

Análise de Rede de Coautoria Científica do PPGIA UNIFOR

José Freitas Alves Neto

[Link do trabalho no GitHub](#)

19 de maio de 2025

1 O Que é o Meu Grafo: Detalhamento, Incertezas e Potenciais Vieses

O grafo analisado neste trabalho representa uma tentativa de mapear a rede de coautoria científica entre pesquisadores que possuem publicações relacionadas ao Programa de Pós-Graduação em Informática Aplicada (PPGIA) da Universidade de Fortaleza (UNIFOR). A construção deste grafo teve como ponto de partida a coleta de dados de publicações científicas indexadas exclusivamente na plataforma Google Scholar, utilizando o termo de pesquisa “ppgia unifor”. Diante disso, vale salientar a ausência de alguns autores e a não completude dos dados, sendo uma amostra não segmentada de todo o conjunto e não sendo representativa em relação ao todo.

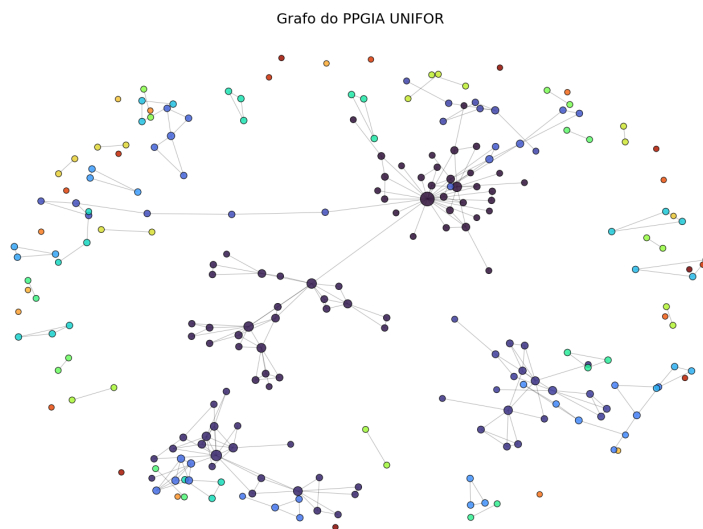


Figura 1: Visualização do Grafo extraído a partir do Google Scholar.

Características Fundamentais do Grafo

- **Nós (Vértices):** Cada nó no grafo simboliza um autor/pesquisador identificado nas publicações.

- **Arestas (Ligações):** Uma aresta é estabelecida entre dois nós (autores) se eles compartilham a autoria de pelo menos uma publicação científica dentro do conjunto de dados coletado.
- **Tipo de Grafo:** O grafo é inerentemente não direcionado, refletindo a natureza mútua da coautoria: se A colaborou com B, então B colaborou com A.

Incertezas, Limitações e Enviesamento dos Dados

Apesar do esforço em pré-processar e limpar os dados, é crucial reconhecer as limitações intrínsecas ao processo de coleta e à natureza da fonte de dados, que introduzem um certo grau de enviesamento e incerteza na representação da rede:

- **Fonte Única de Dados (Google Scholar):** A utilização exclusiva do Google Scholar como fonte de dados significa que o grafo reflete apenas as publicações indexadas por esta plataforma específica. Outras bases de dados acadêmicas (como Scopus, Web of Science, DBLP, Lattes) poderiam apresentar um conjunto diferente ou complementar de publicações e, conseqüentemente, uma estrutura de rede distinta. Trabalhos não indexados pelo Google Scholar, por qualquer motivo, estão ausentes.
- **Método de Extração Inicial e Cobertura:** Os dados foram inicialmente extraídos a partir dos resultados listados na primeira página da busca (e possivelmente algumas subsequentes, totalizando 185 artigos iniciais), sem uma varredura exaustiva de cada link de artigo para capturar a lista completa de coautores. Isso implica que artigos com muitos autores podem ter tido apenas os primeiros nomes listados capturados, ou que a profundidade da coleta pode não ter sido suficiente para abranger todas as publicações relevantes mesmo dentro do Google Scholar.
- **Representatividade e Ausência de Pesquisadores:** Como consequência direta dos pontos anteriores, nem todos os professores e pesquisadores afiliados ao PPGIA UNIFOR podem estar representados no grafo. Alguns podem ter suas principais publicações em plataformas não priorizadas pelo Google Scholar, ou suas publicações podem não ter sido capturadas pela string de busca específica ou pelo método de extração. A ausência de certos indivíduos pode distorcer a percepção da estrutura da rede e da centralidade de outros.
- **Enviesamento da Amostra:** O conjunto de dados, portanto, constitui uma amostra específica e não um censo completo de todas as colaborações do PPGIA UNIFOR. Este enviesamento é inerente ao processo de coleta e significa que as conclusões tiradas, especialmente sobre a importância relativa dos pesquisadores, são válidas dentro do contexto desta amostra específica. A generalização para a totalidade da rede de colaboração do PPGIA deve ser feita com cautela.
- **Variações de Nomes e Desambiguação:** Embora um esforço significativo tenha sido empregado na normalização de nomes (remoção de acentos, padronização de maiúsculas/minúsculas, tratamento de variações), a desambiguação de autores com nomes comuns ou com múltiplas formas de abreviação continua sendo um desafio em qualquer análise bibliométrica. Podem persistir casos onde um mesmo autor é representado por múltiplos nós ou, inversamente, diferentes autores com nomes semelhantes são fundidos.

Processo de Refinamento e Estado Final do Grafo

Apesar dessas limitações, um processo de pré-processamento foi aplicado. Este incluiu a remoção de nós não representativos de autores (como anos), a padronização de nomes e o agrupamento de variações identificáveis do mesmo autor, além da remoção de informações textuais excedentes. Após este tratamento, o grafo final utilizado para a análise de centralidade consolidou-se com 239 nós (autores distintos) e 267 arestas (relações de coautoria únicas).

Em suma, o grafo construído é um modelo da rede de coautoria baseado em uma amostra específica de dados do Google Scholar. As análises subsequentes, embora reveladoras sobre as dinâmicas dentro desta amostra, devem ser interpretadas à luz das incertezas e potenciais vieses inerentes à sua construção, incluindo a possível sub-representação de alguns pesquisadores e a dependência de uma única fonte de dados.

2 Métricas de Análise e o Que é Cada Uma Delas

Para compreender a importância e o papel de cada pesquisador dentro da rede de coautoria, foram aplicadas diversas métricas de centralidade. A centralidade é um conceito fundamental na análise de redes, que busca identificar os nós mais “importantes” ou “influentes” em um grafo. As seguintes métricas foram utilizadas:

2.1 Degree Centrality (Centralidade de Grau)

- **O que é:** É a medida mais direta de centralidade. Para um dado nó, a centralidade de grau é simplesmente o número de arestas conectadas a ele. Em uma rede de coautoria, isso representa o número de pesquisadores diferentes com os quais um autor colaborou diretamente.
- **Interpretação:** Autores com alta centralidade de grau são considerados “hubs” locais, com um grande número de colaboradores diretos, indicando alta atividade colaborativa.

2.2 Closeness Centrality (Centralidade de Proximidade)

- **O que é:** Mede quão próximo, em média, um nó está de todos os outros nós alcançáveis na rede. É calculada como o inverso da soma das distâncias dos caminhos mais curtos entre o nó em questão e todos os outros nós no grafo.
- **Interpretação:** Autores com alta centralidade de proximidade têm a capacidade de alcançar outros membros da rede rapidamente (através de poucos intermediários). Eles estão bem posicionados para disseminar informações ou influência de forma eficiente pela rede.

2.3 Betweenness Centrality (Centralidade de Intermediação)

- **O que é:** Quantifica a frequência com que um nó aparece nos caminhos mais curtos entre todos os outros pares de nós na rede. Um nó com alta centralidade de intermediação atua como uma “ponte” ou “intermediário” entre diferentes partes do grafo.

- **Interpretação:** Autores com alta centralidade de intermediação são cruciais para o fluxo de informação e colaboração entre subgrupos de pesquisadores que, de outra forma, poderiam estar desconectados ou mais distantemente conectados. A remoção de tais nós pode fragmentar a rede.

2.4 Eigenvector Centrality (Centralidade de Autovetor)

- **O que é:** Atribui pontuações relativas a todos os nós na rede com base no princípio de que conexões com nós de alta pontuação contribuem mais para a pontuação do nó em questão do que conexões iguais com nós de baixa pontuação. É uma medida de influência que considera não apenas o número de conexões, mas a “qualidade” ou “importância” dessas conexões.
- **Interpretação:** Autores com alta centralidade de autovetor são influentes porque estão conectados a outros pesquisadores também influentes. Eles estão inseridos em “bons bairros” da rede.

2.5 Katz Centrality

- **O que é:** Semelhante à centralidade de autovetor, a Katz Centrality também mede a influência de um nó com base nas suas conexões. No entanto, ela considera todos os caminhos que conectam os nós, atribuindo pesos menores a caminhos mais longos através de um fator de atenuação (beta). Ela também adiciona uma pequena quantidade de centralidade a cada nó por padrão, garantindo que mesmo nós periféricos tenham alguma pontuação.
- **Interpretação:** Autores com alta Katz Centrality são influentes devido a uma combinação de muitas conexões diretas e indiretas com outros autores na rede, mesmo que algumas dessas conexões sejam através de caminhos mais longos.

A aplicação dessas cinco métricas permite uma análise multifacetada da estrutura da rede e dos papéis desempenhados pelos diferentes pesquisadores.

3 Resultados da Minha Análise

A aplicação das métricas de centralidade ao grafo de coautoria do PPGIA UNIFOR revelou padrões interessantes sobre a estrutura da rede e o papel de seus membros. A seguir, são apresentados os resultados detalhados para cada métrica, incluindo os 10 pesquisadores com maiores pontuações e uma interpretação baseada nas características de cada medida.

3.1 Degree Centrality (Centralidade de Grau)

Esta métrica reflete o número de colaboradores diretos. Um alto grau sugere um pesquisador com muitas colaborações.

Tabela 1: Resultados para Degree Centrality.

| Rank | Nó (Autor) | Degree Centrality (Normalizada) | Grau Absoluto |
|------|----------------|---------------------------------|---------------|
| 1 | MAF RODRIGUES | 0.1092 | 26 |
| 2 | JJPC RODRIGUES | 0.0546 | 13 |
| 3 | YR SERPA | 0.0420 | 10 |
| 4 | A SAMPAIO | 0.0420 | 10 |
| 5 | NC MENDONCA | 0.0420 | 10 |
| 6 | N MENDONCA | 0.0336 | 8 |
| 7 | A BRAYNER | 0.0336 | 8 |
| 8 | RH FILHO | 0.0294 | 7 |
| 9 | C CAMINHA | 0.0294 | 7 |
| 10 | JVV SOBRAL | 0.0294 | 7 |

Interpretação Detalhada (Degree Centrality)

A análise da Centralidade de Grau (Degree Centrality) identificou a professora MAF RODRIGUES como a autora com o maior número de ligações diretas, totalizando 26 colaborações distintas. Este valor elevado é um forte indicador de que ela desempenha um papel excepcionalmente ativo e central em termos de volume de colaborações, desenvolvendo projetos e estudos com um número significativamente maior de indivíduos dentro desta rede em comparação com os demais. Em seguida, JJPC RODRIGUES também se destaca com 13 colaborações diretas. Pesquisadores como YR SERPA, A SAMPAIO e NC MENDONCA, com 10 colaborações cada, também demonstram uma rede colaborativa considerável. Esta métrica, portanto, aponta para os pesquisadores mais “conectados” em termos de parcerias diretas.

3.2 Closeness Centrality (Centralidade de Proximidade)

Esta métrica indica quão “próximo” um autor está, em média, de todos os outros na rede, facilitando a disseminação de informações.

Tabela 2: Resultados para Closeness Centrality.

| Rank | Nó (Autor) | Closeness Centrality |
|------|---------------|----------------------|
| 1 | MAF RODRIGUES | 0.1361 |
| 2 | NC MENDONCA | 0.1212 |
| 3 | YR SERPA | 0.1000 |
| 4 | HP PONTES | 0.0995 |
| 5 | DV MACEDO | 0.0985 |
| 6 | VM CARVALHO | 0.0976 |
| 7 | JH FONTELES | 0.0971 |
| 8 | PPM NETO | 0.0966 |
| 9 | DV DE MACEDO | 0.0966 |
| 10 | RG BARBOSA | 0.0962 |

Interpretação Detalhada (Closeness Centrality)

A partir da análise utilizando o algoritmo de Closeness Centrality, identificamos que a professora MAF RODRIGUES ocupa novamente a posição de maior centralidade de proximidade no grafo (0.1361). Isso significa que, em média, ela está a uma distância mais curta de todos os outros autores da rede. Em termos práticos, essa métrica sugere que a professora possui fácil acesso a qualquer outro pesquisador da rede, sendo uma ponte eficiente entre diferentes partes do grafo. Consequentemente, pesquisadores conectados a ela tendem a estar mais próximos dos demais, o que pode facilitar a disseminação de informações, a formação de novas colaborações e a ampliação da influência acadêmica. NC MENDONCA (0.1212) também demonstra uma excelente posição em termos de proximidade com os demais.

3.3 Betweenness Centrality (Centralidade de Intermediação)

Esta métrica identifica autores que atuam como “pontes” na rede, sendo cruciais para o fluxo de informação entre diferentes grupos.

Tabela 3: Resultados para Betweenness Centrality.

| Rank | Nó (Autor) | Betweenness Centrality |
|------|---------------|------------------------|
| 1 | MAF RODRIGUES | 0.0667 |
| 2 | NC MENDONCA | 0.0456 |
| 3 | A SAMPAIO | 0.0173 |
| 4 | HP PONTES | 0.0134 |
| 5 | VM CARVALHO | 0.0113 |
| 6 | N MENDONCA | 0.0094 |
| 7 | JB F DUARTE | 0.0092 |
| 8 | ES FURTADO | 0.0092 |
| 9 | YR SERPA | 0.0087 |
| 10 | CM ADERALDO | 0.0071 |

Interpretação Detalhada (Betweenness Centrality)

Ao aplicar a Betweenness Centrality no grafo, identificamos que a professora MAF RODRIGUES (0.0667) é o nó com a maior centralidade de intermediação, ou seja, é aquela que mais frequentemente aparece nos caminhos mais curtos entre outros pares de nós. Isso indica que ela tem um papel estratégico na rede: atua como uma ponte crucial entre diferentes partes ou subgrupos do grafo, facilitando a comunicação e a conexão entre professores que, de outra forma, poderiam estar isolados ou mais distantes. Ela não necessariamente tem o maior número de conexões diretas (embora também se destaque nisso), mas é fundamental para o fluxo de informação e para o acesso indireto entre membros da rede. Pesquisadores que não estão diretamente conectados entre si muitas vezes “passam por ela” para se conectarem, o que reforça sua importância estrutural. NC MENDONCA (0.0456) também se destaca significativamente como um importante intermediador.

3.4 Eigenvector Centrality (Centralidade de Autovetor)

Esta métrica mede a influência de um autor com base na influência de seus colaboradores. Estar conectado a outros autores influentes aumenta a própria influência.

Tabela 4: Resultados para Eigenvector Centrality.

| Rank | Nó (Autor) | Eigenvector Centrality |
|------|------------------|------------------------|
| 1 | MAF RODRIGUES | 0.6169 |
| 2 | YR SERPA | 0.2908 |
| 3 | JH FONTELES | 0.2121 |
| 4 | RG BARBOSA | 0.1872 |
| 5 | DV MACEDO | 0.1803 |
| 6 | NC MENDONCA | 0.1800 |
| 7 | HP PONTES | 0.1708 |
| 8 | PPM NETO | 0.1693 |
| 9 | DV DE MACEDO | 0.1532 |
| 10 | PATTERNS AND ... | 0.1517 |

Interpretação Detalhada (Eigenvector Centrality)

Com base na métrica de Eigenvector Centrality, observamos novamente que a professora MAF RODRIGUES (0.6169) se destaca como a autora mais influente da rede. Esse resultado indica que sua importância não se dá apenas pelo número de conexões diretas ou pela proximidade com os demais nós, mas principalmente pela “qualidade” dessas conexões. Ou seja, ela está conectada a outros autores que também são altamente conectados e influentes, o que amplifica ainda mais sua relevância na estrutura global do grafo. Esta centralidade evidencia um papel de liderança e influência acadêmica significativa dentro da rede de coautorias. YR SERPA (0.2908) e JH FONTELES (0.2121) também apresentam alta centralidade de autovetor, indicando que estão bem conectados dentro de “clusters” influentes da rede. (Nota: “PATTERNS AND ...” como décimo colocado sugere um nó que, apesar de um nome incompleto, está conectado a outros nós importantes).

3.5 Katz Centrality

Esta métrica também mede influência, considerando todos os caminhos (diretos e indiretos) com um fator de atenuação para caminhos mais longos.

Interpretação Detalhada (Katz Centrality)

A análise por meio da Katz Centrality confirmou novamente a posição de destaque da professora MAF RODRIGUES (0.2618) na rede de coautorias. Essa métrica leva em consideração não apenas o número de conexões diretas, mas também a influência dos nós aos quais um autor está conectado, atribuindo um peso adicional às conexões indiretas com autores influentes, mesmo que por caminhos mais longos. O fato de a professora liderar também nesse indicador reforça sua posição central e estratégica na rede acadêmica, demonstrando que sua influência se estende além dos vínculos diretos, alcançando de forma relevante diferentes camadas da rede. Isso sugere que ela atua como um importante elo

Tabela 5: Resultados para Katz Centrality.

| Rank | Nó (Autor) | Katz Centrality |
|------|----------------|-----------------|
| 1 | MAF RODRIGUES | 0.2618 |
| 2 | JJPC RODRIGUES | 0.1479 |
| 3 | NC MENDONCA | 0.1430 |
| 4 | YR SERPA | 0.1388 |
| 5 | A SAMPAIO | 0.1245 |
| 6 | JVV SOBRAL | 0.1071 |
| 7 | RH FILHO | 0.1053 |
| 8 | JH FONTELES | 0.1050 |
| 9 | N MENDONCA | 0.1042 |
| 10 | A BRAYNER | 0.1022 |

de disseminação de conhecimento e de fortalecimento das colaborações dentro da comunidade científica analisada. JJPC RODRIGUES (0.1479), NC MENDONCA (0.1430) e YR SERPA (0.1388) também apresentam valores elevados, indicando uma influência considerável que se propaga pela rede.

Grafo da professora MAF RODRIGUES

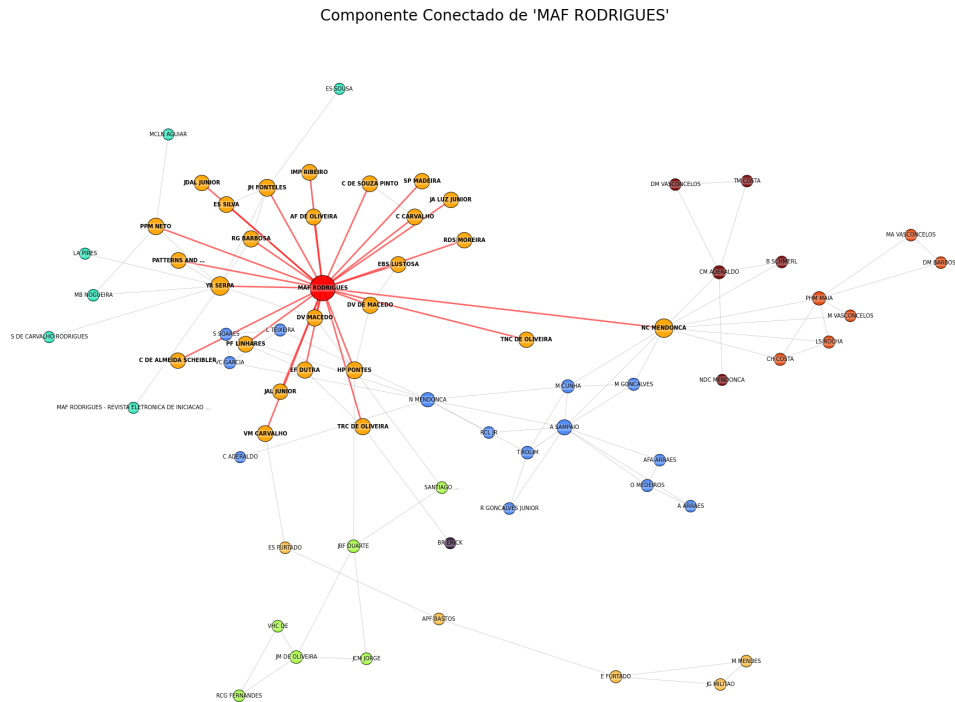


Figura 2: Visualização do grafo da professora MAF Rodrigues onde temos em destaque os nós que possuem relação direta com a autora.

Conclusão Geral dos Resultados

A análise multifacetada utilizando cinco diferentes métricas de centralidade aponta de forma consistente para MAF RODRIGUES como a pesquisadora mais proeminente na rede de coautoria do PPGIA UNIFOR, considerando o conjunto de dados utilizado. Sua alta pontuação em Degree Centrality indica um grande volume de colaborações; em Closeness Centrality, sua capacidade de rápido alcance na rede; em Betweenness Centrality, seu papel crucial como intermediadora; e em Eigenvector e Katz Centrality, sua significativa influência, derivada tanto de suas conexões diretas quanto da importância de seus colaboradores. Outros pesquisadores, como NC Mendonça, YR Serpa, JJPC Rodrigues e A Sampaio, também demonstram papéis importantes e variados na dinâmica da rede. É fundamental reiterar que estes resultados são um reflexo da amostra de dados coletada e podem não representar integralmente a totalidade das interações de coautoria do programa.

Visualizações Adicionais do Grafo

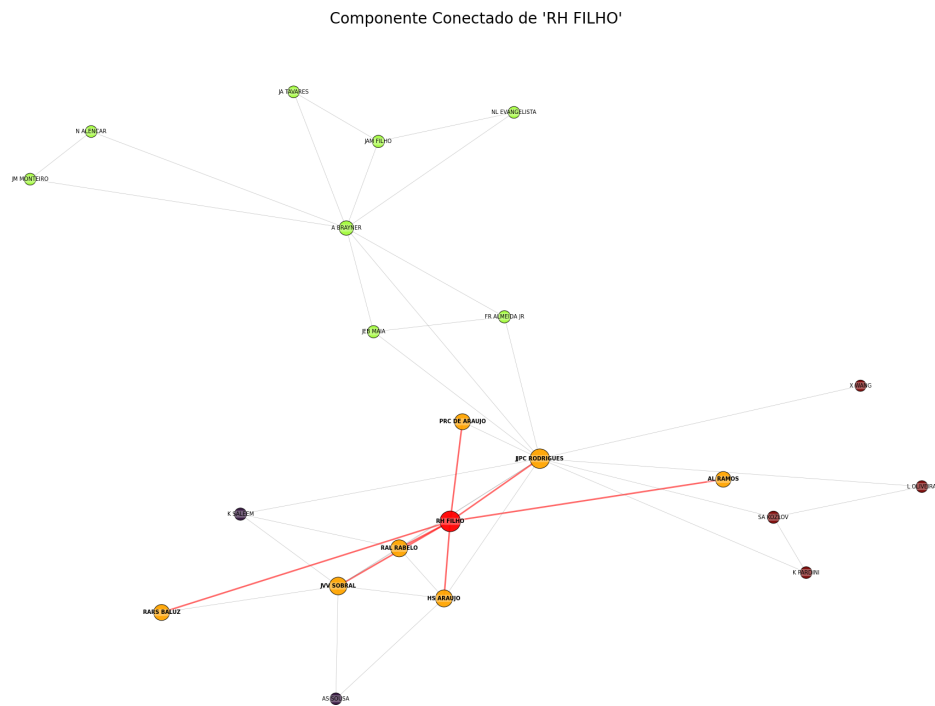


Figura 3: Visualização do grafo das ligações do professor Raimir Filho, com destaque para as ligações direta com o professor.

Visualização de Grafo (239 nós, 267 arestas)

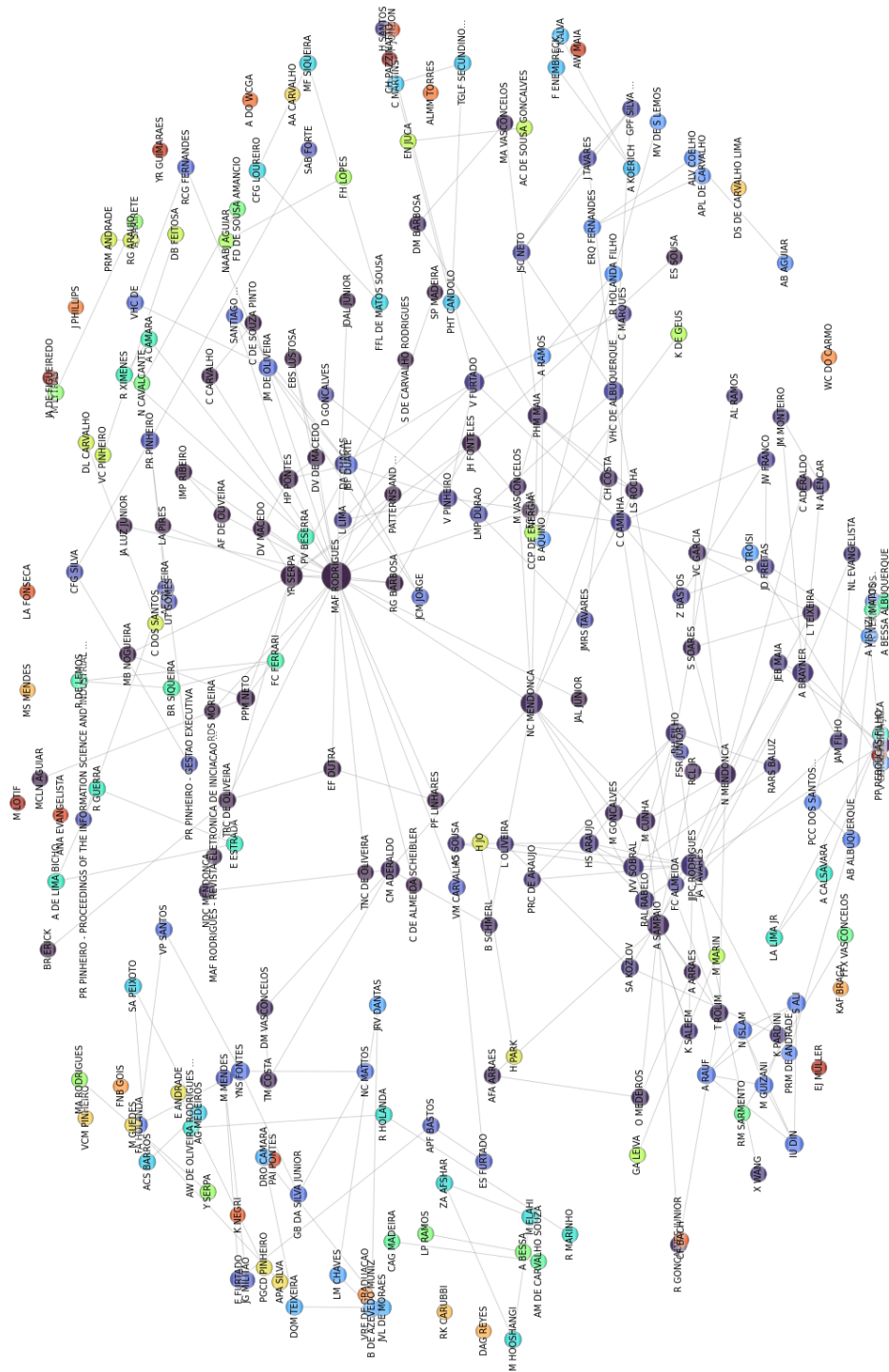


Figura 4: Visão geral do grafo com todos os nós que compõe ele (imagem girada 90 graus).