados a pesquisadores mais centrais, de acordo com o grau, têm maiores centralidades. Desta forma: (1) pela CANP um pesquisador será mais central se estiver ligado a pesquisadores com maiores CGNP; (2) pela CAPW um pesquisador será mais central se estiver ligado a pesquisadores com maiores CGPW; (3) e pela CGPZ um pesquisador será mais central se estiver ligado a pesquisadores com maiores CGPZ. As correlações dessas métricas com o nível de produtividade foram significativas, principalmente a CANP e a CAPZ.

5.5.7 PageRank

A Tabela 5.19 apresenta os resultados dos 10 autores mais bem posicionados, de acordo com o PageRank: não-ponderado (PRNP); ponderado com os pesos das arestas (PRPW) e; ponderado com os pesos das arestas e dos nós (PRPZ). Os pesquisadores mais influentes são os de nível 1A. Eles estão presentes principalmente no PRPW e no PRPZ. Nota-se também que o *status* de um mesmo pesquisador nessas três métricas são iguais ou bem próximos.

Tabela 5.19 - Os 10 mais bem posicionados no PageRank

Posição	PRNP	PRPW	PRPZ
1	PQ124 (0,026) 1A	PQ82 (0,025) 1A	PQ124 (0,032) 1A
2	PQ84 (0,018) 1D	PQ124 (0,024) 1A	PQ82 (0,026) 1A
3	PQ114 (0,017) 1C	PQ0 (0,021) 1A	PQ0 (0,025) 1A
4	PQ62 (0,016) 1D	PQ62 (0,021) 1D	PQ62 (0,025) 1D
5	PQ65 (0,016) 1B	PQ12 (0,019) 1A	PQ12 (0,024) 1A
6	PQ0 (0,015) 1A	PQ9 (0,019) 2	PQ111 (0,022) 1A
7	PQ111 (0,014) 1A	PQ111 (0,018) 1A	PQ9 (0,019) 2
8	PQ9 (0,014) 2	PQ117 (0,017) 1D	PQ117 (0,018) 1D
9	PQ82 (0,013) 1A	PQ114 (0,015) 1C	PQ114 (0,016) 1C
10	PQ85 (0,013) 1D	PQ102 (0,014) 2	PQ41 (0,015) 1D

Para avaliar e comparar as distribuições das métricas de PageRank entre os níveis das bolsas são apresentados os gráficos boxplots nas Figuras 5.35, 5.36 e 5.37. Na Figura 5.35 nota-se os níveis 2 e 1A como os de maiores variabilidades. Nas Figuras 5.36 e 5.37 a maior variabilidade é apresentada pelo nível 1A. O Nível 1A apresenta também a maior mediana nas três métricas de PageRank.

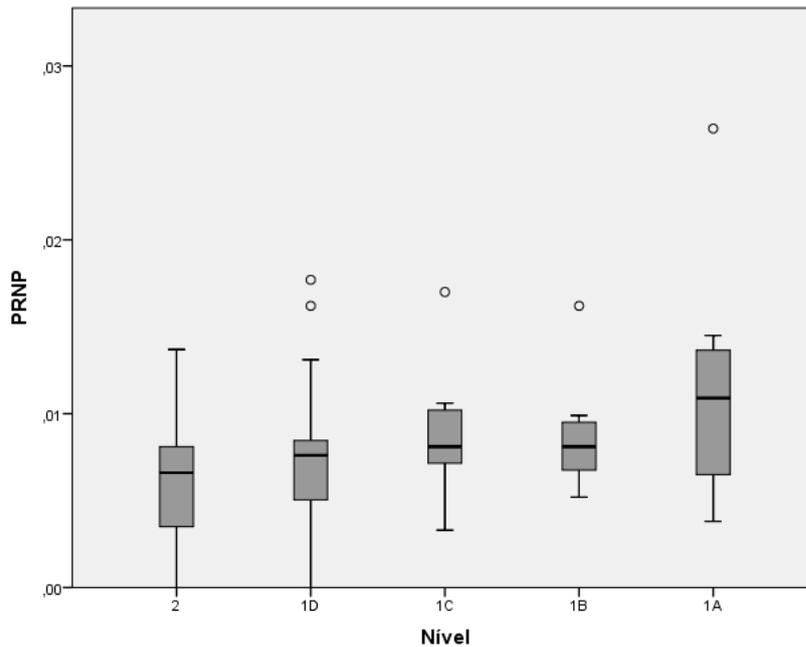


Figura 5.35 – Gráfico boxplot do PRNP

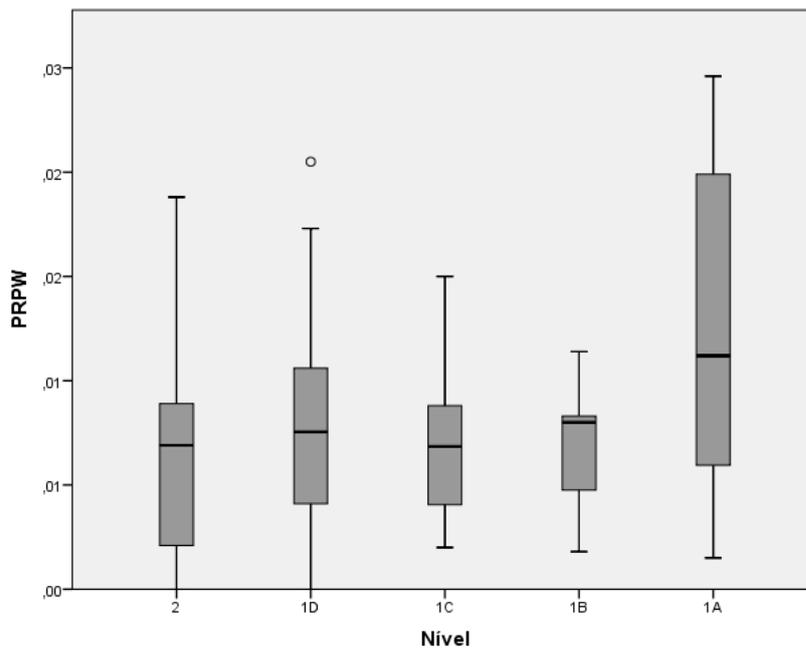


Figura 5.36 – Gráfico boxplot do PRPW

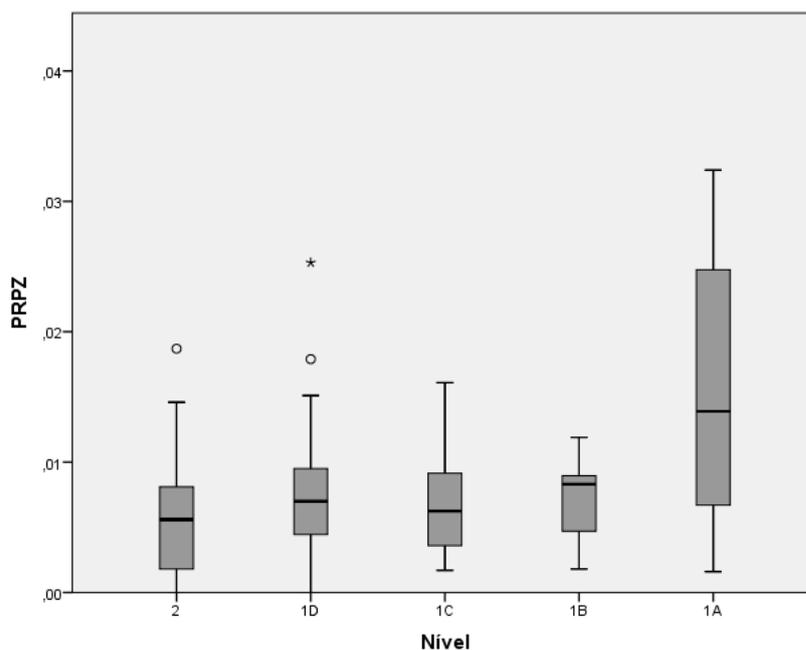


Figura 5.37 – Gráfico boxplot do PRPZ

A Tabela 5.20 mostra que as correlações entre essas três métricas são fortes. A maior correlação obtida foi entre o PRPW e o PRPZ. O que justifica a similaridade das posições dos pesquisadores nessas duas métricas.

Tabela 5.20 - Correlações entre os PageRank

	Correlações		
	PRNP	PRPW	PRPZ
PRNP	1		
PRPW	0,785**	1	
PRPZ	0,783**	0,979**	1

** . A correlação é significativa no nível 0,01 (teste bilateral)

Nas Figuras 5.38, 5.39 e 5.40 é possível visualizar as dispersões e as correlações entre as métricas de PageRank. Observa-se que os pontos estão mais concentrados em torno das medianas na Figuras 5.38 e 5.39. A relação entre PRPW e PRPZ, na Figura 5.40, aparenta ser linear e os pesquisadores nível 1A se destacam com os maiores valores.

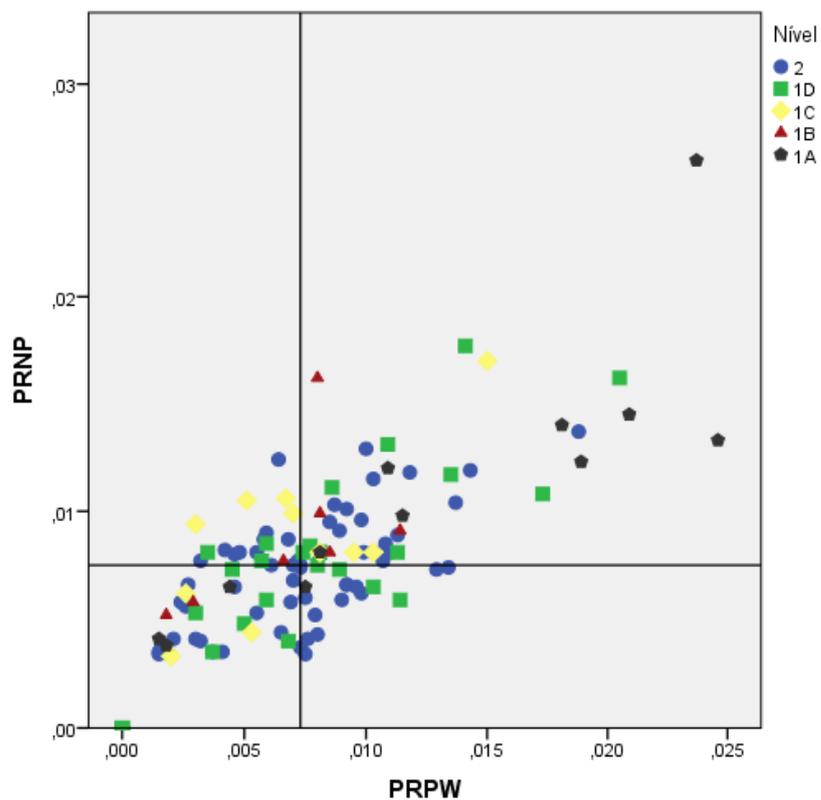


Figura 5.38 – Relação entre o PRNP e o PRPW

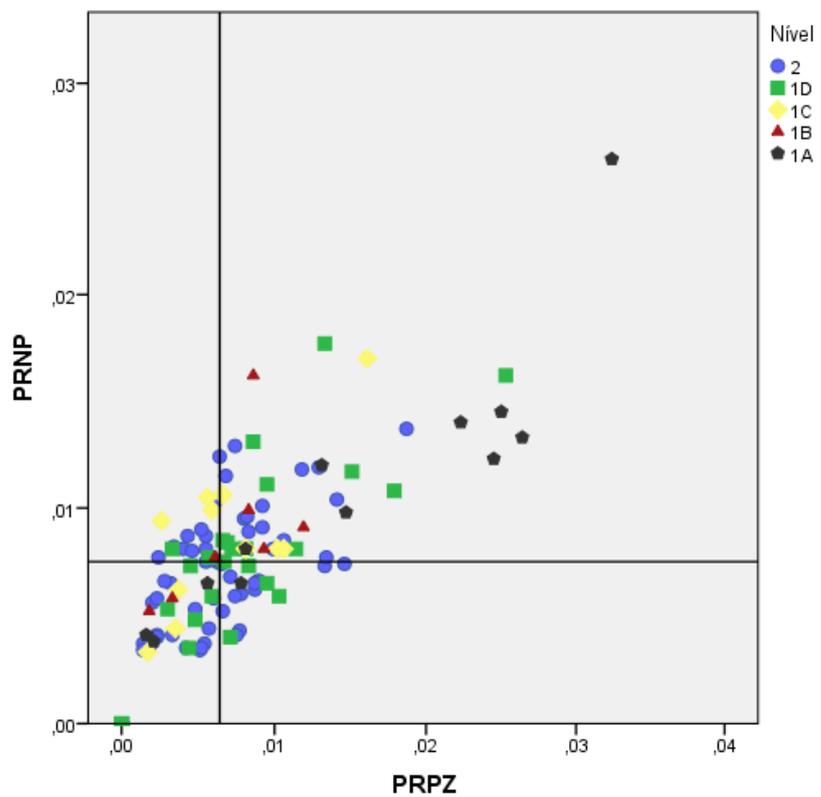


Figura 5.39 – Relação entre o PRNP e o PRPZ

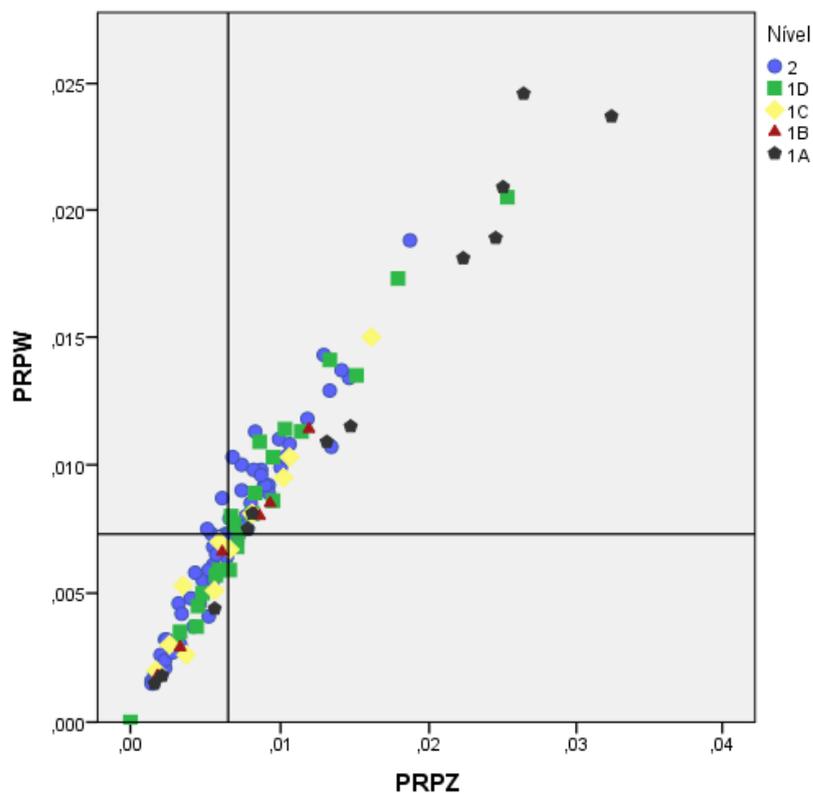


Figura 5.40 – Relação entre o PRPW e o PRPZ

A Tabela 5.21 aponta que as correlações dos PageRank com o nível de produtividade foram significativas. No entanto o *status* do pesquisador pelo PRNP tem maior impacto no nível de produtividade.

Tabela 5.21 - Correlações dos PageRank com os níveis das bolsas

	Correlações		
	PRNP	PRPW	PRPZ
Nível de produtividade	0,306**	0,199*	0,271**

*. A correlação é significativa no nível 0,05 (teste bilateral)

** . A correlação é significativa no nível 0,01 (teste bilateral)

5.5.8 Coeficiente de Cluster

A Tabela 5.22 expõe as classificações dos 10 pesquisadores mais bem posicionados de acordo com o coeficiente de *cluster*: não-ponderado (CLNP); ponderado com os pesos das arestas (CLPW) e; ponderada com os pesos das arestas e dos nós (CLPZ). A composição dessa tabela é formada basicamente por pesquisadores nível 2. E as posições deles nas três métricas são quase as mesmas.

Tabela 5.22 - Os 10 mais bem posicionados no Cluster

Posição	CLNP			CLPW			CLPZ		
1	PQ56	(1,00)	2	PQ56	(0,89)	2	PQ56	(0,31)	2
2	PQ108	(1,00)	2	PQ108	(0,89)	2	PQ108	(0,31)	2
3	PQ74	(1,00)	2	PQ74	(0,31)	2	PQ74	(0,28)	2
4	PQ134	(1,00)	2	PQ134	(0,31)	2	PQ134	(0,14)	2
5	PQ33	(1,00)	2	PQ107	(0,29)	2	PQ107	(0,10)	2
6	PQ137	(1,00)	1D	PQ137	(0,15)	1D	PQ41	(0,10)	1D
7	PQ50	(1,00)	2	PQ50	(0,14)	2	PQ95	(0,09)	1B
8	PQ70	(1,00)	2	PQ70	(0,14)	2	PQ50	(0,08)	2
9	PQ77	(1,00)	1C	PQ77	(0,14)	1C	PQ70	(0,08)	2
10	PQ66	(1,00)	2	PQ54	(0,13)	2	PQ77	(0,08)	1C

Os gráficos boxplots, nas Figuras 5.41, 5.42 e 5.43, foram utilizados para avaliar as distribuições dos coeficientes de *cluster* dos pesquisadores nos diferentes níveis de bolsa. A maior mediana e variabilidade é apresentada pelo nível 1D nas três métricas de coeficiente de *cluster*. Em termos medianos, os pesquisadores nível 1D apresentam coeficiente de *cluster*, pelas três métricas, superiores aos apresentados pelos níveis 1C, 1B e 1A.

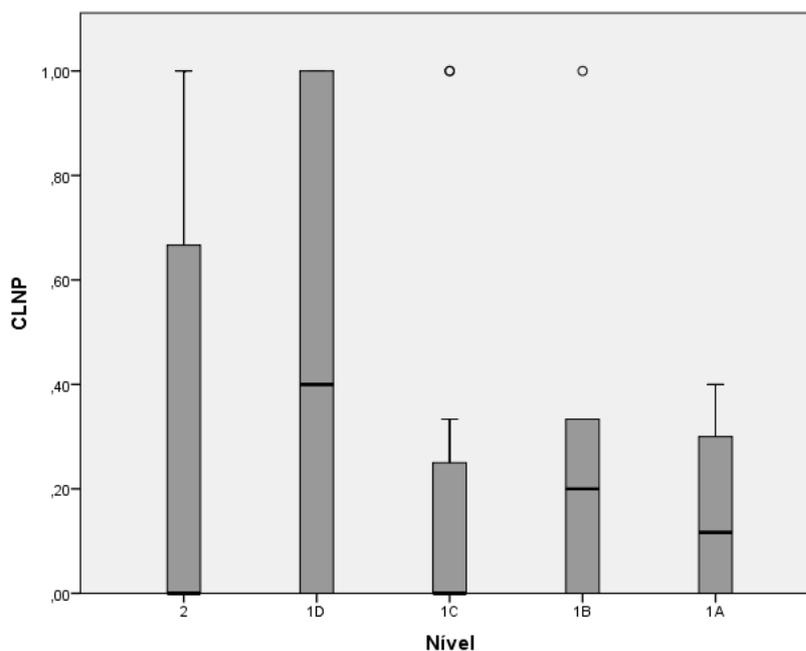


Figura 5.41 – Gráfico boxplot do CLNP

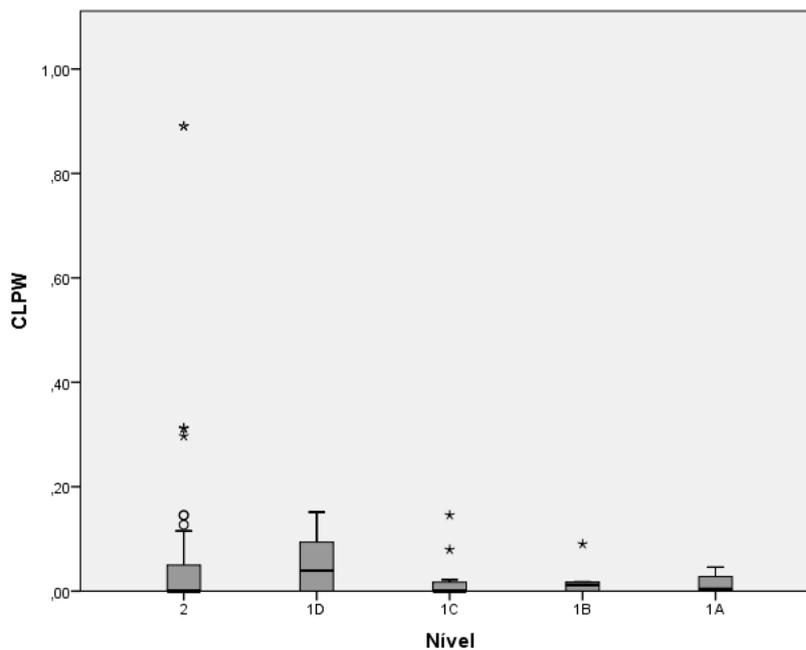


Figura 5.42 – Gráfico boxplot do CLPW

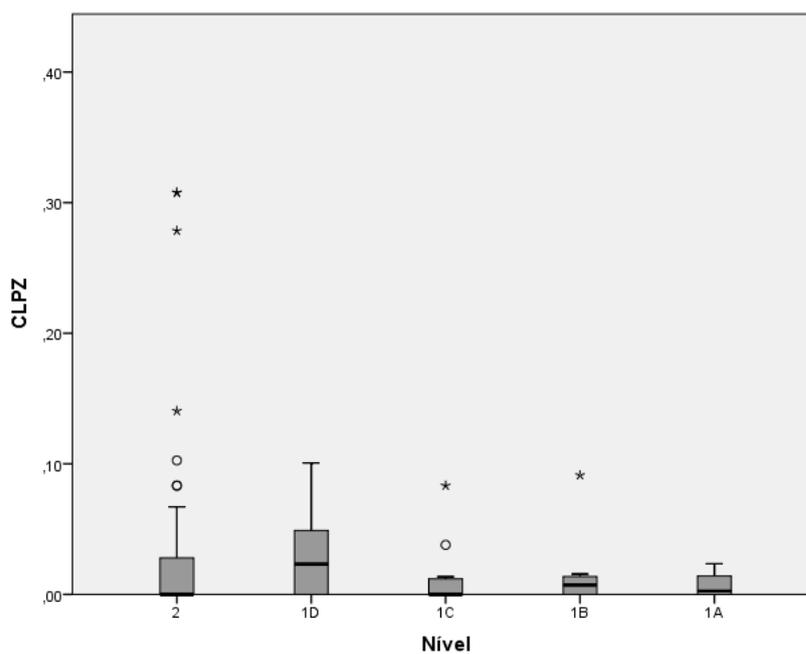


Figura 5.43 – Gráfico boxplot do CLPZ

Observando as posições dos pesquisadores na Tabela 5.22, espera-se correlações fortes entre as métricas de coeficiente de *cluster* e de fato são, como apontam a Tabela 5.23.

Tabela 5.23 - Correlações entre os Coeficientes de Cluster

	Correlações		
	CLNP	CLPW	CLPZ
CLNP	1		
CLPW	0,972**	1	
CLPZ	0,969**	0,994**	1

** . A correlação é significativa no nível 0,01 (teste bilateral)

As dispersões e as relações entre as métricas de coeficiente de cluster são apresentadas nas Figuras 5.44, 5.45 e 5.46. As medianas dessas métricas são todas iguais a zero. Os CLPW e o CLPZ, nas Figuras 5.45 e 5.46, estão concentrados entre 0 (zero) e 0,20 e 0 (zero) e 0,10, respectivamente.

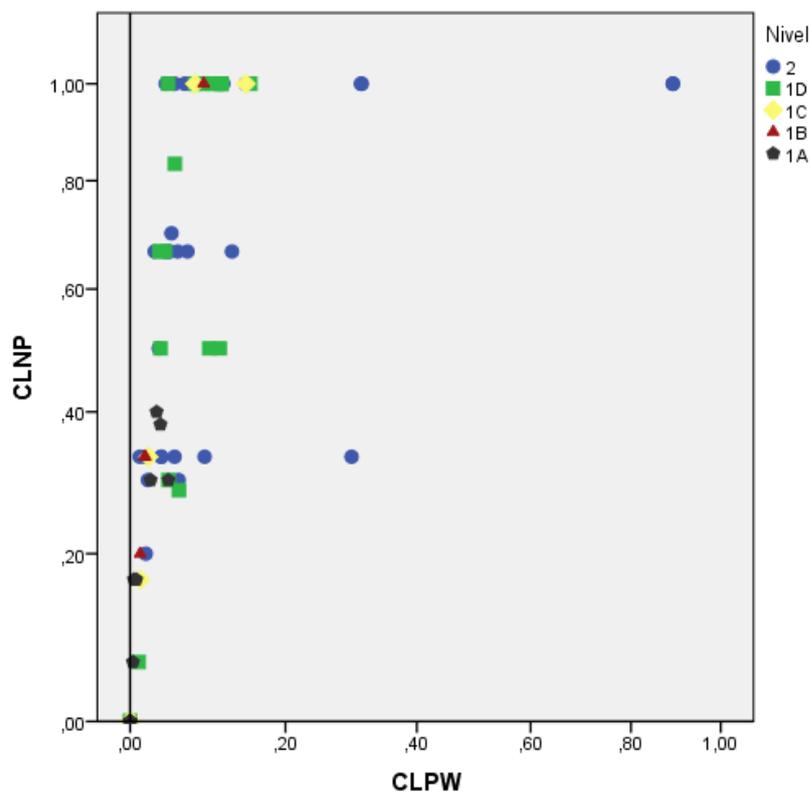


Figura 5.44 – Relação entre o CLNP e o CLPW

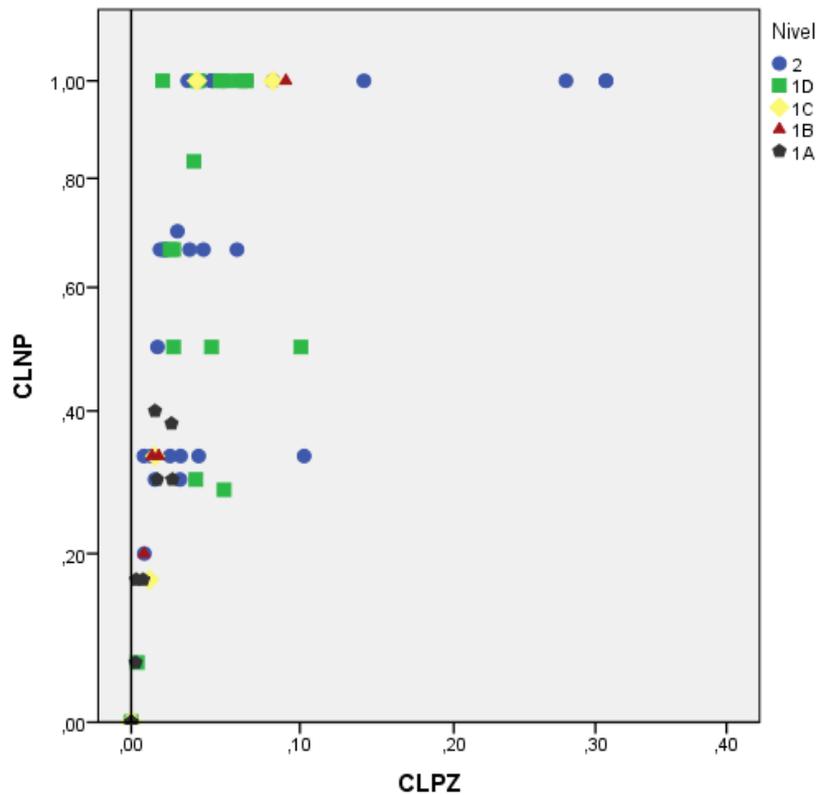


Figura 5.45 – Relação entre o CLNP e o CLPZ

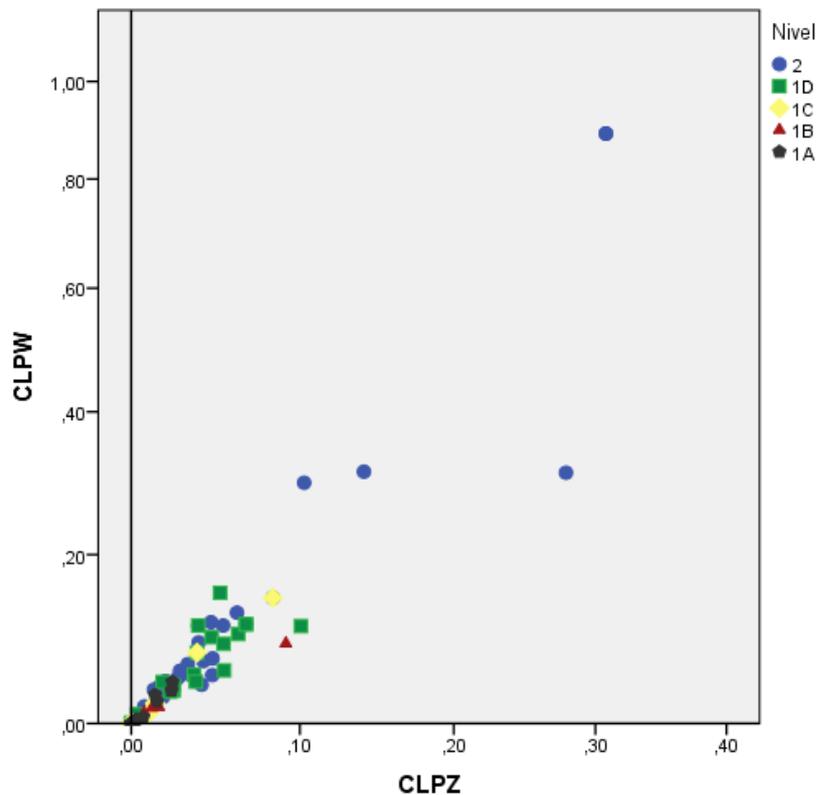


Figura 5.46 – Relação entre o CLPW e o CLPZ

O coeficiente de agrupamento de um determinado pesquisador indica o quanto os colaboradores dele estão colaborando entre si. No entanto essa característica não impacta o nível de produtividade do pesquisador, conforme apontam as correlações na Tabela 5.24.

Tabela 5.24 - Correlações das métricas de Coeficientes de Cluster com os níveis das bolsas

	Correlações		
	CLNP	CLPW	CLPZ
Nível de produtividade	0,053	0,034	0,056

5.5.9 Excentricidade

A Tabela 5.25 exhibe os 10 pesquisadores com maiores excentricidades e foram obtidas de três formas distintas: não-ponderada (ENP); ponderada com os pesos das arestas (EPW) e; ponderada com os pesos das arestas e dos nós (EPZ). Os pesquisadores nível 2 são os que mais figuram nesta tabela, portanto são os pesquisadores com maiores distâncias máxima a partir deles para qualquer outro pesquisador na rede. Como os pesos das arestas e os pesos das arestas combinadas com os pesos dos nós encurtam os caminhos, o pesquisador PQ4 que figura a primeira posição na ENP não figura na EPW possivelmente porque o caminho que liga ele ao pesquisador mais distante é formado por ligações de maior frequência. O pesquisador volta a figurar na tabela na EPZ (posição 3), desta forma, os pesquisadores que estão no maior caminho entre o PQ4 e o pesquisador mais distante são em média menos importantes.

Tabela 5.25 - Os 10 mais bem posicionados na excentricidade

Posição	ENP	EPW	EPZ
1	PQ4 (13) 2	PQ13 (8,62) 2	PQ7 (1,36) 2
2	PQ7 (13) 2	PQ34 (8,62) 1A	PQ13 (1,36) 2
3	PQ34 (13) 1A	PQ130 (8,24) 1D	PQ4 (1,33) 2
4	PQ56 (13) 2	PQ3 (8,16) 1A	PQ3 (1,32) 1A
5	PQ108 (13) 2	PQ118 (8,12) 1A	PQ29 (1,31) 2
6	PQ130 (13) 1D	PQ7 (7,74) 2	PQ18 (1,27) 2
7	PQ3 (12) 1A	PQ56 (7,64) 2	PQ113 (1,26) 2
8	PQ5 (12) 2	PQ108 (7,64) 2	PQ119 (1,24) 2
9	PQ29 (12) 2	PQ119 (7,64) 2	PQ59 (1,20) 2
10	PQ39 (12) 2	PQ51 (7,62) 1B	PQ129 (1,17) 2

Os gráficos boxplots, nas Figuras 5.47, 5.48 e 5.49, foram utilizados para avaliar as distribuições das excentricidades dos pesquisadores nos diferentes níveis de bolsa.

Considerando as três métricas, os níveis 2 e 1A apresentam as maiores variações. O nível 2 e 1D as maiores medianas.

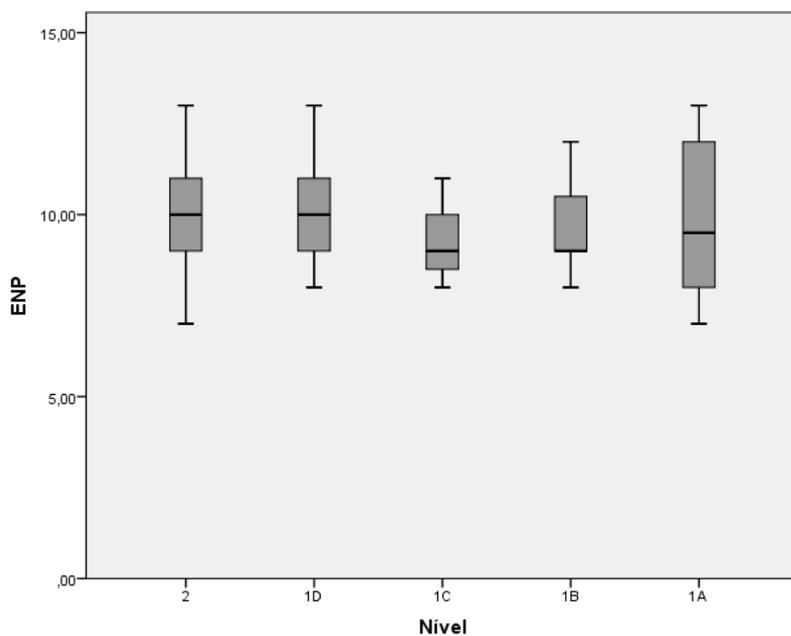


Figura 5.47 – Gráfico boxplot da ENP

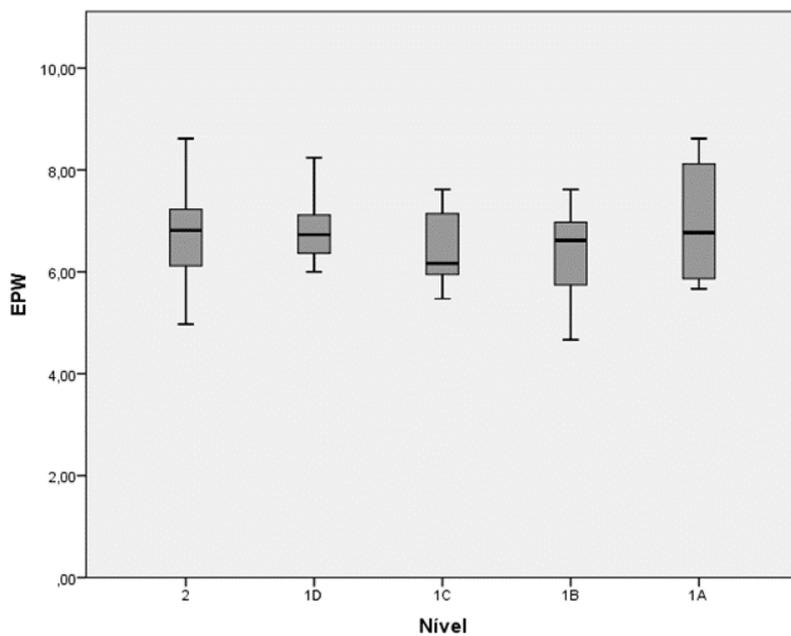


Figura 5.48 – Gráfico boxplot da EPW

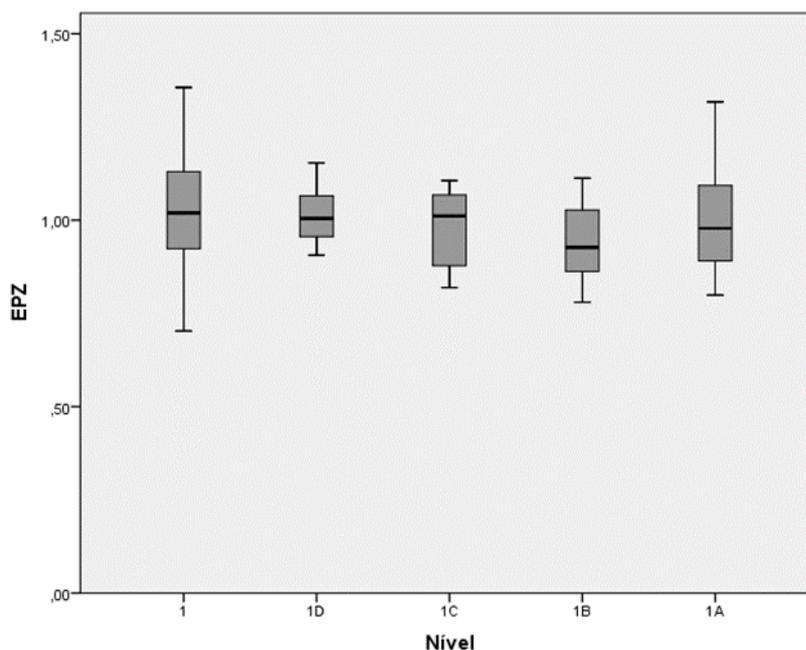


Figura 5.49 – Gráfico boxplot da EPZ

A Tabela 5.26 apontam as correlações entre as métricas de excentricidades. A maior correlação ocorreu entre a EPW e a EPZ, a menor deu-se entre ENP e a EPZ. Desta forma a EPZ foi a que mais alterou as posições dos pesquisadores em relação a ENP.

Tabela 5.26 - Correlações entre as Excentricidades

	Correlações		
	ENP	EPW	EPZ
ENP	1		
EPW	0,656**	1	
EPZ	0,440**	0,790**	1

** . A correlação é significativa no nível 0,01 (teste bilateral)

Para visualizar as dispersões das métricas de excentricidades são mostradas as Figuras 5.50, 5.51 e 5.52. Nas três figuras não é possível evidenciar uma tendência de comportamento das métricas nos diferentes níveis de bolsas.

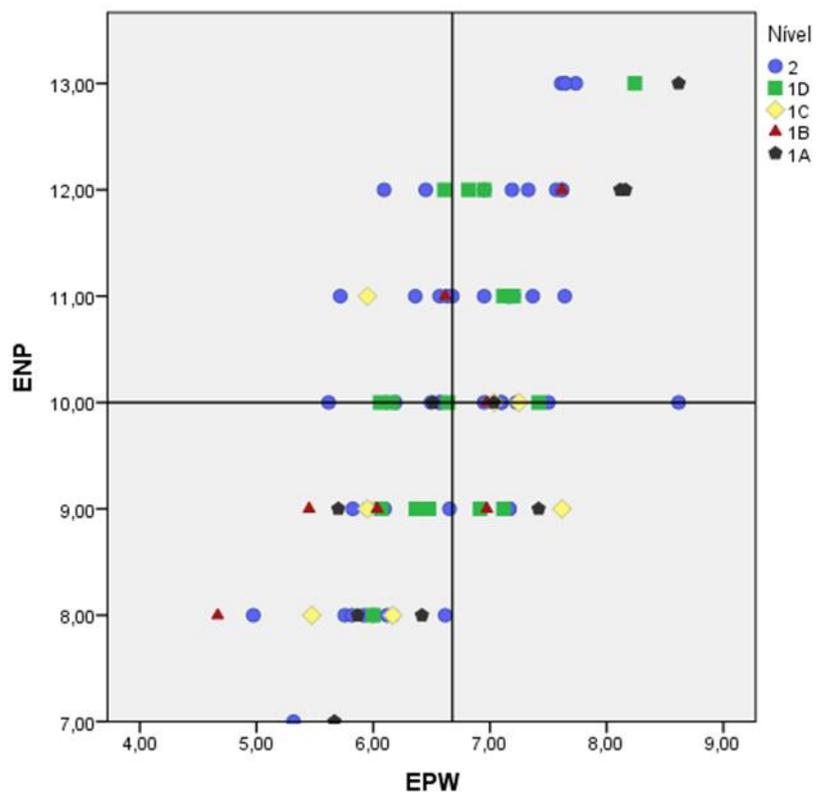


Figura 5.50 – Relação entre a ENP e a EPW

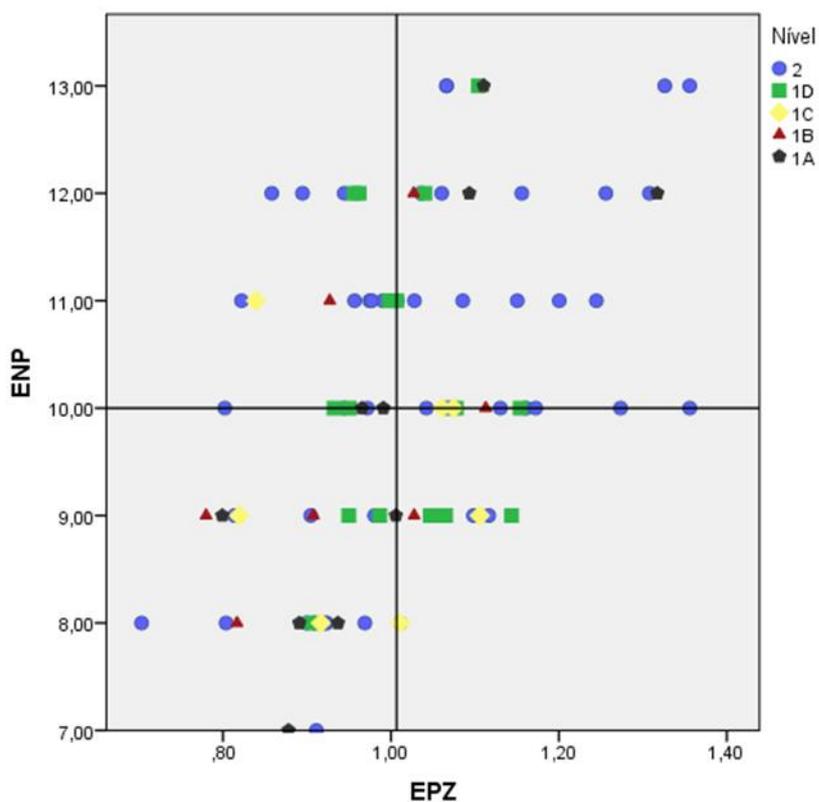


Figura 5.51 – Relação entre a ENP e a EPZ

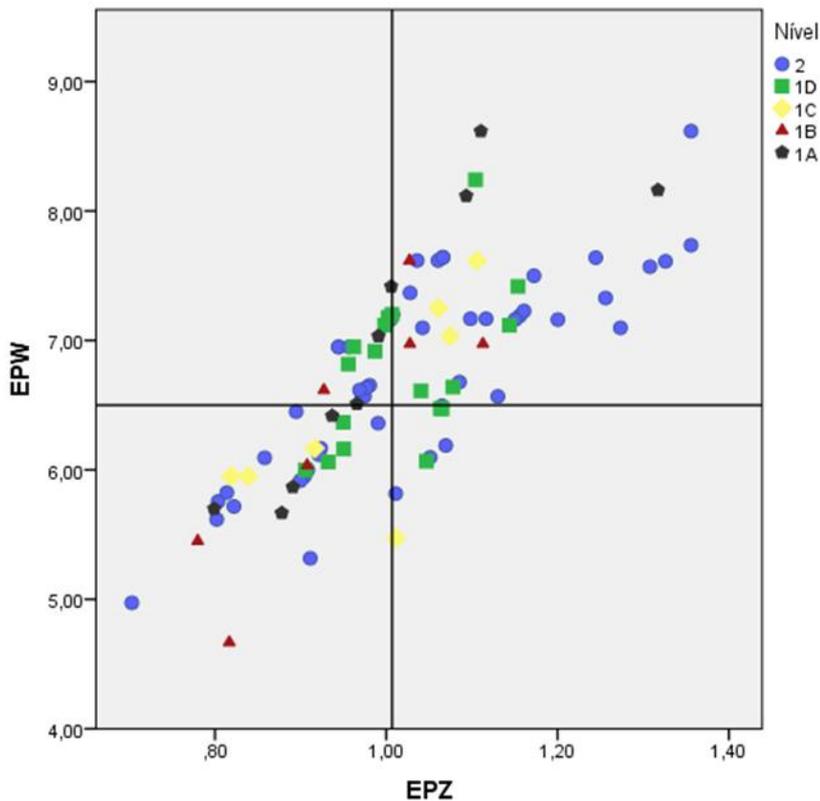


Figura 5.52 – Relação entre a EPW e a EPZ

O relacionamento de um pesquisador com outros pesquisadores é melhor quanto menor for a excentricidade. Contudo, nenhuma das métricas de excentricidade impacta significativamente o nível de produtividade, conforme Tabela 5.27.

Tabela 5.27 - Correlações das métricas de excentricidades com os níveis das bolsas

	Correlações		
	ENP	EPW	EPZ
Nível de produtividade	-0,163	-0,063	-0,154

5.5.10 Utilidade

A utilidade ou benefício de um pesquisador pertencer a estrutura da rede estudada é apresentada na Tabela 5.28, para os 10 mais bem posicionados. A utilidade foi obtida de três maneiras distintas: não-ponderada (UNP); ponderada com os pesos das arestas (UPW); e ponderada com os pesos das arestas e dos nós (UPZ). Os pesquisadores de menores níveis (2 e 1D) são os que possuem maiores benefícios na UNP, já os pesquisadores nível 1A ocupam 50% das 10 primeiras posições na UPW e UPZ. Todos os pesquisadores que figuram na UPW figuram também na UPZ.

Tabela 5.28 - Os 10 mais bem posicionados na Utilidade

Posição	UNP			UPW			UPZ		
1	PQ124	(6,32)	1A	PQ124	(6,12)	1A	PQ124	(7,49)	1A
2	PQ114	(5,50)	1C	PQ82	(5,43)	1A	PQ0	(6,32)	1A
3	PQ85	(4,21)	1D	PQ114	(5,43)	1C	PQ82	(6,01)	1A
4	PQ9	(4,10)	2	PQ0	(5,00)	1A	PQ114	(5,64)	1C
5	PQ143	(4,00)	2	PQ9	(4,80)	2	PQ111	(5,44)	1A
6	PQ84	(3,96)	1D	PQ111	(4,34)	1A	PQ12	(4,86)	1A
7	PQ62	(3,65)	1D	PQ12	(4,11)	1A	PQ9	(4,78)	2
8	PQ82	(3,64)	1A	PQ143	(4,00)	2	PQ62	(4,63)	1D
9	PQ52	(3,44)	1C	PQ62	(3,98)	1D	PQ143	(4,00)	2
10	PQ65	(3,43)	1B	PQ135	(3,97)	2	PQ135	(3,97)	2

As Figuras 5.53, 5.54 e 5.55 apresentam os gráficos boxplots das métricas de utilidade, por meio desses gráficos é possível observar a posição central e a dispersão. Considerando a UNP, Figura 5.53, o nível 1D apresenta a maior variabilidade e o nível 1A a maior mediana.

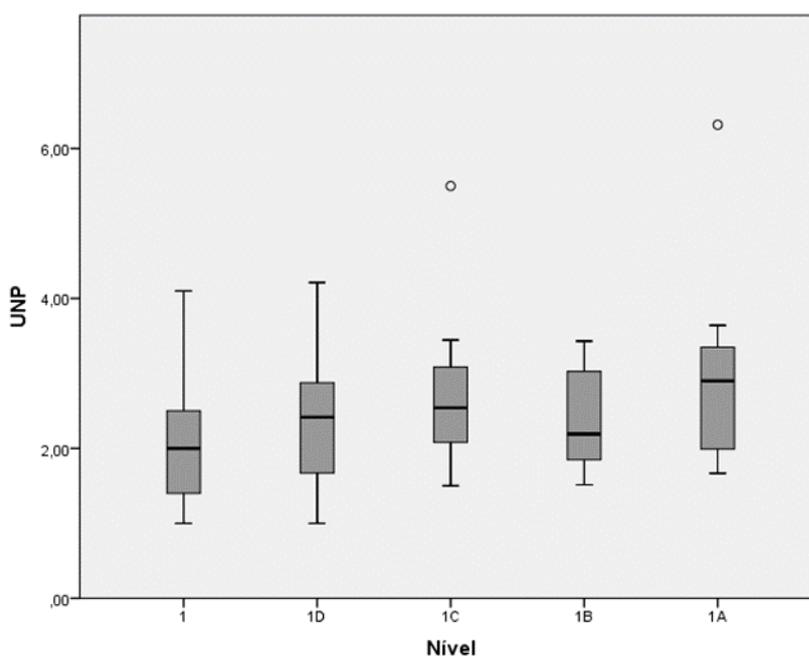


Figura 5.53 – Gráfico boxplot da UNP

Na UPW e na UPZ, Figuras 4.54 e 5.55, respectivamente, o nível 1A apresenta a maior variabilidade e a maior mediana. Observa-se que as medianas do UPW e do UPZ dos pesquisadores nível 1A são superiores aos terceiros quartis destas métricas dos pesquisadores dos outros níveis.

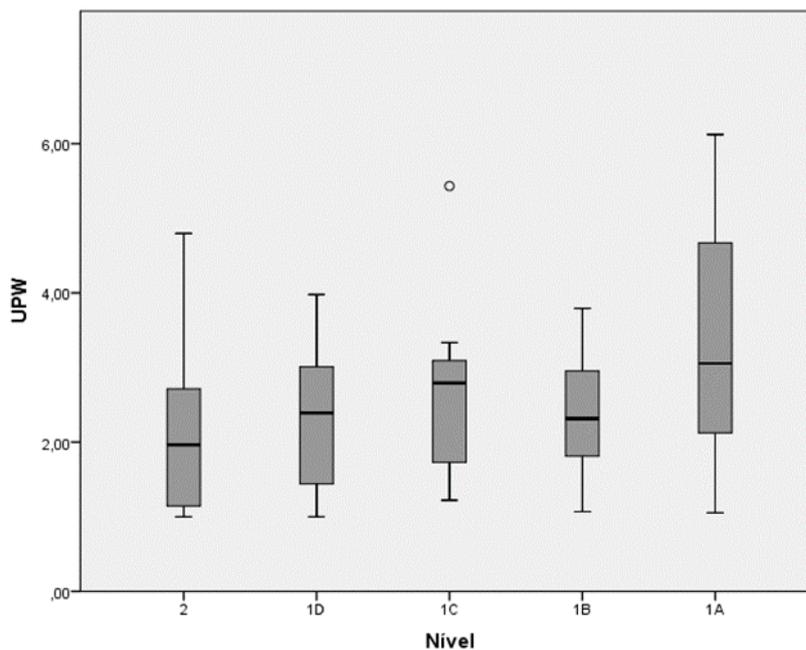


Figura 5.54 – Gráfico boxplot da UPW

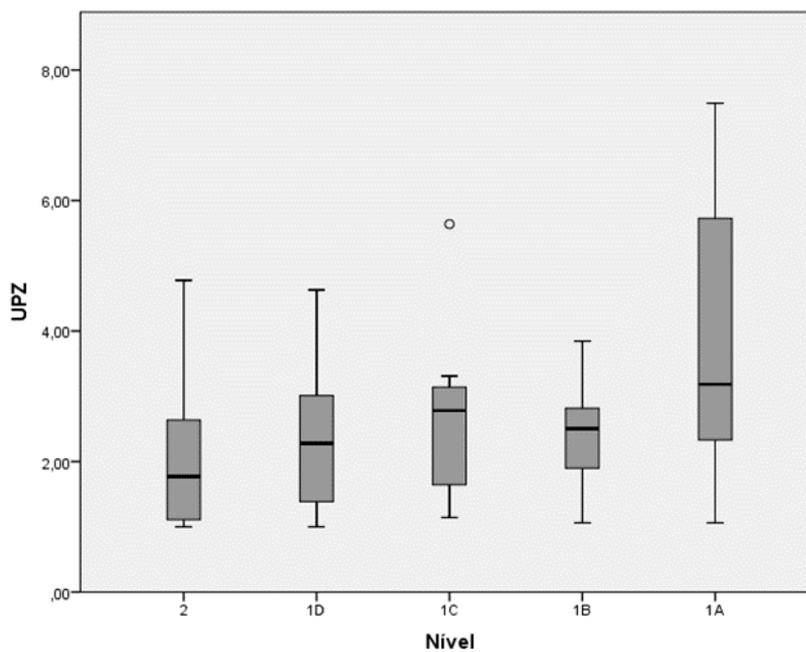


Figura 5.55 – Gráfico boxplot da UPZ

A Tabela 5.29 aponta que as correlações entre as métricas de utilidade são fortes, principalmente a UPW e a UPZ.

Tabela 5.29 - Correlações entre as utilidades

	Correlações		
	UNP	UPW	UPZ
UNP	1		
UPW	0,880**	1	
UPZ	0,874**	0,989**	1

** . A correlação é significativa no nível 0,01 (teste bilateral)

Para verificar as dispersões e as relações entre as métricas de utilidade são apresentadas as Figuras 5.56, 5.57 e 5.58. Os gráficos das Figuras 5.56 e 5.57 apresentam comportamentos similares. A linearidade é mais evidente na Figura 5.58.

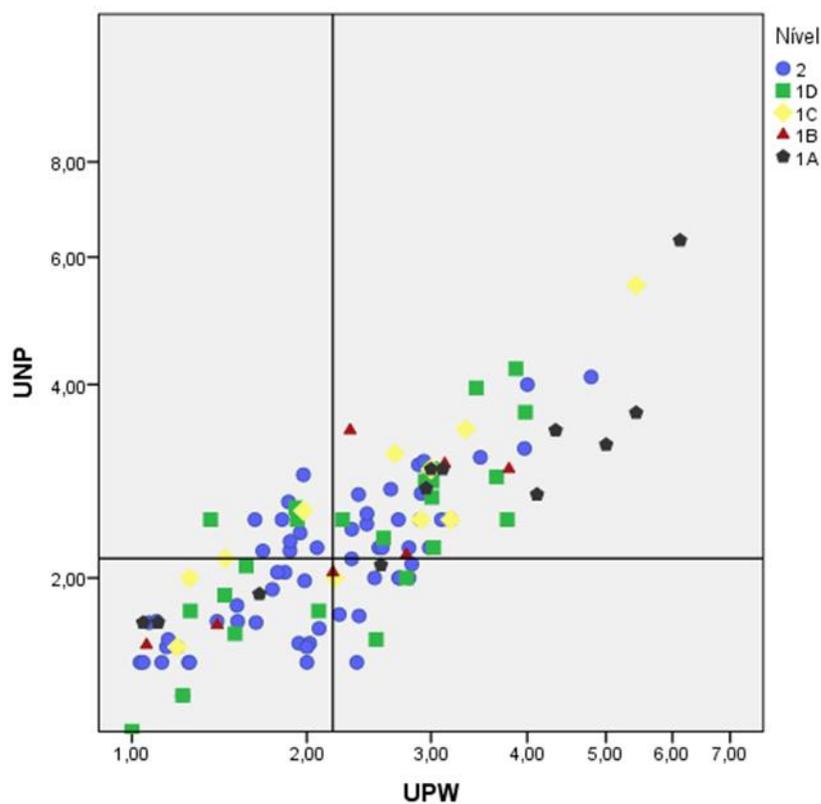


Figura 5.56 – Relação entre a UNP e a UPW

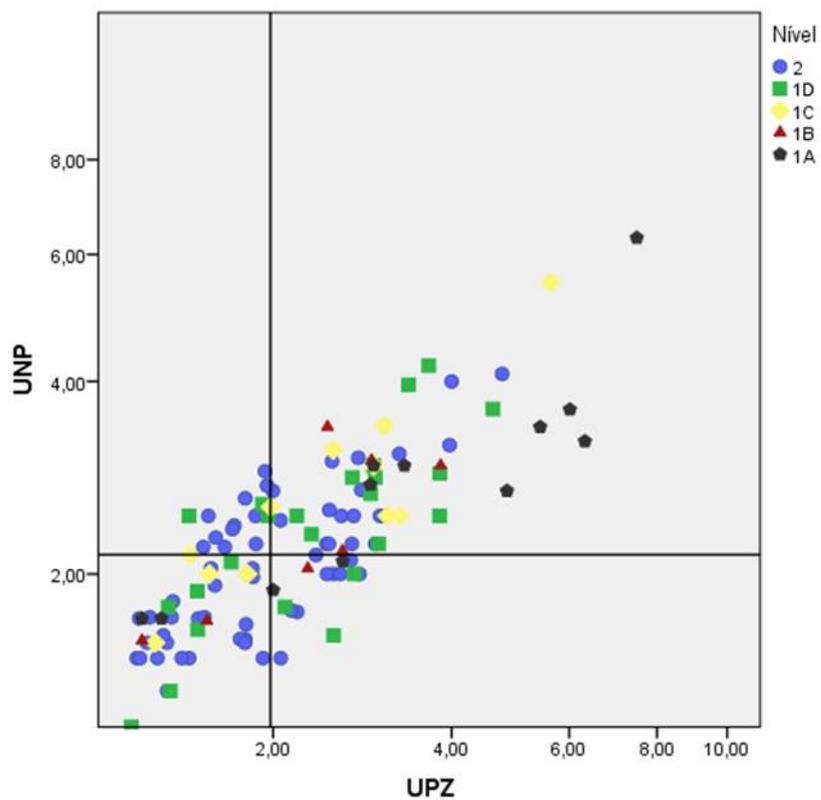


Figura 5.57 – Relação entre a UNP e a UPZ

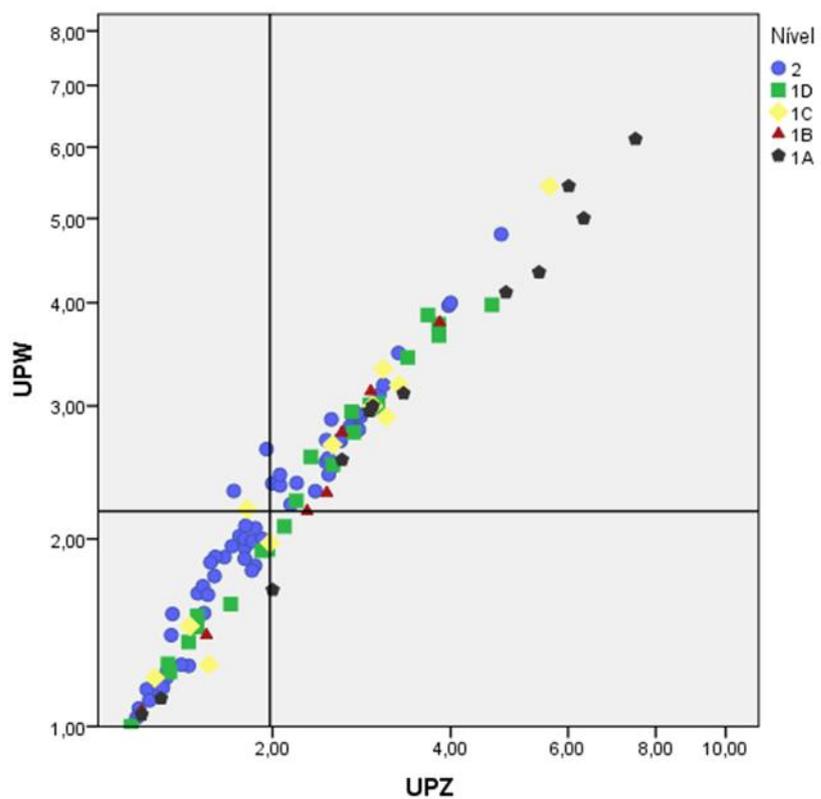


Figura 5.58 – Relação entre a UPW e a UPZ

O benefício de o pesquisador pertencer a rede impacta de maneira significativa e moderada o nível de produtividade dele, como expõe a Tabela 5.30. A UPZ obteve a maior correlação seguida pela UNP.

Tabela 5.30 - Correlações dos níveis das bolsas com a utilidade

	Correlações		
	UNP	UPW	UPZ
Nível de produtividade	0,319**	0,292**	0,327**

** . A correlação é significativa no nível 0,01 (teste bilateral)

5.6 CONCLUSÕES

A rede de coautoria construída com 145 bolsistas de produtividade em pesquisa do CNPq da área de engenharia de produção teve como resultado uma rede desconexa, com 92 pesquisadores pertencentes a maior componente conexa e separados em média por 6,0 ligações. A densidade da rede é considerada baixa, ocorrem somente 1,5% das ligações possíveis. No entanto é justificável, pois a rede considera somente um grupo restrito da produção dos pesquisadores (apenas artigos publicados em periódicos e artigos aceitos para publicações, entre 2005 a 2014) e que são bolsistas de produtividade da área de engenharia de produção, desconsiderando a colaboração com outros pesquisadores.

Foram calculadas ao todo 29 métricas individuais de ARS, divididas entre não-ponderadas, ponderadas só com os pesos das arestas e ponderadas com os pesos das arestas e dos nós. As métricas analisadas são: o *E-I index*, a centralidade de grau, a média das ligações fortes, a centralidade de proximidade, a centralidade de intermediação, a centralidade de autovetor, o PageRank, a excentricidade, o coeficiente de *cluster* e a utilidade.

As métricas não-ponderadas que apontaram maiores impactos no nível da bolsa foram: o *E-I index* (0,424), a utilidade (0,319), o PageRank (0,306), o segundo e; a terceira foi a centralidade de grau (0,287). As métricas ponderadas, só com os pesos das arestas, que apresentaram maiores impactos no nível da bolsa foram: o *E-I index_W* (0,354), a centralidade de intermediação (0,354); seguido pela utilidade (0,292) e; pela centralidade de proximidade (0,221). As métricas ponderadas, com os pesos das arestas e dos nós, que impactaram mais no nível de produtividade foram: a centralidade de intermediação (0,421); o *E-I index_Z* (0,390) obteve a segunda maior correlação, a utilidade (0,327); seguida pela centralidade de proximidade (0,284).

Desta forma, conclui-se que:

- os pesquisadores que assumem um papel de intermediador, controlando o fluxo de informações, tendem a ter maiores níveis de produtividade, principalmente, aqueles que estão entre nós cujos caminhos são formados por ligações mais frequentes e ou apresentam pesquisadores mais importantes;
- os pesquisadores de maiores *status*, pelo PageRank não-ponderado, também apresentam maiores níveis de produtividade. Se o PageRank for ponderado com os pesos das arestas e dos nós o impacto no nível de produtividade é menor;
- os pesquisadores que apresentam as maiores centralidades de grau não-ponderada, ou seja, maiores números de coautores, tendem a ter maiores níveis de produtividade. Se a centralidade de grau for ponderada essa tendência diminui;
- os pesquisadores que apresentam maiores possibilidades de estabelecer parcerias de publicações são aqueles com maiores centralidades de proximidade e também com maiores níveis de produtividade, principalmente se os parceiros forem pesquisadores com os maiores índices h .
- os pesquisadores com maiores benefícios de pertencerem a rede também apresentam os maiores níveis de produtividade, principalmente aqueles que têm índice h alto ou colaboram com pesquisadores com os maiores índices h .
- os pesquisadores que estabelecem relações com pesquisadores de níveis diferentes tendem a ter maiores níveis de bolsas. Se as relações levam em consideração os pesos das arestas, a correlação com o nível da bolsa tem uma pequena queda. Considerando os pesos das arestas e dos nós, a correlação volta a subir, mas fica abaixo da não-ponderada.

Não existiram correlações significativas entre o nível de produtividade e a CGPW, e a MLFW, e as métricas de coeficiente de cluster, e as métricas de excentricidade. Em geral, as correlações entre as mesmas métricas (ponderadas e não-ponderadas) foram elevadas, não ocorrendo modificações significativas nas posições dos pesquisadores (nas tabelas apresentadas), com exceção a centralidade de autovetor.

6 REDES DE COAUTORIA ENTRE OS PROFESSORES PERMANENTES

6.1 INTRODUÇÃO

As redes de coautoria evoluem ao longo do tempo e o estudo dos aspectos estruturais e das dinâmicas (evoluções) dessas redes estão, segundo Digiampietri *et al.* (2012), recebendo grande interesse pela comunidade científica. Permitindo, aprofundar e desvendar conhecimento pertinente ao comportamento social entre pesquisadores ou grupos acadêmicos.

Trabalhos recentes buscaram analisar o desempenho de grupos acadêmicos com base na estrutura e dinâmica de relacionamento entre os pesquisadores. O estudo de Mello (2010) analisou a evolução de redes de coautoria entre professores vinculados a Programas de Pós-Graduação em Administração, e os resultados apontaram aumento na cooperação de coautoria de um triênio para outro (2001-2003 e 2004-2006), provavelmente induzidos pelo aumento da aquiescência dos programas em relação às exigências do órgão de avaliação, CAPES.

Os resultados do trabalho de caracterização dos Programas de Pós-Graduação em Ciência da Computação no Brasil e as relações entre eles, de Digiampietri *et al.* (2014), apontam que os programas com maiores densidades nas redes de coautoria também possuem maiores produtividade em pesquisa, entre 2004 e 2009. Martins *et al.* (2014) analisou a influência da rede de coautoria na performance dos programas de pós-graduação no sistema de avaliação da CAPES, concluindo que há uma relação positiva entre a formação de grupos de coautoria e a pontuação obtida pelos programas estudados.

Amorim (2014) abordou análises das redes de colaboração entre Programas de Doutorado, na área de Estatística no Brasil e mostrou que de forma concomitante a evolução das redes dos programas, entre 2004 a 2012, ocorreram também uma evolução nas medidas das redes, bem como o número de colaborações.

O objetivo deste capítulo é analisar se existe uma relação entre a evolução da performance dos Programas de Pós-graduação em Engenharia de Produção (PPGEP) e a evolução das redes de coautoria entre os docentes dos programas, das redes de coautoria entre esses programas de pós-graduação e das redes de coautoria entre todos os professores dos programas.

O presente capítulo está estruturado em cinco subcapítulos. O primeiro a introdução, onde é abordado uma revisão de trabalhos que analisaram as evoluções de redes de coautorias. Em seguida, o Subcapítulo 6.2 faz uma breve descrição sobre o sistema de avaliação dos programas de pós-graduação. O Subcapítulo 6.3 descreve os procedimentos metodológicos. O

Subcapítulo 6.4 apresenta as métricas de ARS aplicadas as redes de coautoria e os impactos nos conceitos dos programas. O Subcapítulo 6.5 apresenta as conclusões e finaliza o capítulo.

6.2 SISTEMA DE AVALIAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO

A Avaliação do Sistema Nacional de Pós-Graduação (SNPG), atividade essencial para assegurar e manter a qualidade dos cursos de Mestrado e Doutorado no país, entrou em vigor a partir de 1998. A avaliação do SNPG é orientada pela Diretoria de Avaliação/CAPES e realizada com a participação da comunidade acadêmico-científica por meio de consultores *ad hoc* (CAPES, 2014).

O sistema de avaliação abrange dois processos (CAPES, 2014):

- Acompanhamento Anual – realizado no período compreendido entre os anos de realização das avaliações trienais. O acompanhamento anual tem por objetivo promover um diálogo entre a CAPES e as instituições promotoras de cursos de mestrado e doutorado com o intuito de orientar a atuação dos programas de forma que possam elevar a qualidade do desempenho e identificar possíveis problemas que podem ser corrigidos antes da Avaliação Trienal subsequente.
- Avaliação Trienal – é realizada ao final de cada triênio, e a realização é estabelecida pela sequência histórica do processo de avaliação CAPES. Os resultados da avaliação de cada programa são apresentados na “Ficha de Avaliação” definida pelo Conselho Técnico-Científico (CTC), os vários quesitos e itens avaliados, os atributos a ele consignados, como os respectivos comentários e justificativas da comissão organizadora, e, no final, é atribuído o conceito correspondente ao desempenho no triênio, os conceitos variam de 1 a 7 ordem crescente.

No dia 11 de dezembro de 2014, o Conselho Superior do CAPES decidiu que a avaliação do Sistema Nacional de Pós-Graduação (SNPG), a partir da próxima edição, passa a ser quadrienal. O Conselho também aprovou a proposta de realização, no meio do período quadrienal, de uma análise que aponte tendências dos programas de pós-graduação.

Deve ser descredenciado do SNPG o curso que obtiver conceito inferior a 3, enquanto os programas que obtiverem conceito igual ou superior a 3 recebem renovação do “reconhecimento” garantido pelo triênio subsequente. Os conceitos superiores a 5 são atribuídos somente aos programas de mestrado e doutorando, sendo que os programas que

obtiverem os conceitos 6 ou 7, devem representar a excelência na Área, reconhecidos como programas de nível internacional.

Entre os critérios avaliados pelo SNPG: Proposta do Programa (10%); Corpo Docente (20%); Corpo Discente, Teses e Dissertações (25%); Produção Intelectual (35%) e Inserção Social (10%). A Produção Intelectual apresenta o maior percentual e são observados e avaliados em termos de atributos: publicações qualificadas do Programa por docentes permanentes (50%); distribuição de publicações qualificadas em relação ao corpo docente permanente do Programa (30%) e; produções técnicas, patentes e outras produções consideradas relevantes (20%). Deste modo, percebe-se que a produção científica é um importante critério na avaliação dos programas de pós-graduação.

6.3 METODOLOGIA

Para construção das redes de coautoria para cada PPGEP avaliados pela CAPES nos triênios 2007-2009 e 2010-2012, foram consideradas somente as publicações em periódicos dos professores permanentes em cada programa. Também, foram construídas uma rede para cada triênio contendo todos os professores permanentes dos PPGEP e uma rede formada por estes programas, considerando que dois programas são relacionados se existe alguma coautoria entre docentes permanentes destes programas.

Para construir essas redes de coautoria foram adotadas as seguintes etapas:

- Etapa 1 - identificação dos PPGEP avaliados pela CAPES nos triênios 2007-2009 e 2010-2012 e os respectivos conceitos a cada triênio;
- Etapa 2 - identificação dos professores permanentes, associados a cada PPGEP;
- Etapa 3 - identificação dos currículos Lattes de cada professor associado a cada programa e identificação dos índices h;
- Etapa 4 - extração das produções acadêmicas de cada professor;
- Etapa 5 - identificação das publicações realizadas em coautoria;
- Etapa 6 – produção das redes de coautoria;
- Etapa 7 – cálculo das métricas de ARS.

Os PPGEP avaliados nos triênios 2007-2009 e 2010-2012 foram extraídos das Relação de Cursos Recomendados e Reconhecidos pela CAPES correspondentes à área de Engenharia de Produção e que tenha a nomenclatura Engenharia de Produção, excluindo programas que embora relacionados a engenharia de produção possuam outras nomenclaturas e extensões, a

exemplos de Engenharia de Produção e Sistemas e de Engenharia Industrial. Ao todo, foram identificados 20PPGEP avaliados nos triênios 2007-2009 e 20 programas avaliados no triênio 2010-2012 (CAPES, 2015).

Para cada triênio e cada PPGEP foram identificados os nomes completos de todos os professores, na modalidade de professores permanentes, registrados nos Indicadores de Corpo Docente - Atuação (DA) pertencentes aos Cadernos de Indicadores da CAPES. Ao todo, foram identificados 334 professores correspondentes aos PPGEP avaliados no triênio 2007-2009, e também 334 professores correspondentes ao triênio 2010-2012 (CAPES, 2015).

A Etapa 3, identificação dos currículos Lattes, os IDs Lattes foram obtidos diretamente nos currículos Lattes dos professores. Os índices h foram obtidos no banco de dados da *Scopus*. Na Etapa 4 e na Etapa 5 foi utilizado o scriptLattes, cuja descrição se encontra no Subcapítulo 2.1. No arquivo texto, fonte de entrada para a ferramenta, incluía em cada linha, nesta ordem e separado por vírgula: o ID; o nome do professor; o tempo de permanência do professor no Programa, em cada triênio; e o Programa que ele pertence como rótulo.

Na Etapa 6, as relações de coautoria da Etapa 5 foram consideradas: na construção das redes para cada PPGEP, as relações de coautoria dos professores permanentes dos PPGEP; na construção da rede dos PPGEP, as relações de coautoria entre os professores permanentes aos PPGEP e; na construção da rede contendo todos os professores permanentes dos PPGEP, as relações de coautoria entre todos os professores permanentes.

No total foram construídas 44 redes de coautoria, sendo 22 redes por cada triênio. As métricas globais de análise de redes sociais, dessas redes, foram encontradas pelo software Gephi, as métricas individuais pelo NetworkX. O *E-I index* (Krackhardt e Stern, 1998) foi calculado no UNICET por meio da rede de coautoria entre todos os professores permanentes, considerando como relações externas, as coautorias com professores de PPGEP diferentes.

6.4 REDES DE COAUTORIAS

Com as relações de coautoria dos professores permanentes foram construídas três conjuntos de redes: 40 redes entre professores (uma para cada programa/triênio); 2 redes de coautoria entre os programas (uma por triênio) e; 2 redes entre os professores de todos os programas (uma por triênio). A Tabela 6.1 mostra os conceitos dos PPGEP nos triênios 2007-2009 e 2010-2012 e a variação.

Tabela 6.1- Conceitos dos PPGEP

UF	INSTITUIÇÃO	SIGLA	2007-2009		2010-2012		VARIAÇÃO
RJ	PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO	PUC-RIO	M/D	4	M/D	5	1
SP	UNIVERSIDADE DE SÃO PAULO	USP/SC	M/D	5	M/D	5	0
SP	UNIVERSIDADE EST.PAULISTA JÚLIO DE MESQUITA FILHO/BAURU	UNESP/BAU	M	3	M	4	1
RJ	UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY RIBEIRO	UENF	M	3	M	3	0
PB	UNIVERSIDADE FEDERAL DA PARAÍBA/JOÃO PESSOA	UFPB/J.P.	M	4	M	3	-1
MG	UNIVERSIDADE FEDERAL DE ITAJUBÁ	UNIFEI	M	4	M/D	5	1
MG	UNIVERSIDADE FEDERAL DE MINAS GERAIS	UFMG	M/D	4	M/D	5	1
PE	UNIVERSIDADE FEDERAL DE PERNAMBUCO	UFPE	M/D	5	M/D	6	1
SC	UNIVERSIDADE FEDERAL DE SANTA CATARINA	UFSC	M/D	4	M/D	5	1
RS	UNIVERSIDADE FEDERAL DE SANTA MARIA	UFSM	M	3	M	3	0
SP	UNIVERSIDADE FEDERAL DE SÃO CARLOS	UFSCAR	M/D	5	M/D	4	-1
RJ	UNIVERSIDADE FEDERAL DO RIO DE JANEIRO	UFRJ	M/D	6	M/D	5	-1
RN	UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE	UFRN	M	3	M	3	0
RS	UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL	UFRGS	M/D	5	M/D	6	1
RJ	UNIVERSIDADE FEDERAL FLUMINENSE	UFF	M/D	4	M/D	4	0
SP	UNIVERSIDADE METODISTA DE PIRACICABA	UNIMEP	M/D	3	M/D	3	0
SP	UNIVERSIDADE NOVE DE JULHO	UNINOVE	M	3	M	4	1
SP	UNIVERSIDADE PAULISTA	UNIP	M/D	4	M/D	5	1
PR	UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ	UTFPR	M	4	M/D	4	0
SP	UNIVERSIDADE DE SÃO PAULO	USP	M/D	5	M/D	4	-1

M – Mestrado; D – Doutorado

No triênio 2007-2009 existiam 12 programas de doutorado, no triênio seguinte 14 programas tinham doutorados. Nove programas tiveram aumento nos conceitos, sete permaneceram com os mesmos conceitos e quatro foram rebaixados. A Tabela 6.2 apresenta a quantidade de programas com os mesmos conceitos por triênio.

Tabela 6.2 - Distribuições dos conceitos dos programas

2007-2009		2010-2012	
Número de programas	Conceito	Número de programas	Conceito
6	3	5	3
8	4	6	4
5	5	7	5
1	6	2	6

A maioria dos programas (40%) tiveram conceito 4 no primeiro triênio estudado e no segundo triênio a maioria (35%) obtiveram conceito 5. Em média o conceito dos programas no triênio de 2007-2009 foi 4,05, no triênio seguinte a média passou a ser 4,3. No primeiro triênio, somente, um programa obteve conceito 6 (UFRJ) no segundo triênio dois programas obtiveram conceito 6 (UFPE e UFRGS).

6.4.1 Redes de Coautorias dos Programas de Pós-Graduação em Engenharia de Produção

Antes de verificar se a evolução das redes de coautoria de cada programa impacta os conceitos dos mesmos, foi verificado o impacto da quantidade de publicações na evolução. As seguintes medidas foram correlacionadas com os conceitos dos programas: número de artigos totais; número de artigos em coautoria; número de artigos por autor; número de artigos em coautoria por autor.

As Tabelas 6.3 e 6.4 exibem dados referentes as produções bibliográficas dos programas. As três instituições cujos PPGE mais publicaram no triênio 2007-2009 foram a UFSCAR, a UFSC e a USP, os que mais colaboraram (colaboração entre os professores do mesmo programa), nesse triênio, foram a USP, a UNIFEI e a UFSCAR. No triênio seguinte os três programas que mais publicaram foram a UFSC, a UFSCAR e a UTFPR. Os mais colaborativos foram a UFSC, UNINOVE, UNIFEI e UNIP.

Em média cada PPGE publicou 83,05 artigos no primeiro triênio e 139,4 no segundo triênio. A média de publicações em coautoria também aumentou de 9,55 para 24,74. No triênio 2010-2012 o PPGE da UFSC publicou 251 artigos a mais que o triênio anterior, a maior variação entre os PPGE. A maior variação de artigos publicados em coautoria também foi da UFSC.

Tabela 6.3 - Dados de publicações

Instituição	Total de Artigos			Artigos em coautoria		
	07-09	10-12	Variação	07-09	10-12	Variação
PUC/RIO	36	72	36	4	8	4
UENF	30	54	24	2	1	-1
UFF	102	161	59	10	20	10
UFMG	55	83	28	4	22	18
UFPB	66	118	52	4	19	15
UFPE	72	77	5	13	15	2
UFRGS	93	178	85	15	24	9
UFRJ	90	86	-4	1	4	3
UFRN	27	76	49	0	33	33
UFSC	192	443	251	13	94	81
UFSCAR	198	199	1	19	8	-11
UFSM	80	169	89	4	12	8
UNESP/BAU	61	110	49	3	29	26
UNIFEI	42	95	53	21	48	27
UNIMEP	26	64	38	0	2	2
UNINOVE	33	161	128	4	54	50
UNIP	66	144	78	15	48	33
USP	156	163	7	24	8	-16
USP/SC	140	138	-2	16	10	-6
UTFPR	96	197	101	19	36	17

Ao analisar a relação entre o número de publicações e a quantidade de professores, as maiores relações são encontradas na USP/SC, na UFSC e na UTFPR, no triênio 2007-2009. No triênio seguinte as maiores relações foram encontradas na UTFPR, na UFSC e na UFRGS. Considerando o número de artigos em coautoria por autor, encontram-se a UNIFEI, a UTFPR, UFRGS e a UNIP entre as maiores no primeiro triênio. No segundo as maiores relações são encontradas na UNIP, na UNIFEI e na UFSC. A maioria das variações dessas medidas foram positivas. Mostrando um aumento dos números de publicações totais e em coautorias por professores.

Tabela 6.4 - Relação entre publicações e número de professores

Instituição	N. de professores			Artigos/Autor			Coautoria/Autor		
	07-09	10-12	Variação	07-09	10-12	Variação	07-09	10-12	Variação
PUC/RIO	13	12	-1	2,77	6,00	3,23	0,31	0,67	0,36
UENF	12	11	-1	2,50	4,91	2,41	0,17	0,09	-0,08
UFF	20	21	1	5,10	7,67	2,57	0,50	0,95	0,45
UFMG	17	19	2	3,24	4,37	1,13	0,24	1,16	0,92

UFPB	14	15	1	4,71	7,87	3,15	0,29	1,27	0,98
UFPE	15	14	-1	4,80	5,50	0,70	0,87	1,07	0,20
UFRGS	13	13	0	7,15	13,69	6,54	1,15	1,85	0,69
UFRJ	23	19	-4	3,91	4,53	0,61	0,04	0,21	0,17
UFRN	13	13	0	2,08	5,85	3,77	0,00	2,54	2,54
UFSC	25	27	2	7,68	16,41	8,73	0,52	3,48	2,96
UFSCAR	32	25	-7	6,19	7,96	1,77	0,59	0,32	-0,27
UFSM	18	19	1	4,44	8,89	4,45	0,22	0,63	0,41
UNESP/BAU	17	15	-2	3,59	7,33	3,75	0,18	1,93	1,76
UNIFEI	11	13	2	3,82	7,31	3,49	1,91	3,69	1,78
UNIMEP	10	12	2	2,60	5,33	2,73	0,00	0,17	0,17
UNINOVE	11	16	5	3,00	10,06	7,06	0,36	3,38	3,01
UNIP	13	12	-1	5,08	12,00	6,92	1,15	4,00	2,85
USP	29	29	0	5,38	5,62	0,24	0,83	0,28	-0,55
USP/SC	17	18	1	8,24	7,67	-0,57	0,94	0,56	-0,39
UTFPR	13	12	-1	7,38	16,42	9,03	1,46	3,00	1,54

A Tabela 6.5 apresenta as correlações entre os números de publicações e os conceitos dos programas.

Tabela 6.5 - Correlações entre Conceito e publicações

	Conceito		
	07-09	10-12	Varição
Número de artigos totais	0,689**	0,119	0,421
Artigos em coautoria	0,570**	0,247	0,574**
Artigos Totais/Autor	0,651**	0,064	0,453*
Artigos em Coautoria/Autor	0,514*	0,286	0,583**

** . A correlação é significativa no nível 0,01 (teste bilateral).

* . A correlação é significativa no nível 0,05 (teste bilateral).

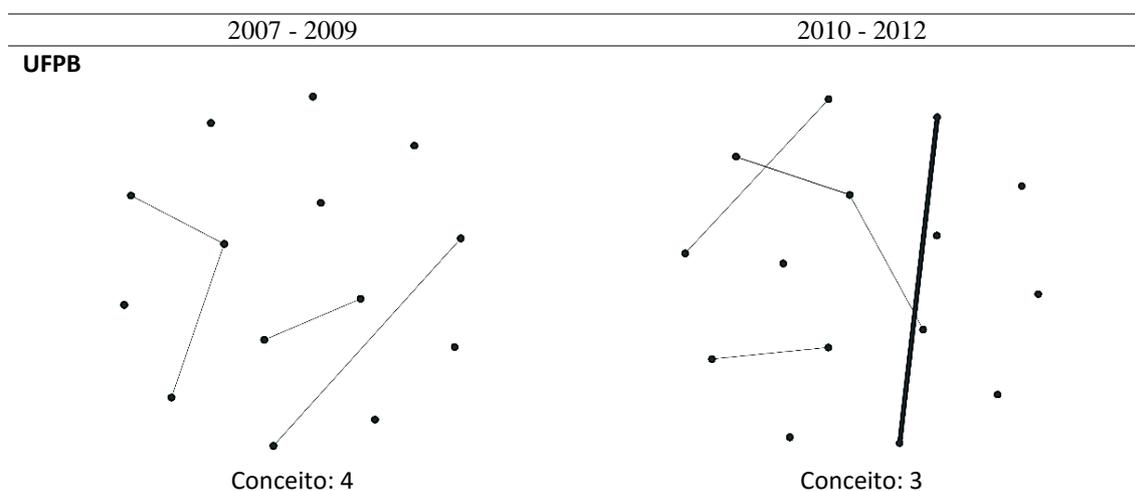
No triênio 2007-2009 todas as medidas correlacionadas com os conceitos dos programas tiveram fortes correlações. O contrário ocorreu no triênio 2010-2012, não houve correlações significativas. Portanto no primeiro triênio as medidas de publicações tiveram fortes influências nas definições dos conceitos, mas perderam importância no segundo triênio. Apesar da quantidade expressiva do aumento das publicações pela maioria dos programas. Talvez uma relação entre quantidade e qualidade das publicações explicasse esse resultado. As medidas que mais impactaram os conceitos dos programas no primeiro triênio foram o número de artigos e artigos por autor.

Quando analisadas as variações das medidas estudadas de um triênio ao outro e as correlações com as variações dos conceitos dos programas, somente a variação do número de artigos não impactou a variação dos conceitos dos programas. A variação da medida que mais influenciou a variação dos conceitos foi a de coautoria por autor. Evidenciando a importância da colaboração entre os professores dos programas para alcançar melhores conceitos. Apesar de um artigo realizado em coautoria entre professores de um mesmo programa ser “pontuado” somente uma vez, na avaliação do CAPES, o impacto maior da variação de coautoria por autor nos conceitos dos programas à frente do impacto da variação no número de artigos totais por autor, pode ser explicada, como várias pesquisas apontam, pelo aumento da qualidade das publicações.

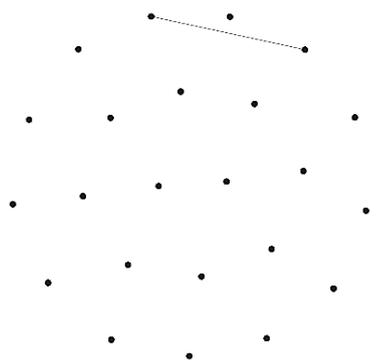
Para verificar se as evoluções das redes dos programas influenciaram os conceitos obtidos em cada triênio, as seguintes medidas foram analisadas: número de nós, número de arestas, densidade, coeficiente de *cluster*, tamanho da componente principal e o grau médio, a proximidade média, a intermediação média, o autovetor médio, o PageRank médio e a utilidade média.

As Figuras 6.1, 6.2 e 6.3 apresentam as evoluções das redes dos programas. Cada nó indica um professor permanente vinculado ao PPGEP avaliado pela CAPES no triênio 2007-2009 e 2010-2012. As arestas representam as relações de coautorias e a espessura é proporcional a frequência de colaboração.

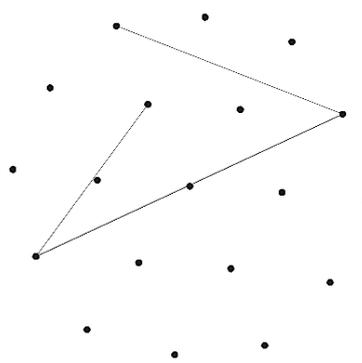
A Figura 6.1 apresenta as redes que obtiveram um conceito menor em relação ao triênio anterior.



UFRJ

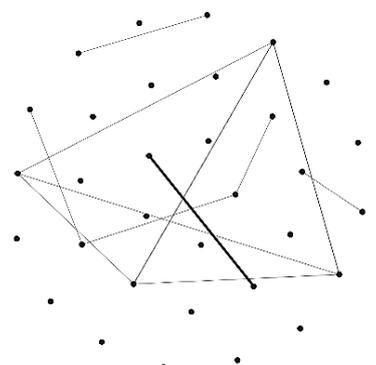


Conceito: 6

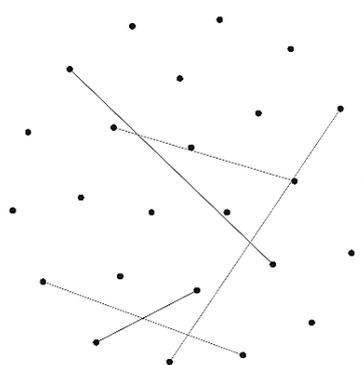


Conceito: 5

UFSCAR

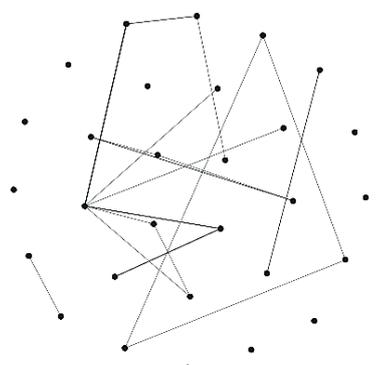


Conceito: 5

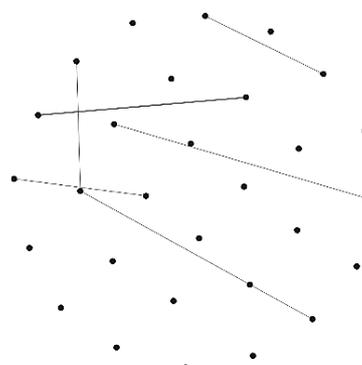


Conceito: 4

USP



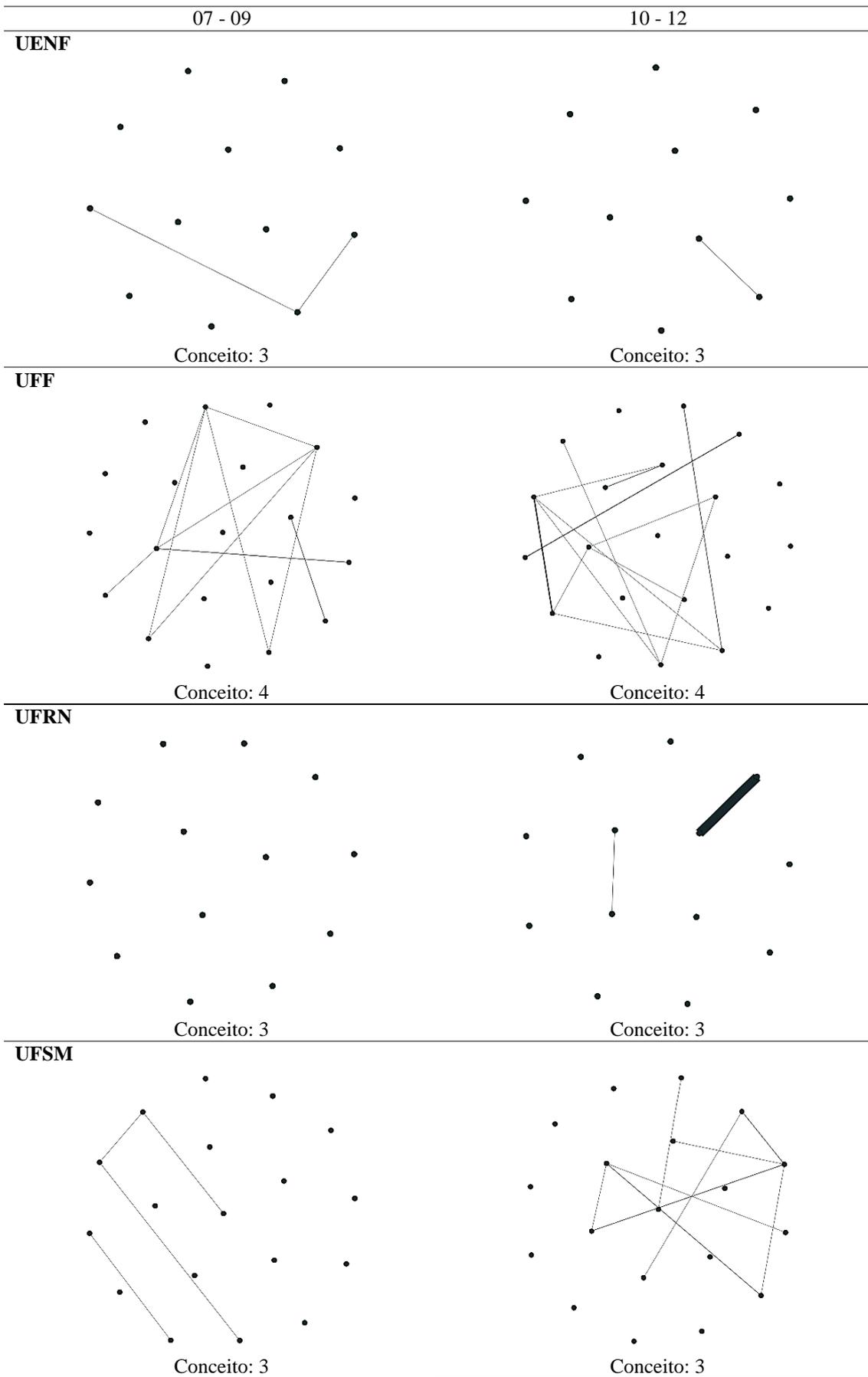
Conceito: 5



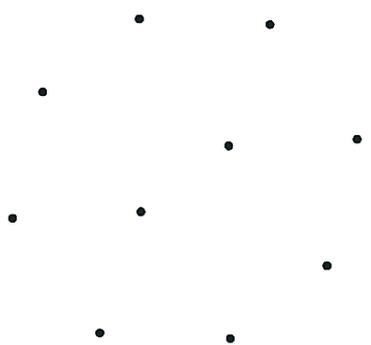
Conceito: 4

Figura 6.1– Evolução das redes que diminuíram o conceito

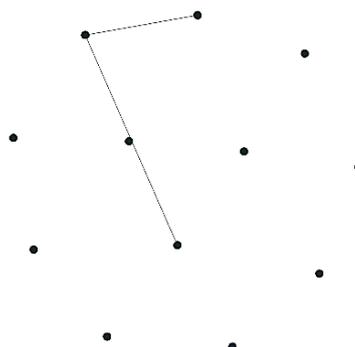
A Figura 6.2 apresenta as redes que mantiveram os conceitos.



UNIMEP

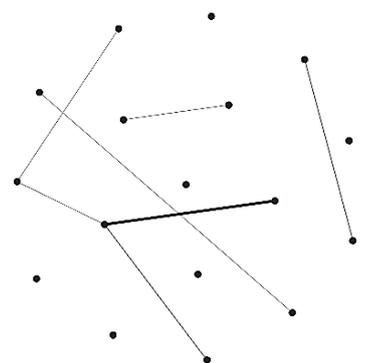


Conceito: 3

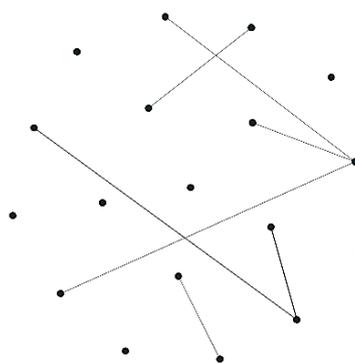


Conceito: 3

USP/SC

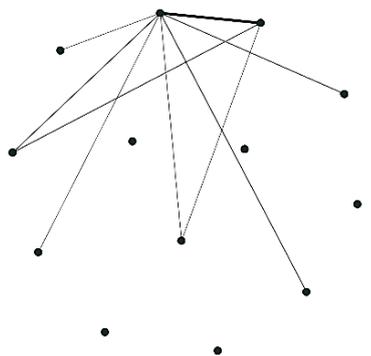


Conceito: 5

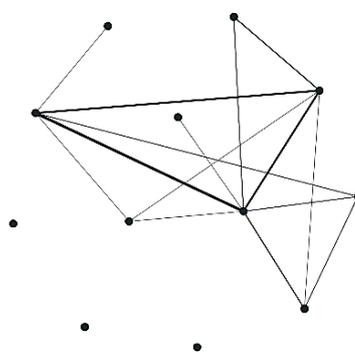


Conceito: 5

UTFPR



Conceito: 4



Conceito: 4

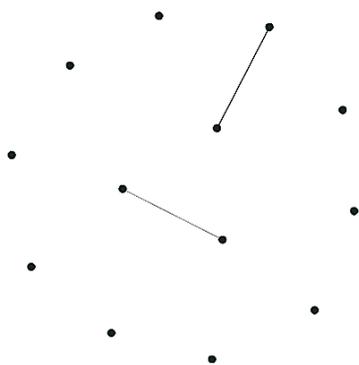
Figura 6.2 – Evolução das redes que mantiveram o conceito

A Figura 6.3 apresenta as redes dos programas cujos conceitos aumentaram.

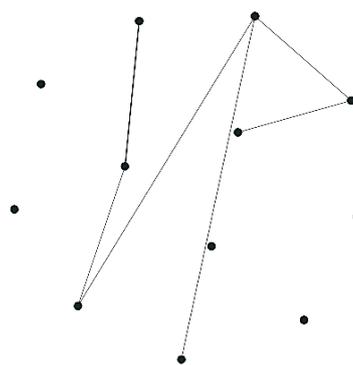
07 - 09

10 - 12

PUC/RIO

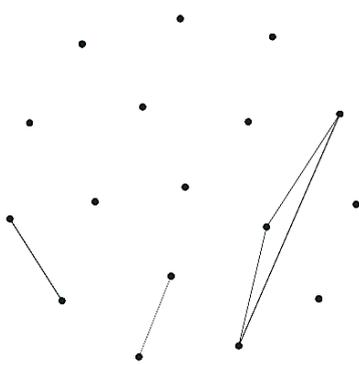


Conceito: 4

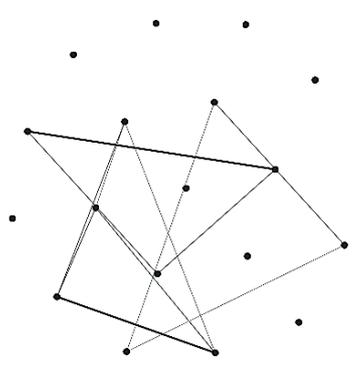


Conceito: 5

UFMG

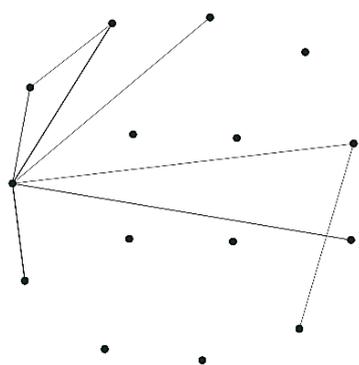


Conceito: 4

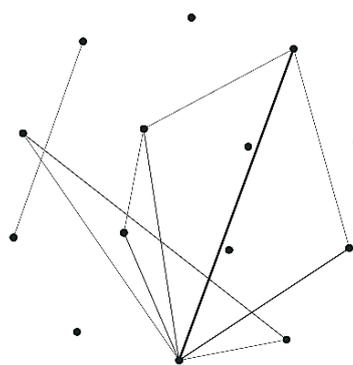


Conceito: 5

UFPE

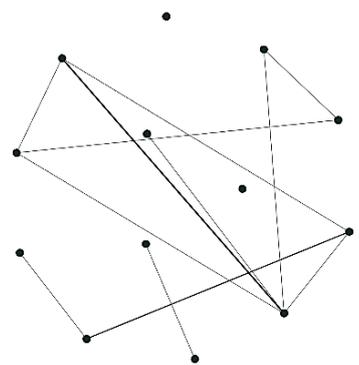


Conceito: 5

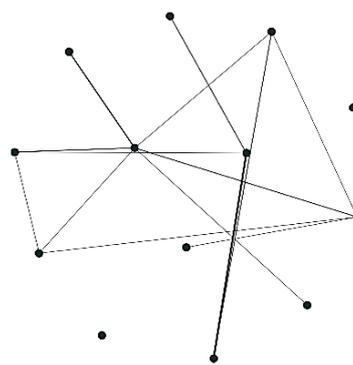


Conceito: 6

UFRGS

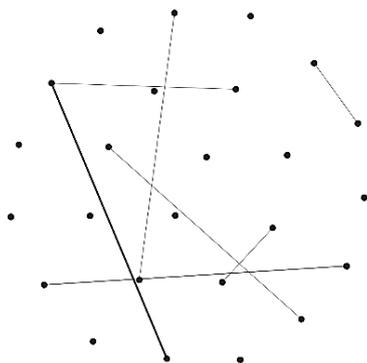


Conceito: 5

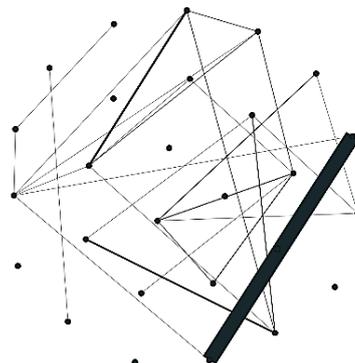


Conceito: 6

UFSC

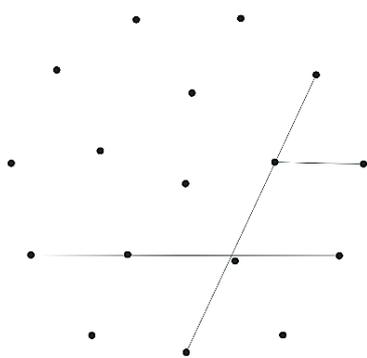


Conceito: 4

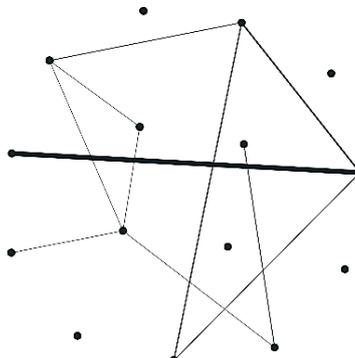


Conceito: 5

UNESP/BAU

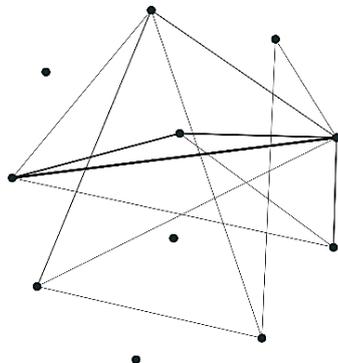


Conceito: 3

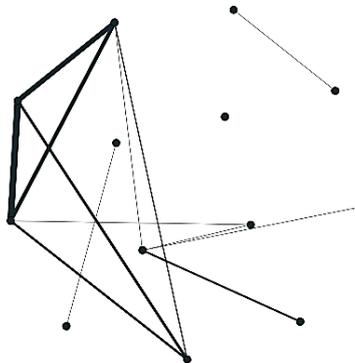


Conceito: 4

UNIFEI

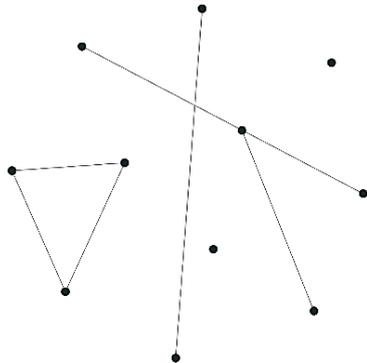


Conceito: 4

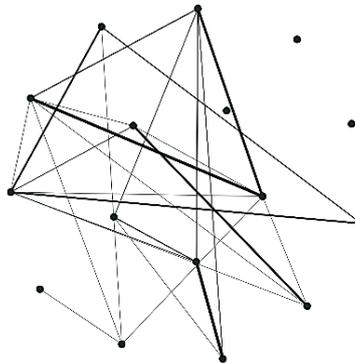


Conceito: 5

UNINOVE

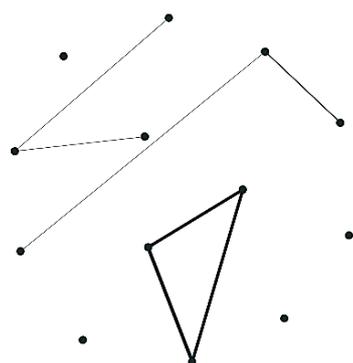


Conceito: 3

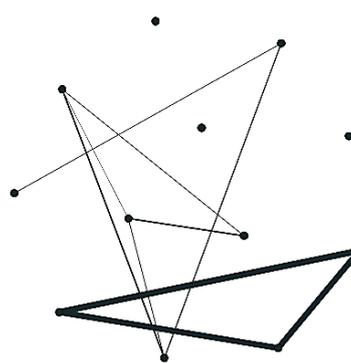


Conceito: 4

UNIP



Conceito: 4



Conceito: 5

Figura 6.3 - Evolução das redes que aumentaram o conceito

Tabela 6.6 – Dados sobre nós e arestas

Instituição	Número de nós			Arestas			Média das Arestas		
	07-09	10-12	Variação	07-09	10-12	Variação	07-09	10-12	Variação
PUC/RIO	13	12	-1	2	6	4	0,15	0,50	0,35
UENF	12	11	-1	2	1	-1	0,17	0,09	-0,08
UFF	20	21	1	10	13	3	0,50	0,62	0,12
UFMG	17	19	2	5	12	7	0,29	0,63	0,34
UFPB	14	15	1	4	5	1	0,29	0,33	0,05
UFPE	15	14	-1	8	11	3	0,53	0,79	0,25
UFRGS	13	13	0	12	14	2	0,92	1,08	0,15
UFRJ	23	19	-4	1	3	2	0,04	0,16	0,11
UFRN	13	13	0	0	2	2	0,00	0,15	0,15
UFSC	25	27	2	7	28	21	0,28	1,04	0,76
UFSCAR	32	25	-7	12	5	-7	0,38	0,20	-0,18
UFSM	18	19	1	4	9	5	0,22	0,47	0,25
UNESP/BAU	17	15	-2	3	11	8	0,18	0,73	0,56
UNIFEI	11	13	2	14	13	-1	1,27	1,00	-0,27
UNIMEP	10	12	2	0	2	2	0,00	0,17	0,17
UNINOVE	11	16	5	6	26	20	0,55	1,63	1,08
UNIP	13	12	-1	7	10	3	0,54	0,83	0,29
USP	29	29	0	18	6	-12	0,62	0,21	-0,41
USP/SC	17	18	1	7	7	0	0,41	0,39	-0,02
UTFPR	13	12	-1	9	15	6	0,69	1,25	0,56

A Tabela 6.6 apresenta o número de nós (professores permanentes), o número de arestas e a média das arestas. As redes que tiveram maiores quantidades de nós foram a UFSCAR, a USP e a UFSC no triênio 2007-2009 e a USP, a UFSC e a UFSCAR foram as que tiveram as maiores quantidades de nós no triênio 2010-2012. O programa da UFSCAR foi o que mais

perdeu professores (sete), a UNINOVE a que mais ganhou (cinco). Em média o número de professores de um triênio a outro se manteve constante e igual a 16,8.

As arestas representam relações de coautorias, os programas com as maiores quantidades de arestas foram os pertencentes a USP, a UNIFEI, a UFRGS e a UFSCAR compartilham a terceira posição no triênio 2007-2009. No segundo triênio quem ocupava essas posições eram a UFSC, a UNINOVE e a UTFPR, respectivamente. O número médio de arestas aumentou de um triênio a outro, de 6,55 a 9,95. A média das arestas é a relação entre a quantidade de arestas e a quantidade de nós.

Tabela 6.7 – Dados sobre densidade, cluster e componente gigante

Instituição	Densidade			Coeficiente de <i>Cluster</i>			Tamanho da componente principal (%)		
	07-09	10-12	Variação	07-09	10-12	Variação	07-09	10-12	Variação
PUC/RIO	0,03	0,09	0,07	0,00	0,00	0,00	0,15	0,58	0,43
UENF	0,03	0,02	-0,01	0,00	0,00	0,00	0,25	0,18	-0,07
UFF	0,05	0,06	0,01	0,16	0,04	-0,12	0,35	0,52	0,17
UFMG	0,04	0,07	0,03	0,18	0,53	0,35	0,18	0,21	0,03
UEPB	0,04	0,05	0,00	0,00	0,00	0,00	0,21	0,20	-0,01
UFPE	0,08	0,12	0,05	0,14	0,40	0,26	0,53	0,50	-0,03
UFRGS	0,15	0,18	0,03	0,12	0,14	0,03	0,69	0,85	0,15
UFRJ	0,00	0,02	0,01	0,00	0,00	0,00	0,09	0,21	0,12
UFRN	0,00	0,03	0,03	0,00	0,00	0,00	0,00	0,15	0,15
UFSC	0,02	0,08	0,06	0,00	0,31	0,31	0,12	0,67	0,55
UFSCAR	0,02	0,02	-0,01	0,13	0,00	-0,13	0,13	0,08	-0,05
UFSM	0,03	0,05	0,03	0,00	0,00	0,00	0,22	0,42	0,20
UNESP/BAU	0,02	0,11	0,08	0,00	0,21	0,21	0,12	0,67	0,55
UNIFEI	0,26	0,17	-0,09	0,41	0,23	-0,18	0,73	0,62	-0,11
UNIMEP	0,00	0,03	0,03	0,00	0,00	0,00	0,00	0,25	0,25
UNINOVE	0,11	0,22	0,11	0,27	0,49	0,22	0,27	0,81	0,54
UNIP	0,09	0,15	0,06	0,23	0,47	0,24	0,23	0,50	0,27
USP	0,04	0,02	-0,03	0,28	0,00	-0,28	0,34	0,10	-0,24
USP/SC	0,05	0,05	-0,01	0,00	0,00	0,00	0,29	0,22	-0,07
UTFPR	0,12	0,23	0,11	0,21	0,38	0,17	0,62	0,75	0,13

A Tabela 6.7 apresenta a densidade, o cluster médio e o tamanho da componente gigante das redes dos programas. A rede mais densa é a da UNIFEI e três programas tiveram densidades zero no triênio 2007-2009, no triênio seguinte a UTFPR foi que apresentou a maior densidade e a menor densidade (0,02) ficou com a UFRJ, a USP e a UENF. Em média a densidade dos

programas aumentou de 0,06 no primeiro triênio para 0,09 no segundo triênio. Evidenciando um aumento das relações de coautorias.

A UNIFEI, a USP e a UNINOVE no triênio 2007-2009 e a UFMG, a UNINOVE e a UNIP no triênio 2010-2012 foram as que obtiveram maiores *clusters* médio. A maioria dos programas que tiveram *cluster* médio igual a zero no triênio 2007-2009 permaneceram com *cluster* zero no triênio seguinte. A médio dos coeficientes de *cluster* dos programas passou de 0,11 para 0,16. O que revela o fortalecimento das relações com as vizinhanças.

Com relação a percentagem do tamanho da componente gigante a UNIFEI apresentou a maior porcentagem e, tanto os professores da UFRN com os da UNIMEP não estabeleceram relações de coautoria no triênio 2007-2009. No triênio 2010-2012, 85% dos professores do programa da UFRGS estavam na componente gigante e somente 8% dos professores do programa da UFSCAR pertenciam à componente gigante. A média da proporção de nós da componente gigante dos programas passou 0,28 para 0,42, demonstrando a diminuição da segregação de professores.

Tabela 6.8 – Grau médio, proximidade média e intermediação média dos programas

Instituição	Grau			Proximidade			Intermediação		
	07-09	10-12	Variação	07-09	10-12	Variação	07-09	10-12	Variação
PUC/RIO	0,31	1,00	0,69	0,10	0,26	0,15	0,00	0,16	0,16
UENF	0,33	0,18	-0,15	0,19	0,18	-0,01	0,08	0,00	-0,08
UFF	1,00	1,24	0,24	0,17	0,19	0,02	0,17	0,06	-0,11
UFMG	0,59	1,26	0,68	0,10	0,14	0,04	0,00	0,00	0,00
UFPB	0,57	0,67	0,10	0,10	0,09	-0,01	0,00	0,00	0,00
UFPE	1,07	1,57	0,50	0,29	0,27	-0,02	0,08	0,03	-0,05
UFRGS	1,85	2,15	0,31	0,28	0,40	0,12	0,07	0,11	0,04
UFRJ	0,09	0,32	0,23	0,09	0,13	0,04	0,00	0,07	0,07
UFRN	0,00	0,31	0,31	0,00	0,10	0,10	0,00	0,00	0,00
UFSC	0,56	2,07	1,51	0,05	0,23	0,18	0,00	0,06	0,06
UFSCAR	0,75	0,40	-0,35	0,06	0,04	-0,02	0,00	0,00	0,00
UFSM	0,44	0,95	0,50	0,11	0,16	0,06	0,02	0,05	0,03
UNESP/BAU	0,35	1,47	1,11	0,07	0,26	0,19	0,00	0,14	0,14
UNIFEI	2,55	2,00	-0,55	0,47	0,25	-0,22	0,07	0,03	-0,04
UNIMEP	0,00	0,33	0,33	0,00	0,19	0,19	0,00	0,08	0,08
UNINOVE	1,09	3,24	2,15	0,14	0,43	0,29	0,00	0,07	0,07
UNIP	1,08	1,67	0,59	0,15	0,24	0,09	0,01	0,04	0,03
USP	1,24	0,41	-0,83	0,10	0,04	-0,06	0,01	0,00	-0,01
USP/SC	0,82	0,78	-0,05	0,10	0,09	-0,01	0,01	0,00	-0,01
UTFPR	1,39	2,50	1,12	0,38	0,47	0,09	0,07	0,07	0,00

A Tabela 6.8 apresenta as medidas do grau médio, da proximidade média, da intermediação média. As médias dessas medidas são obtidas somando as centralidades dos professores da rede e dividindo pela quantidade de professores do programa.

O maior grau médio no triênio 2007-2009 foi obtido pela UNIFEI, no triênio 2010-2012 a UNINOVE. De maneira geral, os programas tiveram variações positivas no grau médio, a exceção que se destaca foi a USP que perdeu 0,83 grau na média. O aumento da média do grau médio evidencia que os professores estão criando novas relações ou professores isolados estão estabelecendo relações.

A UNIFEI também no primeiro triênio obteve a maior proximidade média, ou seja, em média os professores do PPGE da UNIFEI estão mais próximos que os professores de outros PPGE. No triênio seguinte a UNINOVE conseguiu a maior proximidade média. A similaridade com o parágrafo anterior pode ser explicada pelo fato que a proximidade média é fortemente correlacionada com grau médio, conforme aponta a Tabela 6.10. A maioria dos programas registrou variações positivas, o PPGE da UNIFEI foi a que obteve a maior variação negativa. O aumento da proximidade média sinaliza que as distâncias entre os professores estão diminuindo e, desta forma, aumentando a eficiência da comunicação e a possibilidade de estabelecer parcerias de publicações.

A centralidade de intermediação destaca o nó com maior centralidade um papel de intermediador, o que detém o fluxo de informação. Porém ao analisar a rede, quanto menor a centralidade de intermediação melhor, pois a descentralização aumentaria a autonomia dos nós. No entanto, como se pode ver na Tabela 6.10 a intermediação média é fortemente relacionada com a proximidade média. Então conclui-se que a diminuição das distâncias entre os nós ocorreu por ligações indiretas, por intermediadores, e não por ligações diretas. Desta forma, ainda existe o controle do fluxo de informação. A maior intermediação média no triênio 2007-2009 foi obtida pela UFF, no triênio seguinte pela PUC/RIO. Em relação a variação, a UFF teve a variação mais negativa e a PUC/RIO a maior variação positiva. Isto é, ocorreu pela UFF, uma perda média do controle do fluxo de informação, e pela PUC/RIO, um aumento médio do controle.

A Tabela 6.9 apresenta as medidas do autovetor médio, do PageRank médio e da utilidade média. O maior autovetor médio no triênio 2007-2009 foi obtido pela UNIFEI, no triênio 2010-2012 pela UTFPR. A maioria dos programas tiveram variações positivas no autovetor médio. O aumento do autovetor médio evidencia que em média os professores estão criando novas

relações ou professores isolados estão estabelecendo relações com professores de maiores centralidades de grau.

O *status* médio do programa é medido pelo PageRank médio. A UFF apresentou o maior PageRank médio no primeiro triênio e a UENF o maior no segundo triênio. Os programas apresentaram pequenas variações de um triênio para o outro, destaca-se a UFF com a maior perda de *status* médio.

Tabela 6.9 – Média da centralidade de autovetor, do PageRank e da utilidade

Instituição	Autovetor			PageRank			Utilidade		
	07-09	10-12	Variação	07-09	10-12	Variação	07-09	10-12	Variação
PUC/RIO	0,15	0,20	0,05	0,08	0,08	0,01	1,62	1,90	0,29
UENF	0,14	0,13	-0,01	0,08	0,09	0,01	1,42	1,36	-0,05
UFF	0,17	0,14	-0,03	0,17	0,05	-0,12	1,57	1,95	0,38
UFMG	0,10	0,11	0,00	0,06	0,05	-0,01	1,74	1,75	0,02
UFPB	0,12	0,11	-0,01	0,07	0,07	0,00	1,93	2,13	0,20
UFPE	0,17	0,18	0,01	0,07	0,07	0,00	1,73	1,95	0,22
UFRGS	0,20	0,22	0,02	0,08	0,08	0,00	2,28	2,13	-0,15
UFRJ	0,06	0,10	0,04	0,04	0,05	0,01	1,17	1,34	0,17
UFRN	0,00	0,11	0,11	0,00	0,08	0,08	1,00	1,62	0,62
UFSC	0,07	0,13	0,06	0,04	0,04	0,00	2,00	2,04	0,04
UFSCAR	0,06	0,06	-0,01	0,03	0,04	0,01	1,74	1,80	0,06
UFSM	0,11	0,14	0,03	0,06	0,05	0,00	1,58	1,82	0,24
UNESP/BAU	0,08	0,19	0,11	0,06	0,07	0,01	1,71	1,95	0,25
UNIFEI	0,25	0,19	-0,05	0,09	0,08	-0,01	1,93	2,43	0,50
UNIMEP	0,00	0,14	0,14	0,00	0,08	0,08	1,00	1,42	0,42
UNINOVE	0,16	0,20	0,04	0,09	0,06	-0,03	2,50	2,02	-0,48
UNIP	0,13	0,19	0,05	0,08	0,08	0,01	2,12	2,10	-0,02
USP	0,09	0,06	-0,03	0,03	0,03	0,00	2,08	1,72	-0,36
USP/SC	0,13	0,11	-0,02	0,06	0,06	0,00	2,16	2,06	-0,10
UTFPR	0,20	0,23	0,03	0,08	0,08	0,01	1,78	1,94	0,15

A UNINOVE apresentou o maior benefício ou utilidade média no triênio 2007-2009, a UNIFEI o maior no segundo triênio. A maior variação positiva, ganho de utilidade média, foi apresentada pela UFRN e a variação mais negativa, perda da utilidade média, foi apontada pela UNINOVE.

A Tabela 6.10 mostra as correlações entre as médias das medidas estudadas.

Tabela 6.10 – Correlações entre as médias das medidas

2007-2009	Grau	Proximidade	Intermediação	Autovetor	PageRank	Utilidade
Grau	-	0,724**	0,532*	0,720**	0,550*	767**
Proximidade	-	-	0,825**	0,936**	0,826**	0,336
Intermediação	-	-	-	0,715**	0,558*	0,067
Autovetor	-	-	-	-	0,902**	0,425
PageRank	-	-	-	-	-	0,330
Utilidade	-	-	-	-	-	-
10-12	Grau	Proximidade	Intermediação	Autovetor	PageRank	Utilidade
Grau	-	0,777**	0,434	0,737**	0,038	0,701**
Proximidade	-	-	0,715**	0,947**	0,475*	0,378
Intermediação	-	-	-	0,704*	0,225	0,098
Autovetor	-	-	-	-	0,579**	0,461*
PageRank	-	-	-	-	-	0,108
Utilidade	-	-	-	-	-	-
Variação	Grau	Proximidade	Intermediação	Autovetor	PageRank	Utilidade
Grau	-	0,794**	0,590**	0,692**	0,018	0,011
Proximidade	-	-	0,811**	0,857**	0,276	0,044
Intermediação	-	-	-	0,800**	0,376	0,047
Autovetor	-	-	-	-	0,606**	0,240
PageRank	-	-	-	-	-	0,274
Utilidade	-	-	-	-	-	-

** . A correlação é significativa no nível 0,01 (teste bilateral).

* . A correlação é significativa no nível 0,05 (teste bilateral).

As correlações entre o número de nós, o número de arestas, média das arestas, densidade, coeficiente de *cluster* médio, tamanho da componente principal, grau médio, proximidade média, intermediação média, autovetor médio, PageRank médio e da utilidade média com os conceitos dos programas são apresentadas na Tabela 6.11.

No triênio 2007-2009 ocorreram correlações significativas entre duas medidas (número de nós e média das arestas) e os conceitos dos programas. Desta forma a quantidade de professores e a quantidade de relações de coautoria impactaram os conceitos dos programas.

No triênio 2010-2012 ocorreram correlações significativas entre cinco medidas (número de arestas, média das arestas, densidade, *cluster* médio e a média da centralidade de grau) e os conceitos dos programas. Diferente do período anterior, número de nós não impactou os conceitos dos programas nesse triênio. Professores entraram e saíram dos programas, deste modo os professores que constituem a rede de um programa no triênio 2007-2009 podem ser, e foram na grande parte dos programas, diferentes no triênio 2010-2012.

Tabela 6.11 – Correlações entre as métricas globais de rede e os conceitos dos programas

	Conceito		
	07-09	10-12	Variação
Número de nós	0,512*	0,074	0,259
Número de arestas	0,539*	0,492*	0,605**
Média das arestas	0,388	0,522*	0,626**
Densidade	0,220	0,470*	0,570**
Cluster médio	0,216	0,488*	0,649**
T. da componente gigante (%)	0,306	0,414	0,477*
Grau	0,388	0,522*	0,619**
Proximidade	0,145	0,359	0,414
Intermediação	0,233	0,224	0,313
Autovetor	0,133	0,207	0,300
PageRank	-0,088	-0,074	-0,275
Utilidade	0,343	0,405	-0,026

** . A correlação é significativa no nível 0,01 (teste bilateral).

* . A correlação é significativa no nível 0,05 (teste bilateral).

A densidade, medida relacionada com o nível geral das ligações entre dois nós na rede, impactou os conceitos dos programas. Desta forma, quanto mais densos os programas maiores os conceitos.

O *cluster* médio que é uma medida relacionada à probabilidade de que quaisquer dois colaboradores de um professor tenham colaborado entre si, também impactou os conceitos dos programas. Assim, quanto maior o coeficiente de *cluster* médio maior a probabilidade de relacionamento, entre os professores, e maior o conceito do programa.

No triênio 2010-2012, de maneira diferente ao triênio anterior, os conceitos dos programas sofreram impacto do grau médio. A proximidade média, intermediação média, o autovetor médio, o PageRank médio e a utilidade média não impactaram os conceitos dos programas em nenhum dos triênios analisados. Nas variações essas métricas e o número de nós não impactaram os conceitos dos programas.

Foram obtidas também, das redes dos programas, o grau médio, a proximidade média, a intermediação média, o autovetor médio, o PageRank médio, o cluster médio e a utilidade média ponderada com os pesos das arestas (W) e com os pesos das arestas e dos nós (Z), assim como as correlações com os conceitos dos programas.

A Tabela 6.12 e 6.13 apresentam as correlações das métricas ponderadas por W e por Z, respectivamente, com os conceitos dos programas.

Tabela 6.12 – Correlações das métricas ponderadas por W e com os conceitos dos programas

Métricas	Conceito		
	07-09	10-12	Variação
Grau	0,486*	0,328	0,613**
Proximidade	0,209	-0,06	0,212
Intermediação	0,255	0,274	0,32
Autovetor	0,056	0,184	0,083
PageRank	-0,101	-0,074	-0,318
Cluster	0,17	0,495*	0,29
Utilidade	0,343	0,405	0,009

** . A correlação é significativa no nível 0,01 (teste bilateral).

* . A correlação é significativa no nível 0,05 (teste bilateral).

Tabela 6.13 – Correlações das métricas ponderadas por Z com os conceitos dos programas

Métricas	Conceito		
	07-09	10-12	Variação
Grau	0,507*	0,607**	0,826**
Proximidade	0,414	0,703**	0,398
Intermediação	0,236	0,284	0,342
Autovetor	0,084	0,148	-0,043
PageRank	-0,101	-0,241	-0,412
Cluster	0,218	0,455*	0,339
Utilidade	0,322	0,407	0,009

** . A correlação é significativa no nível 0,01 (teste bilateral).

* . A correlação é significativa no nível 0,05 (teste bilateral).

Ocorreram correlações significativas entre o grau médio ponderado por W e por Z com os conceitos dos programas no triênio 2007-2009 e no triênio 2010-2012. Portanto quanto maior o grau médio dos professores, mais ativos eles são, maior o conceito do programa. A exceção do grau médio ponderado por W no triênio 2007-2009 cuja correlação não foi significativa. O crescimento do grau médio do primeiro período para o segundo também impactou na melhora dos conceitos dos programas. Observa-se também que o grau médio que considera a importância dos nós, definido pelo índice h, ou seja, o peso Z, causa um impacto maior nos conceitos dos programas.

O *cluster* médio, no triênio 2010-2012, impactou os conceitos dos programas, assim quanto maior a probabilidade de relacionamento dos professores, que possuem um coautor em comum, maiores os conceitos dos programas.

A proximidade média, a intermediação média, o autovetor médio e utilidade média, ponderada por W e por Z, não impactaram significativamente os conceitos dos programas. Com

exceção da proximidade média ponderada por Z que causou um forte impacto nos conceitos no triênio 2010-2012.

6.4.2 Redes de Coautorias entre os Programas de Pós-Graduação em Engenharia de Produção

Na rede de coautoria entre PPGEp, cada programa é um nó da rede e a coautoria entre programas são dadas pelas relações de coautorias entre professores permanentes dos programas.

A Figura 6.4 ilustra a rede de coautoria, gerada pelo software Gephi, entre os programas no triênio 2007-2009. A espessura das arestas é proporcional ao peso (total de artigos em coautoria) e o diâmetro do nó é proporcional ao grau. A Tabela 6.14 apresenta as características da rede.

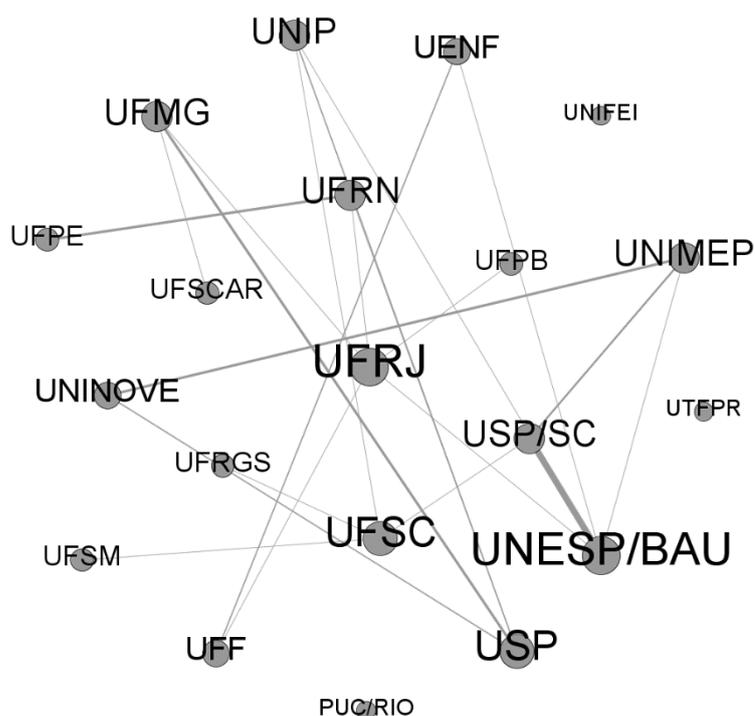


Figura 6.4– Rede de coautoria entre programas no triênio 2007-2009

A rede está dividida em 4 componentes, sendo a componente gigante composta por 17 vértices, representando 85% dos nós da rede; a segunda possui 1 vértice (5%), são três os programas isolados na rede.

Tabela 6.14 - Descrição geral da rede de coautoria entre programas no triênio 2007-2009

Número de programas:	20
Número de arestas:	22
Número de componentes:	4
Números de programas no componente gigante (%):	85,0
Coefficiente de <i>Cluster</i> médio:	0,038
Densidade:	0,116
Diâmetro*:	5
Raio*:	3
Distância média*:	2,66
Número de caminhos mais curtos*:	272

*. Relativo a Componente Principal

Ao analisar a densidade nota-se que a rede é dispersa, pois somente possui densidade igual a 0,116, ou seja, apenas estão ocorrendo 11,6% das ligações possíveis na rede. A densidade é considerada muito baixa se comparada com a densidade máxima que é igual a 1.

O diâmetro da rede é igual a 5 e o raio 0, representando a excentricidade máxima e a mínima, respectivamente, sendo o raio da componente gigante igual a 3. O coeficiente de *cluster* médio é igual a 0,038, sabendo que esse coeficiente pode variar de 0 a 1, então a probabilidade de que quaisquer dois programas colaboradores de um outro programa venham a colaborar é pequena.

A distância média de um caminho entre um par de vértices é aproximadamente 2,7. Esse valor refere-se à componente gigante, e significa que, em média, 2,7 ligações separam dois programas quaisquer no componente. O número de caminhos mais curtos é de 272.

A Figura 6.5 ilustra a rede de coautoria, gerada pelo software Gephi, entre os programas no triênio 2010-2012. A espessura das arestas é proporcional ao peso (total de artigos em coautoria) e o diâmetro do nó é proporcional ao grau. A Tabela 6.15 apresenta as características da rede.

Tabela 6.15 - Descrição geral da rede de coautoria entre programas no triênio 2010-2012

Número de programas:	20
Número de arestas:	29
Número de componentes:	2
Números de programas na componente gigante (%):	95,0
Coefficiente de <i>Cluster</i> médio:	0,119
Densidade:	0,153
Diâmetro*:	6
Raio*:	3
Distância média*:	2,569
Número de caminhos mais curtos*:	342

*. Relativo a Componente Principal

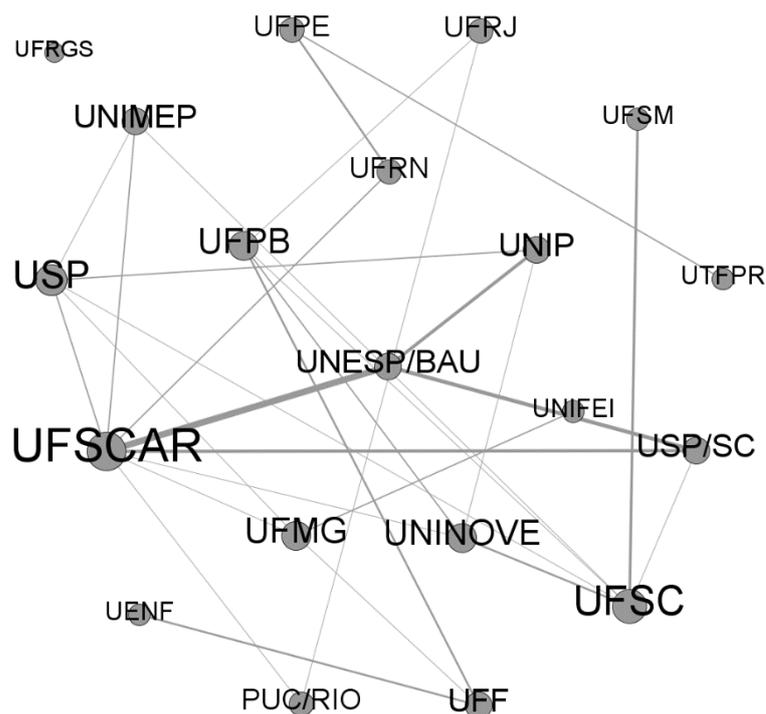


Figura 6.5 - Rede de coautoria entre programas no triênio 2010-2012

A rede está dividida em 2 componentes, sendo a componente gigante composta por 19 vértices, representando 95% dos nós da rede; a segunda possui 1 vértices (5%), único programa isolado na rede.

Ao analisar a densidade constata-se que a rede é dispersa, pois somente possui densidade igual a 0,153, ou seja, apenas estão ocorrendo 15,3% das ligações possíveis na rede. A densidade é considerada muito baixa se comparada com a densidade máxima que é igual a 1.

O diâmetro da rede é igual a 6 e o raio 0, representando a excentricidade máxima e a mínima, respectivamente, sendo o raio da componente gigante igual a 3. O coeficiente de *cluster* médio é igual a 0,119, sabendo que esse coeficiente pode variar de 0 a 1, então a probabilidade de que quaisquer dois programas colaboradores de um outro programa venham a colaborar é pequena.

A distância média de um caminho entre um par de vértices é aproximadamente 2,6. Esse valor refere-se à componente gigante, e significa que, em média, 2,6 ligações separam dois programas quaisquer no componente. O número de caminhos mais curtos é de 342.

6.4.2.1 Impactos da métricas de análises de redes sociais nos conceitos dos programas

As métricas calculadas da rede entre os PPGEP e usadas para correlacionar com os conceitos dos programas são: centralidade de grau e grau ponderado por W (Tabela 6.16); centralidade de proximidade e centralidade de intermediação (Tabela 6.17); centralidade de autovetor e PageRank (Tabela 6.18) e; excentricidade, coeficiente de *cluster* e utilidade (Tabela 6.19).

Tabela 6.16 – Centralidade de grau e grau ponderado

Instituição	Grau			Grau ponderado		
	07-09	10-12	Variação	07-09	10-12	Variação
PUC/RIO	0	2	2	0	2	2
UENF	2	1	-1	3	3	0
UFF	2	3	1	3	7	4
UFMG	3	4	1	6	5	-1
UFPB	1	4	3	1	7	6
UFPE	1	2	1	4	5	1
UFRGS	1	0	-1	1	0	-1
UFRJ	5	2	-3	5	2	-3
UFRN	3	2	-1	6	5	-1
UFSC	4	6	2	4	11	7
UFSCAR	1	8	7	1	24	23
UFSM	1	1	0	1	4	3
UNESP/BAU	5	3	-2	14	21	7
UNIFEI	0	1	1	0	2	2
UNIMEP	3	3	0	8	4	-4
UNINOVE	2	4	2	6	7	1
UNIP	3	3	0	4	8	4
USP	4	5	1	9	7	-2
USP/SC	3	3	0	14	12	-2
UTFPR	0	1	1	0	2	2

O PPGEP da UNESP/BAU e da UFRJ foram os que mais estabeleceram relações com outros PPGEP no triênio 2007-2009. A PUC/RIO, a UNIFEI e a UTFPR não estabeleceram relações de coautoria nesse triênio. Também nesse triênio a UNESP/BAU e a USP/SC foram as que tiveram maiores frequências de coautorias. No triênio 2010-2012 a UFSCAR foi a que mais estabeleceu relações de coautoria e as relações mais frequentes. A UFSCAR foi o programa que mais evoluiu na centralidade de grau e no grau ponderado.

Tabela 6.17 – Centralidade de proximidade e centralidade de intermediação

Instituição	Proximidade			Intermediação		
	07-09	10-12	Variação	07-09	10-12	Variação
PUC/RIO	0,00	0,40	0,40	0,00	0,04	0,04
UENF	0,36	0,29	-0,07	0,02	0,00	-0,02
UFF	0,35	0,40	0,05	0,02	0,12	0,10
UFMG	0,42	0,49	0,07	0,10	0,19	0,09
UFPB	0,33	0,45	0,12	0,00	0,11	0,11
UFPE	0,30	0,31	0,00	0,00	0,10	0,10
UFRGS	0,29	0,00	-0,29	0,00	0,00	0,00
UFRJ	0,48	0,36	-0,12	0,24	0,01	-0,22
UFRN	0,42	0,41	-0,01	0,10	0,19	0,09
UFSC	0,39	0,46	0,07	0,17	0,15	-0,02
UFSCAR	0,30	0,58	0,28	0,00	0,44	0,44
UFSM	0,29	0,32	0,04	0,00	0,00	0,00
UNESP/BAU	0,52	0,42	-0,10	0,22	0,02	-0,21
UNIFEI	0,00	0,33	0,33	0,00	0,00	0,00
UNIMEP	0,41	0,44	0,03	0,04	0,01	-0,02
UNINOVE	0,37	0,49	0,11	0,02	0,07	0,05
UNIP	0,48	0,39	-0,09	0,16	0,01	-0,15
USP	0,47	0,50	0,03	0,17	0,09	-0,08
USP/SC	0,42	0,44	0,02	0,07	0,03	-0,04
UTFPR	0,00	0,24	0,24	0,00	0,00	0,00

A UNESP/BAU apresentou a maior centralidade de proximidade, ou seja, é o programa mais próximo de todos os programas, no triênio 2007-2009. Nesse triênio, A UFRJ apresentou a maior centralidade de intermediação, o que faz deste programa um importante intermediador entre os PPGEP. No triênio seguinte, a UFSCAR assumiu o lugar da UNESP/BAU e da UFRJ como o programa de maior centralidade de proximidade e centralidade de intermediação, respectivamente. Os programas que mais evoluíram nessas métricas foram a PUC/RIO na centralidade de proximidade e a UFSCAR na centralidade de intermediação.

Pela centralidade de autovetor, programas que estão ligados a programas mais centrais, de acordo com o grau, têm maiores centralidades. No triênio 2007-2009 a UFRJ apresentou a maior centralidade de autovetor. No triênio seguinte a UFSCAR obteve a maior centralidade de autovetor. O PageRank analisa o *status* ou a importância de um programa pela atenção dada por outros programas. Neste caso, os programas mais importantes da rede foram a UNESP/BAU e a USP/SC no triênio 2007-2009 e UFSCAR no triênio 2010-2012.

Tabela 6.18 – Centralidade de autovetor e PageRank

Instituição	Autovetor			PageRank		
	07-09	10-12	Variação	07-09	10-12	Variação
PUC/RIO	0,00	0,29	0,29	0,01	0,02	0,01
UENF	0,43	0,07	-0,36	0,04	0,03	-0,01
UFF	0,39	0,26	-0,13	0,04	0,06	0,02
UFMG	0,52	0,53	0,01	0,07	0,05	-0,03
UFPB	0,26	0,46	0,20	0,02	0,06	0,04
UFPE	0,17	0,08	-0,09	0,05	0,05	0,00
UFRGS	0,17	0,00	-0,17	0,02	0,01	-0,01
UFRJ	0,84	0,18	-0,66	0,06	0,03	-0,04
UFRN	0,52	0,26	-0,26	0,07	0,04	-0,03
UFSC	0,53	0,78	0,25	0,06	0,08	0,01
UFSCAR	0,17	1,00	0,83	0,02	0,14	0,12
UFSM	0,17	0,19	0,02	0,02	0,03	0,01
UNESP/BAU	1,00	0,48	-0,52	0,12	0,11	-0,01
UNIFEI	0,00	0,13	0,13	0,01	0,02	0,02
UNIMEP	0,63	0,62	-0,02	0,07	0,03	-0,04
UNINOVE	0,40	0,65	0,25	0,06	0,05	-0,01
UNIP	0,68	0,46	-0,22	0,05	0,05	0,00
USP	0,67	0,81	0,14	0,09	0,05	-0,05
USP/SC	0,67	0,54	-0,13	0,12	0,07	-0,05
UTFPR	0,00	0,02	0,02	0,01	0,03	0,02

A excentricidade é a distância máxima a partir de um programa para qualquer outro, assim, quando menor a distância melhor o relacionamento. Três programas no triênio 2007-2009 tiveram as menores e as mesmas distâncias máximas: UNESP/BAU, UNIP e USP. Já no triênio 2010-2012 a menor distância máxima foi apresentada, somente, pela UFSCAR.

Os colaboradores dos programas UNIMEP e USP/SC foram os que mais colaboraram entre si, no triênio 207-2009, por isto esses programas apresentaram os maiores coeficientes de *cluster*. No triênio seguinte foram os colaboradores da UNIMEP que mais colaboraram entre si.

A utilidade refere ao benefício de um programa pertencer a uma dada estrutura de rede. O maior benefício no triênio 2007-2009 foi apresentado pela UFSC e no triênio seguinte pela UFSCAR. A PUC/RIO, UNIFEI e UTFPR tiveram utilidades iguais a um por serem nós isolados, no primeiro triênio. No segundo triênio a UFRGS é único programa isolado.

Tabela 6.19 – Excentricidade, cluster e utilidade

Instituição	Excentricidade			Cluster			Utilidade		
	07-09	10-12	Variação	07-09	10-12	Variação	07-09	10-12	Variação
PUC/RIO	6	4	-2	0,00	0,00	0,00	1,00	1,94	0,94
UENF	4	6	2	0,00	0,00	0,00	2,05	1,67	-0,38
UFF	5	5	0	0,00	0,00	0,00	2,05	3,00	0,95
UFMG	4	4	0	0,00	0,17	0,17	2,93	3,07	0,14
UFPB	5	5	0	0,00	0,17	0,17	1,40	2,57	1,16
UFPE	5	5	0	0,00	0,00	0,00	1,67	3,25	1,58
UFRGS	5	7	2	0,00	0,00	0,00	1,50	1,00	-0,500
UFRJ	4	5	1	0,00	0,00	0,00	3,84	2,13	-1,72
UFRN	4	4	0	0,00	0,00	0,00	2,93	1,94	-0,99
UFSC	4	5	1	0,00	0,13	0,13	4,33	3,76	-0,57
UFSCAR	5	3	-2	0,00	0,11	0,11	1,67	4,04	2,37
UFSM	5	6	1	0,00	0,00	0,00	1,50	1,33	-0,17
UNESP/BAU	3	4	1	0,10	0,33	0,23	3,04	2,05	-0,98
UNIFEI	6	5	-1	0,00	0,00	0,00	1,00	1,50	0,50
UNIMEP	4	4	0	0,33	0,67	0,33	2,38	1,65	-0,72
UNINOVE	4	4	0	0,00	0,17	0,17	1,88	2,09	0,22
UNIP	3	5	2	0,00	0,00	0,00	1,93	2,04	0,11
USP	3	4	1	0,00	0,30	0,30	2,88	2,45	-0,43
USP/SC	4	4	0	0,33	0,33	0,00	2,04	1,83	-0,21
UTFPR	6	6	0	0,00	0,00	0,00	1,00	2,00	1,00

A Tabela 6.20 apresenta as correlações das métricas apresentadas com os conceitos dos programas.

Tabela 6.20 – Correlações com os conceitos dos programas

Métricas	Conceito		
	07-09	10-12	Variação
Grau	-0,012	-0,090	0,004
Excentricidade	0,151	0,097	0,051
Proximidade	-0,027	-0,193	-0,016
Intermediação	-0,040	-0,017	-0,050
Cluster	-0,175	-0,209	-0,097
Autovetor	-0,047	-0,154	0,021
PageRank	-0,009	-0,089	-0,030
Utilidade	-0,073	0,150	0,010

Nenhuma das métricas e variações, das redes de coautorias entre os PPGE, obtiveram correlações significativas com os conceitos dos programas e as variações dos conceitos. E ainda

que insignificantes as relações entre as medidas e os conceitos foram inversamente proporcionais. O que leva a induzir que as relações externas dos PPGEP não impactam os conceitos. Essa conclusão fica mais evidente na seção seguinte pela métrica *E-I Index*, que avalia as relações externas e internas.

Foram obtidas também, das redes entre as instituições, as métricas ponderadas com os pesos das arestas (W) e com os pesos das arestas e dos nós⁵ (Z), e as correlações com os conceitos dos programas. A Tabela 6.21 e 6.22 apresentam as correlações das métricas ponderadas por W e por Z, respectivamente, com os conceitos dos programas.

Tabela 6.21 – Correlações das métricas ponderadas por W com os conceitos dos programas

Métricas	Conceito		
	0709	1012	Variação
Grau	-0,085	-0,125	0,165
Excentricidade	0,102	0,058	-0,089
Proximidade	-0,04	-0,237	-0,028
Intermediação	-0,012	-0,085	-0,011
Cluster	-0,175	-0,207	-0,109
Autovetor	-0,217	-0,088	0,003
PageRank	-0,047	-0,126	0,066
Utilidade	-0,001	0,053	0,112

Tabela 6.22 – Correlações das métricas ponderadas por Z com os conceitos dos programas

Métricas	Conceito		
	0709	1012	Variação
Grau ponderado	-0,035	0,013	0,158
Excentricidade	0,002	0,187	0,075
Proximidade	0,07	-0,112	-0,002
Intermediação	0,138	0,051	-0,037
Cluster	-0,175	-0,207	-0,109
Autovetor	-0,115	-0,071	-0,04
PageRank	0,018	0,083	0,066
Utilidade	0,068	0,152	0,184

⁵ Os pesos dos nós das redes entre os programas foram definidos como a média dos índices h dos professores pertencentes aos programas.

Comparando os resultados da Tabela 6.21 e 6.22 com a Tabela 6.20 verifica-se que não houve alterações relevantes. Portanto as métricas (sem peso, com peso W e com o peso Z) aplicadas nas redes entre as instituições não têm nenhuma correlação significativa com os conceitos dos programas.

6.4.3 Redes de Coautorias entre os Professores Permanentes dos PPGE

Na rede de coautoria entre os professores permanentes dos PPGE, cada professor é um nó da rede e as relações entre professores são dadas pelas coautorias.

A Figura 6.6 ilustra a rede de coautoria, gerada pelo software Gephi, entre os professores no triênio 2007-2009. A espessura das arestas é proporcional ao peso (total de artigos em coautoria) e o diâmetro do nó é proporcional ao grau. A Tabela 6.23 apresenta as características da rede.

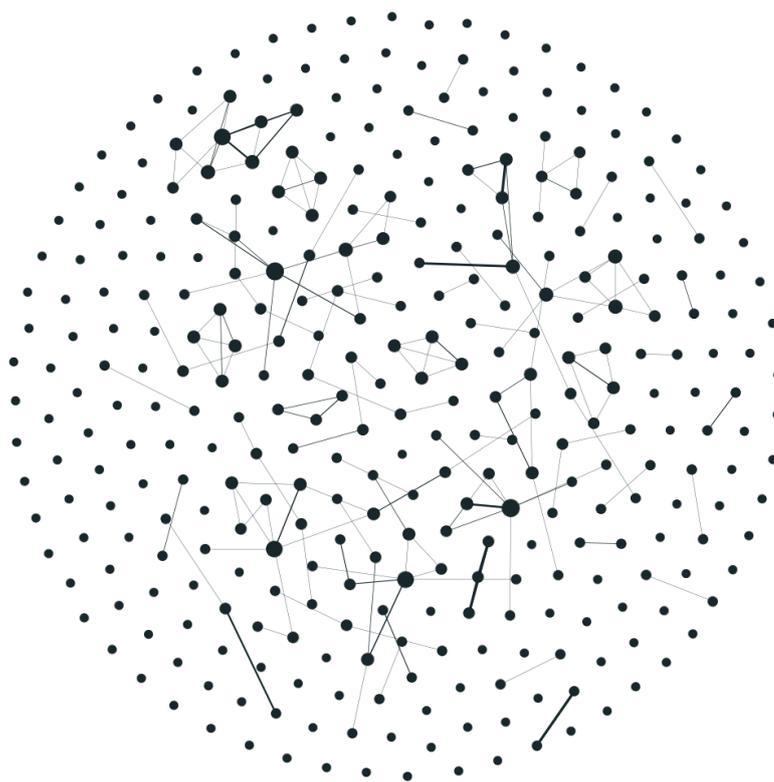


Figura 6.6- Rede de coautoria entre professores permanentes dos programas no triênio 2007-2009

A rede está dividida em 212 componentes, sendo a componente gigante composta por 12 vértices, representando 3,59% dos nós da rede; a segunda possui 11 vértices (3,29%), são 161 professores isolados na rede.

Tabela 6.23 – Descrição geral da rede de coautoria entre professores permanentes dos programas (2007-2009)

Número de professores:	334
Número de arestas:	157
Número de componentes:	212
Números de professores na componente gigante (%):	3,59
Coefficiente de <i>Cluster</i> médio:	0,140
Densidade:	0,003
Diâmetro*:	5
Raio*:	3
Distância média*:	1,87
Número de caminhos mais curtos*:	742

*. Relativo a componente gigante

Ao analisar a densidade nota-se que a rede é dispersa, pois somente possui densidade igual a 0,003, ou seja, apenas estão ocorrendo 0,3% das ligações possíveis na rede. A densidade é considerada muito baixa se comparada com a densidade máxima que é igual a 1. O que demonstra poucas relações entre os professores dos programas.

O diâmetro da rede é igual a 5 e o raio 0, representando a excentricidade máxima e a mínima, respectivamente, sendo o raio do componente gigante igual a 3. O coeficiente de *cluster* médio é igual a 0,140, sabendo que esse coeficiente pode variar de 0 a 1, então a probabilidade de que quaisquer dois professores colaboradores de um outro professor venham a colaborar entre si é pequena.

A distância média de um caminho entre um par de vértices é aproximadamente 1,87. Esse valor refere-se à componente gigante, e significa que, em média, 1,87 ligações separam dois professores quaisquer na componente. O número de caminhos mais curtos é de 742.

A Figura 6.7 ilustra a rede de coautoria, gerada pelo software Gephi, entre os programas no triênio 2010-2012. A espessura das arestas é proporcional ao peso (total de artigos em coautoria) e o diâmetro do nó é proporcional ao grau. A Tabela 6.24 apresenta as características da rede.

Tabela 6.24 – Descrição geral da rede de coautoria entre professores permanentes dos programas (2010-2012)

Número de professores:	334
Número de arestas:	257
Número de componentes:	155
Números de professores na componente gigante (%):	22,09
Coefficiente de <i>Cluster</i> médio:	0,174
Densidade:	0,005
Diâmetro*:	14
Raio*:	7
Distância média*:	5,236
Número de caminhos mais curtos*:	6924

*. Relativo a componente gigante

A rede está dividida em 155 componentes, sendo a componente gigante composta por 76 vértices, representando 22,09% dos nós da rede; a segunda possui 24 vértices (6,98%), são 155 professores isolados na rede.

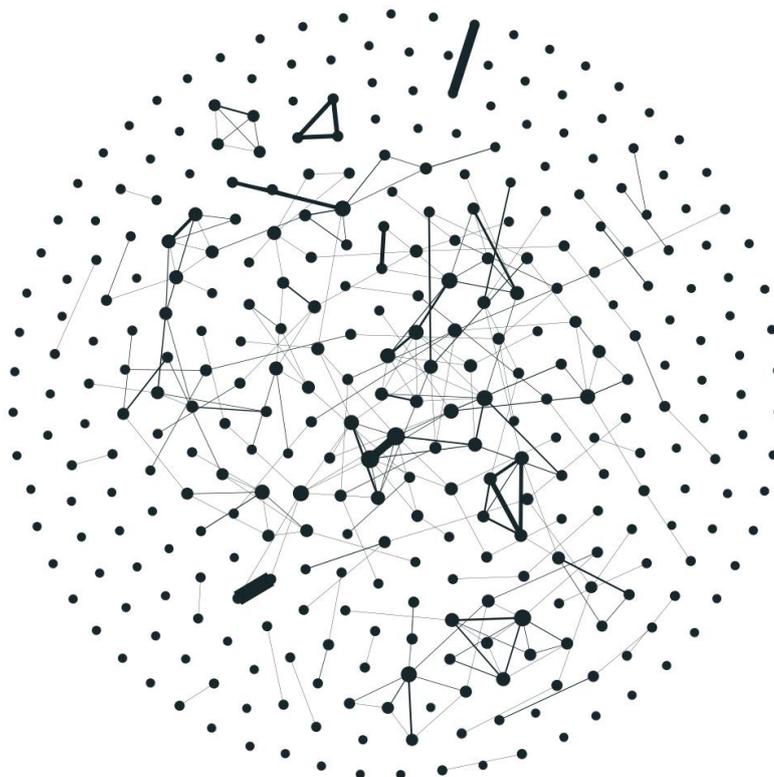


Figura 6.7 - Rede de coautoria entre professores permanentes dos programas no triênio 2010-2012

Ao analisar a densidade nota-se que a rede é dispersa, pois somente possui densidade igual a 0,005, ou seja, apenas estão ocorrendo 0,5% das ligações possíveis na rede. A densidade é considerada muito baixa se comparada com a densidade máxima que é igual a 1. O que demonstra poucas relações entre os professores dos programas.

O diâmetro da rede é igual a 14 e o raio 0, representando a excentricidade máxima e a mínima, respectivamente, sendo o raio da componente gigante igual a 7. O coeficiente de *cluster* médio é igual a 0,174, sabendo que esse coeficiente pode variar de 0 a 1, então a probabilidade de que quaisquer dois colaboradores de um mesmo professor venham a colaborar entre si é pequena.

A distância média de um caminho entre um par de vértices é aproximadamente 5,236. Esse valor refere-se a à componente gigante, e significa que, em média, 5,236 ligações separam dois professores quaisquer na componente. O número de caminhos mais curtos é de 6924.

Observa-se uma evolução da rede do triênio 2007-2009 ao triênio 2010-2012. O número de nós permaneceu constante, mas número de arestas cresceu de 157 para 257, foram 100 relações de coautoria a mais. Essas condições elevaram também a densidade da rede. O número de professores que passaram a pertencer a componente gigante no triênio 2010-2012 foi 6,33 vezes a mais que o triênio anterior. As relações entre as vizinhanças de também se intensificaram como mostra o aumento do coeficiente de *cluster* médio.

6.4.3.1 Impactos da métricas de análises de redes sociais nos conceitos dos programas

As métricas obtidas das redes de professores permanentes foram: centralidade de grau e grau ponderado com os pesos das arestas, centralidade de proximidade, centralidade de intermediação, centralidade de autovetor, PageRank, coeficiente de *cluster* e utilidade. Essas métricas são individuais, relativas aos nós (professores). Portanto, para representar os programas foram calculadas as médias dos valores obtidos pelos professores de cada programa. A Tabela 6.25 apresenta a média da centralidade de grau e grau ponderado dos programas.

Tabela 6.25 - Média da centralidade de grau e grau ponderado

Instituição	Centralidade de Grau			Grau Ponderado		
	07-09	10-12	Variação	07-09	10-12	Variação
PUCRIO	0,3077	1,1667	0,8590	0,6154	1,5000	0,8846
UENF	0,5833	0,4545	-0,1288	0,8333	0,4545	-0,3788
UFF	1,1500	1,5238	0,3738	1,5000	2,3333	0,8333
UFMG	0,8824	1,5263	0,6440	1,6471	3,2105	1,5635
UFPB	0,6429	1,0667	0,4238	0,6429	3,0000	2,3571
UFPE	1,3333	1,8571	0,5238	2,1333	3,4286	1,2952
UFRGS	1,9231	2,1538	0,2308	2,6923	3,6923	1,0000
UFRJ	0,2609	0,4211	0,1602	0,3043	0,5263	0,2220
UFRN	0,4615	0,5385	0,0769	0,5385	5,4615	4,9231
UFSC	0,7600	2,5926	1,8326	1,2400	9,4074	8,1674
UFSCAR	0,7813	0,9200	0,1388	1,4063	1,5200	0,1138
UFSM	0,5000	1,1579	0,6579	0,5000	1,5263	1,0263
UNESPBAU	0,7647	2,2000	1,4353	1,4118	6,4000	4,9882
UNIFEI	2,5455	2,1538	-0,3916	5,2727	10,4615	5,1888
UNIMEP	0,6000	0,6667	0,0667	1,1000	0,7500	-0,3500
UNINOVE	1,4000	3,6250	2,2250	1,6000	9,6250	8,0250
UNIP	1,3846	2,2500	0,8654	5,2308	12,4167	7,1859
USP	1,4483	0,8966	-0,5517	2,2069	2,2759	0,0690
USPSC	1,0588	1,2222	0,1634	2,7647	1,8333	-0,9314
UTFPR	1,3846	2,6667	1,2821	2,9231	6,4167	3,4936

A UNIFEI e a UNINOVE apresentam os programas com maiores centralidades de grau no triênio 2007-2009 e 2010-2012, respectivamente. A UNINOVE foi que teve maior aumento da centralidade de grau. No grau ponderado os programas que se destacaram foram a UNIFEI no primeiro triênio e a UNIP no segundo triênio. Desta forma, a UNIFEI desponta como um dos programas cujos professores mais estabeleceram relações de coautorias e também as mais frequentes. A maioria dos programas tiveram variações positivas em ambas as métricas. A Tabela 6.26 apresenta a centralidade de proximidade e a centralidade de intermediação dos programas.

Tabela 6.26 - Média da centralidade de proximidade e de intermediação

Instituição	Proximidade			Intermediação		
	07-09	10-12	Variação	07-09	10-12	Variação
PUCRIO	0,0018	0,0095	0,0077	0,0000	0,0002	0,0002
UENF	0,0040	0,0058	0,0018	0,0000	0,0000	0,0000
UFF	0,0088	0,0183	0,0095	0,0001	0,0010	0,0010
UFMG	0,0052	0,0110	0,0058	0,0000	0,0002	0,0002
UFPB	0,0041	0,0225	0,0184	0,0000	0,0006	0,0006
UFPE	0,0141	0,0162	0,0022	0,0001	0,0004	0,0003
UFRGS	0,0190	0,0193	0,0003	0,0003	0,0002	-0,0001
UFRJ	0,0017	0,0036	0,0019	0,0000	0,0000	0,0000
UFRN	0,0045	0,0040	-0,0006	0,0000	0,0000	0,0000
UFSC	0,0061	0,0549	0,0487	0,0000	0,0062	0,0061
UFSCAR	0,0049	0,0132	0,0083	0,0000	0,0001	0,0001
UFSM	0,0035	0,0247	0,0212	0,0000	0,0026	0,0026
UNESP/BAU	0,0049	0,0417	0,0368	0,0000	0,0099	0,0099
UNIFEI	0,0190	0,0159	-0,0031	0,0001	0,0002	0,0001
UNIMEP	0,0052	0,0183	0,0130	0,0000	0,0011	0,0011
UNINOVE	0,0080	0,0569	0,0489	0,0000	0,0065	0,0065
UNIP	0,0089	0,0264	0,0175	0,0000	0,0024	0,0024
USP	0,0133	0,0111	-0,0023	0,0002	0,0017	0,0015
USPSC	0,0082	0,0250	0,0168	0,0001	0,0033	0,0033
UTFPR	0,0153	0,0233	0,0079	0,0001	0,0003	0,0002

No triênio 2007-2009 a UNIFEI e a UFRGS foram as mais centrais, pela centralidade de proximidade e no triênio seguinte a UNINOVE e a UFSC foram as mais próximas dos outros programas. O papel de intermediador foi exercido principalmente pela UFRGS no primeiro triênio e pela UNESP/BAU no segundo triênio. Obtiveram variações negativas na centralidade de proximidade a UFRN, a UNIFEI e a USP. Somente a UFRGS teve variação negativa na

centralidade de intermediação. A Tabela 6.27 apresenta o coeficiente de *cluster*, o PageRank e a centralidade de autovetor. De modo geral no triênio 2010-2012 os programas ficaram mais próximos.

Tabela 6.27 - Média do coef. de cluster, do PageRank e centralidade de autovetor

Instituição	Cluster			PageRank			Autovetor		
	07-09	10-12	Variação	07-09	10-12	Variação	07-09	10-12	Variação
PUC/RIO	0,0000	0,0000	0,0000	0,0018	0,0031	0,0013	0,0000	0,0000	0,0000
UENF	0,0556	0,0909	0,0354	0,0017	0,0014	-0,0002	0,0000	0,0000	0,0000
UFF	0,2083	0,0810	-0,1274	0,0031	0,0033	0,0002	0,0000	0,0000	0,0000
UFMG	0,2353	0,4474	0,2121	0,0031	0,0028	-0,0003	0,0000	0,0000	0,0000
UFPB	0,0000	0,1333	0,1333	0,0030	0,0026	-0,0003	0,0000	0,0026	0,0026
UFPE	0,3206	0,3265	0,0059	0,0029	0,0032	0,0003	0,0001	0,0000	-0,0001
UFRGS	0,1128	0,1436	0,0308	0,0051	0,0040	-0,0010	0,0000	0,0000	0,0000
UFRJ	0,0000	0,0000	0,0000	0,0014	0,0013	-0,0001	0,0000	0,0000	0,0000
UFRN	0,0385	0,0769	0,0385	0,0013	0,0020	0,0007	0,0000	0,0000	0,0000
UFSC	0,0000	0,2965	0,2965	0,0032	0,0038	0,0006	0,0000	0,0870	0,0870
UFSCAR	0,1250	0,0400	-0,0850	0,0025	0,0030	0,0005	0,0000	0,0020	0,0020
UFSM	0,0000	0,0175	0,0175	0,0020	0,0025	0,0005	0,0000	0,0037	0,0037
UNESPBAU	0,0784	0,2517	0,1733	0,0030	0,0036	0,0006	0,0000	0,0055	0,0055
UNIFEI	0,4091	0,1974	-0,2117	0,0042	0,0045	0,0003	0,2476	0,0000	-0,2476
UNIMEP	0,2000	0,0139	-0,1861	0,0014	0,0013	0,0000	0,0000	0,0182	0,0182
UNINOVE	0,3000	0,4565	0,1565	0,0046	0,0044	-0,0002	0,0000	0,0574	0,0574
UNIP	0,4359	0,5694	0,1335	0,0039	0,0042	0,0003	0,0000	0,0002	0,0002
USP	0,2552	0,0220	-0,2331	0,0043	0,0024	-0,0019	0,0000	0,0222	0,0222
USPSC	0,0490	0,0833	0,0343	0,0037	0,0029	-0,0008	0,0000	0,0131	0,0131
UTFPR	0,2125	0,4599	0,2475	0,0035	0,0036	0,0001	0,0000	0,0000	0,0000

Com relação a inserção de um programa na vizinhança a UNIP apresenta o maior coeficiente de *cluster* tanto no primeiro triênio quanto no segundo. A UFF, a UFSCAR, a UNIFEI, a UNIMEP e a USP tiveram variações negativas, ou seja, perderam participações da vizinhança.

O *status* da UNINOVE e da UNIFEI foram os maiores no primeiro e segundo triênio, respectivamente, pois apresentaram os maiores PageRank. A PUC/RIO foi a que mais aumentou o *status* de um triênio a outro e a USP foi a que mais perdeu.

A maior centralidade de autovetor no triênio 2007-2009 foi apresentada pela UNIFEI e no triênio 2010-2012 a UFSC teve a maior centralidade. A UNIFEI e a UFSC foram também a

que mais variou de um triênio a outro. No entanto a variação da UNIFEI foi negativa e da UFSC positiva. A Tabela 6.28 apresenta as duas últimas métricas analisadas.

Tabela 6.28 – Utilidade e E-I Index

Instituição	Utilidade			E-I Index		
	07-09	10-12	Varição	07-09	10-12	Varição
PUC/RIO	1,6154	2,0093	0,3939	-1,0000	-0,5000	0,5000
UENF	1,4630	1,4955	0,0325	0,2000	0,5000	0,3000
UFF	1,8188	2,0278	0,2090	-0,5385	-0,3684	0,1700
UFMG	1,9461	1,8300	-0,1160	0,0000	-0,4118	-0,4118
UFPB	1,9286	1,9302	0,0017	-0,6000	0,0909	0,6909
UFPE	1,6548	1,9855	0,3307	-0,3333	-0,4667	-0,1333
UFRGS	2,2970	2,1303	-0,1667	-0,8462	-1,0000	-0,1538
UFRJ	1,4348	1,4298	-0,0050	0,6667	-0,2000	-0,8667
UFRN	1,3471	1,7473	0,4002	1,0000	0,2000	-0,8000
UFSC	2,0100	2,0056	-0,0044	-0,1667	-0,5135	-0,3468
UFSCAR	1,7231	2,1262	0,4031	-0,8462	0,4444	1,2906
UFSM	1,6111	1,7953	0,1842	-0,6000	-0,3846	0,2154
UNESP/BAU	1,9346	2,0330	0,0984	0,3333	-0,1579	-0,4912
UNIFEI	1,9306	2,4272	0,4967	-1,0000	-0,7333	0,2667
UNIMEP	1,3000	1,4745	0,1745	1,0000	0,2000	-0,8000
UNINOVE	2,4545	2,1270	-0,3276	0,1111	-0,6774	-0,7885
UNIP	2,0021	2,2351	0,2330	-0,4286	-0,3333	0,0952
USP	2,1858	1,8673	-0,3185	-0,3913	0,0909	0,4822
USP/SC	2,1046	1,9090	-0,1956	-0,2727	0,0000	0,2727
UTFPR	1,7839	1,9222	0,1384	-1,0000	-0,7647	0,2353

A utilidade representa o benefício de estabelecer relações de coautoria. As utilidades dos programas apresentados na tabela são as médias dos benéficos dos professores pertencentes aos programas. As utilidades dos programas em ambos os períodos foram maiores que 1, portanto foi preferível estabelecer relações a ficar isolado. A UNINOVE obteve a maior utilidade no triênio 2007-2009 e a UNIFEI foi a mais beneficiada no segundo triênio. A maior perda de utilidade foi apresentada pela UNINOVE o maior ganho pela UNIFEI.

O E-I Index métrica que avalia o grau de externalidade e internalidade das relações em uma rede em que os atores são rotulados ou particionados por uma ou várias de suas características. Na rede dos professores dos PPGE cada professor foi rotulado pela instituição que ele pertencia. Desta forma, as relações estabilidade por professores de diferentes

instituições são relações externas e as internas são as estabelecidas por professores do mesmo programa.

As relações estabelecidas pelos professores das instituições PUC/RIO, UNIFEI e UTFPR foram totalmente externas no triênio 2007-2009. Nesse triênio a UFRN e a UNIMEP estabeleceram relações totalmente internas. No triênio 2010-2012 somente a UFRGS concentrou suas relações internamente e nenhum programa estabeleceu relações puramente externas.

Apresentadas as métricas analisadas na rede dos professores e a evolução dessas, buscase agora verificar os impactos delas nos conceitos dos programas. A Tabela 6.29 apresenta as correlações das métricas analisadas com os conceitos dos programas.

Tabela 6.29 – Correlações das métricas não-ponderadas com os conceitos dos programas

Métricas	Conceito		
	07-09	10-12	Variação
Grau	0,280	0,441	0,534*
Proximidade	0,302	0,459*	0,147
Intermediação	0,398	0,005	0,195
Cluster	0,074	0,308	0,372
Autovetor	0,103	-0,250	-0,182
PageRank	0,280	0,499*	0,300
Utilidade	0,257	0,441	0,069
E-I index	-0,335	-0,607**	-0,303

** . A correlação é significativa no nível 0,01 (teste bilateral).

* . A correlação é significativa no nível 0,05 (teste bilateral).

No triênio 2007-2009, pela classificação de Cohen (1988), a métrica de centralidade de grau apresentou correlação fraca com os conceitos dos programas. O *cluster*, o autovetor, o PageRank, a utilidade, a centralidade de proximidade, a centralidade de intermediação e o E-I Index não apresentam correlações significativas.

No triênio 2010-2012, as correlações da centralidade de intermediação, da centralidade de autovetor e da utilidade com os conceitos dos programas, a centralidade de grau, o coeficiente de *cluster* não apresentaram correlações com os conceitos dos programas. Mas somente a centralidade de proximidade e o PageRank apontaram correlações significativas. Desta forma é possível afirmar que a centralidade de proximidade impacta os conceitos dos programas, ou seja, os programas cujos os professores apresentam as maiores médias de

centralidades de proximidade tendem a terem maiores conceitos, e que o *status* ou a importância dos programas, definidos pelo PageRank, também impacta os conceitos.

O *E-I Index*, no triênio 2020-2012, apresentou uma correlação forte e significativa com os conceitos dos programas. Quanto maior as relações de coautorias entre professores de um mesmo programa maior é o conceito do programa. Em outras palavras, quando as relações internas superam as externas impactam os conceitos dos programas de maneira positiva.

Quanto a variação das métricas e as correlações com as variações dos conceitos dos programas. A métrica centralidade de grau apresentou correlação forte e significativa. Desta forma o número de ligações impacta os conceitos dos programas. Não é possível afirmar que as variações das outras métricas impactam as variações dos conceitos dos programas, pois as correlações não foram significativas.

Foram obtidas também, das redes entre os professores, as métricas ponderadas com os pesos das arestas (W) e com os pesos das arestas e dos nós (Z). E as correlações com os conceitos dos programas. A Tabela 6.30 apresenta as correlações das métricas ponderadas por W com os conceitos dos programas.

Tabela 6.30 – Correlações das métricas ponderadas por W com os conceitos dos programas

Métricas	Conceito		
	07-09	10-12	Variação
Grau	0,304	0,295	0,585**
Proximidade	0,319	-0,031	0,191
Intermediação	0,410	-0,031	0,388
Cluster	0,001	0,29	-0,267
Autovetor	0,092	-0,119	-0,252
PageRank	0,264	0,471*	0,312
Utilidade	0,284	0,341	0,134
E-I Index	-0,162	-0,358	-0,239

** . A correlação é significativa no nível 0,01 (teste bilateral).

* . A correlação é significativa no nível 0,05 (teste bilateral).

A Tabela 6.31 apresenta as correlações das métricas ponderadas por Z com os conceitos dos programas. Comparando os resultados das Tabelas 6.30 e 6.31 com a Tabela 6.29 verifica-se que no primeiro triênio nenhuma das métricas apresentaram correlações significativas com os conceitos dos programas. No triênio seguinte, a centralidade de grau e a utilidade, ambas ponderadas por Z, obtiveram correlações significativas. Portanto, a importância dos professores (definida pelo índice h) combinadas com a frequência de relacionamento levou

essas métricas a impactarem os conceitos dos programas. O status ou a importância dos programas, definidos pelo PageRank (ponderado por W ou por Z), também impacta os conceitos.

Tabela 6.31 – Correlações ponderadas por Z com os conceitos dos programas

Métricas	Conceito		
	07-09	10-12	Variação
Grau	0,311	0,540*	0,786**
Proximidade	0,300	0,234	0,296
Intermediação	0,410	-0,077	0,243
Cluster	0,151	0,411	-0,003
Autovetor	0,115	-0,082	-0,154
PageRank	0,300	0,568**	0,408
Utilidade	0,261	0,473*	0,203
E-I Index	-0,290	-0,673	-0,375

** . A correlação é significativa no nível 0,01 (teste bilateral).

* . A correlação é significativa no nível 0,05 (teste bilateral).

Quanto a variação das métricas ponderadas e as correlações com as variações dos conceitos dos programas. A métrica centralidade de grau (ponderada por W ou por Z) apresentou correlação forte e significativa. Não é possível afirmar que as variações das outras métricas ponderadas impactam as variações dos conceitos dos programas, pois as correlações não foram significativas.

6.5 CONCLUSÕES

Neste capítulo foram construídas e analisadas três tipos de redes de coautorias: rede de coautoria entre professores permanentes para cada programa, rede de coautoria entre os programas e rede de coautoria entre professores permanentes de todos os programas. O objetivo foi analisar se a variação (evolução) das redes, nos triênios 2007-2009 e 2010-12, impactavam os conceitos dos programas.

A princípio foram analisadas as redes dos programas em separado. Foram 20 redes no triênio 2007-2010 e 20 redes no triênio 2010-2012 correspondentes aos vinte programas de pós-graduação em engenharia de produção avaliados pelo CAPES. Medidas referentes à produtividade, métricas globais e as médias das métricas individuais de ARS, calculadas de sem peso, com peso W e com peso Z, foram correlacionadas com os conceitos dos programas. Verificou-se que a variação da quantidade de artigos publicados (com ou sem coautoria)

impacta nas variações dos conceitos dos programas. Em relação às variações das métricas globais de rede, somente a variação do número de nós não impactou os conceitos dos programas. As variações do número de arestas, da média das arestas, da densidade, do coeficiente médio apontaram correlações fortes e significativas com os conceitos dos programas. Em relação a média das métricas individuais, o grau médio e o grau médio ponderado por W e por Z também apontaram correlações fortes e significativas com os conceitos dos programas. A variação do tamanho da componente gigante impactou moderadamente os conceitos dos programas.

Nas redes entre os programas o objetivo foi verificar a evolução das métricas individuais. Nessa rede cada programa representava um nó e a coautoria entre programas foram dadas pelas relações de coautorias entre professores permanentes dos programas. A maioria dos programas tiveram variações positivas em todas medidas analisadas. Foram correlacionadas com os conceitos dos programas métricas não ponderadas, ponderadas por W e ponderadas por Z: a centralidade de grau, a centralidade de intermediação, a centralidade de proximidade, a centralidade de autovetor, o PageRank, o coeficiente de cluster, a excentricidade e a utilidade. Nenhuma dessas métricas apresentou correlações significativas com os conceitos dos programas.

A última rede analisada foi composta por todos os professores permanentes dos programas. Foi construída uma rede para cada triênio, 2007-2009 e 2010-2012. Os valores médios, de uma métrica, obtidos por um determinado grupo de professores pertencente a um dado programa foram considerados com o valor médio do programa. As métricas (não ponderadas, ponderadas por W e por Z) analisadas foram a centralidade de grau, a centralidade de proximidade, a centralidade de intermediação, a centralidade de autovetor, o PageRank, o coeficiente de cluster e a utilidade. Dessas métricas, somente, a variação da centralidade de grau e a variação do grau ponderado (por W e por Z) impactaram nos conceitos dos programas. Foram calculadas, também de três maneiras distintas, para as redes o grau de externalidade e internalidade, pela métrica conhecida como *E-I Index*. Considerou-se como partições os programas de origem dos professores. A variação do *E-I Index* não impactou os conceitos dos programas.

7 CONCLUSÕES E TRABALHOS FUTUROS

O presente trabalho analisou como a estrutura social influenciou o nível de produtividade dos bolsistas de produtividade do CNPq na área de engenharia de produção e os conceitos dos programas de pós-graduação em engenharia de produção por meio da análise de redes sociais. Foi desenvolvido também um novo método que insere a importância individual do nó na estrutura topológica da rede.

Para analisar como a estrutura social impactou o nível de produtividade dos bolsistas, foi construída uma rede entre eles e calculadas métricas de ARS de três maneiras distintas: não-ponderadas; ponderadas com os pesos das arestas e; ponderadas com os pesos das arestas e dos nós. Em seguida, as métricas foram correlacionadas com o nível de produtividade. De maneira geral as métricas apresentaram correlações moderadas e significativas com o nível de produtividade. As métricas que mais impactaram os níveis das bolsas foram, nesta ordem: o *E-I index* não-ponderado, a centralidade de intermediação ponderada com os pesos das arestas e dos nós, o *E-I index_Z*, a centralidade de intermediação ponderada só com os pesos das arestas, o *E-I index_W*, a utilidade ponderada com os pesos das arestas e dos nós, a utilidade não-ponderada, e o PageRank não-ponderado. Observa-se que três dessas métricas estão relacionados ao grau de internalidade e externalidade, duas são as que determinam um papel de intermediador aos nós, outras duas são as que medem os benefícios de pertencer à rede e a última a que mede o *status* de um nó levando em consideração os *status* de seus vizinhos. Também fica evidente que dentre essas métricas três foram calculadas com os pesos das arestas e dos nós, duas somente com os pesos das arestas e três não-ponderadas.

Os cálculos das métricas com os pesos das arestas e com os pesos das arestas e dos nós são um diferencial deste trabalho, pois não foi identificado nenhum outro que apresente essa abordagem, aplicadas a várias métricas. O trabalho de LIU *et al.* (2014) utiliza os pesos das arestas e dos nós apenas no cálculo do PageRank. A inserção do peso dos nós na estrutura da rede não alterou significativamente a posição dos nós. Como pode ser percebido pelas tabelas de classificações apresentadas e pelas correlações entre as métricas, que foram em sua grande maioria altas. No entanto, geraram impactos diferentes nos níveis de produtividade, para algumas métricas o peso das arestas ou das arestas e dos nós combinadas promoveu o aumento das correlações com os níveis das bolsas em outras provocou a diminuição.

Vale ressaltar que a rede foi formada somente entre bolsistas de produtividade do CNPq da área de engenharia de produção. Desta forma, não foram considerados as relações de

coautorias entre esses e outros autores não bolsistas ou de outras áreas de conhecimento. Para trabalho futuros, poder-se-ia desenvolver uma rede que envolva além das relações entre os bolsistas as suas demais relações. E comparar os resultados das correlações das métricas dessa rede com as estudadas neste trabalho.

Com o objetivo de analisar se a variação (evolução) das redes, nos triênios 2007-2009 e 2010-12, impactavam os conceitos dos programas, foram construídas e analisadas três tipos de redes de coautorias: rede de coautoria entre professores permanentes para cada programa, rede de coautoria entre os programas e rede de coautoria entre professores permanentes de todos os programas.

Para o primeiro conjunto de redes foram analisadas as variações das medidas globais entre os triênios e os impactos dessas variações nas definições dos conceitos e as seguintes conclusões foram obtidas: a variação da quantidade de artigos publicados (com ou sem coautoria) impactou nas variações dos conceitos dos programas; as variações do número de arestas, da média das arestas, da densidade, do coeficiente médio, do grau médio e do grau médio ponderado (por W e por Z) apontaram correlações fortes e significativas com os conceitos dos programas.

Para o segundo conjunto de redes, onde cada programa é um nó da rede, foram analisadas as variações das métricas individuais de ARS entre os triênios e as correlações com os conceitos dos programas, nenhuma das métricas calculadas apresentou correlação significativa.

No último conjunto de redes, os valores médios, de uma métrica, obtidos por um determinado grupo de professores pertencente a um dado programa foram considerados com o valor médio do programa. As variações das métricas entre os triênios foram correlacionadas com as variações dos conceitos dos programas, concluindo-se que as variações da centralidade de grau e do grau ponderado impactaram os conceitos dos programas.

Pelo grau de externalidade e internalidade obtidos pelo *E-I Index*, nota-se que as relações internas, ou seja, entre os professores de um mesmo programa, foram mais favoráveis a definições dos conceitos dos programas que as relações externas.

Desta forma, com base nos resultados apresentados, o fortalecimento ou desenvolvimento do programa de pós-graduação em engenharia de produção depende do estreitamento das relações entre os professores dos mesmos programas, por meio da coautoria. A diversificação e a frequência das relações de coautoria contribuíram, praticamente com a mesma intensidade, para os conceitos dos programas. No entanto, a frequência associada a importância dos

professores (definidos pelo índice h) tiveram um impacto muito maior na definição dos conceitos.

Para trabalhos futuros poder-se-ia analisar a evolução das redes de coautorias entre os professores dos programas desde a origem dos programas, ou seja, da primeira avaliação realizada pela CAPES. Desta forma ter-se-ia toda a história bibliográfica dos professores que participaram desses programas e a contribuição dessas publicações nos conceitos. Analisando concomitantemente a tendência de evolução das publicações e a tendência de evolução dos conceitos. Poder-se-ia identificar também quais os professores e ou programas que mais contribuíram para o desenvolvimento de um programa. Também seria interessante analisar a qualidade das publicações e uma forma de inserir essa informação na rede.

REFERÊNCIAS

- ABBASI, A., ALTMANN, J. “On the correlation between research performance and social network analysis measures applied to research collaboration networks”. In Hawaii International Conference on System Sciences, Proceedings of the 41st Annual. Waikoloa, HI: IEEE, 2011.
- ABBASI, A.; ALTMANN, J.; HOSSAIN, L. Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics*, vol. 5, p. 594-607, 2011.
- ABBASI, A.; HOSSAIN, L.; LEYDESDORFF, L. Betweenness centrality as a driver of preferential attachment in the evolution of research collaboration networks. *Journal of Informetrics*.vol. 6, p. 403– 412, 2012.
- ALBERT, R.; BARABASI, A. L. Statistical mechanics of complex networks. *Reviews of Modern Physics*, vol. 74, 47–97.2002.
- AMORIN, C. “Curriculum vitae organization: the Lattes software platform,” *Pesquisa Odontologica Brasileira*, vol. 17, no. 1, p. 18–22, 2003.
- AMORIM, R. M. de. Uma Análise das Redes de Colaboração entre Bolsistas de Produtividade e entre Programas de Doutorado em Estatística no Brasil. Recife, 2014. 86p. (Mestrado – Pós-Graduação em Estatística / UFPE).
- ANASTASIOS, T.; SGOIROPOULOU, C.; PAPAGEORGIOU, E.; TERRAZ, O.; MIAOULIS, G. Co-authorship networks in academic research communities: the role of network strength. *16th Panhellenic Conference on Informatics*, 2012.
- ANDRADE, R. L.; RÊGO, L. C. Conhecendo a rede de coautoria dos bolsistas de produtividade em pesquisa da área de engenharia de produção e a sua influência no nível de produtividade. *In. XLVII Simpósio Brasileiro de Pesquisa Operacional – SBPO, 2015a*.
- ANDRADE, R. L.; RÊGO, L. C. A influência da rede de coautoria no nível das bolsas de produtividade da área de engenharia de produção. *In. XXXV Congresso da Sociedade Brasileira de Computação – CSBC, 2015b*.
- AZEVEDO, T. B.; RODRIGUEZ, M. V. R. Softwares para análise de redes sociais – ARS. *VI Congresso Nacional de Excelência em Gestão*, Niterói, RJ, 2010.
- BARABASI, A. L.; JEONG, H.; NEDA, Z.; RAVASZ, E.; SCHUBERT, A; TICSEK, T. Evolution of the social network of scientific collaboration. *Physica A: Statistical Mechanics and Applications*, vol. 311, p. 590-614, 2002.
- BARNETT, A. H.; AULT, R. W.; KASERMAN, D. L. The Rising Incidence of Co-authorship in Economics: Further Evidence. *Review of Economics and Statistics*, vol, 70, p. 539-543, 1988.
- BASTIAN, M.; HEYMANN, S.; JACOMY, M. Gephi: An open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*, 2009.
- BONACICH, P. Power and centrality: a family of measures. *The American Journal of Sociology*, vol, 92, nº 5, p. 1170–1182, 1987.

- COHEN, J. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates. New York, 1988.
- BORGATTI, S.P.; EVERETT, M.G.; FREEMAN, L.C. *Ucinet for Windows: Software for Social Network Analysis*. Analytic Technologies. Harvard, 2002.
- CAPES. Avaliação da Pós-graduação. Disponível em: <http://www.capes.gov.br/avaliacao/avaliacao-da-pos-gradua>. Acesso em: 26 de janeiro de 2015.
- CAPES. Caracterização do Sistema de avaliação da pós-graduação. Disponível em: <http://www.capes.gov.br/avaliacao/resultado-da-avaliacao-de-programas/2316>. Acesso em: 26 de janeiro de 2015.
- CNPq. Critérios de Julgamento. Disponível em: <http://www.cnpq.br/web/guest/criterios-de-julgamento>. Acesso em: 28 de janeiro de 2015.
- DIGIAMPIETRI, L. A.; MENA-CHALCO, J. P.; SILVA, G. S.; OLIVEIRA, L. B.; MALHEIROS, A. P.; MEIRA, D. Dinâmica das Relações de Coautoria nos Programas de Pós-graduação em Computação no Brasil. In. *CSBC BraSNAM*, 2012.
- DIGIAMPIETRI, L.; SANTIAGO, C.; ALVES, C. Predição de coautorias em redes sociais acadêmicas: um estudo exploratório em ciência da computação. In. *CSBC BraSNAM*, 2013.
- DIGIAMPIETRI, L. A.; MENA-CHALCO, J. P.; VAZ DE MELO, P. O. S.; MALHEIRO, A. P. R.; MEIRA, D. N. O.; FRANCO, L. F.; OLIVEIRA, L. B. BraX-Ray: An X-Ray of the Brazilian Computer Science Graduate Programs. *PLOS ONE*, vol. 9, Issue 4, April 2014.
- EATON, J. P.; WARD, J. C.; KUMAR, A. Structural Analysis of Co-Author Relationships and Author Productivity in Selected Outlets for Consumer Behavior Research. *Journal of Consumer Psychology*, vol. 8, p. 39-59, 1999.
- EGGHE, L. Theory and practise of the g-index. *Scientometrics*, vol. 69, 131–152, 2006.
- FEOFILOFF, P.; KOHAYAKAWA, Y.; WAKABAYASHI, Y. Uma Introdução Sucinta à Teoria dos Grafos. 2011. Disponível em <http://www.ime.usp.br/~pf/teoriadosgrafos/>
- FREEMAN, L. C. Centrality in Social Networks Conceptual Clarification. *Social Networks*, vol. 1, p. 215-239, 1979.
- FURTADO, A. L. Teoria dos Grafos: Algoritmos. *Livros Técnicos e Científicos*. Rio de Janeiro, 1973.
- JACKSON, M. O. Social and economic networks. *Princeton University Press*, 2008.
- GIBBONS, M.; LIMOGES, C.; NOWOTNY, H.; SCHWARTZMAN, S.; SCOTT, P.; TROW, M. The New Production of Knowledge. *Sage*, London, 1994.
- GRANOVETTER, M. S. The strength of weak ties. *American Journal of Sociology*, vol. 78, p. 1360-1380, 1973.
- HART, R. Co-authorship in the academic library literature: A survey of attitudes and behaviors. *Journal of Academic of Librarianship*, vol. 26, p. 339-345, 2000.
- HAUKE, J.; KOSSOWSKI, T. Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones geographicae*, vol. 30, p. 87-93, 2011.

- HANNEMAN, R. A.; RIDDLE, M. Introduction to social network methods. *Riverside: University of California*. 2005.
- HIRSCH, J. E. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, vol. 102, p. 16569-16572, 2005.
- HUANG, P.-Y., LIU, H.-Y.; CHEN, C.-H.; Cheng, P.-J. The impact of social diversity and dynamic influence propagation for identifying influencers in social networks, in: Web Intelligence (WI) and Intelligent Agent Technologies (IAT), *2013 IEEE/WIC/ACM International Joint Conferences on*, vol. 1, p. 410–416, 2013.
- HUDSON, John. Trends in Multi-Authored Papers in Economics. *Journal of Economic Perspectives*, vol. 10, p. 153–158, 1996.
- JACKSON, M. O. Social and Economic Networks. *Princeton University Press*. Stanford University, February 2008.
- JACKSON, M. O.; WOLINSKY, A. A Strategic Model of Social and Economic Networks. *Journal of economic theory*, vol. 71, p. 44-74, 1996.
- KEMPE, D.; KLEINBERG, J. É Tardos, Influential nodes in a diffusion model for social networks. In: Automata, Languages and Programming, *Springer*, vol. 3580, p. 1127–1138, 2005.
- KENNEDY, D. P.; JACKSON, G. L.; GREEN, H. D., BRADBURY, T N.; KARNEY, B. R. The Analysis of Duocentric Social Networks: A Primer. *Journal of Marriage and Family*, vol. 77: 295–311, 2015.
- KRACKHARDT, D.; STERN, R. Informal networks and organizational crises: An experimental simulation. *Social Psychology Quarterly*, vol. 51, p. 123-140. 1988.
- KRACKHARDT, D. The strength of strong ties: The importance of philos in organizations. *In Networks and Organizations: Structure, Form, and Action*, p. 216–239, 1992.
- KUMAR, S. Co-authorship networks: a review of the literature. *Aslib Journal of Information Management*, Kuala Lumpur, Malaysia. Vol. 67, p. 55-73, 2015.
- LAZEGA, E.; JOURDA, M-T.; MOUNIER, L.; STOFER, R. Catching up with big fish in the big pond? Multi-level network analysis through linked design. *Social Networks*, vol. 30, p. 159–176, 2008.
- LEE, SOOHO; BOZEMAN, BARRY. The impact of research collaboration on scientific productivity. *Social Studies of Science*, vol. 35, p. 673-702, 2005.
- LEHMANN, S.; JACKSON, A. D.; LAUTRUP, B. E. Measures for measures. *Nature*, vol. 444, p. 1003–1004, 2006.
- LEYDENA, D. P.; LINKA, A. N.; SIEGE, D. S. A theoretical analysis of the role of social networks in entrepreneurship. *Research Policy*, vol. 43, p. 1157–1163. 2014.
- LIU, J.; LI, Y.; RUAN, Z.; FU, G.; CHEN, X.; SADIQ, R.; DENG, Y. A new method to construct co-author networks. *Physica A*, vol. 419, p. 29-39. 2014.

- LIU, X.; BOLLEN, J.; NELSON, M. L.; SOMPEL, H. VAN DE. Co-authorship networks in the digital library research community. *Information Processing and Management*, vol. 41, p. 1461-1480, 2005.
- MARTELETO, R. M. Análise de redes sociais – aplicação nos estudos de transferência de informação. *Ciência de Informação*, vol. 30, p. 71-81, 2001.
- MARTELETO, R. M. Informação, rede e redes sociais – fundamentos e transversalidades. *Informação e Informação*, v. 12, 2007.
- MARTINS, L. D; FERREIRA, S. M. S. P. Proposta de metodologia de mapeamento e avaliação da produção científica da Universidade de São Paulo com foco na estrutura e dinâmica de suas redes de colaboração científica: em busca de novos modelos causais. *Liinc em Revista*, vol. 9, p. 181-195, 2013.
- MARTINS, C. B.; MACCARI, E. A.; MARTINS, S. B.; SERRA, F. A. R. The Influence of Co-Authorship Networks on the Performance of Graduate Programs in the Brazilian Evaluation System. *Portland International Conference on Management of Engineering & Technology (PICMET)*, p. 1569-1584, July, 2014.
- MELLO, C. M. de; CRUBELLATE, J. M.; ROSSONI, L. Dinâmica de Relacionamento e Prováveis Respostas Estratégicas de Programas Brasileiros de Pós-Graduação em Administração avaliação da Capes: Proposições Institucionais a partir da Análise de Redes de Coautoria. *RAC*, vol. 14, p. 434-457, 2010.
- MENA-CHALCO, J. P.; CESAR JUNIOR, R. M. ScriptLattes: An open-source knowledge extraction system from the Lattes platform. *Journal of the Brazilian Computer Society*, vol. 15, p. 31-39, 2009.
- MENA-CHALCO, J. P.; CESAR JUNIOR, R. M. Towards automatic Discovery of co-authorship networks in the Brazilian academic areas. *IEEE Seventh International Conference on e-Science Workshops*, p. 53-60, 2011.
- MENA-CHALCO, J. P.; DIGIAMPJETRI, L. A.; CESAR JUNIOR, R. M. Caracterizando as redes de coautoria de currículos Lattes. *Brazilian Workshop on Social Network Analysis and Mining (BraSNAM)*, 2012.
- MENA-CHALCO, J. P.; CESAR JUNIOR, R. M. Prospecção de dados acadêmicos de Currículos Lattes através de ScriptLattes. In: Hayashi, M.C.P.I.; Mugnaini, C.R.M. (Orgs.). *Bibliometria e Cientometria: reflexões teóricas e interfaces*. São Carlos, Pedro e João, p. 109-128, 2013.
- NETTO, P. O. B.; JURKIEWICZ. Grafos: introdução e práticas. *Blucher*. São Paulo, 2009.
- NEWMAN, M. E. J. The structure of scientific collaboration networks. *Proceeding of the National Academy of Sciences*, vol. 98, p. 404-409, 2001a.
- NEWMAN, M. E. J. Scientific collaboration networks. I. Network constructions and fundamental results. *Physical Review E*, vol. 64, 2001b.
- NEWMAN, M. E. J. Who is the best connected scientist? A study of scientific coauthorship networks. *Complex Networks*, vol. 650, p. 337-370, 2001c.

- NEWMAN, M. E. J. Coauthorship networks and patterns of scientific collaboration. *Proceeding of the National Academy of Sciences*, vol. 101, p. 5200-5204, 2004.
- ONEL, S.; ZEID, A.; Kamarthi, S. The structure and analysis of nanotechnology co-author and citation networks. *Scientometrics*, vol. 89, p. 119–138, 2011.
- PEREZ-CERVANTES, E.; MENA-CHALCO, J. P.; CESAR-JR, R. M. Towards a quantitative academic internationalization assessment of Brazilian research groups. *IEEE 8th International Conference*, 2012.
- SARAMÄKI, J.; KIVELÄ, M.; ONNELA, J.-P.; KASKI, K.; KERTÉSZ, J. Generalizations of the clustering coefficient to weighted complex networks. *Physical Review E*, vol. 75, 2007.
- SANTOS, ANDRÉA MARIA DOS. Aplicações de modelos de grafos na análise de conceitos e de redes sociais. Recife, 2014. 162p. (Doutorado - Programa de Pós-graduação em Estatística / UFPE).
- 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, vol. 35, p. 671-676, 2012.
- SPEARMAN, C. "General intelligence," objectively determined and measured. *American Journal of Psychology*, vol. 15, p. 201-293, 1904.
- WANDERLEY, A. J.; DUARTE, A. N.; BRITO, A. V. de; PRESTES, M. A. S.; FRAGOSO, F. C. Identificando correlações entre métricas de Análise de Redes Sociais e o h-index de pesquisadores de Ciência da Computação. In. *XXXIV Congresso da Sociedade Brasileira de Computação – CSBC, 2014*.
- WAGNER, C., LEYDESDORFF, L. Mapping global science using international coauthorships: A comparison of 1990 and 2000. In J. Guohua, R. Rousseau, & W. Yishan (Eds.), *Proceedings of the 9th International Conference on Scientometrics and Informetrics*, p. 330–340. Dalian, China: Dalian University of Technology Press, 2003.
- WAINER, J.; VEIRA, P. Correlation between bibliometrics and peer evaluation for all disciplines: the avaluation of Brazilian scientists. *Scientometrics (online)*, vol. 96, p. 395-410, 2013.
- WASSERMAN S.; FAUST, K. Social Networks Analysis: Methods and Applications. *Cambridge University Press. Structural analysis in social the social sciences series*, vol. 8, 1994.
- YAN E, DING Y. Applying centrality measures to impact analysis: A coauthorship network analysis. *Journal of the American Society for Information Science and Technology*, vol. 60, p.2107–2118, 2009.