

Dissertação de Mestrado

Redes de Coautoria da Base SciELO: Avaliação da Colaboração Científica por Medidas de Centralidade de Redes Complexas

Vitor Luiz Pereira Ribeiro
vitor.ribeiro@laccan.ufal.br

Orientador:
Dr. Alejandro C. Frery

Maceió, Julho de 2019

Vitor Luiz Pereira Ribeiro

Redes de Coautoria da Base SciELO: Avaliação da Colaboração Científica por Medidas de Centralidade de Redes Complexas

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Modelagem Computacional de Conhecimento do Instituto de Computação da Universidade Federal de Alagoas.

Orientador: Dr. Alejandro C. Frery

**Catalogação na fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico
Bibliotecária:**

T000n Ribeiro, Vitor Luiz Pereira.

Redes de Coautoria da Base SciELO: Avaliação da Colaboração Científica por Medidas de Centralidade de Redes Complexas / Vitor Luiz Pereira Ribeiro. – 2019

00 f. : il.

Orientador: Alejandro C. Frery.

Dissertação (Mestrado em Modelagem Computacional de Conhecimento) – Universidade Federal de Alagoas. Instituto de Computação. Maceió, 2019.

Bibliografia: f. 00–00.

1. Bibliometria. 2. Redes de Coautoria. I. Título.

CDU: 000.000:000.000

Membros da Comissão Julgadora da Dissertação de Mestrado de Vitor Luiz Pereira Ribeiro intitulada “Redes de Coautoria da Base SciELO: Avaliação da Colaboração Científica por Medidas de Centralidade de Redes Complexas”, apresentada ao Programa de Pós-Graduação em Modelagem Computacional de Conhecimento da Universidade Federal de Alagoas, em ..., às ...min, na sala de capacitação do Laboratório de Computação Científica e Análise Numérica (LaCCAN) da UFAL. Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Curso de Mestrado em Modelagem Computacional de Conhecimento do Instituto de Computação da Universidade Federal de Alagoas, aprovada pela comissão examinadora que abaixo assina.

Dr. Alejandro C. Frery – Orientador
Instituto de Computação
Universidade Federal de Alagoas

Dr. Osvaldo Anibal Rosso – Examinador
Instituto de Física
Universidade Federal de Alagoas

Dr. Leonardo Melo de Medeiros – Examinador
Instituto ou Departamento
Instituto Federal de Alagoas

Maceió, Julho de 2019

RESUMO

Na área de bibliometria, estudos vêm sendo realizados para mensurar quantitativa e qualitativamente a evolução da produção do conhecimento científico. Nesta direção, as redes de coautoria é uma das formas de indicar a interação entre autores, pelo esforço colaborativo entre pessoas, instituições e países para a geração e publicação de um trabalho científico. Este trabalho busca a propositura de uma análise investigativa a partir de base de dados de indexação de artigos científico SciELO, utilizando para tal, propriedades e métricas da teoria dos grafos e redes complexas, tais como *degree centrality*, *betweenness centrality* e *closeness centrality*, propondo uma avaliação da rede global de coautoria das Universidades Federais do Brasil e da rede local, definida como vértice focal a Universidade Federal de Alagoas, para série temporal de 2008 à 2017 nas áreas de *Health Sciences* - Ciências da Saúde, *Agricultural Science* - Ciências Agrárias, *Exact and Earth Sciences* - Ciências Exatas e da Terra. Este trabalhou baseou-se em uma metodologia que consistiu nas etapas: a) definição da rede, b) tratamento de dados, c) determinação das características estruturais, d) inspeção visual. Os resultados foram gerados numérica e graficamente, permitindo avaliações e interpretações do comportamento das redes estudadas com base nas medidas de centralidade utilizadas neste trabalho.

Palavras-chave: Bibliometria, Redes de Coautoria, Medidas de Centralidade.

ABSTRACT

In bibliometrics field, studies have been carried out to quantitatively and qualitatively measure the evolution of the production of scientific knowledge. In this direction, co-authorship networks is one way of indicating the interaction between authors, through the collaborative effort between people, institutions and countries for the production and publication of a scientific work. This work seeks the proposition of an investigative analysis based on the index database of SciELO scientific articles, using for that purpose, properties and metrics of graph theory and complex networks, such as degree centrality, betweenness centrality and closeness centrality, proposing a evaluation of the global co-authors network of the Federal Universities of Brazil and the local network, defined as focal point at the Federal University of Alagoas, for a time series from 2008 to 2017 in the areas of Health Sciences, Agricultural Sciences, Exact and Earth Sciences. This work was based on a methodology that consisted of the steps: a) definition of the network, b) data treatment, c) determination of structural characteristics, d) visual inspection. The results were generated numerically and graphically, allowing evaluations and interpretations of the behavior of the networks studied based on the centrality measures used in this work.

Keywords: Bibliometrics, Networks Co-authorship, Centrality Measures

AGRADECIMENTOS

Agradeço primeiramente a Deus pela permissão de cursar Mestrado com a proteção e a iluminação para a conclusão do mesmo. Ao apoio incondicional da minha família, em especial, meus pais Luiz e Magaly, que nos momentos difíceis me acalentaram com conforto e apoio. Gostaria de registrar os agradecimentos a muitos familiares que de perto viram a dedicação e motivação para a realização desta formação. A minha esposa Patrícia que a todos os momentos esteve ao meu lado dando força e energia para que o trabalho seguisse em frente e pudesse então chegar a sua fase final. Não posso deixar de mencionar o apoio de muitos amigos e colegas de sala pelas conversas formais e informais, aos colegas de trabalho e ao meu chefe Marcel, pelo incentivo e reconhecimento da importância da formação em grau de Mestre. A todos os professores das disciplinas do programa, que possibilitaram o engrandecimento do processo de minha formação. Agradeço pela oportunidade de ter sido aluno do Prof. Alejandro Frery, pelo exemplo da vida de um cientista, sua contribuição e influência na minha formação acadêmica, também pela generosidade e paciência na compreensão das minhas dificuldades e limitações, e ter sido o principal incentivador para a superação destas.

Vitor Luiz Pereira Ribeiro

LISTA DE FIGURAS

2.1	Grafo com representação de Rede de Coautoria intra grupos	14
2.2	Grafo com representação de Rede de Coautoria inter grupos	14
2.3	Pequena Rede de Coautoria de Newman	15
2.4	Redes com ilustração das medidas de centralidade (Fonte: Claudio Rocchini - Adaptado)	17
4.1	Matriz de Coautorias <i>Health Sciences</i> (2008-2017)	25
4.2	Matriz de Coautorias <i>Agricultural Sciences</i> (2008-2017)	26
4.3	Matriz de Coautorias <i>Exact and Earth Sciences</i> (2008-2017)	27
4.4	Rede de Coautoria das Universidades Federais do Brasil - 2015 (<i>Agricultural Sciences</i>)	28
4.5	Rede de Coautoria das Universidades Federais de Alagoas - 2015 (<i>Agricultural Sciences</i>)	29
4.6	Centralidade do Grau (<i>Health Sciences</i>)	31
4.7	Centralidade do Grau (<i>Agricultural Sciences</i>)	31
4.8	Centralidade do Grau (<i>Exact and Earth Sciences</i>)	32
4.9	Histograma da Centralidade do Grau (<i>Health Sciences</i>)	33
4.10	Histograma da Centralidade do Grau (<i>Agricultural Sciences</i>)	33
4.11	Histograma da Centralidade do Grau (<i>Exact and Earth Sciences</i>)	34
4.12	Centralidade de Proximidade (<i>Health Sciences</i>)	35
4.13	Centralidade de Proximidade (<i>Agricultural Sciences</i>)	36
4.14	Centralidade de Proximidade (<i>Exact and Earth Sciences</i>)	36
4.15	Histograma da Centralidade de Proximidade (<i>Health Sciences</i>)	37
4.16	Histograma da Centralidade de Proximidade (<i>Agricultural Sciences</i>)	38
4.17	Histograma da Centralidade de Proximidade (<i>Exact and Earth Sciences</i>)	39
4.18	Centralidade de Proximidade (<i>Health Sciences</i>)	40
4.19	Centralidade de Proximidade (<i>Agricultural Sciences</i>)	41
4.20	Centralidade de Proximidade (<i>Exact and Earth Sciences</i>)	41
4.21	Histograma da Centralidade de Proximidade (<i>Health Sciences</i>)	42
4.22	Histograma da Centralidade de Proximidade (<i>Agricultural Sciences</i>)	43
4.23	Histograma da Centralidade de Proximidade (<i>Exact and Earth Sciences</i>)	44
A.1	Redes de Coautoria Universidades Federal do Brasil (<i>Health Sciences</i>)	53
A.2	Redes de Coautoria da Universidade Federal de Alagoas	54
A.3	Redes de Coautoria Universidades Federal do Brasil: Área: <i>Agricultural Sciences</i> .	56
A.4	Redes de Coautoria da Universidade Federal de Alagoas - Área <i>Agricultural Sciences</i>	57
A.5	Redes de Coautoria Universidades Federal do Brasil - Área <i>Exact and Earth Sciences</i>	59
A.6	Redes de Coautoria da Universidade Federal de Alagoas Área <i>Exact and Earth Sciences</i>	60

LISTA DE TABELAS

4.1	Topologia das Redes (Vértices e Arestas) por Ano	23
4.2	Características Estruturais Rede Brasil (2008-2017)	24
4.3	Medidas de Centralidade do Grau por Área (2008-2017)	30
4.4	Medidas de Centralidade de Proximidade por Área (2008-2017)	35
4.5	Medidas de Centralidade de Intermediação por Área (2008-2017)	40
4.6	Ranking Medidas de Centralidade (<i>Health Sciences</i>)	45
4.7	Ranking Medidas de Centralidade (<i>Agricultural Sciences</i>)	46
4.8	Ranking Medidas de Centralidade (<i>Exact and Earth Sciences</i>)	47

SUMÁRIO

1 INTRODUÇÃO	9
1.1 Definição do Problema	9
1.2 Revisão da literatura	10
1.3 Contribuições	11
1.4 Objetivo	11
1.5 Organização da Dissertação	11
2 FUNDAMENTAÇÃO TEÓRICA	13
2.1 Teoria dos Grafos	13
2.2 Análise de Redes Sociais	14
2.3 Redes Complexas	15
2.4 Colaboração Científica	17
3 METODOLOGIA	19
3.1 Definição da rede	20
3.2 Tratamento de dados da rede	20
3.3 Determinação das características estruturais	20
3.4 Inspeção visual	21
4 RESULTADOS E ANÁLISES	22
4.1 Redes de Coutorias - Base SciELO	22
4.1.1 Características Estruturais	23
4.2 Medidas de Centralidade	29
4.2.1 Centralidade do Grau	30
4.2.2 Centralidade de Proximidade	34
4.2.3 Centralidade de Intermediação	39
4.3 Ranking das Medidas de Centralidade	44
4.4 Discussão dos Resultados	47
5 CONCLUSÕES	49
5.1 Considerações Finais	49
5.2 Trabalhos futuros	50
A REDES DE COAUTORIA	51
A.1 Health Sciences	53
A.2 Agricultural Sciences	56
A.3 Exact and Earth Sciences	59
Referências Bibliográficas	61

1

INTRODUÇÃO

1.1 Definição do Problema

A colaboração científica é um fenômeno social que tem por objeto a produção de conhecimento e os cientistas como os principais atores. As redes de coautoria é uma das formas de mensuração e indicação da interação entre esses atores, pelo esforço colaborativo entre pessoas, instituições e países para a geração e publicação de um trabalho científico.

Nos últimos anos vários estudos tem sido realizado para a compreensão deste fenômeno (Maia & Caregnato 2008), buscando entender como ocorre. Avançar nos estudos e no entendimento da colaboração científica no Brasil é fundamental para que tenhamos uma ideia mais clara de como este fenômeno vem acontecendo na comunidade científica brasileira, possibilitando a definição e o direcionamento de políticas científicas mais adequadas (Vanz & Stumpf 2010).

O Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq¹ é o mantenedor da base de Currículos Lattes², a principal fonte de registro dos trabalhos e publicação científica do país. Entretanto, a plataforma Lattes não vem apresentando avanços significativos, principalmente pela não disponibilização de dados abertos, assim, impossibilitando a realização de pesquisas que possam descrever o comportamento bibliométrico da ciência no país. Uma evidência da desatualização do Lattes é a própria plataforma Painel Lattes³ que não possui dados atualizados.

A Scientific Electronic Library Online – SciELO⁴ é uma biblioteca eletrônica criada em 1998. Ela realiza a indexação de um conjunto de periódicos visando o desenvolvimento de uma metodologia comum para a preparação, armazenamento, disseminação e avaliação de literatura científica em formato eletrônico.

¹<http://www.cnpq.br>

²<http://lattes.cnpq.br>

³<http://estatico.cnpq.br/painelLattes/>

⁴<http://www.scielo.br>

No entanto, apesar dos esforços da SciELO em disponibilizar uma ferramenta para o maior conhecimento de métricas bibliométricas e a observação do comportamento da colaboração científica, sua plataforma SciELO Analytics⁵ ainda carece de funcionalidades para a visualização de redes de coautoria, considerando o aspecto interinstitucional.

Diante do contexto supra explanado, essa pesquisa visa a proposta de um modelo da avaliação da colaboração científica com base na análise de redes de coautorias inter instituições do Brasil. Buscamos responder precipuamente as seguintes questões:

- Como se caracterizam as redes de coautoria na base SciELO?
- Quais as propriedades e métricas que indicam a evolução das redes de coautoria ao longo do tempo?
- Como se observa as comunidades de colaboração científica a partir das redes de coautoria?

1.2 Revisão da literatura

A bibliometria é uma campo de conhecimento multidisciplinar com origem na biblioteconomia e na ciência da informação, que aplica métodos estatísticos e matemáticos para analisar e construir indicadores sobre a dinâmica e evolução da informação científica e tecnológica de determinadas disciplinas, áreas, organizações ou países.

(Pritchard et al. 1969) descrevem que bibliometria é o tratamento quantitativo das propriedades da escrita científica publicada e do comportamento que lhe é inerente. (Osareh 1996) define a bibliometria como o estudo dos padrões das publicações científicas aplicando análises quantitativas e estatísticas.

Estudos e análises bibliométricas podem ser empregadas a diversas áreas do conhecimento com abordagens como avaliação da produtividade científica, análise de citações e cocitações, redes de coautoria, análises de fatores de impacto, dentre outros. A bibliometria possui estreitas relações com as áreas de cienciometria (ou cientometria), infometria, webometria, patentometria, e altmetria.

Para (Barabási et al. 2001) 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. Por outro lado, em outra obra de (Barabasi 2003) reconhecem que em redes de coautoria (*co-authorship network*), estudos vêm apresentando evoluções significativas, entretanto de maneira muito fragmentada, se concentrando em uma característica da rede por vez. Os autores ressaltam também a complexidade envolvida em virtude da velocidade de crescimento das redes, comparando-as como o comportamento da rede *World Wide Web*, visto que há de se considerar os vários

⁵<http://analytics.scielo.br>

formatos das bases de dados e suas indexações. Fica, assim, configurado o desafio para estudos relacionados a este tema.

1.3 Contribuições

Neste trabalho propomos desenvolver um estudo para avaliação da colaboração científica por meio de redes de coautoria, com as seguintes características:

- Reprodutibilidade com a definição da instituição vértice de referência.
- Definição do intervalo temporal em anos para análises comparativas.
- Explanar as características e métricas quantitativas das redes pelo aspecto da colaboração científica.

1.4 Objetivo

O objetivo desta pesquisa é a proposta da criação de um estudo modelo avaliativo do comportamento das redes de coautoria na base de dados objeto desse estudo, considerando alguns objetivos secundários, são eles:

- Avaliação das propriedades e indicadores de redes de coautoria inter instituição (unidades federativas).
- Visualização por comunidades definidas pelas regiões geográficas do Brasil.
- Caracterização da colaboração científica a partir dos resultados de medidas de centralidade de redes complexas aplicadas a redes de coautoria.

1.5 Organização da Dissertação

Esta dissertação está organizada da seguinte forma:

- No Capítulo 2, referendamos a fundamentação dos conhecimentos utilizados para a proposta realizada neste trabalho, teoria dos grafos, redes complexas e as medidas de centralidade.
- No Capítulo 3, explanamos a metodologia utilizada, baseada na obra de (De Nooy et al. 2018) que consiste em quatro etapas fundamentais: a) definição da rede, b) tratamento de dados da rede, c) características estruturais da rede e d) inspeção visual.

- No Capítulo 4, apresentamos os resultados produzidos com suas respectivas análises e interpretações a respeito do uso das medidas de centralidade para avaliação da colaboração científica em redes de coautoria.
- No capítulo 5, manifestamos as considerações finais, com comentários dos resultados obtidos e a sugestão de trabalhos futuros contextualizando com o problema deste trabalho.

O capítulo 1, versou sobre a contextualização, motivação e objetivos deste trabalho. O capítulo 2, seguinte, trata da fundamentação teórica utilizada.

2

FUNDAMENTAÇÃO TEÓRICA

A Fundamentação deste trabalho encontra-se respaldado na teoria dos grafos, na análise de redes sociais e medidas de centralidade de redes complexas.

2.1 Teoria dos Grafos

Em teoria de grafos, temos que um grafo é o conjunto de $G = (V, E)$ (G = Grafos, V = Vértices e E = Arestas), onde V é um conjunto não vazio de objetos denominados vértices (ou nós) e E é um subconjunto de pares não ordenados de V . Neste entendimento temos que grafos pode serem caracterizados no prisma de algumas propriedades, como por exemplo o comportamento da ligação das arestas pode ser classificado como não direcional ou direcional, onde cada aresta \vec{E} de $\vec{G} = (V, \vec{E})$ possui um sentido ou orientação definido por um par ordenado (x_1, x_2) , sendo x_1 o ponto inicial e x_2 o ponto final da aresta em questão.

O estudo de redes complexas é um tema que abrange diversas áreas de conhecimento, tais como a ciência da computação, matemática, física, biologia e sociologia. O termo “redes complexas” refere-se a um grafo que apresenta uma estrutura topográfica não trivial, composto por um conjunto de vértices (nós) que são interligados por meio de arestas (Barabasi 2003). O estudo de redes na forma de grafos é um dos pilares da matemática discreta e teve início em 1735, quando Euler propôs uma solução para o problema das pontes de Königsberg, originando a teoria dos grafos.

Redes de coautoria se caracteriza por ligação não direcional, ou seja, que independe de quem seja o ponto de origem para o destino da ligação, pois devemos para a proposta deste trabalho, considerar que se uma instituição publicou um artigo em coautoria com outra instituição, a recíproca é verdadeira e não há distinção de qual tem propriedades valoradas, como a iniciativa ou mesmo o grau de participação do coautor na publicação em tela.

As figuras abaixo representam a plotagem de grafos apresentando a visão da abordagem de coautorias intra-regional conforme a Figura 2.1, e coautorias inter-regional vide

Figura 2.2, onde S é a representação da região da instituição, X os autores (vértices ou nós) e as arestas/ligações representam a relação de coautoria entre eles.

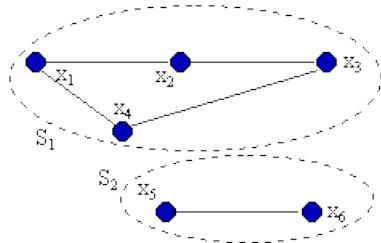


Figura 2.1: Grafo com representação de Rede de Coautoria intra grupos

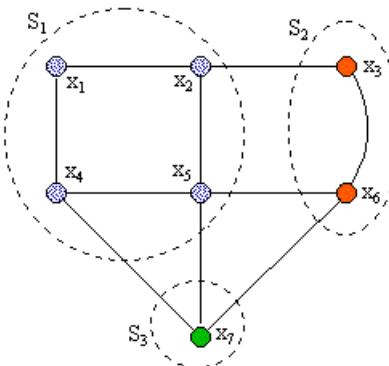


Figura 2.2: Grafo com representação de Rede de Coautoria inter grupos

2.2 Análise de Redes Sociais

Conforme (Silva et al. 2006), a análise de redes sociais interessa a pesquisadores de vários campos do conhecimento que, na tentativa de compreender o seu impacto sobre a vida social, deram origem a diversas metodologias de análise que têm como base as relações entre os indivíduos, em uma estrutura em forma de redes.

A análise de redes sociais (ARS ou SNA, da expressão em inglês *Social Network Analysis*) é uma abordagem oriunda da sociologia, da psicologia social e da antropologia. (Freeman 1996, Wasserman & Faust 1994). ARS não é algo recente, há estudos, como o trabalho de (Otte & Rousseau 2002) que buscaram desde a década de 1970, análises de redes de informação, de pesquisadores e de citações, como também de redes de coautoria.

Redes de coautoria utilizam-se da análise de redes sociais embasada nas propriedades e aplicações de redes complexas, a qual vemos a seguir.

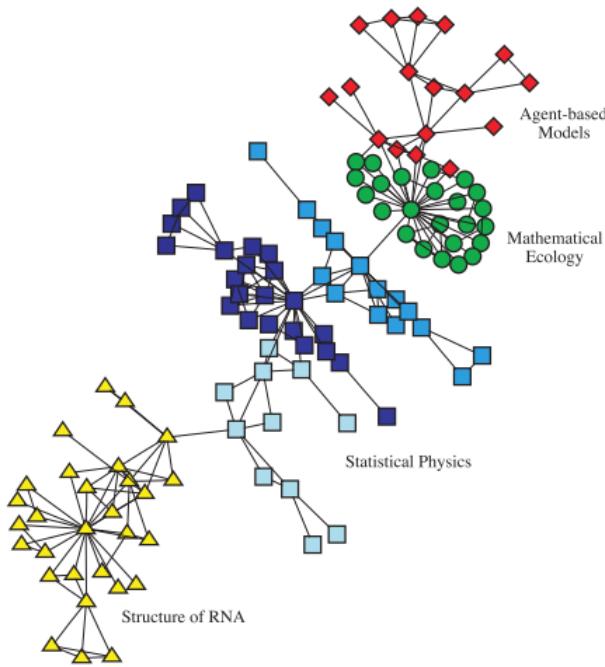


Figura 2.3: Pequena Rede de Coautoria de Newman

2.3 Redes Complexas

No contexto da teoria de redes complexas, uma rede corresponde a um grafo, que se representa por um conjunto de nós ligados por arestas, que em conjunto formam uma rede. Esta rede ou grafo, permite representar relações com fundamento das propriedades dos vários tipos de grafos e como são constituídos, ou seja, como se agrupam os seus nós e são ligadas suas arestas.

As propriedades de redes complexas às quais utilizaremos nesse trabalho se encontram definidas a seguir.

Definição 2.3.1 (Grau de um nó). O número de vértices adjacentes a um determinado vértice x_i chama-se de grau de x_i . Em um grafo direcionado $\vec{G} = (V, \vec{E})$, diferenciamos o grau de saída (output) do grau de entrada (input):

$$k_{x_i}^{out}(\vec{G}) = \sum_{x_j \in V: x_j \neq x_i} A[x_i, x_j] \quad (2.1)$$

$$k_{x_i}^{in}(\vec{G}) = \sum_{x_j \in V: x_j \neq x_i} A[x_j, x_i] \quad (2.2)$$

Desse modo, chamamos de grau de x_i :

$$k_{x_i}(\vec{G}) = k_{x_i}^{out}(\vec{G}) + k_{x_i}^{in}(\vec{G}) \quad (2.3)$$

Definição 2.3.2 (Centralidade de um nó - *Degree centrality*). Definida como o número de ligações incidentes de um vértice, esta medida nos informa qual a probabilidade de um vértice receber alguma informação da rede. Logo, a centralidade de um vértice v de um grafo $G = (V, E)$ é dada por:

$$C_D(v) = \deg(v) \quad (2.4)$$

Uma vez que a direção das arestas podem influenciar o cálculo, em redes direcionadas existem duas medidas distintas de centralidade de grau: indegree, que corresponde ao número de ligações direcionadas para o nó e outdegree, que consiste no número de ligações que o nó encaminha para os demais vértices da rede.

Definição 2.3.3 (Intermediação - *Betweenness centrality*). Trata-se de uma medida de centralidade de um grafo baseada em caminhos mínimos. A intermediação de um nó v é dada do seguinte modo:

$$g(v) = \sum_{s \neq t \neq v} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (2.5)$$

Onde σ_{st} é o número total de caminhos mínimos do nó s para o nó t e $\sigma_{st}(v)$ é o número de caminhos mínimos de s para t que utilizam o nó v como intermediário.

A intermediação também pode ser normalizada, sem perda de precisão, dividindo seu resultado pelo número de pares de nós não incluindo v , de maneira que $g(v) \in [0,1]$. Tal medida é vista como mais poderosa que apenas a conectividade, uma vez que fornece uma informação mais global sobre a rede em questão.

Definição 2.3.4 (Proximidade - *Closeness Centrality*). Em um grafo conectado, a proximidade de um nó é uma medida de centralidade calculada pela inversa da soma do comprimento dos caminhos mínimos entre um dado nó v e todos os demais nós do grafo. Desse modo, quanto mais central for o nó, mais próximo dos demais esse se encontrará. Sua fórmula é definida como:

$$C(v) = \frac{1}{\sum_{\forall u \in V, u \neq v} d_G(u, v)} \quad (2.6)$$

Onde, $d(u, v)$ é a distância do caminho mínimo entre u e v .

Usualmente, na literatura costuma-se usar sua versão normalizada, dada não mais pela soma de seus caminhos mínimos e sim pela média destes, como visto a seguir:

$$C(v) = \frac{N-1}{\sum_{\forall u \in V, u \neq v} d_G(u, v)} \quad (2.7)$$

Onde, N é o número de nós do grafo, ou seja, $N = |V|$. Para grafos extremamente grandes a diferença -1 torna-se irrelevante, podendo ser descartada.

Definição 2.3.5 (Coeficiente de agregação ou aglomeração - *clustering coefficient*). Trata-se

do número de ligações entre os vizinhos mais próximos de um nó.

Definição 2.3.6 (Redes estáticas/dinâmicas). Uma rede é estática quando não há variação do número de nós e, é dinâmica, quando é possível modelar o seu crescimento pela análise da variação da sua estrutura ao longo do tempo.

Os estudos das medidas de centralidade de redes foram evoluindo ao longo do tempo; centralidade, proximidade e intermediação é apontada por (Freeman 1996), como medidas essenciais aos estudos e análises de redes.

A figura abaixo ilustra a topologia das medidas de centralidade: a) *degree centrality*; b) *closeness centrality* e c) *betweenness centrality*. Com elas podemos perceber por meio dos vértices de cor vermelha a disposição dos mesmos na topologia da rede, e seus respectivos posicionamentos, nos levando a compreender o tipo da sua centralidade e seu papel na rede.

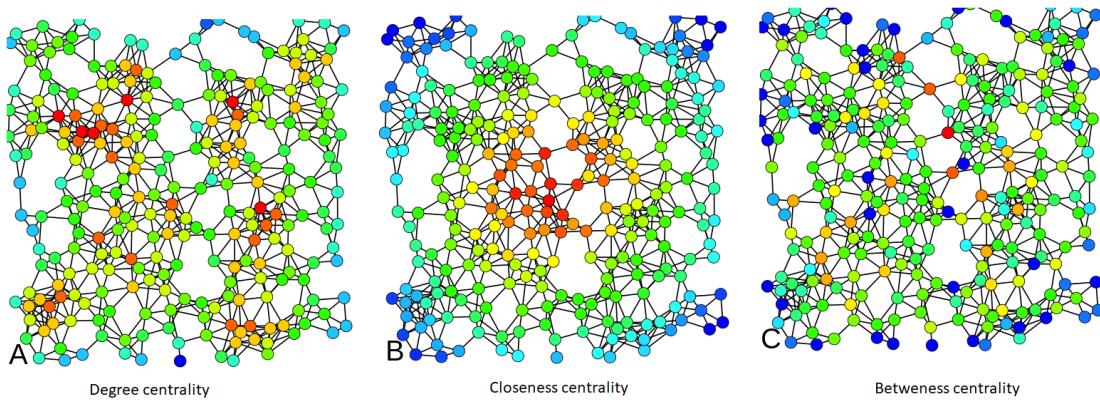


Figura 2.4: Redes com ilustração das medidas de centralidade (Fonte: Claudio Rocchini - Adaptado)

2.4 Colaboração Científica

Colaboração científica pode ser definida como interação que ocorre dentro de um contexto social entre dois ou mais cientistas, que facilita a partilha de significado e conclusão de tarefas em relação a um objetivo mútuo, compartilhado e organizado. Os cientistas que colaboram também podem trazer metas individuais adicionais para uma colaboração. (Sonnenwald & Sonnenwald 2007)

Ainda, segundo a autora, a colaboração ocorre dentro do contexto social da ciência, que inclui elementos como a revisão por pares, sistemas de prêmios, paradigmas científicos, políticas de ciência nacionais e internacionais, bem como normas implícitas ao campo disciplinar das instituições de pesquisa e/ou universidades.

Conforme (Newman 2004), há algum tempo se percebe que a coautoria de artigos em periódicos acadêmicos fornece uma janela sobre padrões de colaboração dentro da comunidade

acadêmica.

Para ([Sonnenwald & Sonnenwald 2007](#)), estudos da colaboração científica está aumentando, por possuir o potencial de auxiliar a resolução de problemas científicos complexos e promover agendas políticas, econômicas e sociais.

Neste Capítulo, explanamos os conhecimentos e fundamentações teóricas necessárias para este trabalho. No Capítulo seguinte é apresentada a metodologia utilizada para o desenvolvimento deste.

3

METODOLOGIA

ESTE capítulo descreve a metodologia utilizada neste trabalho, descrevendo suas respectivas etapas.

A metodologia utilizada para a presente pesquisa utiliza um modelo exploratório que consiste em quatro etapas cíclicas: a) definição da rede, b) tratamentos de dados da rede; c) determinação das características estruturais e d) inspeção visual (De Nooy et al. 2018).

A definição da rede resultou da coleta dos dados realizada no portal da SciELO¹. Foram realizados tratamentos desses dados normalizando a padronização os campos da Unidade Federativa (UF). A base foi complementada pela informação da localidade geográfica associando os estados (UF) às regiões do Brasil.

Os dados foram coletados no sítio eletrônico da SciELO a partir do endereço http://static.scielo.org/tabs/tabs_bra. Os dados associam os documentos publicados em coautoria de um autor com sua respectiva instituição de origem por um ID. O conjunto de dados amostral utilizado para esta pesquisa compreende o período entre 2008 à 2017 e os documentos relacionados as Universidades Federais do país.

Foi utilizada a linguagem de programação R, possibilitando que o resultado deste trabalho seja coberto ponta a ponta com a mesmo código, facilitando assim a reproduzibilidade desta pesquisa.

O total de documentos indexados coletados da base de dados resultou um montante de 1.302.659 documentos, sendo um total de 112.762 para área de *Health Sciences*, 84.437 para área de *Agricultural Sciences* e 15.278 para *Exact and Earth Sciences*.

A produção dos resultados teve por fundamento a teoria dos grafos e na análise de redes sociais, onde buscou a aplicação das métricas de redes complexas, para a compreensão do comportamento das redes de coautoria do presente estudo.

Este trabalho utilizou-se da linguagem R com o uso dos pacotes tidyverse, igraph, ggraph, sna e visNetwork. A seguir descrevemos as etapas cíclicas proposta por (De Nooy

¹<http://static.scielo.org>

et al. 2018) para estudos relacionados ARS.

3.1 Definição da rede

As redes objeto de estudo desta pesquisa é a relação de coautoria existentes nos documentos indexados na base de dados da SciELO, objetivando a criação de um modelo exploratório. Foram definidas as redes de coautoria segregadas pela Unidade Federativa (UF) e Região Geográfica do Brasil para as Universidades Federais do país. Redes de coautoria são classificadas como não-direcionada, ou seja, as ligações existentes entre os vértices independem de uma orientação ou direção.

Os vértices da rede são as Unidades Federativas (UF), agrupadas pelas instituições federais às quais fazem parte da mesma, as arestas as ligações refere-se a ocorrência de coautoria entre as instituições das UF. As arestas são valoradas possuindo pesos pela frequência de ocorrência das coautorias, os vértices são ponderados pelo volume de coautorias, bem como pelos laços, ou seja, as coautorias realizadas entre instituições da mesma UF.

3.2 Tratamento de dados da rede

Nesta etapa foram elaborados alguns tratamentos da rede, tal como limpeza e padronização de alguns dados, de forma mínima e minuciosa, para evitar que a reproduzibilidade reste prejudicada. Buscamos a padronização dos nomes das Unidades Federativas, e a associação de suas respectivas região geográfica.

As Universidades Federais do Brasil foram agrupadas por estados sendo representadas pelo nome no Estado (UF - Unidade Federativa). Foram realizadas rotinas de geração das matrizes e *datasets* necessários aos cálculos das métricas, tal como a plotagem das redes e gráficos.

3.3 Determinação das características estruturais

Conforme (De Nooy et al. 2018) a exploração de uma rede por suas métricas que descreve sua caracterização estrutural é uma forma mais concisa e precisa quando comparada a inspeção visual, no entanto as métricas por vezes são abstratas e de difícil interpretação. Por isso, deve-se utilizar ambas abordagens analíticas as métricas estruturais da rede e sua observação gráfica.

Foram escolhidas as métricas básicas do número de vértices e arestas, grau médio do vértice, diâmetro da rede, e distância média, centralidade do grau, centralidade de proximidade, e centralidade de intermediação, que foram aplicadas ao período em análise de 2008

à 2017 para a área de *Health Science*. As métricas aferidas denotam de forma quantitativa o comportamento da colaboração científica com bases nas redes de coautoria geradas.

3.4 Inspeção visual

Uma forma enriquecedora de análise de redes sociais é a observação gráfica, para isso são gerados gráficos das redes para representar suas relações e permitir por análise visual a o conhecimento da dinâmica existente pela disposição dos vértices e arestas.

Primeiramente insta ressaltar que no estudos iniciais e exploratórios da base SciELO foram encontradas diversos tipos de instituições de ensino e pesquisa, e que para o objeto do estudo deste trabalho os escopos das redes foram definidos pelas coautorias existentes entre as Universidades Federais do país em suas respectivas Unidades Federativas (UF).

A análise das redes de coautoria consistiu em uma perspectiva regional explorando suas propriedades e características. Foram definidos a área de *Health Sciences* e o período compreendido entre 2008 à 2017, considerando a rede global que consiste todas as Universidades Federais do Brasil agrupadas pelo Estado (UF - Unidade Federativa) a que correspondem. E outra rede tendo como vértice local (ou focal) a Unidade Federativa Alagoas, de maneira a se observar a dinâmica de interação e evolução dessas redes em um prisma analítico e comparativo das redes de coautoria ao longo do tempo.

Os resultados obtidos (gráficos e numéricos) permitem-nos obter um conhecimento descritivo das interações existentes na relação de coautoria entre as relações de coautorias estudadas da base SciELO. No aspecto gráfico as arestas possuem espessuras maior ou menor de acordo com o peso definido pelo número de publicações em coautoria, assim, as UF que possuem interações mais fortes ou mais fracas.

Os vértices foram definidos de um mesmo tom de cor para representar a região geográfica do país a qual pertence a UF, a espessura do laço do vértice detona também o volume de coautorias realizadas dentro da mesma UF.

Neste Capítulo foram apresentadas as quatro etapas da metodologia utilizada: determinação da rede, tratamento de dados, determinação das características estruturais, inspeção visual. No capítulo seguinte apresentamos as análises e interpretações dos resultados.

4

RESULTADOS E ANÁLISES

COM a aplicação dos métodos explanados e a fundamentação teórica referendada, apresentamos os resultados, contextualizando com o problema central deste trabalho e objetivos propostos.

Para a apresentação dos resultados subdividimos este capítulo em quatro partes, sendo as três primeiras referindo-se a cada uma das áreas pesquisadas, são elas: Características Estruturais das Redes, Medidas de Centralidade, Ranking das Medidas de Centralidade e Discussão dos Resultados.

Na primeira parte apresentamos os resultados das topologias das redes, número de vértice e arestas, características estruturais, densidade, coeficiente de clusterização. É apresentado também as matrizes de coautoriais de cada área para todo o período em cada área.

Na segunda parte, e as medidas de centralidade de cada rede: centralidade do grau (*degree centrality*), centralidade de proximidade (*closeness centrality*), centralidade de intermediação (*betweenness centrality*) a apresentação dos gráficos comparativos entre o posicionamento da Universidade Federal de Alagoas, quando comparada de forma analítica em relação ao comportamento das medidas médias das Universidades Federais do Brasil, seguindo com as interpretações e análises dos resultados.

Na terceira parte, é apresentado o ranking dos índices aferidos por cada Unidade Federativa (UF) em cada área do conhecimento para cada medida de centralidade.

Na quarta parte, realizamos a discussão dos resultados; contextualizando os resultados das medidas de centralidade de rede com a colaboração científica, entendendo o posicionamento do vértice focal (Alagoas) e o desempenho de cada UF em cada área.

4.1 Redes de Coutorias - Base SciELO

Neste ponto apresentamos a análise do comportamento e dos resultados das medidas de centralidade das redes de coautoria de artigos científicos indexados na base SciELO no

período compreendido entre 2008 à 2017 para as áreas de: *Health Sciences*, *Agricultural Sciences*, *Exact and Earth Sciences*.

4.1.1 Características Estruturais

Nesta seção trazemos a tona as características estruturais das redes objeto de estudo, sendo elas determinadas: Rede Brasil, compreendendo as Universidades Federais relacionadas ao respectivo Estado (UF - Unidade Federativa) às quais estão sediadas; para cada área do conhecimento visto acima.

Neste tópico apresentamos a estrutura da Rede Brasil com a finalidade de compreendermos a dimensão e topologia de cada área analisada dentro do período estudado, a base SciELO. Na figura 4.1.1, verificamos a projeção do número de vértices e arestas para cada ano.

Topologia: Vértices e Areostas por Área						
Área	Health Sciences		Agricultural Sciences		Exact and Earth Sciences	
	Ano	Vértices	Arestas	Vértices	Arestas	Vértices
2008	27	109	25	90	21	63
2009	25	122	26	100	25	87
2010	25	163	25	154	22	91
2011	25	165	27	150	23	88
2012	26	181	26	180	24	88
2013	26	183	26	178	22	86
2014	26	197	27	179	25	83
2015	27	194	26	173	22	77
2016	26	298	26	182	25	99
2017	26	199	25	190	24	94

Tabela 4.1: Topologia das Redes (Vértices e Areostas) por Ano

Pela topologia verificamos que as áreas de *Health Sciences* e *Agricultura Sciences* possuem um maior número de arestas ao longo do tempo quando comparada a *Exact and Earth Sciences*. Destaque para o ano de 2016 em que observamos um pico que eleva o número de arestas no ano de 2016, como visto na tabela 4.1.1 o que possível neste ano maior número de publicações/indexações foram realizadas, o que resultou o aumento da conectividade da rede, que impactou em sua densidade e também no resultados das medidas de centralidade como veremos a frente.

Em *Agricultural Sciences* podemos considerar uma normalidade, considerando o número médio de vértices e uma tendência crescente de arestas (coautoria) ao longo do tempo; *Exact and Earth Sciences* possui uma observação curiosa pela oscilação do número de vértices de

cada ano, ou seja a desconexão e reconexão destes, mostrando que ao longo do tempo as participações dos estados nas coautoriais alguns estados deixaram de estar colaborando um com o outro.

A tabela 4.2 abaixo apresenta a densidade e aglomeração (clustering coefficient) para complementar a compreensão da estrutura e dimensão de cada rede, que caracteriza as relações de coautoriais existentes em cada área.

Características Estruturais da Rede Brasil por Área						
Ano	Densidade			Aglomerado (Clustering Coefficient)		
	Health Sciences	Agricultural Sciences	Exact and Earth Sciences	Health Sciences	Agricultural Sciences	Exact and Earth Sciences
2008	0,121	0,155	0,15	0,523	0,364	0,355
2009	0,165	0,152	0,145	0,625	0,415	0,398
2010	0,23	0,258	0,197	0,726	0,557	0,506
2011	0,235	0,214	0,174	0,744	0,543	0,356
2012	0,238	0,271	0,159	0,773	0,601	0,391
2013	0,242	0,271	0,186	0,772	0,619	0,429
2014	0,265	0,252	0,138	0,745	0,619	0,395
2015	0,238	0,258	0,167	0,803	0,608	0,339
2016	0,418	0,282	0,165	0,943	0,642	0,489
2017	0,266	0,313	0,17	0,798	0,682	0,422

Tabela 4.2: Características Estruturais Rede Brasil (2008-2017)

Matrizes de Coautoriais por Unidades Federativas

A seguir apresentamos as matrizes de coautoriais por Unidades Federativas por área, compreendendo todo o período (2008 à 2017). As matrizes é uma das formas primárias de observar a colaboração científica existente entre os estados, e nos referenciaremos a elas, ao longo das análises dos resultados das medidas de centralidade.

Health Sciences

	AC	AL	AM	AP	BA	CE	DF	ES	GO	MA	MG	MS	MT	PA	PB	PE	PI	PR	RJ	RN	RO	RR	RS	SC	SE	SP	TO	
AC	223																											
AL	1	2689																										
AM	4	15	610																									
AP	1	3	4	134																								
BA	2	27	6	2	4874																							
CE	5	17	15	2	45	7096																						
DF							12																					
ES	1	28	11	2	33	9		2112																				
GO	3	44	6	2	16	27	1	19	4744																			
MA	1	6	4	2	35	80			13	2	2945																	
MG	22	115	31	23	397	107		189	349	70	26387																	
MS	100	1			13			1	21			34	417															
MT	5	50	3	2	5	25		29	30	4	41	10	1369															
PA	2	12	30	16	8	13	1	9	19	11	51	1	6	1144														
PB	4	118	23	2	101	93		23	20	53	65		22	6	3882													
PE	2	219	40	2	175	156		28	28	33	250	1	12	27	440	6410												
PI	2	20	3	2	8	250		2	23	25	81		8	11	56	77	1933											
PR	24	12			59	31		13	27	8	172	2	14	11	15	44		8181										
RJ	7	108	35	35	200	142		166	106	46	710	8	107	62	143	127	123	156	15338									
RN	8	26	13	2	64	139		96	20	66	761	6	11	6	242	199	23	77	213	10292								
RO	32	9	14	2	2	2		2	3	2	14	1	2	2	2	3	2		31	3	166							
RR	1	1	5	1	3	1		2	1	5	15		1	1	1	1	3	3	15	5	1	197						
RS	2	127	58	5	226	86		139	50	41	446	22	17	60	80	125	32	173	451	1153	2	2	24775					
SC	4	21	47	4	133	46		22	52	27	216	2	12	99	71	89	11	260	199	206	5	5	757	6783				
SE	1	154	7	2	68	19		5	5	5	112		6	3	91	190	2	39	45	35	2	1	112	24	3624			
SP	20	193	36	9	205	163	2	174	186	102	915	96	69	53	141	304	18	241	421	274	7	4	619	152	112	27726		
TO						3			31		14							5		1	2	12			2		30	123

Figura 4.1: Matriz de Coautorias *Health Sciences* (2008-2017)

Observamos para todo período analisado que a matriz na área de *Health Sciences*, os valores absolutos de coautorias entre os estados, denota uma forte conectividade desta rede, quando comparada as áreas de *Agricultural Sciences* e *Exact and Earth Sciences*, exceto para alguns estados, da região Norte do país, como Tocantins e Roraima. Podemos perceber o grande volume de coautorias existente dentro do mesmo estado, como São Paulo, Rio Grande do Sul, Minas Gerais e Rio de Janeiro.

Agricultural Sciences

	AC	AL	AM	AP	BA	CE	DF	ES	GO	MA	MG	MS	MT	PA	PB	PE	PI	PR	RJ	RN	RO	RR	RS	SC	SE	SP	TO
AC	292																										
AL		1833																									
AM	14	3	320																								
AP			9																								
BA	6	6		2253																							
CE	9	6	8		36	3208																					
DF							20																				
ES	2	2	1		6	6		1670																			
GO	1	25	3	2	26	13	16	20	3298																		
MA		3			2	85				10	167																
MG	42	116	59	4	193	161	1	302	446	22	32724																
MS	27	97	1		12	17		2	147		86	2254															
MT	3	68	2		11	7		27	80	31	497	26	3122														
PA		10	13		7	40		3	26	1	219	7	29	2286													
PB	1	107	8		315	161		47	2	45	239	1	70	67	6782												
PE	14	207	61		252	139		98	16	11	311	16	59	75	843	7351											
PI	1	9	2		33	66		2	18	69	272	1	57	5	179	115	1930										
PR	7	12	9		31	3		73	11	1	162	53	38	3	4	53	5	3908									
RJ	7	30	1		29	33		114	22	2	472	6	108	260	43	100	30	23	5031								
RN	25	135	20	3	163	268	3	478	92	50	3824	82	326	125	502	618	97	80	217	15816							
RO	8	1	4		3	3		12	1		16	13	4		9		1	1	4	32	35						
RR		2	2								51		4	4	12	12	2	2	10	100	195						
RS	9	105	13		159	8	1	95	234		573	49	182	143	193	246	42	550	138	319	45	9	37794				
SC	3	3			6	9		1	8	1	133	11	3	13	19	10		80	27	27	1	1	395	1343			
SE		14	1		5	16		6	8		114			8	62	58	11	6	11	43			3	5	633		
SP	2	3	1		8	2		18	20		73	3	15	9	15	15	1	4	13	36			8	11	3	548	
TO	1	1			30			4	132	11	266	17	12	45	13	4	6	7	4	109	3		95	15	1	10	1290

Figura 4.2: Matriz de Coautorias *Agricultural Sciences* (2008-2017)

Com a matriz de coautorias na área de *Agricultural Sciences*, percebemos uma densidade menor e muitas lacunas da colaboração entre alguns estados federados, como o Distrito Federal, Roraima, e pouca volume em número de coautorias entre diversos estados. Destaque para o estado do Rio Grande do Sul, e no nordeste o Rio Grande do Norte, que por meio de suas Universidade Federal do Rio Grande do Norte e da Universidade Federal Rural do Semi-Árido, apresentaram grande participação nas coautorias em Ciências Agrárias.

Exact and Earth Sciences

	AC	AL	AM	AP	BA	CE	DF	ES	GO	MA	MG	MS	MT	PA	PB	PE	PI	PR	RJ	RN	RO	RR	RS	SC	SE	SP	TO	
AC	7																											
AL		594																										
AM		4	235																									
AP				3																								
BA		8			766																							
CE		39	16		22	2007																						
DF							3																					
ES		5			4			470																				
GO		29	7		23	14			25	429																		
MA		4				17				120																		
MG	1	63	11		23	29			82	84	18	3728																
MS		28			6				1		10	44																
MT	4					17			1	2	10	1	450															
PA			9	2	4	13		3		3	4			74	781													
PB		44	4		43	11		5	2	20	18					14	970											
PE		29	6	1	29	59			2		27			5	8	174	1004											
PI			3		23	11			5	14	15			9	7	28	35	200										
PR	2	3	64		12	7		11	3		22				22	10	13		1160									
RJ		41	28	3	44	229		20	21	7	142			15	8	11	33	7	66	2539								
RN		2		2	13	23		11	4		201			5	8	33	24	3	11	42	2736							
RO			1											4	1		1				15	8	28					
RR			1								4				1								2					
RS		25	3		14	35		13	1	6	46	4		22	4	35		45	54	325		9	2935					
SC	5	11			7	1		2	6		58			11	2	9	1		53	54	52		4	107	1048			
SE		9	13		25	13		4			16			1		24	44		55	4	5			5	344			
SP	1	40	32		21	81		3	10		78	18	8	28	32	10	33	36	44	36			36	28	83	1091		
TO									1		6							2	1	5							14	

Figura 4.3: Matriz de Coautorias *Exact and Earth Sciences* (2008-2017)

Em *Exact and Earth Sciences* percebemos que pelo volume de coautoria, a Base SciELO não é um referência para a indexação de artigos nesta área, sua matriz apresenta poucas coautorias, realizada entre alguns estados, apresentando muitas lacunas principalmente em estados do Centro-Oeste e Norte do país. Destaque para Minas Gerais e Rio de Janeiro.

Visualização da rede

Para inspeção visual que coaduna com as análises numéricas, postamos a seguir, por seleção aleatória o ano de 2015 a rede de *Agricultural Sciences* para Rede Brasil e Rede Alagoas, trata-se de uma plotagem exemplo uma vez que estamos trabalhando com 60 redes ao total, e como dito anteriormente nesse texto, o mosaico esta disponível no Apêndice A.

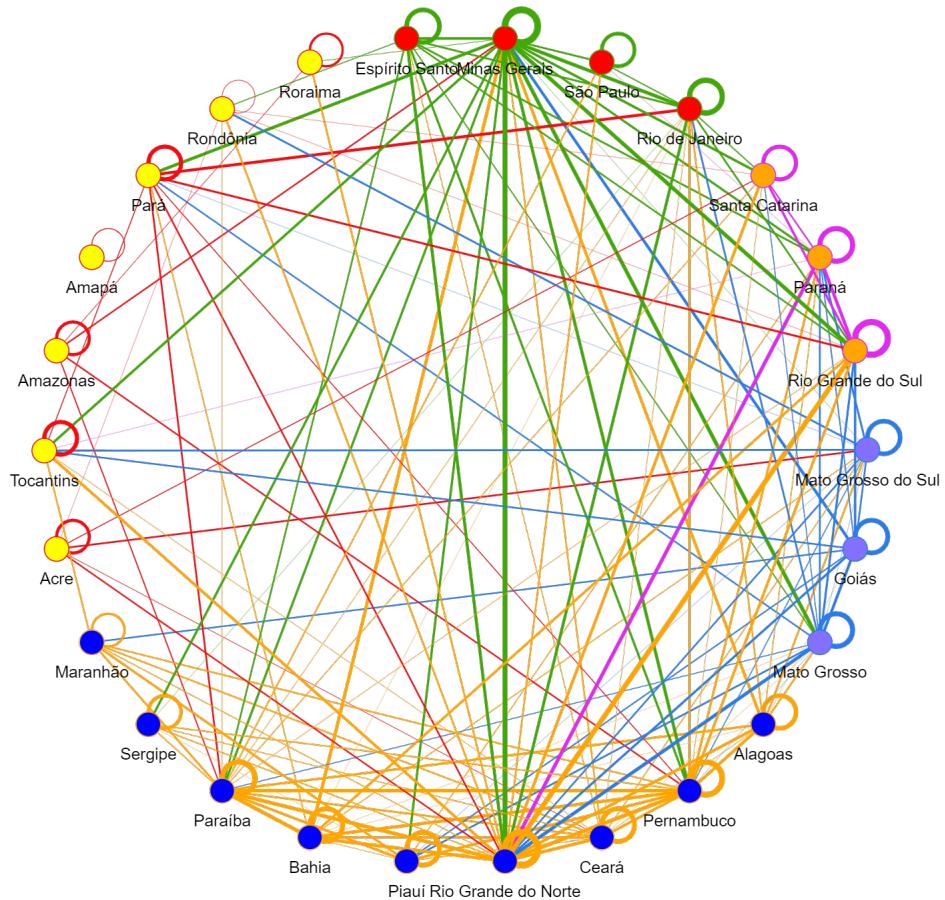


Figura 4.4: Rede de Coautoria das Universidades Federais do Brasil - 2015 (*Agricultural Sciences*)

As características propostas na plotagem da rede é, formato em círculo para fins comparativos, disposição dos vértices com coloração das UF pertencentes a mesma região geográfica, espessura e coloração das arestas indicando o maior volume/peso das coautorias existente entre as UFs, espessuras dos laços, indicando peso das coautorias existentes na mesma UF.

A figura 4.4 apresenta a rede de coautoria das Universidades Federais do Brasil, composta pelos vértices que representam os Estados (Unidades Federativas - UF) do Brasil e suas ligações (arestas) representando as coautorias existentes entre eles. A figura 4.5 representa para o mesmo ano e área a rede plotada a partir do vértice Alagoas.

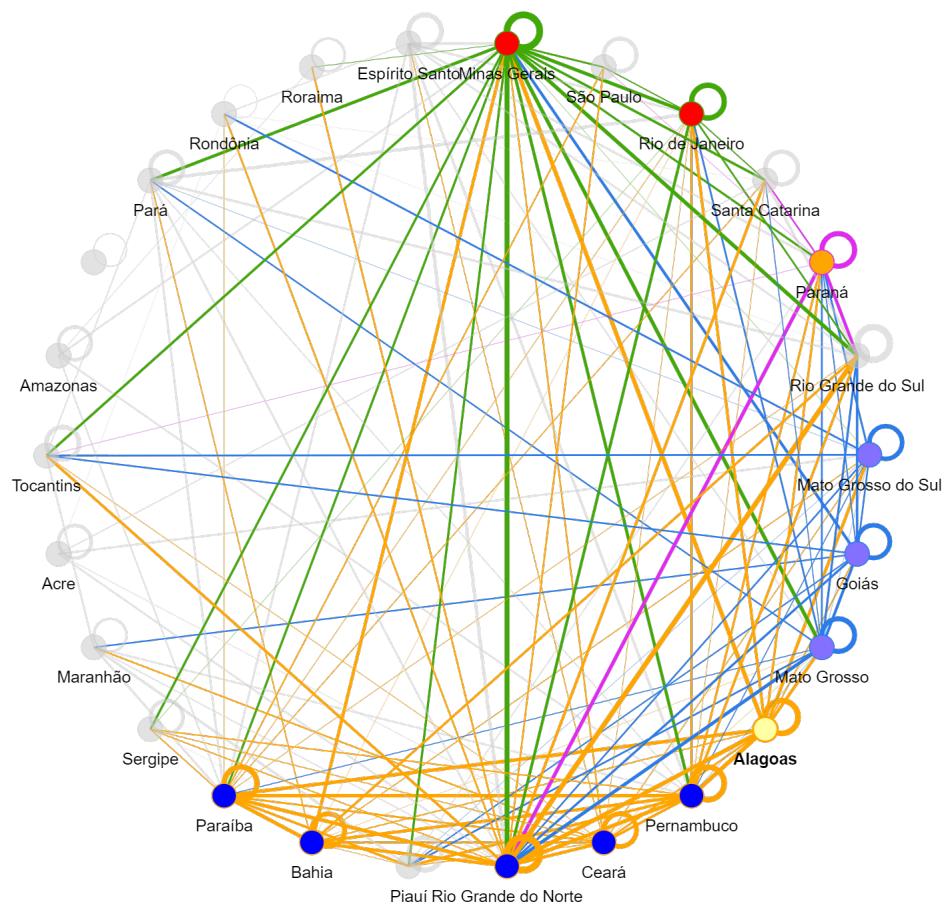


Figura 4.5: Rede de Coautoria das Universidades Federais de Alagoas - 2015 (*Agricultural Sciences*)

4.2 Medidas de Centralidade

Nesta seção, após a análise das características estruturais das Redes Brasil para cada área do conhecimento, determinamos o vértice focal de Alagoas, como Rede Alagoas, representado pela Universidade Federal de Alagoas, como ponto de interesse para uma avaliação do posicionamento e desempenho desta rede/vértice em comparação a Rede Brasil (média). Sendo esta, organizada por subseções para cada medida de centralidade.

Após a análise de cada medida, será apresentado um Ranking por cada área do e o o respectivo índice médio do período (2008-2017) para cada Unidade Federativa (UF).

Válido ressaltar que foge ao escopo desse trabalho o entendimento de fatores exógenos ou específicos que não seja o do estudo das medidas de centralidade aplicadas a avaliação da colaboração científica, considerando que há várias possibilidades de investigações *drill-down* que possam detalhar conhecimento complementares a respeito das coautorias.

4.2.1 Centralidade do Grau

Esta medida indica o melhor posicionamento dos vértices utilizando o parâmetro que possuem o maior número de ligações entre outros vértices, ou seja, a partir desse ponto de vista são vértices mais influentes, observaremos graficamente como se comportou a Rede Alagoas face a média da Rede Brasil.

A tabela 4.3 apresenta os resultados da centralidade do grau para todas áreas comparando a Rede Brasil (média) e Alagoas.

Centralidade do Grau (Degree centrality) por Área						
Ano	Health Sciences		Agricultural Sciences		Exact and Earth Sciences	
Rede	Brasil	Alagoas	Brasil	Alagoas	Brasil	Alagoas
2008	6,29	7	7,44	5	8,19	2
2009	7,92	7	7,62	8	10,24	14
2010	11,04	15	12,40	12	12,55	14
2011	11,28	10	11,11	13	11,30	12
2012	11,92	12	13,54	15	10,67	18
2013	12,07	14	13,54	14	12,00	12
2014	13,23	18	13,11	13	9,44	14
2015	12,37	19	12,92	13	10,36	6
2016	20,92	24	14,08	12	12,32	12
2017	13,30	19	15,04	15	12,17	14

Tabela 4.3: Medidas de Centralidade do Grau por Área (2008-2017)

Health Sciences

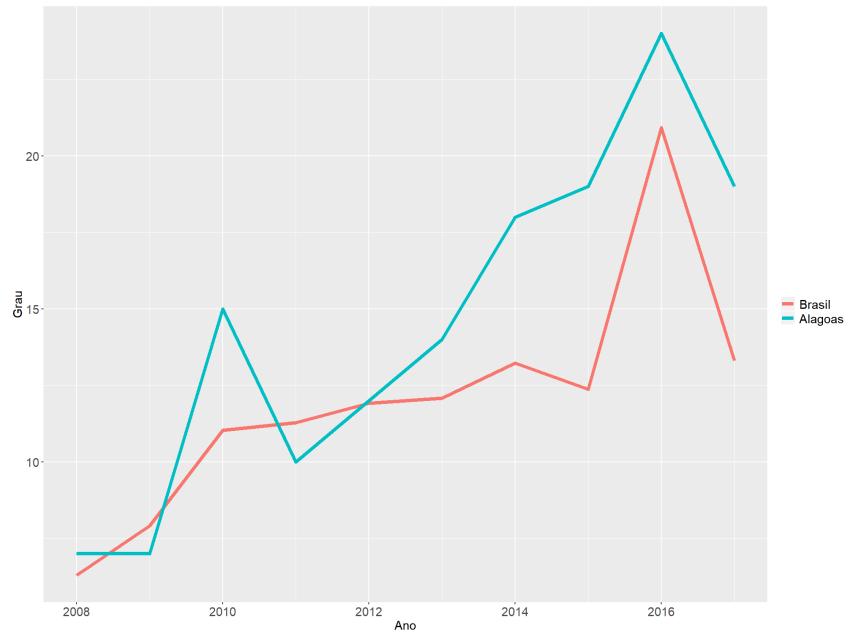


Figura 4.6: Centralidade do Grau (*Health Sciences*)

Agricultural Sciences

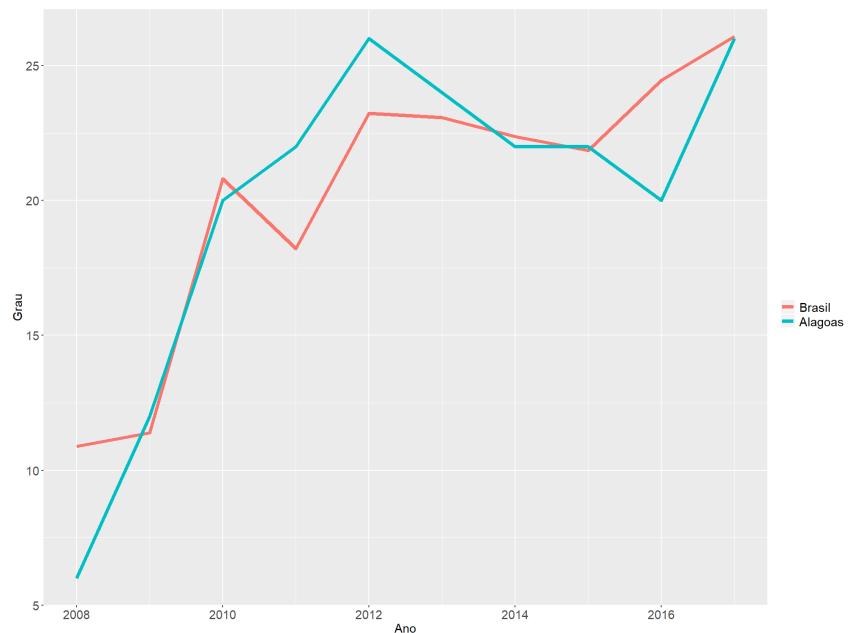


Figura 4.7: Centralidade do Grau (*Agricultural Sciences*)

Exact and Earth Sciences

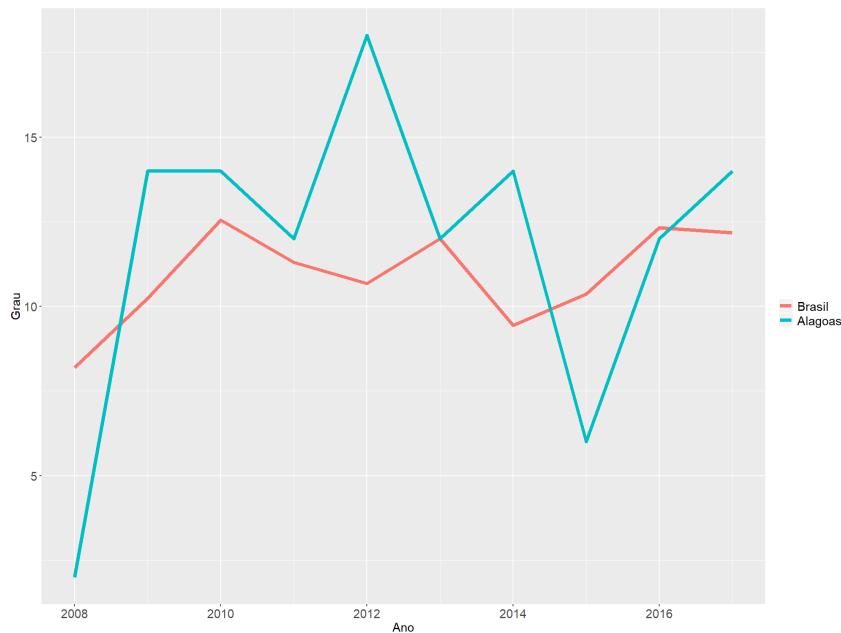


Figura 4.8: Centralidade do Grau (*Exact and Earth Sciences*)

Os gráficos 4.6, 4.7 e 4.8, apresentam respectivamente as curvas na série temporal (2008-2017) da centralidade do grau da média da Rede Brasil e Alagoas.

Na disposição das curvas na área de *Health Sciences*, sendo Rede Brasil (vermelha) e Alagoas (azul), percebemos ao longo do período projeção em direção análoga, e que exceto como aferido no ano de 2009 e 2011, Alagoas esteve com a centralidade do grau maior que a média do Brasil. O que nos levar a inferir que nesse período

O gráfico 4.7 mostra que Alagoas andou muito próximo da média do Brasil para *Agricultural Sciences* para este indicador. Destaque para o ano de 2012, em que obteve o maior índice de centralidade na rede, estas razões podem ser investigadas com maior detalhamento sobre as ocorrências das coautorias neste ano.

A análise da centralidade para área de *Exact and Earth Sciences* mostra que apesar de não ser uma área de grande volume de artigos indexados na base SciELO, a partir de 2009, Alagoas obteve um índice de centralidade maior que a média da rede. Com um queda considerável do seu índice em 2015, como dito uma análise investigatória poderá responder o que ocorreu neste ano, como a desconexão com outros vértices, não ocorrendo coautorias em pares com outros estados.

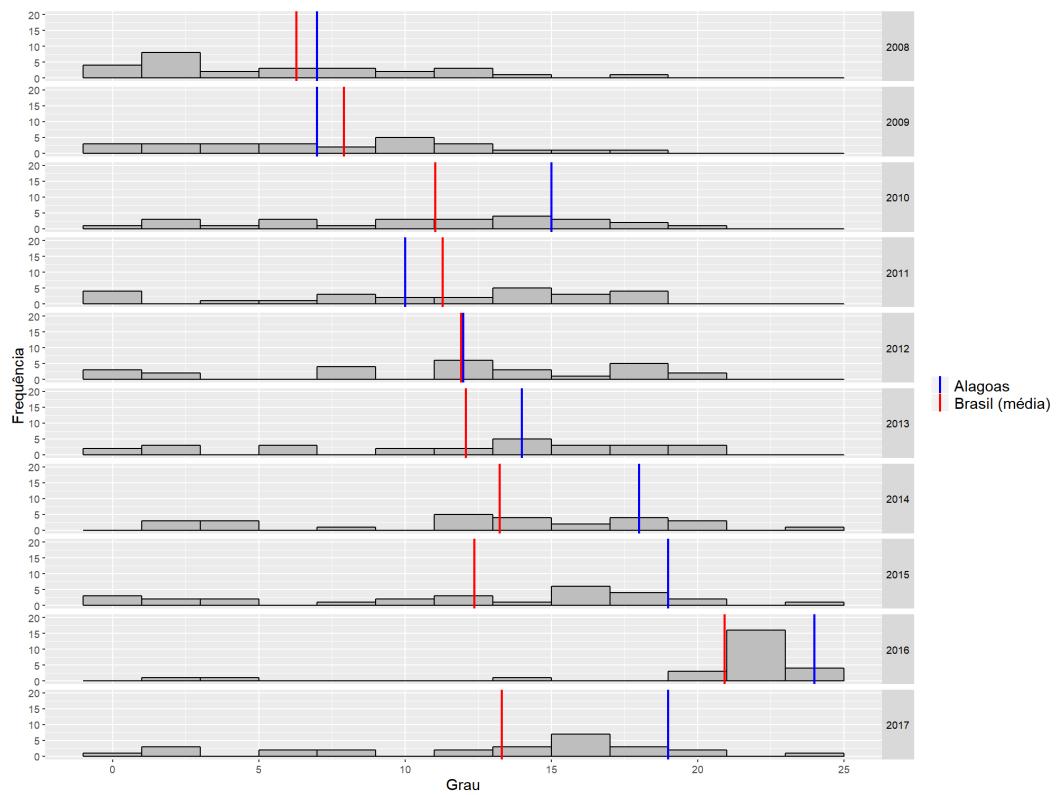


Figura 4.9: Histograma da Centralidade do Grau (*Health Sciences*)

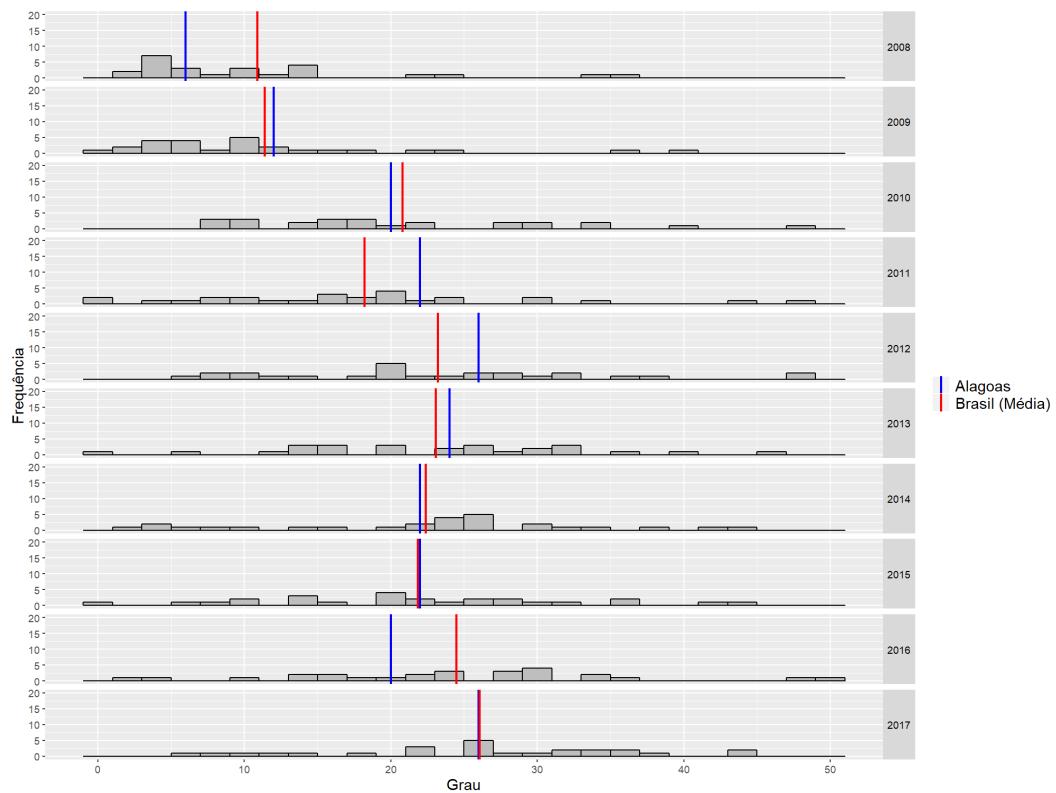


Figura 4.10: Histograma da Centralidade do Grau (*Agricultural Sciences*)

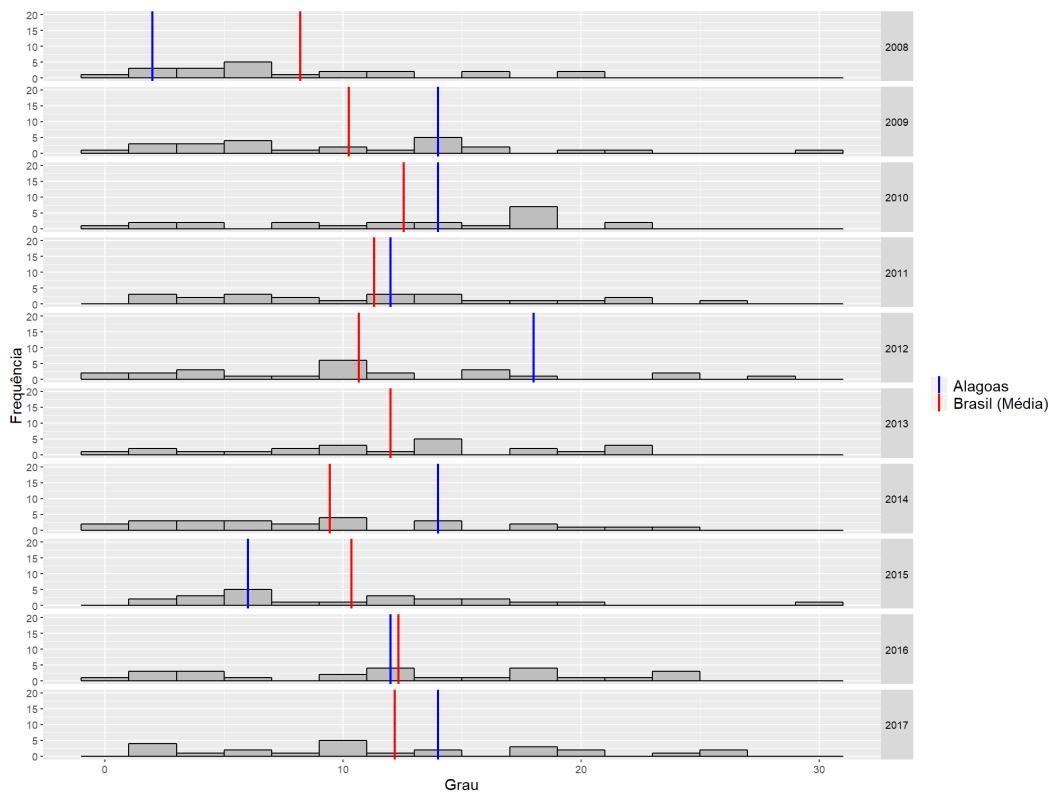


Figura 4.11: Histograma da Centralidade do Grau (*Exact and Earth Sciences*)

O histograma 4.9 indica a distribuição da medida de centralidade do grau para área de *Health Sciences*, maneira a qual percebemos uma grande frequência no ano de 2016, em virtude da clusterização observada na tabela 4.2, corroborando os anos que Alagoas (linha azul) esteve a frente neste marcador quando comparado a média do Brasil.

Na figura 4.10 observamos um posicionamento muito próximo da média da Rede Brasil (linha vermelha) e Alagoas (linha azul), e que exceto nos primeiros anos de 2008 a 2009 a distribuição concentrou-se na região central, de forma que muitos vértices obtiveram índices parecidos, exceto *outliers* que estão presentes em cada ano, dos quais podemos verificar na matriz 4.3 e no ranking apresentando em seção abaixo.

O histograma, figura 4.11 da área de *Exact and Earth Sciences* mostra o posicionamento da centralidade do grau de Alagoas acima do Brasil, e uma distribuição dos índices aferidos mais dispersos.

4.2.2 Centralidade de Proximidade

A figura 4.4 apresenta os resultados da medida de centralidade de proximidade em relação a Rede Brasil (média) e Alagoas, para todas as áreas. Essa medida indica quão próximo um vértice encontra-se um do outro, considerando seu posicionamento na rede. Indicando assim o seu papel de influência e importância pelo aspecto da proximidade.

Centralidade de Proximidade (Closeness Centrality) por Área						
Ano	Health Sciences		Agricultural Sciences		Exact and Earth Sciences	
Rede	Brasil	Alagoas	Brasil	Alagoas	Brasil	Alagoas
2008	0,5303	0,5434	0,5414	0,4528	0,4824	0,3333
2009	0,5789	0,5609	0,5543	0,5455	0,4978	0,561
2010	0,6772	0,7419	0,6531	0,6316	0,5435	0,5263
2011	0,6699	0,6216	0,6444	0,6487	0,5216	0,5366
2012	0,67	0,6666	0,6663	0,6757	0,54	0,6
2013	0,6852	0,7058	0,6747	0,6667	0,5544	0,5556
2014	0,6957	0,7812	0,6312	0,6191	0,5033	0,5641
2015	0,6697	0,7812	0,6676	0,6487	0,5306	0,5122
2016	0,8788	0,9615	0,678	0,625	0,5206	0,5111
2017	0,6993	0,8064	0,6995	0,6857	0,5267	0,575

Tabela 4.4: Medidas de Centralidade de Proximidade por Área (2008-2017)

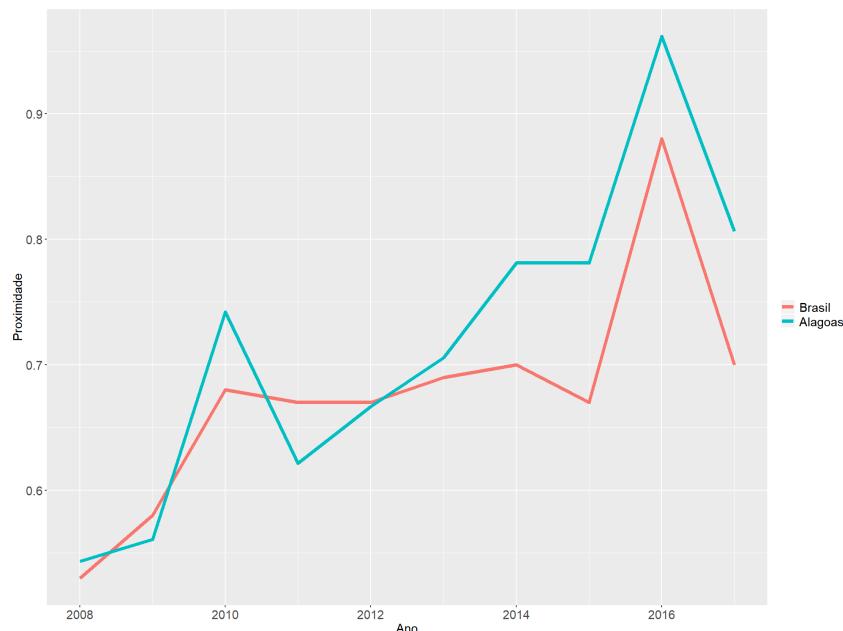
Health Sciences

Figura 4.12: Centralidade de Proximidade (Health Sciences)

Agricultural Sciences

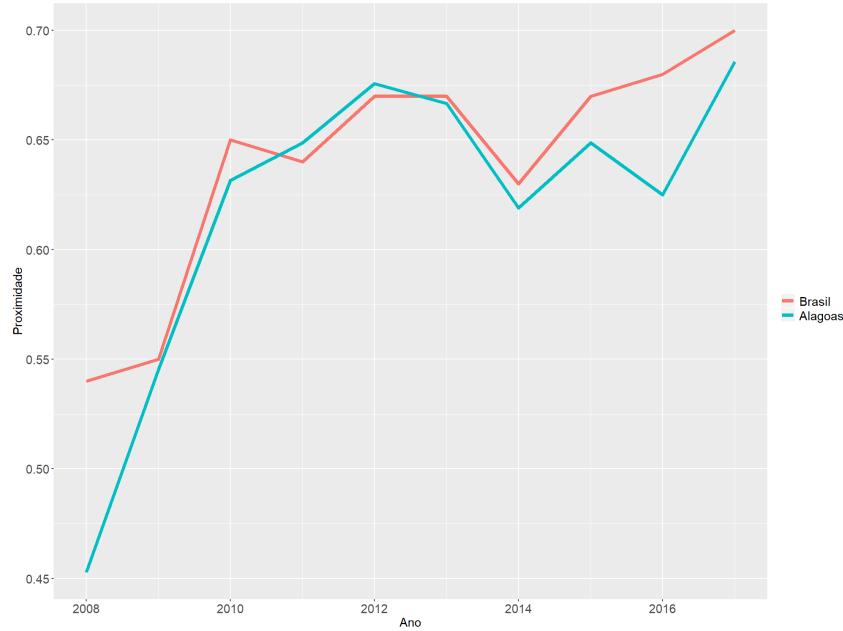


Figura 4.13: Centralidade de Proximidade (*Agricultural Sciences*)

Exact and Earth Sciences

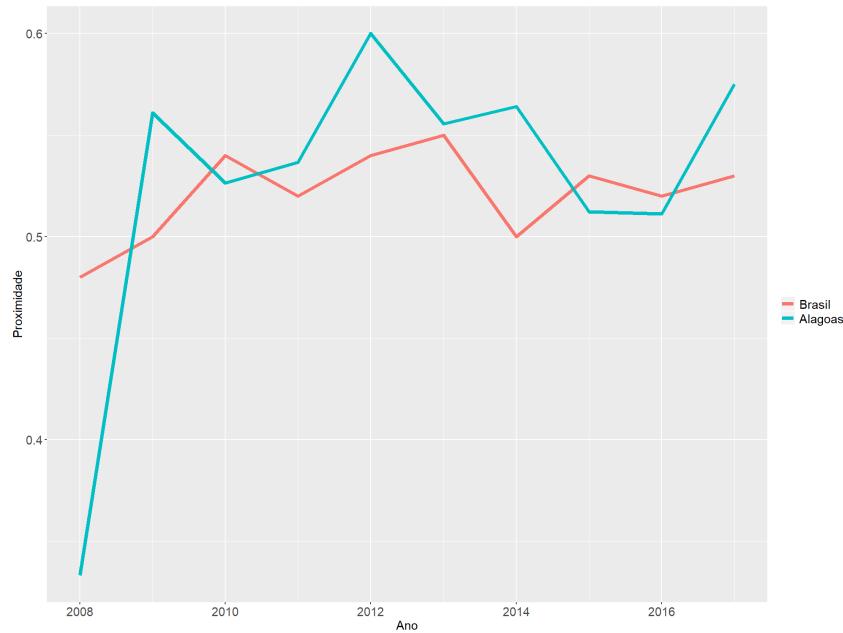


Figura 4.14: Centralidade de Proximidade (*Exact and Earth Sciences*)

As curvas da Rede Brasil (vermelha) e Alagoas (azul), também se projetaram em sentidos análogos, destaque que na área de *Health Sciences* conforme 4.12 que exceto nos anos 2009 e 2011 em todo o período manteve-se acima da média do Brasil.

Na área de *Agricultural Science* conforme figura 4.13, os resultados foram muito próximos, resultando que Alagoas ficou em que alguns anos abaixo da média, como vimos na figura da Matriz de Coautoria 4.3 e o que será corroborado pelo ranking apresentado abaixo, pelo destaque dos estados do Rio Grande do Sul e Rio Grande do Norte em Ciências Agrárias.

Na análise gráfica da centralidade de proximidade da área *Exact and Earth Sciences*, verificamos um baixo índice aferido por Alagoas no ano de 2008 e que a partir de 2009 seguiu em patamar superior a média do Brasil, tendo sido a área de maior destaque para medida de centralidade da proximidade.

Health Sciences

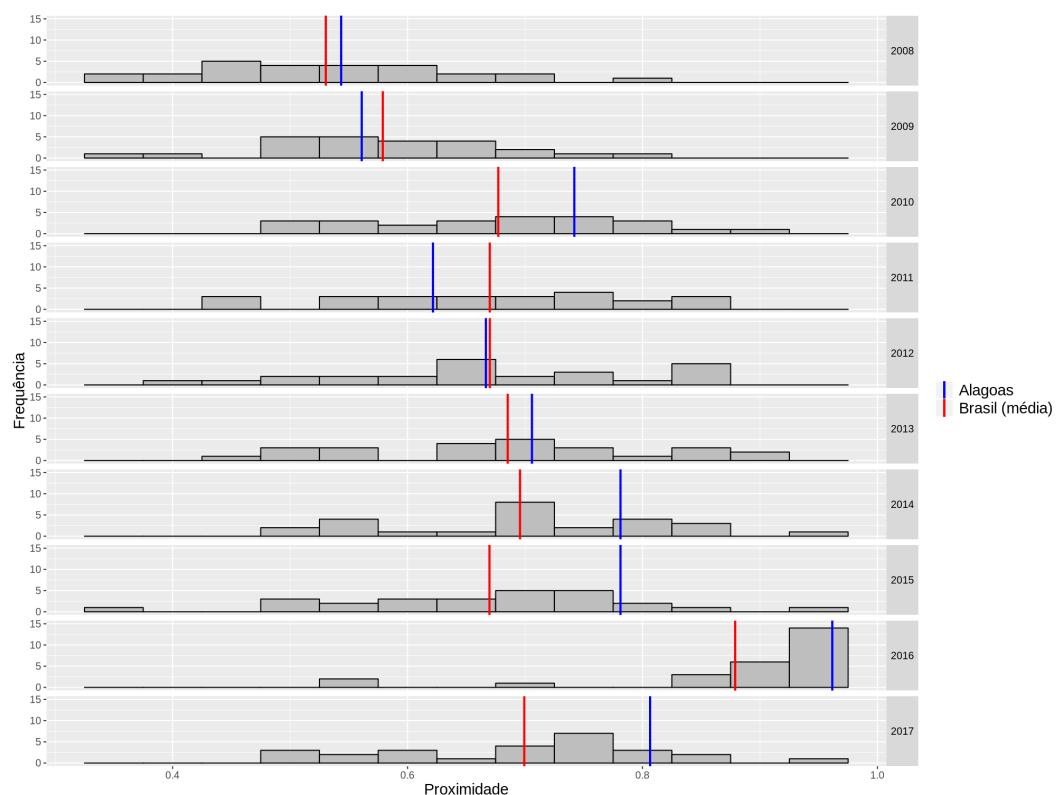


Figura 4.15: Histograma da Centralidade de Proximidade (*Health Sciences*)

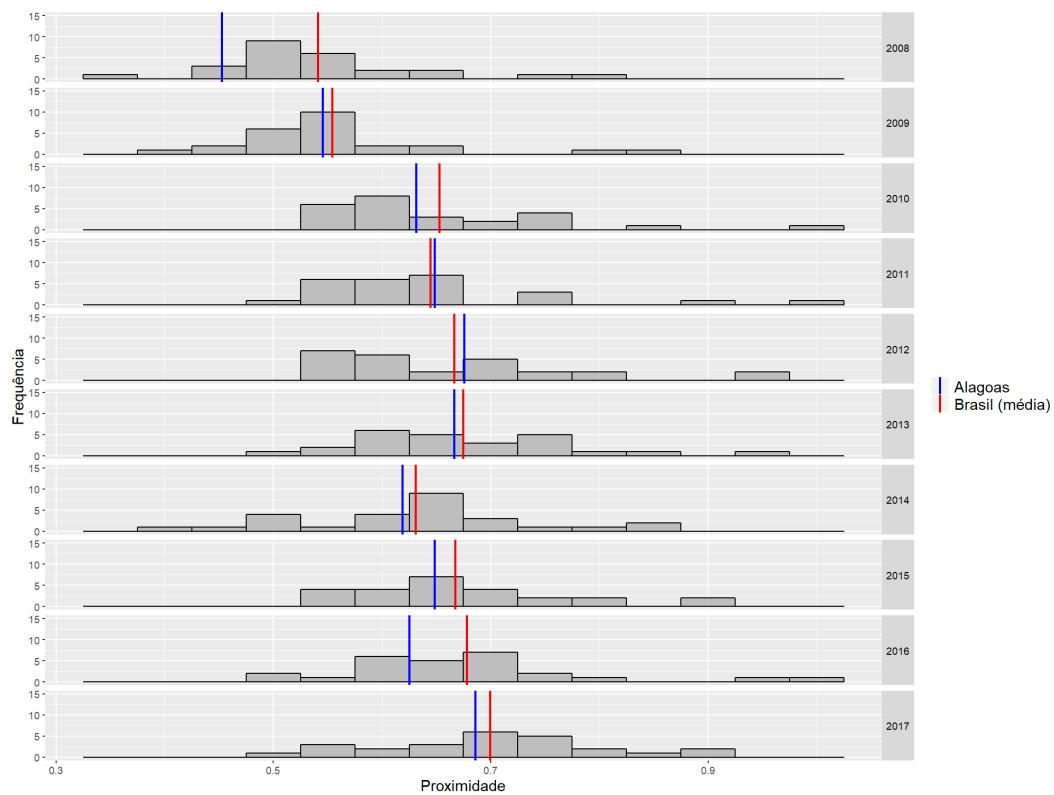
Agricultural Sciences

Figura 4.16: Histograma da Centralidade de Proximidade (*Agricultural Sciences*)

Exact and Earth Sciences

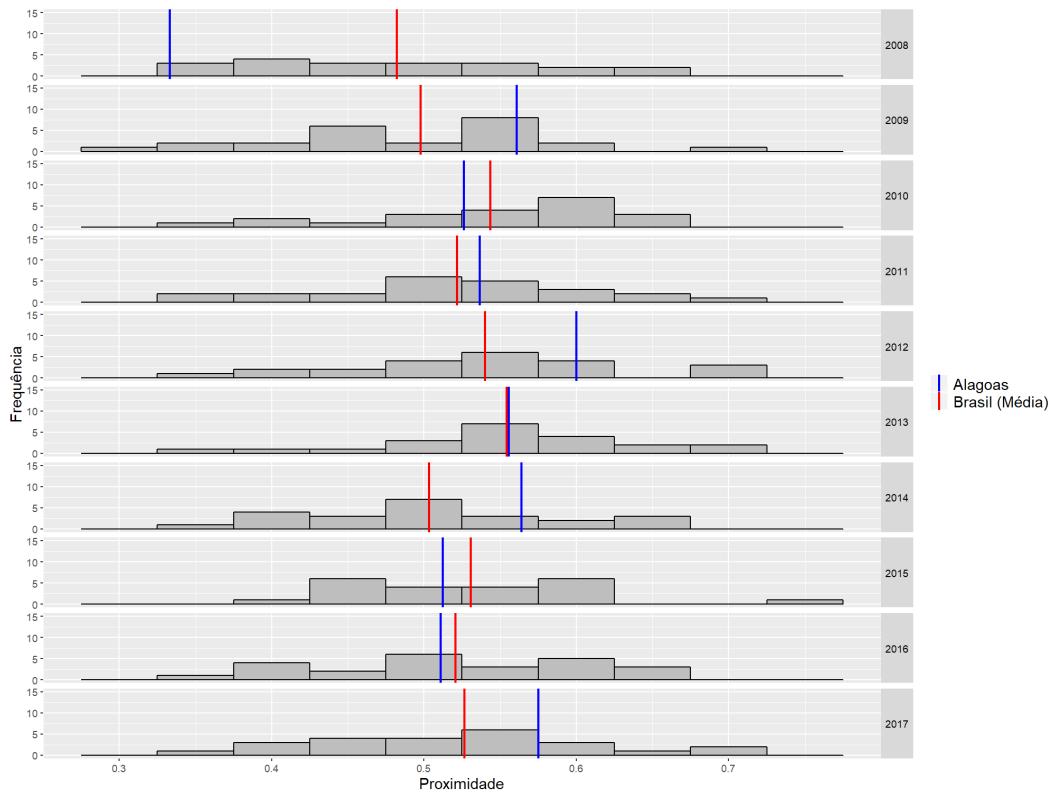


Figura 4.17: Histograma da Centralidade de Proximidade (*Exact and Earth Sciences*)

As figuras 4.15, apresenta o histograma de distribuição para a área de *Health Sciences*, o qual podemos destacar ano de 2016, em que as frequências se concentraram em alto índice comparado aos demais anos, e que Alagoas ficou acima da média para essa medida desde o ano de 2013. Resultando que sua proximidade quando comparada a média dos vértices do Brasil é superior.

Em *Agricultural Sciences*, visto pelo histograma na figura 4.16, a média do Brasil e Alagoas são marcados muito próximos no histograma que possui distribuição em todos anos muito central. De forma que esta medida não apresentou relevante informação nessa área.

O histograma 4.17 apresenta uma distribuição de sua frequência mais espaçada, sendo observador o marcador de Alagoas muito baixo no ano de 2008, muito provável pelos poucos artigos indexados neste ano por autores Alagoanos na área de *Exact and Earth Sciences*, mas que a partir de 2009 a proximidade de Alagoas com a média do Brasil, supera em 2011, 2012, 2014, e 2017,

4.2.3 Centralidade de Intermediação

A figura 4.5 apresenta os resultados da medida de centralidade de intermediação em relação a Rede Brasil (média) e Alagoas, para todas as áreas. Essa medida indica a capacidade

de intermediação de um vértice a outro, considerando seu posicionamento na rede.

Centralidade de Intermediação (Betweenness Centrality) por Área						
Ano	Health Sciences		Agricultural Sciences		Exact and Earth Sciences	
Rede	Brasil	Alagoas	Brasil	Alagoas	Brasil	Alagoas
2008	0,037	0,01	0,04	0,0047	0,0476	0,0021
2009	0,0396	0,01	0,0385	0,013	0,04	0,1036
2010	0,0388	0,04	0,04	0,0138	0,0455	0,0251
2011	0,04	0,01	0,037	0,0154	0,0435	0,0153
2012	0,038	0,02	0,0385	0,0314	0,0417	0,1405
2013	0,0384	0,06	0,0385	0,0289	0,0455	0,0424
2014	0,0384	0,05	0,037	0,0066	0,04	0,0549
2015	0,037	0,04	0,0385	0,0088	0,0455	0,0037
2016	0,0396	0,1	0,0385	0,0099	0,04	0,0212
2017	0,0369	0,04	0,04	0,0216	0,0417	0,0274

Tabela 4.5: Medidas de Centralidade de Intermediação por Área (2008-2017)

Health Sciences

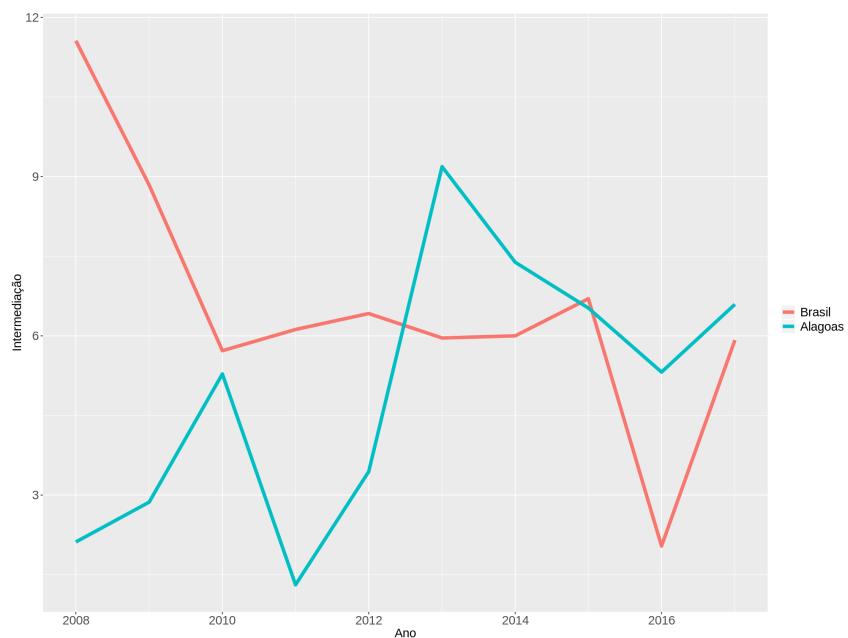


Figura 4.18: Centralidade de Proximidade (Health Sciences)

Agricultural Sciences

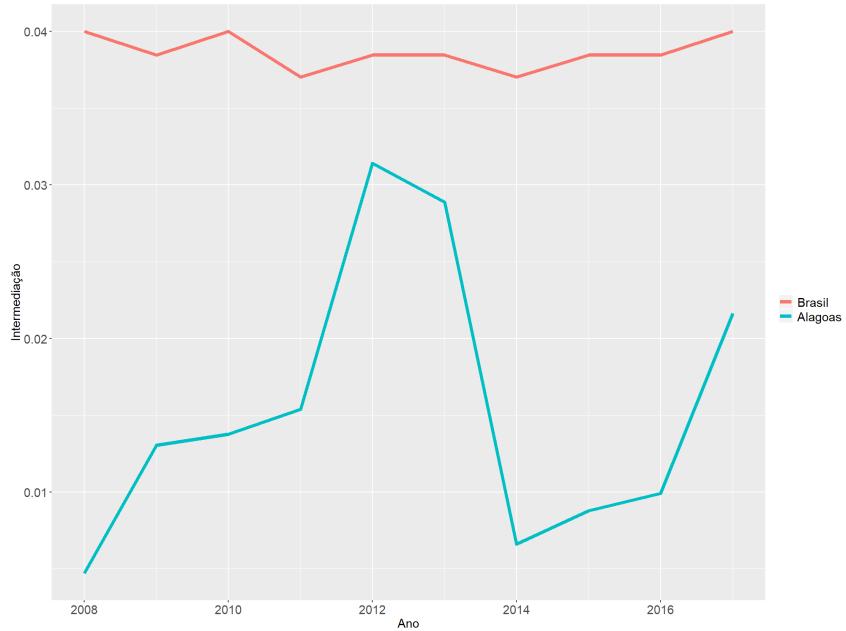


Figura 4.19: Centralidade de Proximidade (*Agricultural Sciences*)

Exact and Earth Sciences

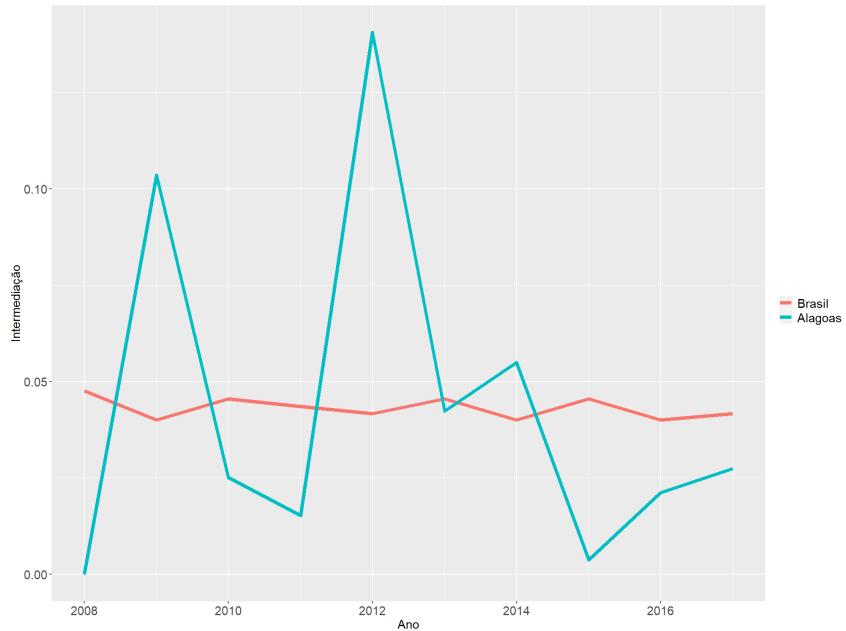


Figura 4.20: Centralidade de Proximidade (*Exact and Earth Sciences*)

Na observação dos gráficos da centralidade de intermediação 4.18, 4.19, 4.20 verificamos que diferentemente do que se observou na centralidade do grau e de proximidade, a relação

encontrada entre a Rede Brasil e Alagoas é de uma função matemática não trivial, destacando a sensibilidade dessa medida de centralidade para avaliação da colaboração científica.

Health Sciences

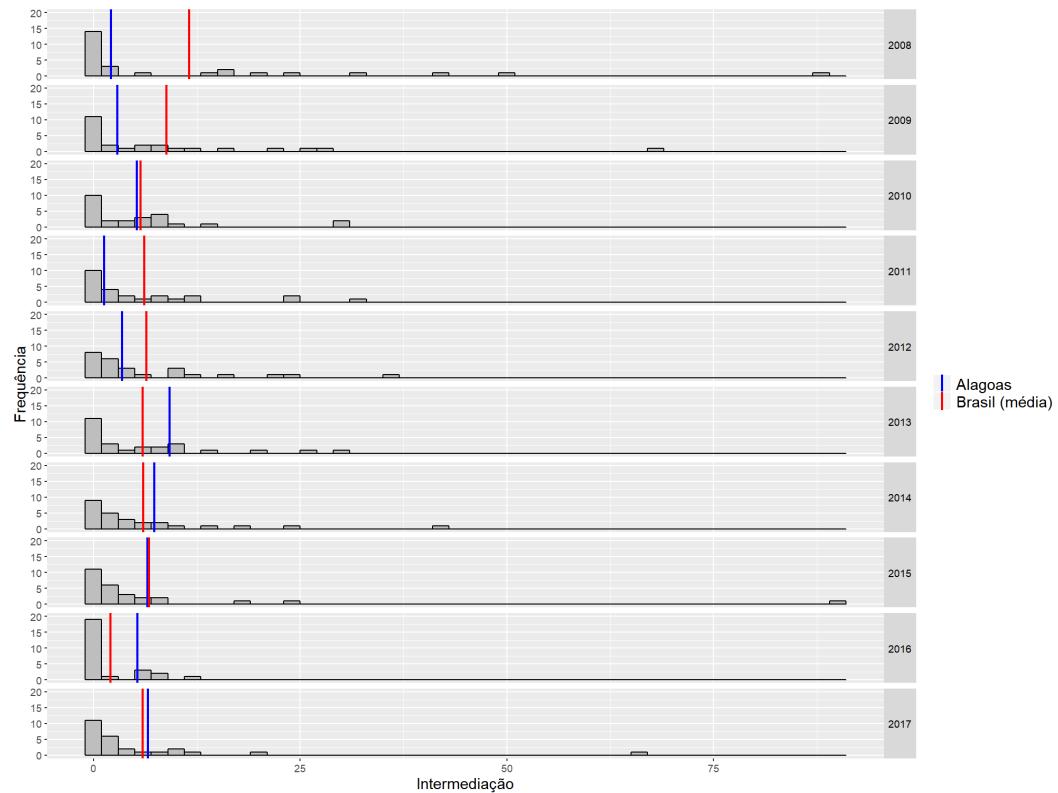


Figura 4.21: Histograma da Centralidade de Proximidade (*Health Sciences*)

Agricultural Sciences

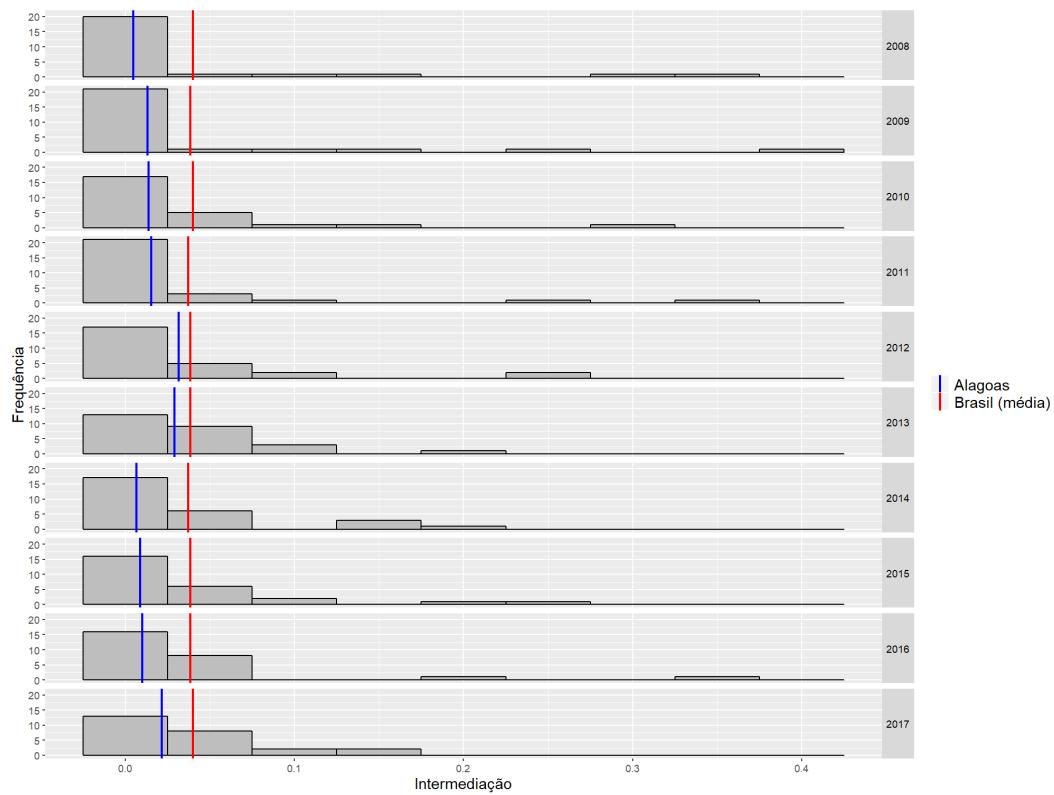


Figura 4.22: Histograma da Centralidade de Proximidade (*Agricultural Sciences*)

Exact and Earth Sciences

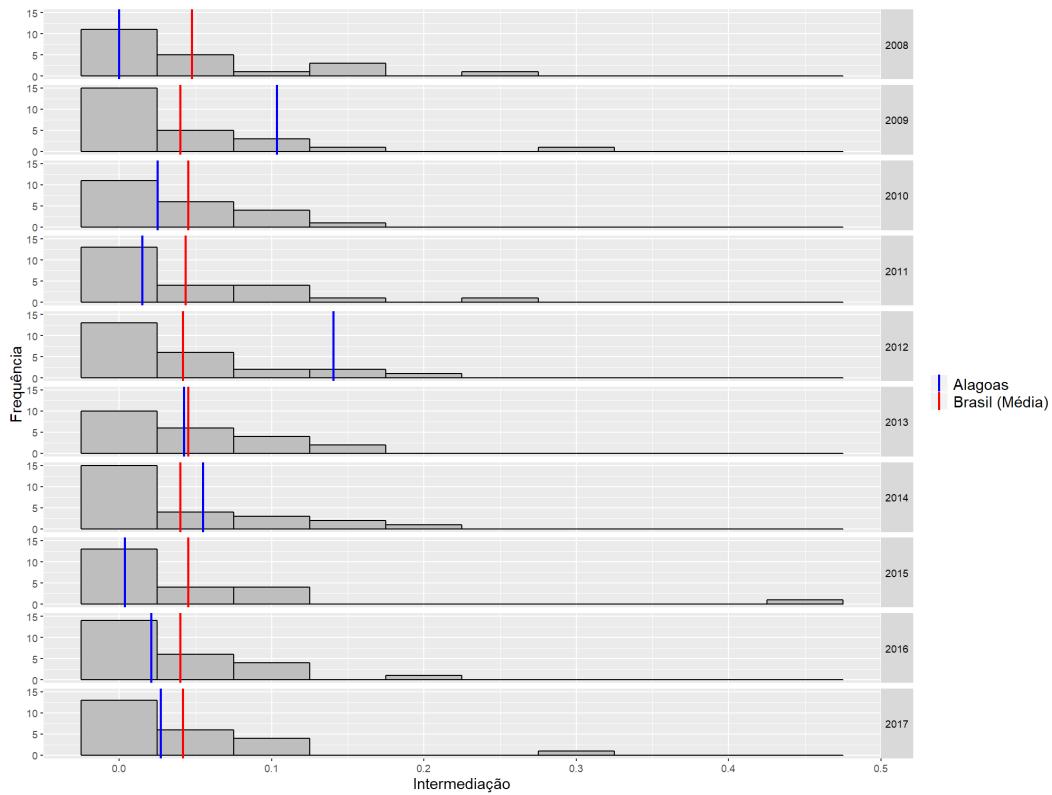


Figura 4.23: Histograma da Centralidade de Proximidade (*Exact and Earth Sciences*)

4.3 Ranking das Medidas de Centralidade

Nesta seção apresentamos os rankings por Estados (UF) em cada área do conhecimento investigadas neste trabalho. Este ranking tem por objetivo apontar o posicionamento de Alagoas frente a rede Brasil, bem como, por oportunidade, mostrar os melhores e piores estados em desempenho para cada medida de centralidade da rede.

Ranking: *Health Sciences*

Ranking das Medidas de Centralidade					
Grau		Proximidade		Intermediação	
UF	Média	UF	Média	UF	Média
SP	20,50	SP	0,870	SP	0,243
MG	19,60	MG	0,839	MG	0,145
RJ	19,00	RJ	0,818	RJ	0,093
RS	17,60	RS	0,784	GO	0,065
SC	16,90	SC	0,770	CE	0,060
PE	16,70	PE	0,765	SC	0,051
CE	15,80	CE	0,745	RS	0,049
RN	15,20	RN	0,725	PE	0,038
BA	14,60	AL	0,717	AL	0,038
AL	14,50	BA	0,715	RN	0,032
PB	14,50	PB	0,707	PR	0,030
PR	13,80	PR	0,700	PA	0,025
GO	13,40	GO	0,690	AM	0,023
ES	12,50	ES	0,678	PB	0,021
PA	12,00	PA	0,660	BA	0,016
SE	11,40	SE	0,650	RO	0,016
MT	10,80	AM	0,641	MA	0,015
AM	10,70	MT	0,641	MT	0,014
PI	9,90	PI	0,616	ES	0,012
MA	9,40	MA	0,614	PI	0,004
MS	5,30	AP	0,557	SE	0,004
AC	4,90	RO	0,552	AC	0,001
RO	4,22	MS	0,538	MS	0,001
RR	3,70	AC	0,518	AP	0,000
AP	3,63	RR	0,500	DF	0,000
TO	2,33	TO	0,495	RR	0,000
DF	1,00	DF	0,436	TO	0,000

Tabela 4.6: Ranking Medidas de Centralidade (*Health Sciences*)

Ranking: Agricultural Sciences

Ranking das Medidas de Centralidade					
Grau		Proximidade		Intermediação	
UF	Média	UF	Média	UF	Média
RN	24,10	MG	0,889	MG	0,213
MG	23,20	RN	0,877	RN	0,205
PE	17,60	RS	0,786	RS	0,115
RS	17,20	PE	0,740	PB	0,067
PB	16,60	PB	0,718	GO	0,067
RJ	15,80	RJ	0,701	PE	0,062
MT	14,40	MT	0,675	RJ	0,042
GO	14,10	GO	0,665	CE	0,031
CE	13,30	CE	0,652	MT	0,031
PA	13,30	PA	0,651	PA	0,020
PR	12,70	BA	0,636	PR	0,020
BA	12,60	PR	0,635	MS	0,018
AL	12,00	AL	0,625	PI	0,017
PI	11,80	ES	0,625	ES	0,016
ES	11,70	PI	0,624	BA	0,016
MS	10,90	MS	0,609	AL	0,015
SC	9,70	TO	0,589	SC	0,009
TO	9,50	SP	0,580	SP	0,008
SP	9,20	SC	0,576	AC	0,008
SE	8,60	SE	0,571	TO	0,008
AM	7,30	AC	0,558	AM	0,004
AC	7,00	AM	0,544	SE	0,004
MA	7,00	RO	0,540	RO	0,003
RO	6,22	MA	0,537	MA	0,002
RR	5,50	RR	0,526	RR	0,000
DF	2,40	AP	0,490	AP	0,000
AP	2,20	DF	0,423	DF	0,000

Tabela 4.7: Ranking Medidas de Centralidade (*Agricultural Sciences*)

Ranking: *Exact and Earth Sciences*

Ranking das Medidas de Centralidade					
Grau		Proximidade		Intermediação	
UF	Média	UF	Média	UF	Média
RJ	13,40	MG	0,657	MG	0,173
MG	13,30	RJ	0,651	RJ	0,144
SP	11,40	SP	0,618	SP	0,091
RS	10,60	RS	0,601	RS	0,071
PB	9,80	RN	0,581	RN	0,059
RN	9,70	PB	0,566	PE	0,055
PE	9,50	PE	0,565	PB	0,054
PR	9,10	BA	0,559	PR	0,049
BA	8,70	SC	0,556	AL	0,043
SC	8,50	PR	0,555	CE	0,043
CE	8,20	CE	0,554	SC	0,039
AL	7,90	AL	0,528	PI	0,036
PA	6,20	GO	0,512	BA	0,029
SE	6,10	ES	0,485	GO	0,029
GO	6,00	SE	0,483	PA	0,025
AM	5,60	AM	0,480	SE	0,018
ES	5,60	PA	0,473	MT	0,013
PI	5,10	PI	0,472	AM	0,012
MT	4,67	RO	0,442	MA	0,010
MS	3,63	MT	0,435	ES	0,006
MA	3,50	MS	0,412	RO	0,002
RO	3,25	TO	0,411	MS	0,001
RR	2,40	MA	0,404	AC	0,000
TO	2,17	RR	0,395	AP	0,000
DF	2,00	AP	0,388	DF	0,000
AP	2,00	AC	0,384	RR	0,000
AC	1,80	DF	0,000	TO	0,000

Tabela 4.8: Ranking Medidas de Centralidade (*Exact and Earth Sciences*)

4.4 Discussão dos Resultados

Esta seção tem a por objetivo a discussão dos resultados obtidos por este trabalho, os estudo de redes de coautoriais por medidas de centralidade de redes complexas, e a con-

textualização destes de ao comportamento da colaboração científica. A metodologia aplica permitiu a compreensão das características estruturais das redes, e a determinação destas para os estudos exploratórios.

Utilizando o vértice focal Alagoas, podemos comparar nas três áreas do conhecimento selecionadas para esse trabalho, o estudo das três medidas de centralidade, elencadas por (Freeman et al. 1991) como essenciais ao estudo de redes.

Numérica e graficamente percebemos que para todas as áreas o comportamento a medida da centralidade de grau e de proximidade para o vértice amostral utilizado e a média parâmetro de estudo que de uma forma amplificada possuem uma relação de uma correlação, que seja em anos positiva ou negativa, o que não se observa para a centralidade de intermediação que em todas as áreas apresentou resultados

As medidas de centralidade utilizadas neste trabalho permitiu a construção de rankings por área, indicando assim os índices aferidos pelos estados para o período de 2008-2017, possibilitando o conhecimento da colaboração científica a partir desses indicadores na base de dados e nas áreas estudadas.

Na área de *Health Sciences* verificamos que conforme o ranking na figura 4.6 os estados em destaque são: São Paulo, Minas Gerais e Rio de Janeiro são os estados que possuem os maiores índices das medidas de centralidade: de grau, de proximidade, de intermediação. Alagoas assume a décima posição junto com Pernambuco para a centralidade do grau. E a nona posição para centralidade de proximidade e de intermediação, nesta última em empate também com Pernambuco.

Em *Agricultural Sciences* o destaque é para os estados do Rio Grande do Norte, Minas Gerais e Rio Grande do Sul (conforme ranking 4.7); tendo Alagoas assumido respectivamente as posições de décimo terceiro para centralidade do grau e de proximidade; décimo sexto para intermediação.

5

CONCLUSÕES

A seguir apresentamos nossas considerações finais, buscando a contextualização do problema de pesquisa em tela com os resultados obtidos e as sugestões para trabalhos futuros.

5.1 Considerações Finais

Com a proposta apresentada, buscamos destacar a importância do estudo de redes de coautoriais relacionando ao arcabouço teórico de redes complexas e a relação com a colaboração científica, neste trabalho a compreensão da interação existente entre as Universidades Federais do Brasil e a Universidade Federal de Alagoas.

Conforme (Barabási et al. 2001), os estudos de redes de coautoria são importantes e possuem sua relevância para apresentar padrões de comportamento e possibilitar modelos preditivos de quais áreas ou campos estão se desenvolvendo, são emergentes, ou quais tendências possuem.

O estudo das redes de coautoria deste trabalho, mostrou um resultado que se aplicado a outras redes poderão descrever o comportamento, dinâmica e características, quando se observa as relações existentes dentro das áreas de *Health Sciences, Agricultural Sciences e Exact and Earth Sciences*, a partir da consideração de uma base de indexação de artigos específica ou de outra.

As medidas de centralidade possibilitam aferir conhecimento a respeito da dinâmica e do comportamento das redes de coautoriais no aspecto temporal. A aplicação destas, em outras áreas do conhecimento e em outras bases de indexação de artigos científicos, mostram-se significativas para a obtenção de resultados específico para avaliação da colaboração científica, possibilitando a compreensão e caracterização dos atores de relevância e de influência da rede.

5.2 Trabalhos futuros

Por trabalhos futuros, consideramos a expansão do uso de outras medidas de centralidade de rede para análises topológica da interação dos vértices e arestas nas redes de coautoria. Sugerimos outras abordagens de uma prisma analítico e probabilístico, como a utilização de técnicas de *network-driven approaches*, de *link prediction*, análises de comunidades de influências na rede, e simulação do comportamento da rede com base em métricas de redes complexas.

Apêndice A

REDES DE COAUTORIA

As características propostas na plotagem da rede é, formato em círculo para fins comparativos, disposição dos vértices com coloração das UF pertencentes a mesma região geográfica, espessura e coloração das arestas indicando o maior volume/peso das coautorias existente entre as UFs, espessuras dos laços, indicando peso das coautorias existentes na mesma UF.

A.1 Health Sciences

Rede de Coautoria das Universidade Federais do Brasil

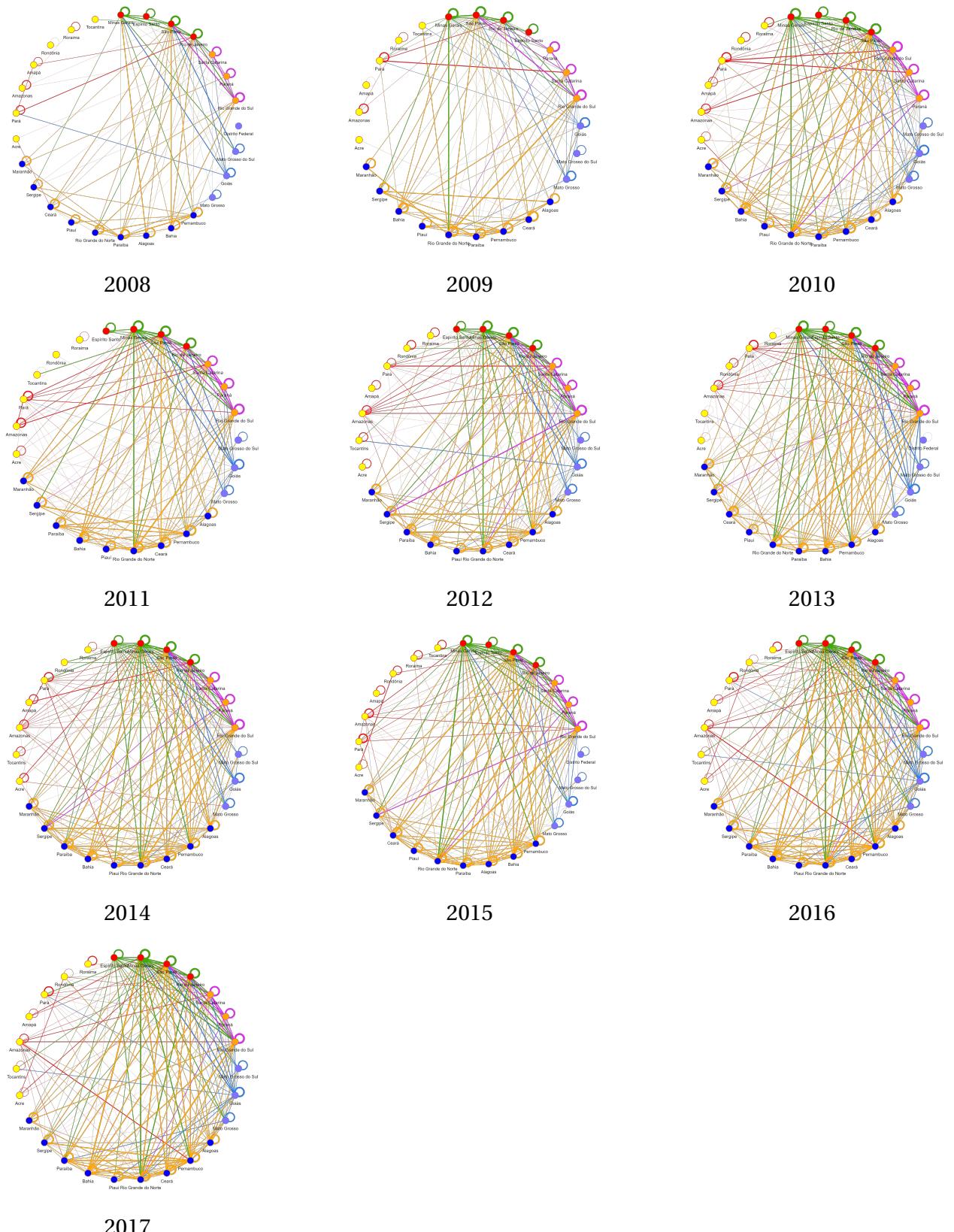


Figura A.1: Redes de Coautoria Universidades Federal do Brasil (*Health Sciences*)

Rede de Coautoria Brasil - Vértice Focal Alagoas

A seguir são apresentadas as redes de coautoria utilizando o vértice focal Alagoas, e então observando as conexões realizadas com os demais vértices da rede (UF). Visualmente é possível notar o crescimento pelo número de novas conexões.

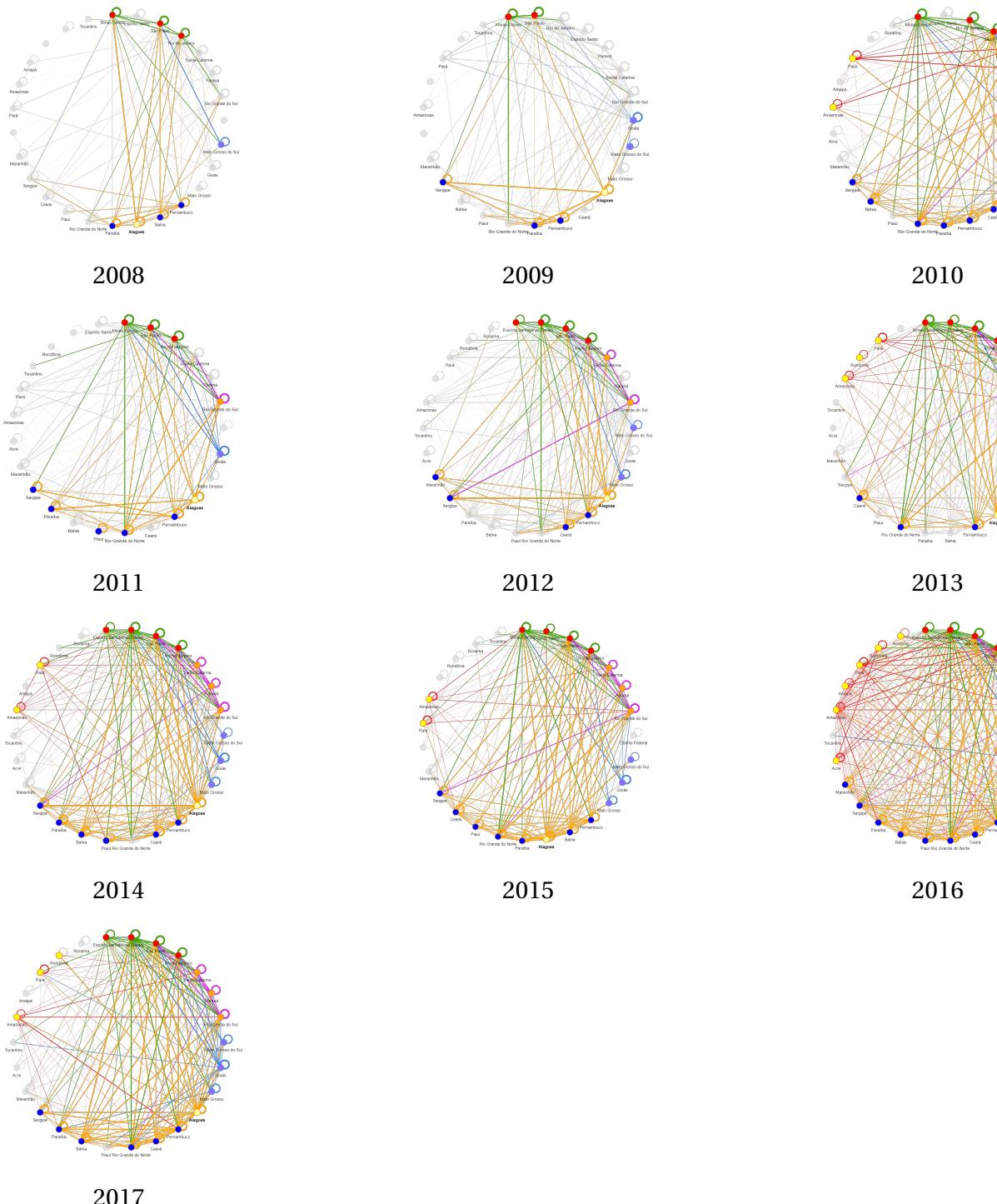


Figura A.2: Redes de Coautoria da Universidade Federal de Alagoas

A.2 Agricultural Sciences

Rede de Coautoria das Universidade Federais do Brasil

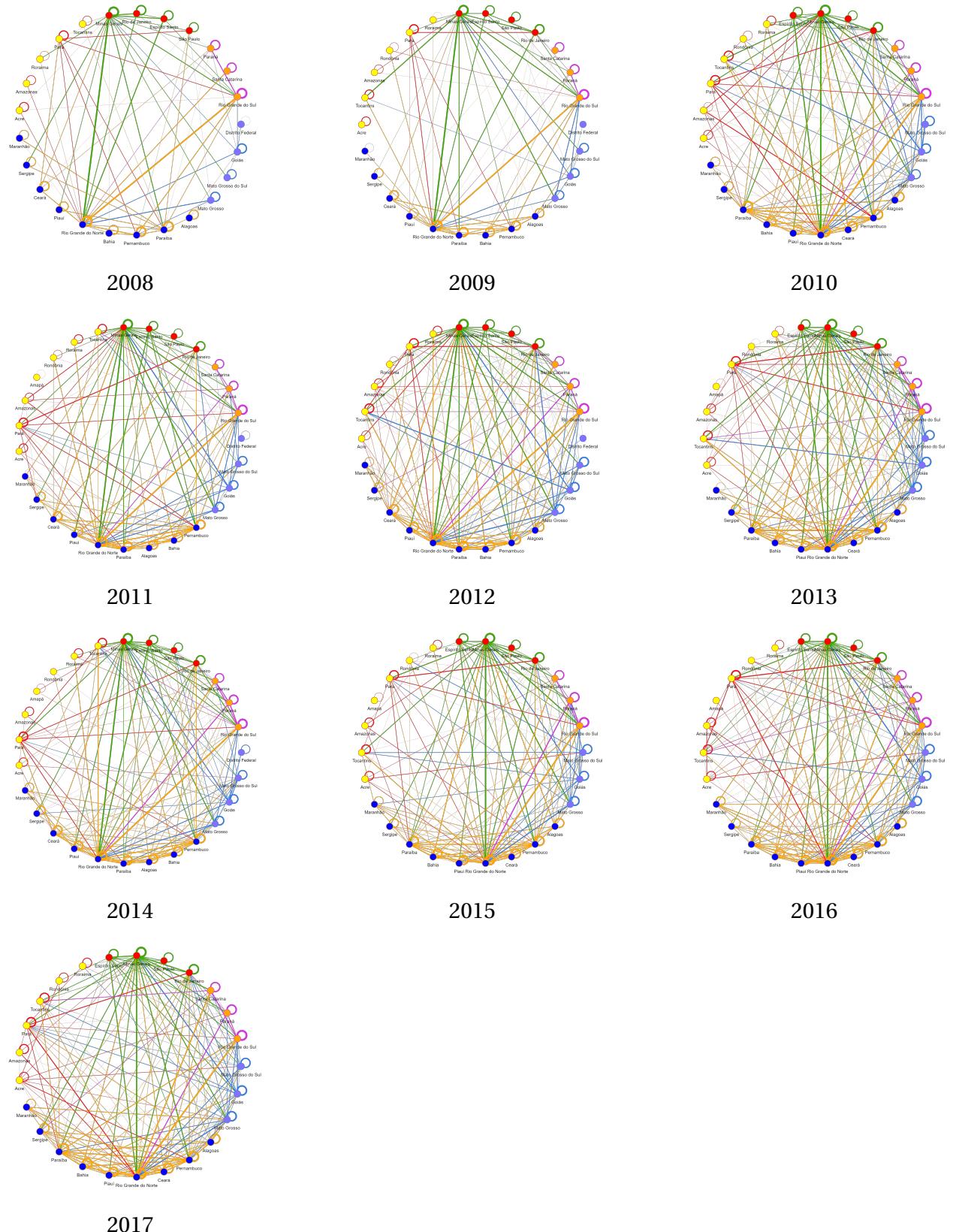


Figura A.3: Redes de Coautoria Universidades Federal do Brasil: Área: *Agricultural Sciences*

Rede de Coautoria Brasil - Vértice Focal Alagoas

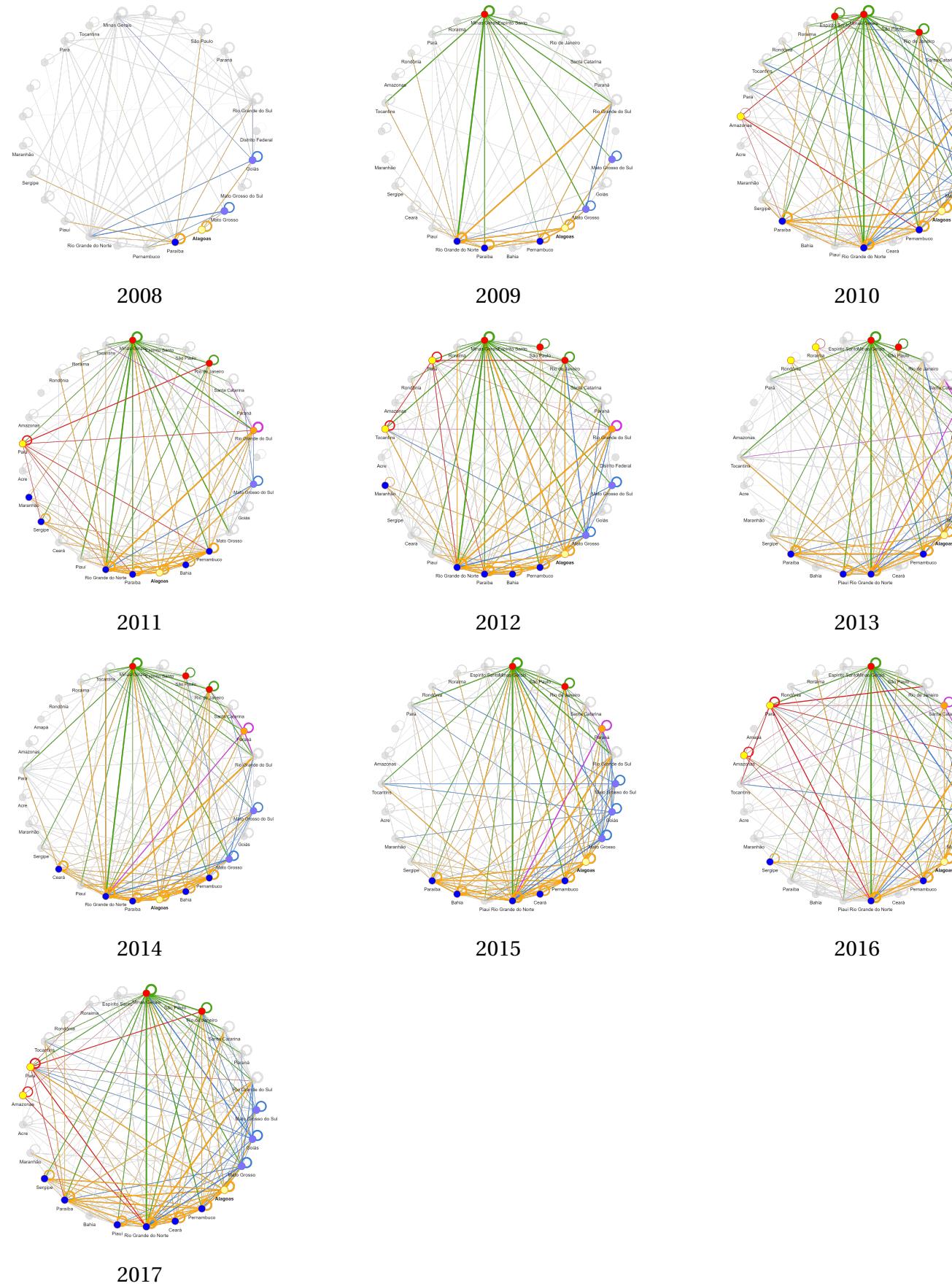


Figura A.4: Redes de Coautoria da Universidade Federal de Alagoas - Área Agricultural Sciences

A.3 Exact and Earth Sciences

Rede de Coautoria das Universidade Federais do Brasil

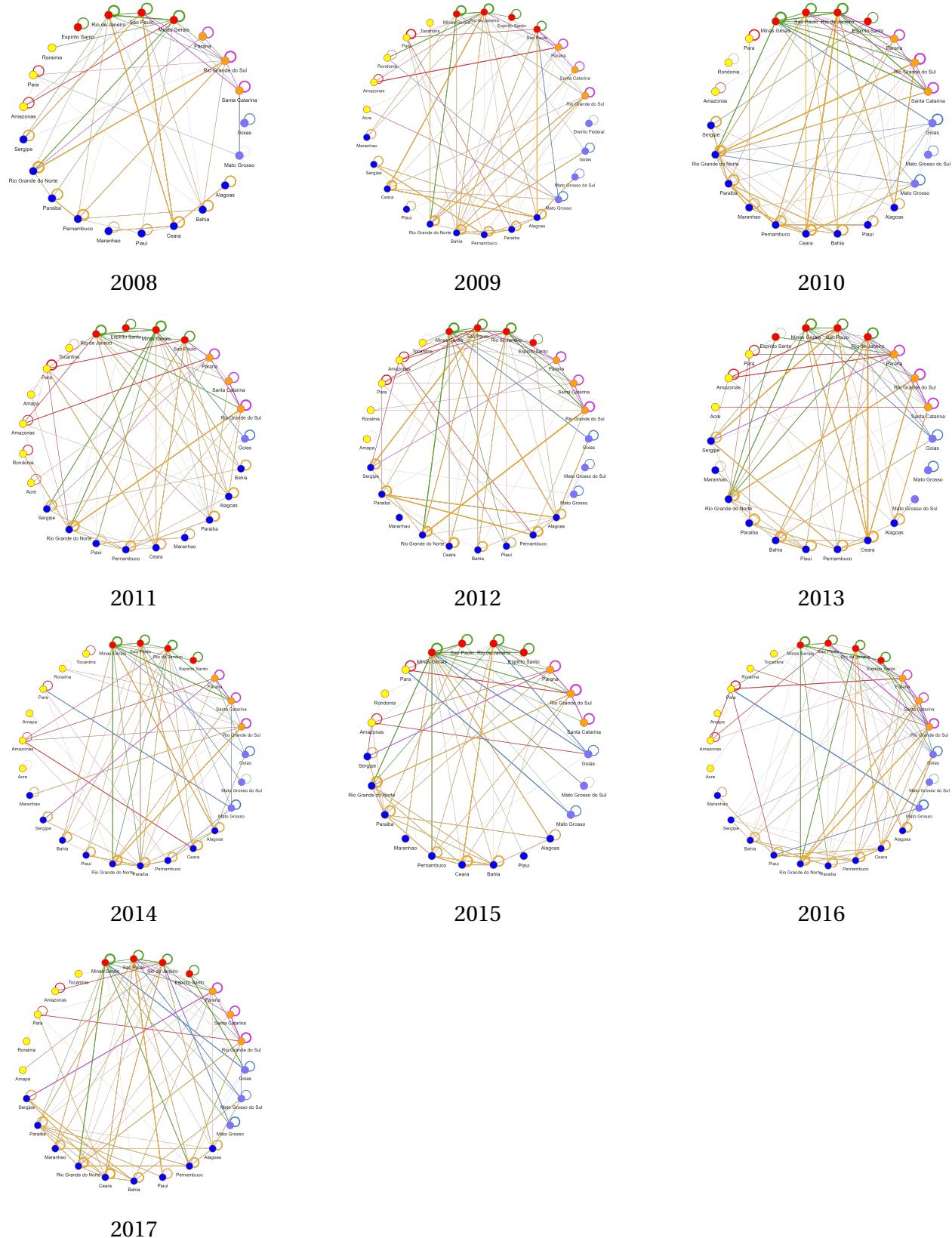


Figura A.5: Redes de Coautoria Universidades Federal do Brasil - Área *Exact and Earth Sciences*

Rede de Coautoria Brasil - Vértice Focal Alagoas

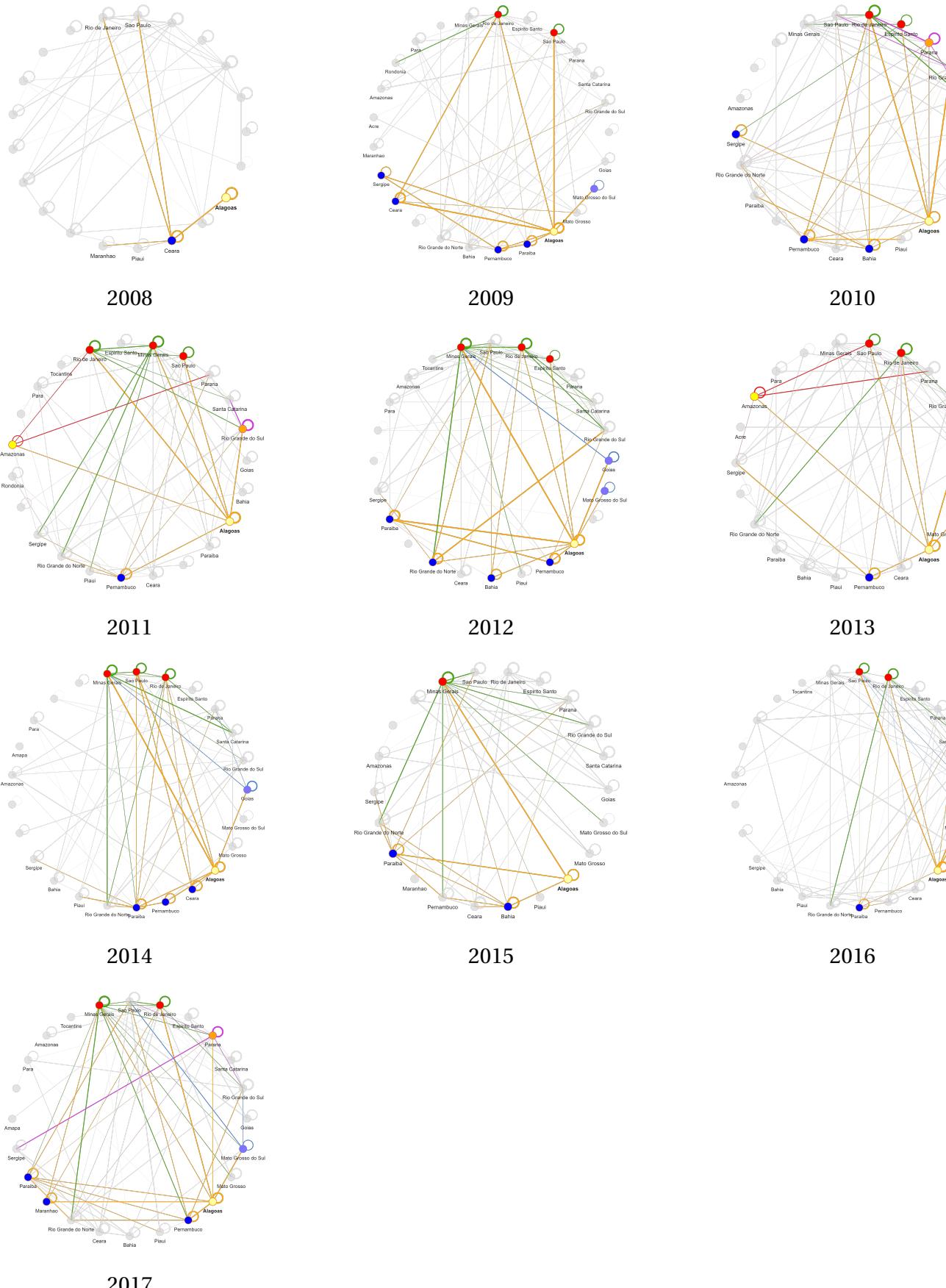


Figura A.6: Redes de Coautoria da Universidade Federal de Alagoas Área *Exact and Earth Sciences*

REFERÊNCIAS BIBLIOGRÁFICAS

- Barabási, A. L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A. & Vicsek, T. (2001), 'Evolution of the social network of scientific collaborations'.
URL: <https://arxiv.org/pdf/cond-mat/0104162.pdf>
- Barabasi, A.-L. L. (2003), 'How everything is connected to everything else and what it means for business, science and everyday life', *New York: Plume*.
- De Nooy, W., Mrvar, A. & Batagelj, V. (2018), *Exploratory social network analysis with Pajek*, Cambridge University Press.
- Freeman, L. C. (1996), 'Some antecedents of social network analysis', *Connections* **19**(1), 39–42.
- Freeman, L. C., Borgatti, S. P. & White, D. R. (1991), 'Centrality in valued graphs: A measure of betweenness based on network flow', *Social networks* **13**(2), 141–154.
- Maia, M. D. F. S. & Caregnato, S. E. (2008), 'Co-autoria como indicador de redes de colaboração científica', *Perspectivas em Ciência da Informação* **13**, 18–31.
- Newman, M. E. J. (2004), Coauthorship networks and patterns of scientific collaboration, Technical report, Center for the Study of Complex Systems and Department of Physics, University of Michigan.
URL: www.pnas.org/cgi/doi/10.1073/pnas.0307545100
- Osareh, F. (1996), 'Bibliometrics, citation analysis and co-citation analysis: A review of literature i', *Libri* **46**(3), 149–158.
- Otte, E. & Rousseau, R. (2002), 'Social network analysis: a powerful strategy, also for the information sciences', *Journal of information Science* **28**(6), 441–453.
- Pritchard, A. et al. (1969), 'Statistical bibliography or bibliometrics', *Journal of documentation* **25**(4), 348–349.
- Silva, A. B. d. O. e., Matheus, R. F., Parreiras, F. S. & Parreiras, T. A. S. (2006), 'Análise de redes sociais como metodologia de apoio para a discussão da interdisciplinaridade na ciência da informação', *Ciência da Informação* **35**(1), 72–93.
URL: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652006000100009&lng=pt&tlang=pt
- Sonnenwald, D. H. & Sonnenwald, D. H. (2007), 'Scientific Collaboration', *Annual Review of Information Science and Technology (ARIST)* **41**, 643–681.
URL: <http://europa.eu.int/comm/>

- Vanz, S. A. d. S. & Stumpf, I. R. C. (2010), 'Colaboração científica: revisão teórico-conceitual', *Perspectivas em Ciência da Informação* 15(2), 42–55.
URL: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362010000200004&lng=pt&tlang=pt
- Wasserman, S. & Faust, K. (1994), *Social network analysis: Methods and applications*, Vol. 8, Cambridge university press.

Este trabalho foi redigido em \LaTeX utilizando uma modificação do estilo IC-UFAL. As referências bibliográficas foram preparadas no Mendeley e administradas pelo $\text{BIB}\text{\TeX}$ com o estilo LaCCAN. O texto utiliza fonte Fourier-GUTenberg e os elementos matemáticos a família tipográfica Euler Virtual Math, ambas em corpo de 12 pontos.