

UNIVERSIDADE DE SÃO PAULO  
FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO  
DEPARTAMENTO DE COMPUTAÇÃO E MATEMÁTICA

PAULO HENRIQUE LIMA DE PAULA

**Competição de Partículas Paralelas para Identificação  
de Comunidades Desbalanceadas com Aplicação em  
Redes Cerebrais Funcionais**

Ribeirão Preto–SP

2022

PAULO HENRIQUE LIMA DE PAULA

**Competição de Partículas Paralelas para Identificação de  
Comunidades Desbalanceadas com Aplicação em Redes Cerebrais  
Funcionais**

Versão Corrigida

Dissertação apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto (FFCLRP) da Universidade de São Paulo (USP), como parte das exigências para a obtenção do título de Mestre em Ciências.

Área de Concentração: Computação Aplicada.

Orientador: Zhao Liang

Coorientador: Renata Ferranti Leoni

Ribeirão Preto–SP

2022

Paulo Henrique Lima de Paula

Competição de Partículas Paralelas para Identificação de Comunidades Desbalanceadas com Aplicação em Redes Cerebrais Funcionais. Ribeirão Preto–SP, 2022.

80p. : il.; 30 cm.

Dissertação apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da USP, como parte das exigências para a obtenção do título de Mestre em Ciências,

Área: Computação Aplicada.

Orientador: Zhao Liang

Coorientador: Renata Ferranti Leoni

1. Redes Cerebrais. 2. Detecção de Comunidades. 3. Inteligência Computacional.

Paulo Henrique Lima de Paula

Competição de Partículas Paralelas para Identificação de Comunidades Desbalanceadas  
com Aplicação em Redes Cerebrais Funcionais

Modelo canônico de trabalho monográfico  
acadêmico em conformidade com as normas  
ABNT.

Trabalho aprovado. Ribeirão Preto–SP, de de 2022:

---

**Orientador:**

Professor Dr. Zhao Liang

---

**Professor**

Dr. Ricardo Marcondes Marcacini

---

**Professor**

Dr. João Roberto Bertini Junior

Ribeirão Preto–SP

2022

*Este trabalho é dedicado às pessoas que sofrem  
com doenças psiquiátricas e/ou neurológicas.*

# Agradecimentos

Em primeiro, eu gostaria de agradecer a Deus, e por todas as coisas boas que ele tem me proporcionado!

A minha família por ter me dado todo apoio e suporte durante esse período, e a união em todas as dificuldades presentes. Principalmente minha irmã, Beatriz Lima de Paula, que desde o começo sempre esteve ao meu lado me incentivando e aconselhando sobre a vida pessoal, profissional e acadêmica.

Ao Prof<sup>o</sup> Dr. Zhao Liang por ser meu guia e conselheiro profissional, sendo meu orientador e as exigências para ser realizado um trabalho de qualidade.

A Prof<sup>a</sup> Dra. Renata F. Leoni e ao aluno de Doutorado Pedro Henrique Rodrigues da Silva por ajudarem com a coorientação e nos auxílios de estudos de neurociência, e também disponibilidade do banco de dados, além de ajudar a desenvolver os métodos de análises e a colaboração desse estudo em aplicação à neurociência. Com essa colaboração, foi possível estender o estudo e obter um melhor desenvolvimento de toda pesquisa.

Todo grupo de pesquisa que apresentou uma forte união e ajuda no trabalho. Principalmente aos meus amigos, Rafael Delalibera Rodrigues e Luan Vinícius de Carvalho Martins por toda ajuda ao longo do processo deste trabalho.

Ao suporte financeiro oferecido pela Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP (processo nº 2019/09319-6).

*Tudo o que temos de decidir  
é o que fazer com o tempo que nos é dado.  
(Gandalf - O Senhor do Anéis)*

# Resumo

PAULA, P. H. L. **Competição de Partículas Paralela para Identificação de Comunidades Desbalanceadas com aplicação em Redes Cerebrais Funcionais**. 2021. 80 p. Dissertação (Mestrado em Ciências - Computação Aplicada) - Faculdade de Filosofia Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto - SP, 2022.

O cérebro humano pode ser considerado uma rede complexa, pois é dividido em regiões estruturais e funcionais que são interconectadas. A conectividade funcional tem sido usada para modelar regiões do cérebro que podem estar fisicamente (ou anatomicamente) separadas, através de dependência temporal dos padrões de ativação neural. Uma tarefa importante na aprendizagem não supervisionada é a detecção de comunidades em redes (SILVA; ZHAO, 2016). Embora muitas técnicas de detecção de comunidade tenham sido propostas, ainda existem alguns problemas, como detecção de comunidade desbalanceada e baixa eficiência. Encontrar estruturas comunitárias em redes é um campo que vem despertando atenção de diversas áreas de estudo. A neurociência inclui um interesse neste estudo, em que comunidades em redes cerebrais apresentam fatores consideráveis no desenvolvimento e na funcionalidade do cérebro, favorável à realização de análises de estruturas de comunidades em várias escalas. A teoria das redes complexas tem sido um mecanismo útil para o estudo do cérebro devido à sua possibilidade de lidar com sistemas que possuem propriedades de alta complexidade. Neste contexto, a teoria de redes complexas junto às técnicas de Imagem por Ressonância Magnética Funcional (do inglês *fMRI - Functional Magnetic Resonance Imaging*) são utilizadas para criar e mapear a rede cerebral. Foram analisados os padrões de conectividade da rede cerebral através de detecção de comunidades. Especificamente, propomos uma técnica de detecção de comunidade inspirada na propagação sequencial de sinais do modelo de competição de partículas e a propagação paralela inspirada nos Mapas Auto-Organizáveis (do inglês *Self-Organizing Map - SOM*). O modelo apresenta duas características salientes: 1) Ele pode detectar comunidades desbalanceadas; 2) É mais eficiente do que o modelo original de competição de partículas devido à introdução da propagação paralela. Nossos resultados demonstram uma alta precisão de detecção dos módulos em redes, além disso, é útil para encontrar comunidades desbalanceadas. Também se mostrou um método eficaz na caracterização de comunidades em redes cerebrais.

**Palavras-chave:** Detecção de Comunidades. Redes Cerebrais. Aprendizado não-supervisionado.



# Abstract

PAULA, P. H. L. **Parallel Particle Competition for the Identification of Unbalanced Communities with Application in Functional Brain Networks.** 2021. 80 p. Dissertação (Mestrado em Ciências - Computação Aplicada) - Faculdade de Filosofia Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto - SP, 2022.

The human brain can be considered a complex network, as it is divided into structural and functional regions that are interconnected. Functional connectivity has been used to model brain regions that may be physically (or anatomically) separated, through temporal dependence on neural activation patterns. An important task in unsupervised learning is the detection of communities in networks (SILVA; ZHAO, 2016). Although many community detection techniques have been proposed, there are still some challenging issues such as unbalanced community detection and low efficiency. Finding community structures in networks is a competence that has been attracting interest from several areas of study. Neuroscience includes an interest in this study, in which communities in brain networks have considerable factors in brain development and functionality, making it favorable to carry out analyzes of community structures at various scales. The theory of complex networks has been a useful mechanism for the study of the brain, due to its possibility to deal with systems in which they have properties of high complexity. In this context, the theory of complex networks together with the techniques of Functional Magnetic Resonance Imaging (fMRI) are used to create and map the brain network. The connectivity patterns of the brain network were analyzed through community detection. Specifically, we proposed a community detection technique that is inspired by the sequential propagation of signals from the particle competition model and by the parallel propagation inspired by the Self-Organizing Maps (SOM). The model has two salient features: 1) It can detect unbalanced communities; 2) It is much more efficient than the original particle competition model due to the introduction of parallel propagation. Our results demonstrate a high detection precision of modules in networks, in addition it is useful to find unbalanced communities. It has also been shown to be an effective method for characterizing communities in brain networks.

**Keywords:** Community Detection. Brain Networks. Unsupervised Learning.

# Lista de figuras

Figura 1 – Ilustração do problema de detecção de comunidades desbalanceadas . . .	17
Figura 2 – Básico de uma rede . . . . .	21
Figura 3 – Ilustração do Modelo de Rede - Rede Aleatória . . . . .	23
Figura 4 – Ilustração do Modelo de Rede - Rede de Mundo Pequeno . . . . .	24
Figura 5 – Ilustração do Modelo de Rede - Rede Livre de Escala . . . . .	25
Figura 6 – Ilustração das comunidades na rede . . . . .	26
Figura 7 – Construção Rede Cerebral Funcional . . . . .	39
Figura 8 – Ilustração do novo modelo de Competição de Partículas . . . . .	49
Figura 9 – Ilustrando o processo de detecção de comunidades pelo Método Proposto	50
Figura 10 – Ilustração de detecção em Redes Artificiais pelo Método Proposto . . .	54
Figura 11 – Nível de Energia do Método Proposto . . . . .	55
Figura 12 – Nível de Energia do Método Proposto em Comunidades Desbalanceadas	55
Figura 13 – Comparando o tempo de iteração . . . . .	56
Figura 14 – Comparando a complexidade com algoritmos de Modularidade . . . . .	57
Figura 15 – Avaliando o impacto do sinal de propagação . . . . .	58
Figura 16 – Método Proposto Aplicado na Rede de GN . . . . .	59
Figura 17 – Resultado obtidos para Benchmark Lancichinetti . . . . .	60
Figura 18 – Método Proposto Aplicado na Rede do clube de caratê Zachary . . . . .	60
Figura 19 – Ilustração do procedimento para detectar comunidades desbalanceadas	61
Figura 20 – Precisão de Desbalanceamento (Tamanho) . . . . .	62
Figura 21 – Precisão de Desbalanceamento (Densidade) . . . . .	63
Figura 22 – Ilustração das Plotagens em 3D do cérebro . . . . .	64
Figura 23 – Numero Ideal de Partículas . . . . .	67
Figura 24 – Ilustração de Comunidades Detectadas com o Número Ideal para a Técnica Proposta . . . . .	68
Figura 25 – Ilustração de Comunidades Detectadas com a técnica de Modularidade Maximizada . . . . .	69
Figura 26 – Análise das regiões que constituem cada comunidade . . . . .	71

# Lista de tabelas

Tabela 1 – Algoritmo de Competição de Partículas . . . . .	33
Tabela 2 – Algoritmo de Treinamento da rede SOM . . . . .	34
Tabela 3 – Parâmetros Utilizados . . . . .	43
Tabela 4 – Algoritmo Proposto . . . . .	48
Tabela 5 – Parâmetros Iniciais . . . . .	52
Tabela 6 – Parâmetros Utilizados . . . . .	53
Tabela 7 – Análise Topológica . . . . .	70

# Lista de abreviaturas e siglas

AAL	Automated Anatomical Labeling
BOLD	Nível de Dependência de Oxigenação do Sangue (do inglês <i>Blood Oxygenation Level Dependence</i> )
fcMap	Mapa de Conectividade Funcional (do inglês <i>Functional Connectivity Map</i> )
fMRI	Imagem por Ressonância Magnética Funcional (do inglês <i>Functional Magnetic Resonance Imaging</i> )
PC	Competição de Partículas (do inglês <i>Particle Competition</i> )
PET	Tomografia por Emissão de Pósitrons (do inglês <i>Positron Emission Tomography</i> )
ROI	Região de Interesse (do inglês <i>Regions of Interest</i> )

# Sumário

<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>15</b>
<b>1.1</b>	<b>Contextualização . . . . .</b>	<b>15</b>
<b>1.2</b>	<b>Objetivo . . . . .</b>	<b>18</b>
<b>1.3</b>	<b>Contribuições . . . . .</b>	<b>19</b>
<b>1.4</b>	<b>Organização do Documento . . . . .</b>	<b>19</b>
<b>2</b>	<b>PESQUISA BIBLIOGRÁFICA . . . . .</b>	<b>20</b>
<b>2.1</b>	<b>Redes Complexas . . . . .</b>	<b>20</b>
2.1.1	Modelos de Redes . . . . .	22
2.1.2	Detecção de Comunidades . . . . .	23
2.1.3	Métodos de Detecção de Comunidades . . . . .	25
<b>2.2</b>	<b>Modelos de Aprendizado Competitivo . . . . .</b>	<b>28</b>
2.2.1	Competição de Partículas para Detecção de comunidades em Redes Complexas . . . . .	28
2.2.2	Mapas Auto Organizáveis de Kohonen . . . . .	32
<b>2.3</b>	<b>Aspectos de Neurociência . . . . .</b>	<b>35</b>
2.3.1	Rede Cerebrais Funcionais . . . . .	35
2.3.2	Imagem por Ressonância Magnética Funcional . . . . .	35
2.3.3	Pré-Processamento em <i>fMRI</i> . . . . .	36
2.3.4	Análise de Conectividade Funcional . . . . .	37
<b>2.4</b>	<b>Detecção de Comunidades em Redes Cerebrais . . . . .</b>	<b>40</b>
2.4.1	Análise Topológica das Comunidades em Rede Cerebrais . . . . .	41
<b>3</b>	<b>MODELO DO COMPETIÇÃO DE PARTÍCULAS PARALELA</b>	<b>43</b>
<b>3.1</b>	<b>Considerações Iniciais . . . . .</b>	<b>43</b>
<b>3.2</b>	<b>Metodologia do Procedimento . . . . .</b>	<b>44</b>
<b>3.3</b>	<b>Descrição do Processo . . . . .</b>	<b>46</b>
<b>3.4</b>	<b>Algoritmo Proposto . . . . .</b>	<b>48</b>
<b>3.5</b>	<b>Condições do Sistema . . . . .</b>	<b>51</b>
3.5.1	Critérios de Inicialização . . . . .	51
3.5.2	Critérios de Finalização . . . . .	51
<b>3.6</b>	<b>Complexidade Computacional . . . . .</b>	<b>51</b>
<b>3.7</b>	<b>Parâmetros do Algoritmo . . . . .</b>	<b>52</b>
<b>4</b>	<b>ANÁLISE EMPÍRICA DO MODELO . . . . .</b>	<b>53</b>
<b>4.1</b>	<b>Considerações Iniciais . . . . .</b>	<b>53</b>

4.2	Propagação de Energia . . . . .	53
4.3	Análise da Complexidade do Algoritmo . . . . .	54
4.4	Impacto do Parâmetro do Sinal de Propagação $\alpha$ . . . . .	56
4.5	Redes de Referências . . . . .	56
4.6	Comunidades Desbalanceadas . . . . .	59
4.7	Considerações Finais . . . . .	61
5	DETECÇÃO DE COMUNIDADES EM REDES CEREBRAIS FUNCIONAIS . . . . .	64
5.1	Considerações Iniciais . . . . .	64
5.2	Conjunto de Dados . . . . .	65
5.3	Processamento de Imagens e Análise da Comunidades . . . . .	65
5.4	Resultado Experimentais . . . . .	67
6	CONCLUSÃO . . . . .	72
6.1	Limitações . . . . .	73
6.2	Publicações . . . . .	73
6.3	Trabalhos Futuros . . . . .	73
	<b>Referências</b> . . . . .	74
	<b>APÊNDICES</b>	79
APÊNDICE A	– DISPONIBILIDADE DO ALGORITMO DE APREN- DIZADO COMPETITIVO PARALELO . . . . .	80

---

# Introdução

## 1.1 Contextualização

Redes complexas são grafos de grande escala com padrões de conexão não triviais, que apresentam aspectos de estruturas agrupadas de vértices (ou por densidade), chamadas de comunidades (SILVA; ZHAO, 2012b; SILVA; ZHAO, 2016; BARABÁSI, 2016; SILVA; ZHAO, 2013). Cada comunidade é uma sub-rede cujos vértices estão densamente conectados dentro dela, enquanto as conexões com o resto da rede são relativamente esparsas (SILVA; ZHAO, 2013; SILVA; ZHAO, 2016). A detecção de comunidade em redes complexas, sendo um aprendizado não-supervisionado (SILVA; ZHAO, 2016), tem o objetivo de encontrar uma estrutura que está escondida nos dados, na qual representa o conhecimento implícito e, também, é utilizado para explicar o conhecimento existente nos dados, diferentemente do aprendizado supervisionado, que requer a existência de um conjunto de treinamento com objetos rotulados, para o algoritmo obter um modelo a partir da análise dos dados (SILVA; ZHAO, 2012b). Assim, a detecção de comunidades tem sido aplicada para resolver muitos problemas reais, como balanceamento de vértices em computação paralela, particionamento de circuitos, desenvolvimento de redes telefônicas e agrupamento de dados (SILVA; ZHAO, 2012a; SILVA; ZHAO, 2012b; SILVA; ZHAO, 2016).

Dentro no estudo da teoria dos grafos, a detecção de comunidades corresponde a divisão do grafo, particionando em subgrafos, e isso tem se apresentado com um problema NP-completo (do inglês *Non-Deterministic Polynomial time*) (FORTUNATO, 2010), ou seja, é um conjunto de problemas que pode ou não ser solucionado em um tempo polinomial não-determinístico (FORTUNATO, 2010). Com isso, diversos trabalhos foram desenvolvidos, para buscar soluções eficientes e robustas na detecção de comunidades (FORTUNATO, 2010; SILVA; ZHAO, 2016), tais como, técnica baseada na medida de *betweenness*, otimização da modularidade, teoria da informação, passeios aleatórios, etc. Para uma revisão abrangente deste tópico, veja (FORTUNATO, 2010).

---

O cérebro humano pode ser descrito como uma rede complexa composta por nós que representam regiões, como um conjunto de *voxels* de imagens de ressonância magnética, e conexões que representam a conectividade estrutural ou funcional entre os nós (FALLANI et al., 2007; SPORNS, 2013; RUBINOV; SPORNS, 2009). Em seguida, a teoria de redes complexas foi aplicada à pesquisa em neurociência (STAM; REIJNEVELD, 2007; SPORNS, 2013), mostrando sua utilidade para avaliar a organização cerebral em doenças neurológicas (BULLMORE; SPORNS, 2009; BULLMORE; SPORNS, 2012).

Um exemplo saliente de organizações modulares em redes complexas, é o cérebro humano (ou rede neural biológica), que está associado à adaptação de redes às múltiplas tarefas e, também, às atividades robustas (HARTWELL et al., 2000; SPORNS, 2013; RUBINOV; SPORNS, 2009). Sporns e Betzel apresentaram um trabalho (SPORNS; BETZEL, 2015) de redes cerebrais modulares, destacando a importância de detectar módulos que são de particular interesse, revelando sub-redes que estão densamente conectadas, geralmente correspondendo a componentes funcionais especializados. O estudo também revela a relação entre os módulos cerebrais estruturais e funcionais (SPORNS; BETZEL, 2015). Em (BETZEL et al., 2019), os autores apresentam um método que modela redes cerebrais multicamadas usando conectividade funcional para descobrir a estrutura modular da rede.

Para detecção da comunidade em redes cerebrais, a maioria das técnicas baseiam-se na maximização de modularidade (SPORNS; BETZEL, 2015; GARCIA et al., 2018; SANCHEZ-RODRIGUEZ et al., 2019; SANCHEZ-RODRIGUEZ et al., 2020). O cérebro apresenta estrutura de comunidades que se demonstra consistente nas suas conexões, nas quais apresentam módulos (e/ou regiões) que possuem funcionalidades similares e que são compartilhadas (BETZEL et al., 2019). A neurociência tem uma questão em aberto com a aplicação da teoria de redes complexas, em como são os agrupamentos e as interconexões das unidades neurais que, por sua vez, fornecem características para percepção, ação e comportamentos adaptativos (GARCIA et al., 2018).

Neste trabalho, é proposta uma nova abordagem. Modificamos o modelo de competição de partículas e construímos uma nova versão para detecção de comunidades. O modelo de competição de partículas foi originalmente proposto em (QUILES et al., 2008) e melhorado em (SILVA; ZHAO, 2012b; SILVA; ZHAO, 2012a; GAO et al., 2019; MARTINS; ZHAO, 2020; LI et al., 2008). A ideia básica consiste em: colocamos algumas partículas em uma dada rede, então, essas partículas viajam na rede e tentam dominar o maior número de vértices possível. Ao mesmo tempo, cada partícula compete com outras partículas para evitar a intrusão em seu território. Ao final, espera-se que cada partícula ocupe uma sub-rede correspondente a uma comunidade. O modelo é inspirado em sistemas reais, como campanhas eleitorais, competição por comida, água ou território entre animais, etc., e pode obter boa precisão de detecção de comunidade.



Entretanto, o modelo de competição de partículas apresenta pelo menos dois problemas: 1) Em uma grande rede, cada partícula pode demorar para retornar a um vértice. Desta forma, há um efeito de “perda de memória”. Com isso, a competição, principal mecanismo deste modelo, não ocorre com frequência. 2) Como todas as partículas têm o mesmo comportamento, o modelo não pode detectar comunidades desbalanceadas (diferentes em tamanho, densidade ou distribuição de grau). A Fig 1 ilustra o problema de detecção desbalanceada abordado com o modelo de competição de partículas, para redes que levam em conta o tamanho e a densidade.

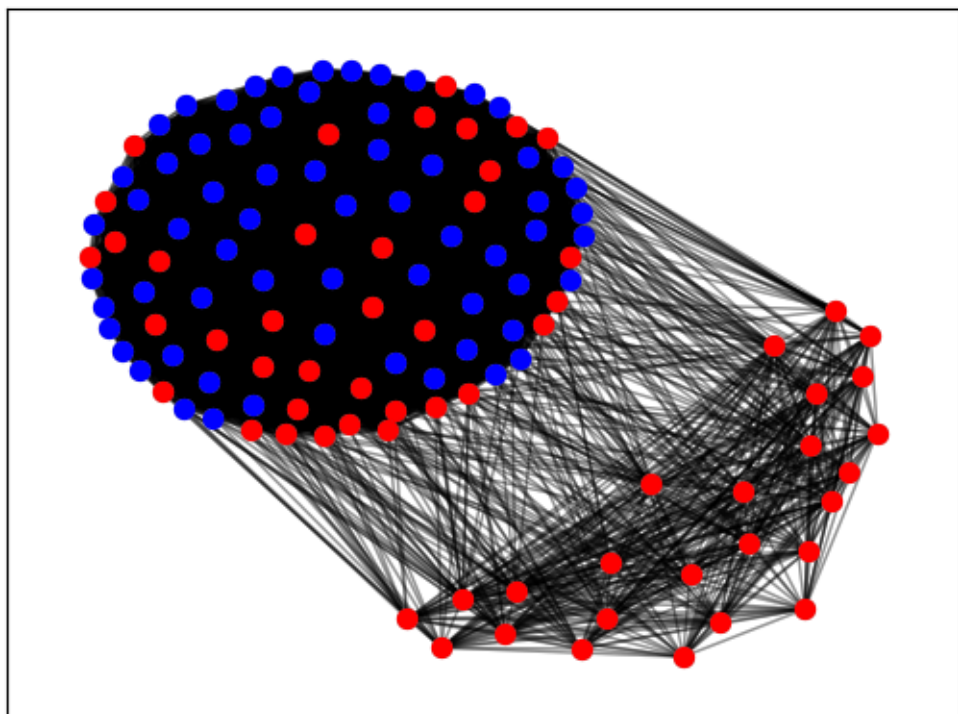


Figura 1 – Ilustração do problema de detecção de comunidades desbalanceadas com a metodologia de competição de partículas (SILVA; ZHAO, 2012b). Construção da rede contou com duas comunidades, uma com 25 vértices e outra com 100.  $P_{out} = 0.1$  e  $P_{in} = 1.0$ . Rede gerada pelo autor através do módulo *NetworkX* na linguagem de programação *Python*.

Com isso, apresentamos a elaboração de uma nova técnica de detecção de comunidades, combinando competição de partículas e propagação de sinal paralela inspirada nos Mapas Auto-Organizáveis (do inglês, *Self-Organizing Map* ou SOM). O primeiro é um processo sequencial, que pode verificar os vértices um a um, enquanto o último é um mecanismo de propagação paralela, que pode capturar rapidamente a estrutura local de cada vértice visitado por uma partícula. Com a ajuda da propagação paralela, o efeito

---

de deslocamento das partículas é amplificado e, conseqüentemente, memorizado. Além disso, cada partícula possui um comportamento distinto, devido à diferença de estruturas locais capturadas pela propagação paralela. Conseqüentemente, espera-se que esse modelo proposto possa também identificar comunidades desbalanceadas em uma rede. Em seguida, aplicamos o novo modelo de competição de partículas para detectar comunidades de formas e tamanhos variados (QUILES et al., 2008; SILVA; ZHAO, 2012b; FORTUNATO, 2010) em redes cerebrais. Por isso, uma proposta para solucionar este problema contribuirá para estudo de mineração de dados e aprendizado de máquina em geral.

## 1.2 Objetivo

Na seção anterior, 1.1, abordamos uma visão geral do trabalho desenvolvido, assim como dos problemas e limitações que são encontrados na detecção de comunidades em rede cerebrais. Dessa forma, foi construída uma questão de pesquisa. Embora muitas técnicas de detecção de comunidades tenham sido propostas, ainda existem alguns problemas desafiadores, como a detecção desbalanceada de comunidades e a baixa eficiência, devido a diversos fatores, tais como organização da rede e densidade de interconexões entre os vértices da rede.

Embora o modelo de competição de partícula consiga obter excelentes resultados na detecção de comunidades, o método pode falhar para identificar comunidades com diferentes tamanhos, pelo fato de as partículas apresentarem o mesmo comportamento de defender os vértices ao qual pertencentes à sua comunidade. Assim, o estudo tem como objetivo propor um modelo modificado de competição de partículas, que permite detectar comunidades desbalanceadas em redes, com um mecanismo paralelo inspirado nos Mapas de Kohonen (KOHONEN, 1990). Além disso, o modelo foi aplicado ao estudo de redes cerebrais e a organização funcional da rede foi analisada, com o objetivo de encontrar a relação entre as comunidades de distintas redes. O desenvolvimento se deu pelas seguintes etapas:

1. Desenvolver uma nova técnica de competição de partículas, para efetuar a detecção de comunidades desbalanceadas em redes, as quais são de tamanhos e densidades diferentes e ocorrem muito em redes reais;
2. Construir redes cerebrais de acordo com a base de dados de imagens de *fMRI* (do inglês *Functional Magnetic Resonance Imaging*) do cérebro;
3. Fazer uma análise modular das comunidades em redes cerebrais e caracterizar as diferenças topológicas de seus módulos para classificação, além de comparar com o método de Maximização da Modularidade que é muito referenciado na literatura.

## 1.3 Contribuições

Como contribuições desse estudo, os resultados obtidos com às duas abordagens de detecção com o novo método formulado, e a sua aplicação a detecção de comunidades em redes cerebrais funcionais, de modo geral, foram:

1. *Detecção de Comunidades Desbalanceadas*: O mecanismo de propagação paralela inspirada nos mapas auto-organizáveis, ajuda a capturar a estrutura local da rede. Como consequência, o novo modelo pode detectar comunidades desbalanceadas;
2. *Redes Cerebrais Funcionais*: Sua aplicação para a detecção de comunidades em redes cerebrais identificou módulos relacionados a padrões de redes funcionais, sugerindo-as como uma ferramenta útil e adicional para realizar mapeamento funcional do cérebro.

## 1.4 Organização do Documento

O Capítulo 2 contém a revisão bibliográfica do projeto, apresentando abordagens de redes complexas, detecção de comunidades, aspectos de neurociência e dissertações a respeito dos algoritmos que usamos como bases para o desenvolvimento da técnica proposta.

O Capítulo 3 esclarece o desenvolvimento e funcionamento da metodologia do processo de detecção da nova técnica que está sendo proposta.

Os Capítulos 4 e 5 apresentam as simulações do experimento, realizado com o método proposto, tanto com uma análise empírica do método (Capítulo 4), quanto aplicado à neurociência (Capítulo 5).

Por fim, o Capítulo 6 contém a conclusão, as contribuições e os trabalhos futuros desse estudo.

---

# Pesquisa Bibliográfica

Neste capítulo serão revisados os conceitos e as técnicas relevantes ao desenvolvimento deste estudo.

## 2.1 Redes Complexas

Redes complexas são ferramentas poderosas para interdisciplinaridades no ramo da ciência, que apresenta uma alta complexidade e versatilidade, e permite um grande espaço para exploração (BARABASI; ALBERT, 1999; BARABÁSI, 2016; SILVA; ZHAO, 2016; BULLMORE; SPORNS, 2012; WATTS, 1998). Além disso, também são procedimentos para a modelagem e aplicação de aprendizado de máquina para análise de dados (SILVA; ZHAO, 2016). A detecção de comunidades em redes complexas tem como propósito, buscar grupos de vértices que estão mais conectados entre si, do que o restante da rede (SILVA; ZHAO, 2012b; FORTUNATO, 2010; NEWMAN; GIRVAN, 2004), e com isso torna-se uma ferramenta útil para o particionamento do grafo, na qual procura verificar os vértices similares na rede (SILVA; ZHAO, 2012b; FORTUNATO, 2010; NEWMAN; GIRVAN, 2004).

Na teoria de redes complexas, uma rede corresponde a um grafo (uma estrutura matemática)  $G$ , em um sistema complexo, apresentando uma topografia não trivial, constituída por um conjunto de  $V$  vértices interligados por  $F$  arestas (SILVA; ZHAO, 2016; BULLMORE; SPORNS, 2012; NEWMAN; GIRVAN, 2004; FORTUNATO, 2010; BARABASI; ALBERT, 1999; BARABÁSI, 2016) (Veja a Fig.2). As conexões entre os vértices podem ser direcionadas, não direcionadas e/ou ponderadas.

- *Direcionada:* Geralmente, é representada por setas. Nesse tipo de conexão pode existir uma conexão do vértice  $i$  ao  $j$ , mas não existir a conexão do vértice  $j$  ao  $i$ ;

- *Não direcionadas*: Cada aresta é denotada por  $(i, j)$ , na qual conecta os vértices  $i$  e  $j$ .
- *Ponderadas*: Também há arestas que conectam os vértices por pesos, na qual as conexões entre os vértices se dão por um peso (algum valor) na aresta. Formalmente, para representar as arestas ponderadas, usa-se a seguinte estrutura:  $(i, j, w)$ , vértices  $i$  e  $j$  e em seguida o peso da conexão  $w$ .

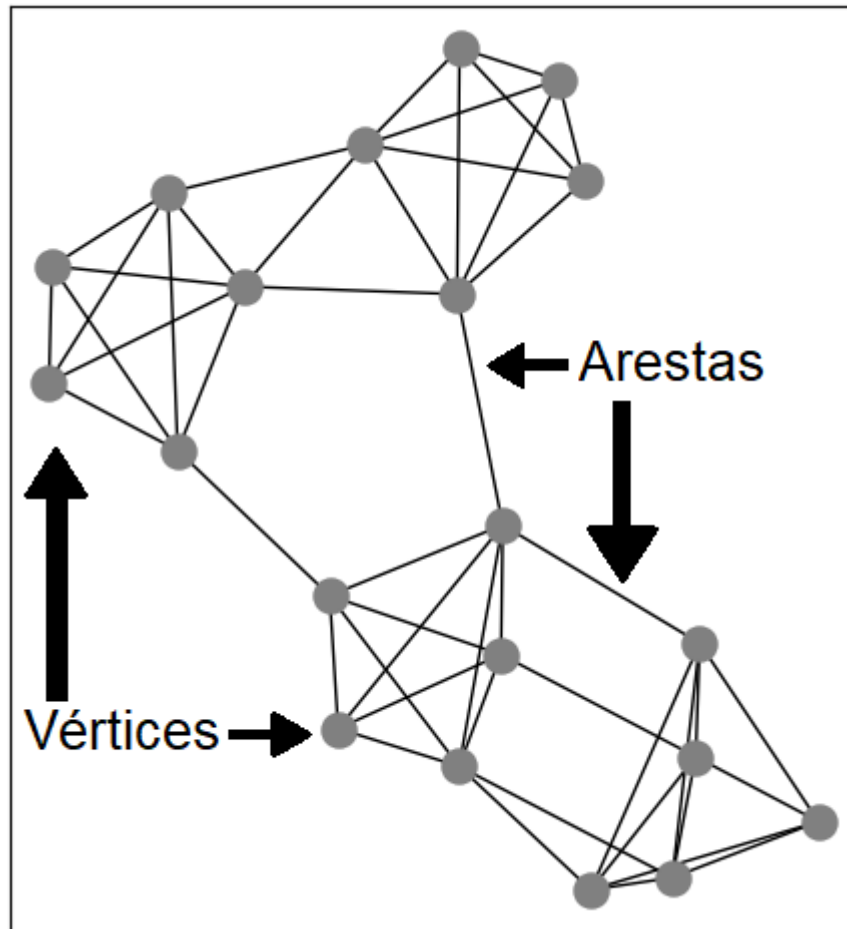


Figura 2 – Ilustrando conceitos básicos de uma rede. Vértices representam uma unidade fundamental em redes, as arestas são conexões entre os vértices. Rede gerada pelo autor através do módulo *NetworkX* na linguagem de programação *Python*.

A dinâmica de rede complexa se apresenta em forma funcional e estrutural, com abstrações naturais e mapeadas (SILVA; ZHAO, 2016; NEWMAN; GIRVAN, 2004; BARABÁSI, 2016). Uma rede com  $|V|$  vértices pode ser representada por uma matriz de adjacência  $A = (a_{i,j})$ ,  $i, j = 1, 2, \dots, |V|$ ,  $a_{i,j} = 1$  se existe uma aresta entre os vértices  $i$  e  $j$ ;  $a_{i,j} = 0$ , caso contrário. A matriz  $A$  é simétrica caso a rede seja não direcionada e, não simétrica, se a rede é direcionada. Em uma rede ponderada,  $a_{i,j}$  é um valor real que representa o peso de conexão entre  $i$  e  $j$ .

Com a teoria de redes complexas, é possível apresentar esquemas reais, como, por exemplo, em redes biológicas para classificação de espécies, genética, etc (SILVA; ZHAO, 2016; BARABÁSI, 2016; BULLMORE; SPORNS, 2009; SPORNS, 2013; RUBINOV; SPORNS, 2009). Dessa forma, as redes são capazes de identificar modelos de classificação (SILVA; ZHAO, 2016; SILVA; ZHAO, 2012b; BARABÁSI, 2016).

### 2.1.1 Modelos de Redes

Serão descritos brevemente alguns modelos de redes e as suas propriedades estruturais.

**Rede Aleatória:** Redes aleatórias são um estudo de redes complexas, dado por Erdős e Rényi (ERDÖS; RÉNYI, 1960). Neste tipo de rede, a probabilidade de conexões entre os vértices é aleatória, fixada em um parâmetro único (ERDÖS; RÉNYI, 1960), sendo que seguem a distribuição de Poisson. Neste modelo, no procedimento para construção da rede, os vértices são conectados uns aos outros de forma aleatória (ERDÖS; RÉNYI, 1960), na qual todos possuem em média a mesma quantidade de conexões, ou seja, contém a mesma probabilidade de receber novas ligações com outros vértices na rede. A Fig. 3 ilustra um modelo de rede aleatória.

**Rede de Pequeno Mundo:** Para a modelagem desta rede, Watts e Strogatz (WATTS, 1998) explicam que ela consiste na alteração das ligações (arestas) de uma rede regular (são redes que contém o mesmo grau para todos os vértices) com uma probabilidade ( $p$ ), e assim se apresentar como um pequeno mundo, sendo  $p = 0$  não há modificação na conexão da aresta, e, se  $p = 1$  as arestas da rede serão todas modificadas entre os vértices (WATTS, 1998).

Com isso, para uma rede de pequeno mundo (caso tiver muitos vértices), a distância média entre dois vértices, não vai ultrapassar de um pequeno número de vértices. Então, é necessário apenas algumas poucas conexões randômicas entre alguns grupos de vértices (WATTS, 1998). Essa rede permite uma comunicabilidade entre longas distâncias, criando ligações entre os vértices que estejam distantes um dos outros (WATTS, 1998). A Fig. 4 ilustra uma rede de pequeno mundo.

**Rede Livre de Escala:** Neste modelo de rede, formulado por Barabási e Albert (BARABÁSI; ALBERT, 1999; BARABÁSI, 2016), as redes não são geradas de forma aleatória. Conforme ilustrado na Fig. 5, são redes que seguem a distribuição de grau pela lei de potência, que podem ser apresentadas com muitos vértices de grau baixo e poucos vértices de grau alto, (RAVASZ; BARABÁSI, 2003; BARABÁSI, 2016), sendo mais

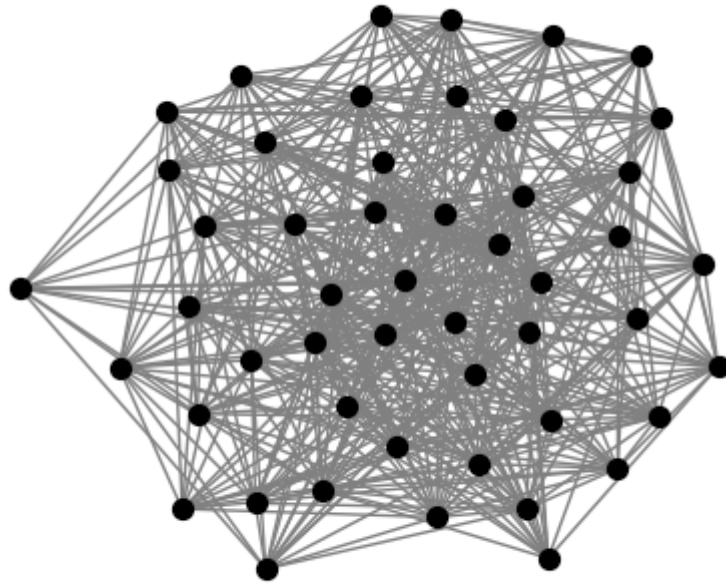


Figura 3 – Ilustração do Modelo de Rede - Rede Aleatória (ERDÖS; RÉNYI, 1960). Rede gerada pelo autor através do módulo *NetworkX* na linguagem de programação *Python*. Parâmetros de construção da rede com 50 vértices e probabilidade  $p$  de conexão de arestas com 0.8.

resistente a falhas, porém é mais vulnerável a ataques (ALBERT; JEONG; BARABASI, 2000).

Desta forma, quanto mais ligações um vértice apresenta, maior a probabilidade de criar ligações, que é a característica desse tipo de rede, pois, um novo vértice possui a tendência de se ligar com outro vértice (existente na rede) que contém mais ligações (BARABASI; ALBERT, 1999; BARABÁSI, 2016). Deste modo, nesse tipo de rede, o grau de conectividade será baixo, pois apenas alguns vértices irão ter muitas conexões com outros vértices, enquanto o restante irá ter poucas conexões (BARABASI; ALBERT, 1999; BARABÁSI, 2016).

### 2.1.2 Detecção de Comunidades

Redes complexas, podem apresentar estruturas agrupadas de vértices, as quais são chamadas de comunidades (SILVA; ZHAO, 2016; SILVA; ZHAO, 2012b; FORTUNATO, 2010; NEWMAN; GIRVAN, 2004; PALLA et al., 2005; FERREIRA; ZHAO, 2015; MARTINS; ZHAO, 2020; SILVA; ZHAO, 2013), como sendo um grupo de vértices que estão altamente conectados, na qual o número de conexões presente nesta estrutura supera a média de

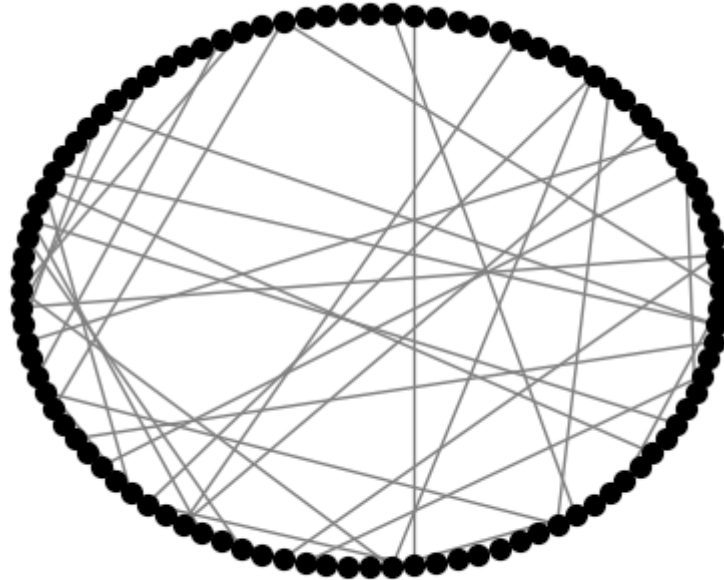


Figura 4 – Ilustração do Modelo de Rede - Rede de Mundo Pequeno (WATTS, 1998). Rede gerada pelo autor, com 100 vértices na forma de topologia em anel, através do módulo *NetworkX* na linguagem de programação *Python*.

conexões de toda a rede (SILVA; ZHAO, 2016; SILVA; ZHAO, 2012b; FORTUNATO, 2010; NEWMAN; GIRVAN, 2004; PALLA et al., 2005; FERREIRA; ZHAO, 2015; MARTINS; ZHAO, 2020; SILVA; ZHAO, 2013). As comunidades são estruturadas em vértices que possuem alguma relação de similaridade entre elas, e/ou são conectados por sua densidade de informações (SILVA; ZHAO, 2016; SILVA; ZHAO, 2012b; SILVA; ZHAO, 2013). Um exemplo de comunidades pode ser visto na Fig. 6.

Com a detecção de comunidades, é possível compreender quais são as relações entre os vértices e quais são suas propriedades e componentes (SILVA; ZHAO, 2012b; SILVA; ZHAO, 2012a; MARTINS; ZHAO, 2020; GARCIA et al., 2018; SILVA; ZHAO, 2013). Na literatura, alguns métodos para detecção de comunidade foram propostas, visando resolver problemas reais (SILVA; ZHAO, 2012b; SILVA; ZHAO, 2016; SILVA; ZHAO, 2012a), tais com agrupamento de dados, partição de circuitos, etc.

Embora diversas técnicas tenham sido propostas, ainda existem alguns problemas desafiadores, como a detecção de comunidades desbalanceadas e a baixa eficiência de detecção em redes (SILVA; ZHAO, 2016; SILVA; ZHAO, 2012b; FORTUNATO, 2010). Dessa forma, a detecção de comunidades em redes complexas é um trabalho fundamental dentro do aprendizado de máquina não supervisionado (SILVA; ZHAO, 2016; FORTU-



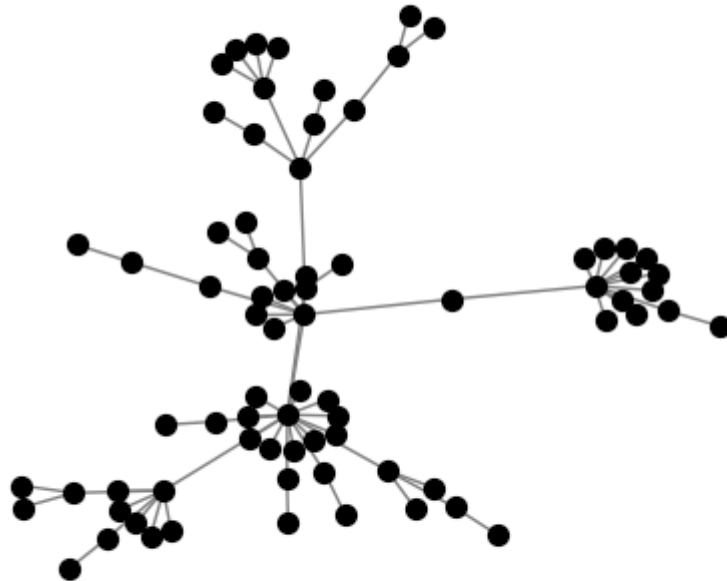


Figura 5 – Ilustração do Modelo de Rede - Rede Livre de Escala (BARABASI; ALBERT, 1999). Rede gerada pelo autor, com 75 vértices, através do módulo *NetworkX* na linguagem de programação *Python*.

NATO, 2010; SILVA; ZHAO, 2013; SILVA; ZHAO, 2012b), e tem como foco buscar um entendimento de que forma os vértices são organizados e estruturados em grupos, com alta conexão (GARCIA et al., 2018; SILVA; ZHAO, 2012b).

### 2.1.3 Métodos de Detecção de Comunidades

Diversas metodologias foram desenvolvidas para a identificação de comunidades em redes complexas (FORTUNATO, 2010). Nesta subseção, será descrita uma revisão resumida de alguns métodos e suas abordagens de detecção de comunidades.

**Métodos Divisivos:** Esses tipos de métodos, dividem a rede em sub-redes cada vez menores, através de eliminação iterativa das arestas que conectam diferentes comunidades na rede (GIRVAN; NEWMAN, 2002; ROCHA, 2017). Um método divisivo desenvolvido por (GIRVAN; NEWMAN, 2002; NEWMAN; GIRVAN, 2004) tem como base o conceito de *betweenness*. Neste conceito, para as arestas (ou conexões) que estão em uma região, na rede, entre comunidades vão receber um valor maior, enquanto as conexões presentes dentro de alguma comunidade, serão desfavorecidas, ou seja, receberam um valor

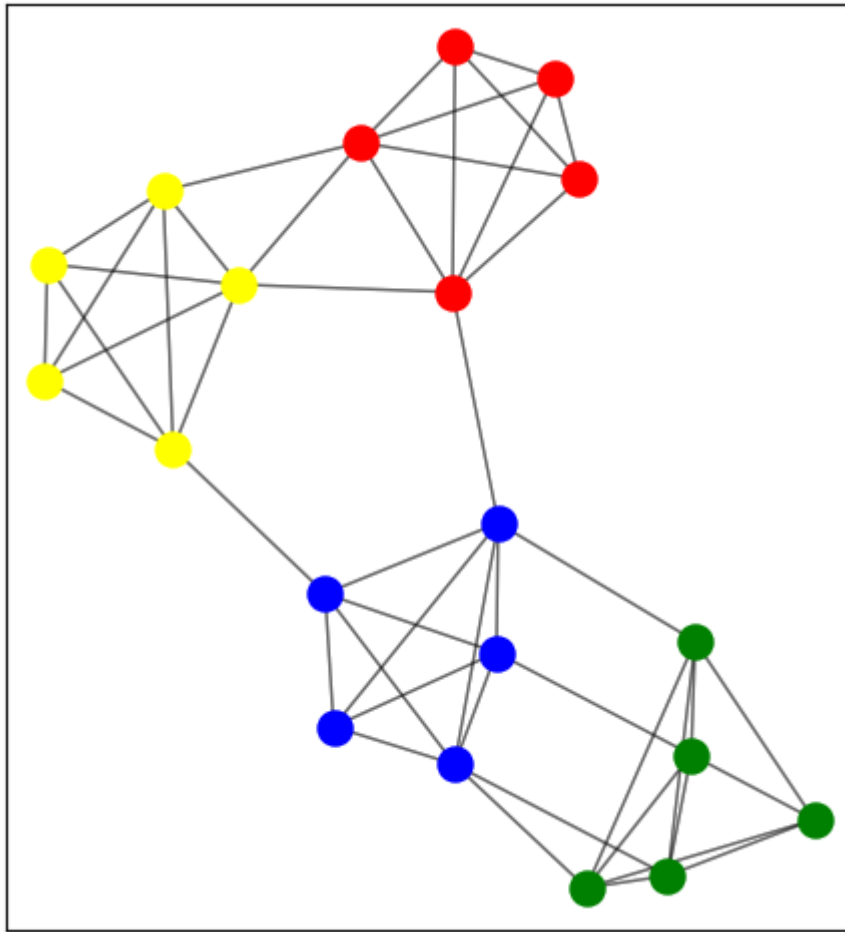


Figura 6 – Ilustração das comunidades na rede (com 4 comunidades). Vértices da mesmas cores representam uma comunidade da rede. Rede gerada pelo autor através do módulo *NetworkX* na linguagem de programação *Python*.

menor (ALONSO, 2008). Assim, remove, a cada iteração, uma aresta na rede com o maior valor *betweenness*, de modo que, no final, as ligações em comunidades permaneçam em sua maioria (ALONSO, 2008; GIRVAN; NEWMAN, 2002).

Esse método é baseado em como o número de aresta interliga duas comunidades distintas na rede, dessa forma, procura saber quais os caminhos que conectam as comunidades, pois o número de arestas é baixo (GIRVAN; NEWMAN, 2002; NEWMAN; GIRVAN, 2004; ALONSO, 2008). Assim, o princípio desse método é contar para as arestas, o número de caminhos que elas pertencem (GIRVAN; NEWMAN, 2002; NEWMAN; GIRVAN, 2004; ALONSO, 2008). Com isso, espera-se que a quantidade de caminho que cada aresta faz parte, seja maior entre diferentes comunidades (ALONSO, 2008; GIRVAN; NEWMAN, 2002).

O conceito de *betweenness* é baseado no caminho mais curto entre dois vértices (GIRVAN; NEWMAN, 2002; NEWMAN; GIRVAN, 2004; ALONSO, 2008). Um exemplo dessa medida é a distância entre dois vértices, que buscar identificar qual é o menor

caminho possível para todo par de vértices, assim para cada aresta é determinado a qual caminho ela faz parte, ou seja, conta para cada aresta na rede, o número de caminhos que ela está presente na rede (GIRVAN; NEWMAN, 2002; NEWMAN; GIRVAN, 2004).

**Método da Modularidade:** A medida de modularidade, foi proposta e formulada por (NEWMAN; GIRVAN, 2004), uma medida que quantifica uma determinada partição da rede em comunidades, dada pela Eq. 2.1:

$$Q = \frac{1}{2m} \sum_{i,j} (A_{i,j} - \frac{k_i k_j}{2m}) \delta(c_i, c_j) \quad (2.1)$$

onde,  $m$  é o número de arestas,  $A_{i,j}$  é um elemento da matriz  $A$ ,  $k_i$  é o grau do vértice  $i$ , assim como  $k_j$  é o grau do vértice  $j$ , e também  $c_i$  é o módulo que contém o vértice  $i$ , e  $c_j$  para o vértice  $j$ , assim  $\delta(c_i, c_j)$  é 1 se  $i$  e  $j$  estão na mesma comunidade e 0 caso contrário. Assim, nota-se que o valor da modularidade vai depender da estrutura da rede, ou seja, o seu conjunto de comunidades e como elas estão organizadas na rede. Com isso, o método de identificar comunidades por meio da modularidade, se dá em maximizar e/ou otimizar o valor da modularidade na rede (NEWMAN; GIRVAN, 2004; CLAUSET; NEWMAN; MOORE, 2004).

Portanto, esse método verifica qual arranjo identificado vai conter a maior modularidade, em relação aos vértices na rede (NEWMAN; GIRVAN, 2004; CLAUSET; NEWMAN; MOORE, 2004; BLONDEL et al., 2008). Assim, a modularidade, vai estar calculando proporção de arestas que conecta os vértices pertencentes ao mesmo módulo, e subtrai (com ligações randômicas entre os vértices) o número de arestas esperadas com a mesma quantia na rede, tendo as mesmas partições de comunidades (SOARES; OLIVEIRA; BRITO, 2014; NEWMAN; GIRVAN, 2004; CLAUSET; NEWMAN; MOORE, 2004).  $Q = 0$  ocorre quando a quantidade de arestas em uma comunidade não é maior que o valor aleatório, enquanto para valores que ficam próximos a  $Q = 1$  (valor máximo), indicam uma estrutura de comunidade forte (SOARES; OLIVEIRA; BRITO, 2014).

Entre os algoritmos de maximização de modularidade conhecidos, há o algoritmo *Fast Greedy* (CLAUSET; NEWMAN; MOORE, 2004), padrão de otimização gulosa, capaz de fornecer resultados rápidos e confiáveis e o algoritmo de *Louvain* (BLONDEL et al., 2008), que se baseou na mesma ideia sendo um método capaz de lidar com redes massivas com otimização de modularidade. No algoritmo da modularidade maximizada de *Louvain* (BLONDEL et al., 2008), a cada iteração, dois grupos de vértices serão unidos em uma única comunidade se esta união leva o maior ganho da modularidade (BLONDEL et al., 2008; ROCHA, 2017). O processo é repetido gerando uma hierarquia de comunidades.

**Método da Propagação de Rótulos:** Também conhecido do inglês como

*Label Propagation* (RAGHAVAN; ALBERT; KUMARA, 2007; CORDASCO; GARGANO, 2010), nesta técnica, um pequeno conjunto de vértices possui rótulos previamente definidos, indicando a qual comunidade cada um desse pequeno grupo de vértices pertence. Ao longo da execução do algoritmo, os rótulos são propagados iterativamente a outros vértices até que todos os vértices da rede recebam rótulos (RAGHAVAN; ALBERT; KUMARA, 2007), finalizando quando nenhum rótulo dos vértices é alterado na iteração.

Um algoritmo de propagação de rótulos, rápido para identificação das comunidades, foi implementando por (RAGHAVAN; ALBERT; KUMARA, 2007). Segue duas etapas: a primeira, depois da inicialização que cada vértice tem o seu rótulo, repete rótulo de um vértice para ser o rótulo que aparece com mais frequência entre os vizinhos desses vértices; como critério de parada do algoritmo, finaliza quando cada vértice possuir o rótulo propagado, ou seja, aquele rótulo que aparecer com mais frequência.

## 2.2 Modelos de Aprendizado Competitivo

Serão descritos dois modelos de aprendizado que são base de estudos para o desenvolvimento do modelo proposto de detecção de comunidades.

### 2.2.1 Competição de Partículas para Detecção de comunidades em Redes Complexas

Um método competitivo para identificar as comunidades na rede, é a Competição de Partículas, que foi originalmente proposto no trabalho (QUILES et al., 2008) e melhorado em (SILVA; ZHAO, 2016; SILVA; ZHAO, 2012b; SILVA; ZHAO, 2012a; GAO et al., 2019; MARTINS; ZHAO, 2020) para detecção de comunidades em redes complexas. Essa técnica possui duas vantagens salientes comparadas com outras técnicas de detecção de comunidades (SILVA; ZHAO, 2012b; GAO et al., 2019).

- 1) é equivalente em mecanismo de competição que é usado por sistemas reais, tais como, competição de territórios e comidas de seres humanos ou animais, campanha eleitoral, etc.;
- 2) a outra vantagem é que a técnica possui uma alta precisão de identificação das comunidades e, simultaneamente, apresenta uma baixa ordem de complexidade computacional.

Nesse procedimento, partículas caminham na rede e tentam conquistar o maior número de vértices possível, ao mesmo tempo, repelem partículas rivais dos seus vértices conquistados. Quando uma partícula visita algum vértice que está em seu domínio, ela aumenta seu nível de dominância nesse vértice visitado e diminui o nível das outras partículas (SILVA; ZHAO, 2012b). O processo segue até convergir, ou até um limiar de repetições estabelecido. No final, a estrutura de comunidade da rede pode ser identificada pelo conjunto de vértices que foram dominados por cada partícula. Este modelo é intrinsecamente um sistema dinâmico estocástico (SILVA; ZHAO, 2012b; GAO et al., 2019), ou seja, um sistema não é determinístico, administrado pelas seguintes equações:

$$p^{(k)}(t+1) = j, j \sim P_{tran}^{(k)}(t) \quad (2.2)$$

$$N_i^{(k)}(t+1) = N_i^{(k)}(t) + \mathbb{1}_{[P^{(k)}(t+1)=i]} \quad (2.3)$$

$$E^{(k)}(t+1) = \begin{cases} \min(\omega_{max}, E^{(k)}(t) + \Delta) \text{se} & \text{proprietária}(k, t), \\ \max(\omega_{min}, E^{(k)}(t) - \Delta) \text{se} & \neg \text{proprietária}(k, t) \end{cases} \quad (2.4)$$

$$S^{(k)}(t) = \mathbb{1}_{[E^{(k)}(t+1)=\omega_{min}]} \quad (2.5)$$

A Eq. 2.2 refere-se à localização de cada uma das partículas nos vértices da rede durante a iteração  $t$ . Na Eq. 2.3,  $N(t)$  é uma matriz armazena o número de visitas de cada partícula para todos os vértices na rede, durante a iteração  $t$ . Eq. 2.4,  $E(t)$  é um vetor que contém o nível de energia da partícula  $k$  durante a iteração  $t$ , onde  $E^{(k)}(t) \in [\omega_{min}, \omega_{max}]$ ,  $\omega_{max} \geq \omega_{min}$  (SILVA; ZHAO, 2012b).  $\Delta$  aumenta quando a partícula visita um vértice que esteja dominando, e  $\Delta$  diminui se a partícula visitar um território rival (SILVA; ZHAO, 2012b; MARTINS; ZHAO, 2020). E por último a Eq. 2.5 é uma função que demonstra se cada partícula está ativa na rede. Para isso, basta olhar o nível de energia das partículas que, caso estejam esgotadas, são reiniciadas em um vértice aleatório com energia máxima, caso contrário a partícula segue com a caminhada na rede.

$N_i^{(k)}(t)$  indica o número de visitas da partícula  $k$  no vértice  $i$  na iteração  $(t)$ , e também é necessário definir o de nível de dominação que a partícula  $k$  exerce no vértice  $i$  (SILVA; ZHAO, 2012b).  $\bar{N}_i^{(k)}(t)$  indica a relativa frequência de visitas de todas as partículas ao vértice  $i$ , na iteração  $t$  (SILVA; ZHAO, 2012b), dada pela Eq. 2.6.

$$\bar{N}_i^{(k)}(t) = \frac{N_i^{(k)}(t)}{\sum_{u=1}^{|V|} N_i^{(u)}(t)} \quad (2.6)$$

onde  $i$  é o vértice,  $k$  a partícula e  $N_i^{(k)}(t)$  a matriz que armazena o número de visitas.

No começo, as partículas são iniciadas randomicamente na rede, e também, é definido as condições iniciais do sistema. Conforme descrito no artigo (SILVA; ZHAO, 2012b), para uma competição justa, os valores iniciais do vetor  $E^{(k)}(0)$  (nível de energia das partículas) serão iguais para todas as partículas, dado na Eq. 2.7. Durante a execução do procedimento, cada partícula percorre a rede com o objetivo de dominar o maior número de vértices possíveis, enquanto repelem as partículas adversárias. Durante a caminhada, o número de visitas das partículas em cada vértice fica armazenado na matriz  $N_i(t)$  (SILVA; ZHAO, 2012b; SILVA; ZHAO, 2016).

$$E^{(k)}(0) = \omega_{min} + \left( \frac{\omega_{max} - \omega_{min}}{K} \right) \quad (2.7)$$

Para percorrer a rede, cada partícula escolhe um vértice para visitar utilizando a combinação de caminhadas (SILVA; ZHAO, 2012b; GAO et al., 2019; MARTINS; ZHAO, 2020). A primeira é a caminhada aleatória  $P_{ale}^{(k)}(i, j)$  (Eq. 2.8), que seleciona aleatoriamente um vértice do conjunto de seus vértices vizinhos, e a segunda caminhada preferencial  $P_{pref}^{(k)}(i, j, t)$  (Eq. 2.9), que tem maior probabilidade de visitar um vizinho cuja partícula atual possui alto nível de dominação.

$$P_{ale}^{(k)}(i, j) = \frac{a_{i,j}}{\sum_{u=1}^{|V|} a_{i,u}} \quad (2.8)$$

$$P_{pref}^{(k)}(i, j, t) = \frac{a_{i,j} \bar{N}_j^{(k)}(t)}{\sum_{u=1}^{|V|} a_{i,u} \bar{N}_u^{(k)}(t)} \quad (2.9)$$

$$P_{rean}^{(k)}(i, j, t) = \frac{\mathbb{1} \left[ \underset{m \in K}{\operatorname{argmax}} (\bar{N}_j^{(m)}(t)) = k \right]}{\sum_{u=1}^{|V|} \mathbb{1} \left[ \underset{m \in K}{\operatorname{argmax}} (\bar{N}_u^{(m)}(t)) = k \right]} \quad (2.10)$$

Para as Eqs. 2.8 e 2.9, um elemento da matriz de adjacência da rede é o  $a_{i,j} \bar{N}_j^{(k)}(t)$  nível de dominação que a partícula  $k$  exerce no vértice  $j$  até a iteração  $t$  (SILVA; ZHAO, 2012b). Já a Eq. 2.10 tem como tarefa permitir que a partícula possa se teletransportar para algum vértice que esteja dominando, ou seja, um vértice que é de sua propriedade, e tem como propriedade a operação  $\mathbb{1}[\cdot]$  que retornar 1 caso o modo da lógica seja atendida, se não retorna 0 (SILVA; ZHAO, 2012b; MARTINS; ZHAO, 2020). No seu funcionamento se cria uma matriz de transição, conforme as outras caminhadas, que possui a mesma probabilidade de caminhar para o vértice na rede na qual tenha dominância (SILVA; ZHAO, 2012b; MARTINS; ZHAO, 2020).

Combinando às regras de caminhada, temos um padrão de equilíbrio  $\lambda$ , na qual obtemos a matriz de transição, dada pela Eq. 2.11, onde  $\lambda \in [0, 1]$  (SILVA; ZHAO,

2012b; GAO et al., 2019). Logo, a caminhada aleatória e a caminhada preferencial são probabilidades de que um vértice vizinho  $j$  seja selecionado, caso a partícula esteja ativa na rede, dado pela Eq. 2.5. Se  $S^{(k)}(t)$  estiver esgotada, o movimento da partícula vai se dá por uma estratégia de defesa, em que se utiliza a matriz  $P_{rean}^{(k)}(t)$ , que leva a partícula a um vértice que seja do seu domínio, e assim possa recarregar a sua energia e defender a sua comunidade (SILVA; ZHAO, 2012b; MARTINS; ZHAO, 2020).

$$P_{tran}^{(k)}(t) = (1 - S^{(k)}(t)) \left[ \lambda P_{pref}^{(k)}(t) + (1 - \lambda) P_{ale}^{(k)}(t) \right] + S^{(k)}(t) P_{rean}^{(k)}(t) \quad (2.11)$$

Na Eq. 2.11, se o valor do padrão de equilíbrio for abaixo de 0.5, a partícula tem a probabilidade de fazer um passeio aleatório, caso ao contrário, se for acima de 0.5, tem de realizar uma caminhada preferencial (SILVA; ZHAO, 2012b). O algoritmo está na Tabela 1, em seguida, o procedimento de iteração do algoritmo é explicado por meio das etapas descritas a seguir.

- Passo 01: Entre com os parâmetros de entrada.  $k$  número de partículas, dados para construção da rede, o padrão de equilíbrio  $\lambda$  e  $\epsilon$ .
- Passo 02: Realiza a construção da Rede.
- Passo 03: As posições iniciais das partículas  $k$  são definidas nos vértices.
- Passo 04: Calcula a caminhada aleatória dada pela Eq. 2.8.
- Passo 05: Inicializa a matriz  $N_i^k(0)$ . Que recebe o valor 2 caso a partícula  $k$  é inicializada no vértice  $i$ , caso contrário recebe o valor de 1, conforme é apresentado em (SILVA; ZHAO, 2012b).
- Passo 06: Inicializa a matriz de dominância  $\bar{N}(0)$ .
- Passo 07: Inicializa o nível de energia no vetor  $E(0)$ .
- Passo 08: Inicializa  $S(0)$ , com o valor 0, conforme descrito no artigo (SILVA; ZHAO, 2012b).
- Passo 09: Inicializa a iteração 1 do procedimento.
- Passo 10: Repita as execuções do procedimento.
- Passo 11: Para cada partícula repita o processo de caminhada.
- Passo 12: Cálculo da caminhada preferencial para as partículas.
- Passo 13: Cálculo do  $P_{rean}^{(k)}(t)$ .

- Passo 14: Cálculo da caminhada para as partículas, combinando as matrizes de caminhada com o padrão de equilíbrio  $\lambda$ .
- Passo 15: A partícula escolhe um novo vértice para visitar e aumentar o seu nível de dominância.
- Passo 16: fim da repetição para cada partícula.
- Passo 17: Atualiza a matriz  $N(t)$ .
- Passo 18: Atualiza a matriz de dominância  $\bar{N}(t)$ .
- Passo 19: Atualiza o vetor do nível de energia  $E(t)$ .
- Passo 20: Atualiza se a partícula continua ativa na rede.
- Passo 21: Atualiza o tempo de execução.
- Passo 22: Repete até o padrão de convergência apresentado em (SILVA; ZHAO, 2012b), que checa se a maior mudança nos níveis de dominação está abaixo de um certo limite  $\epsilon$ .
- Passo 23: Retorna à matriz de dominância  $\bar{N}(t)$ .
- Passo 24: Fim do procedimento.

### 2.2.2 Mapas Auto Organizáveis de Kohonen

A rede SOM (do inglês *Self-Organizing Map*) ou rede auto-organizável, também conhecido como Mapa de Kohonen, é um aprendizado competitivo, que usa o treinamento de aprendizagem não-supervisionada na qual a rede vai encontrar similaridades nos dados de entrada, através de padrões e/ou características, e foi originalmente proposto por Kohonen (KOHONEN, 1990).

A rede SOM, tem uma camada de entrada e uma camada de saída (KOHONEN, 1990). Na organização dessas camadas, os neurônios que estão na camada de saída estão conectados entre si, formando uma vizinhança entre eles, e também estão ligados a todos os outros neurônios da camada de entrada (KOHONEN, 1990; ALONSO, 2008). É utilizado quando se contém um conjunto de dados com muitos atributos, pois ele retorna um mapeamento ordenado, geralmente com duas dimensões, sendo que a posição espacial de um neurônio artificial de saída em um mapa, corresponde a um domínio particular ou à característica dos dados obtidos do espaço de entrada (KOHONEN, 1990).

O funcionamento do algoritmo começa com a inicialização dos neurônios artificiais em um reticulado espacial e a atribuição aleatória de uma matriz de pesos  $w$ , definindo



Tabela 1 – Algoritmo de Competição de Partículas conforme descrito em (SILVA; ZHAO, 2012b)

<b>Algoritmo de Competição de Partículas</b>	
01:	<b>procedimento</b> PARTICLECOMPETITION( $K$ , data, $\lambda$ , $\epsilon$ )
02:	$A = \text{construa\_Grafo}(\text{data})$
03:	$p(0) = \text{gera\_ParticulaInicial}(A)$
04:	$P_{ale} = \text{calcula\_Random\_Matriz}(A)$ pela Eq. 2.8
05:	$N(0) = \text{calcula\_NInicial}(p(0))$
06:	$\bar{N}(t) = \text{calcula\_NBar}(N(0))$ pela Eq. 2.6
07:	$E(0) = \text{calcula\_EInicial}(K)$ pela Eq. 2.7
08:	$S(0) = \text{calcula\_SInicial}()$
09:	$t = 1$
10:	<b>repita</b>
11:	<b>para</b> $k = 1$ to $K$ <b>faça</b>
12:	$P_{pref}^{(k)}(t) = \text{calcula\_P}_{pref}(N(t-1), p(t-1))$ pela Eq. 2.9
13:	$P_{rean}^{(k)}(t) = \text{calcula\_P}_{rean}(N(t-1), p(t-1))$ pela Eq. 2.10
14:	$P_{tran}^{(k)}(t) = \text{seta\_P}_{tran}(\lambda, P_{ale}, P_{pref}^{(k)}(t), P_{rean}^{(k)}(t))$ pela Eq. 2.11
15:	$p^k(t) = \text{seleciona\_No}(P_{tran}^{(k)}(t), p^k(t-1))$
16:	<b>fim-para</b>
17:	$N(t) = \text{atualiza\_N}(N(t-1), p(t))$ pela Eq. 2.3
18:	$\bar{N}(t) = \text{atualiza\_NBar}(N(t))$ pela Eq. 2.6
19:	$E(t) = \text{atualiza\_E}(\Delta, E(t-1), \bar{N}(t), p(t))$ pela Eq. 2.4
20:	$S(t) = \text{atualiza\_S}(E(t))$ pela Eq. 2.5
21:	$t = t + 1$
22:	<b>até que</b> $\ \bar{N}(t) - \bar{N}(t-1)\ _{\infty} \leq \epsilon$
23:	<b>retorna</b> $\bar{N}(t)$
24:	<b>fim-procedimento</b>

também os parâmetros da vizinhança  $H_{i,j}$ , e os parâmetros de aprendizagem  $\eta$  (KOHONEN, 1990; ALONSO, 2008). Na Tabela 2 é descrito o algoritmo de treinamento da rede SOM. Assim, cada neurônio da camada de saída possui um vetor de peso (que é o peso da conexão com um neurônio da camada de entrada), e o treinamento da rede vai consistir em ajustar os pesos das conexões (ALONSO, 2008; KOHONEN, 1990). Então, para cada entrada de dados (padrão de treinamento), seu funcionamento se passa por três etapas:

1. Competição: Na etapa de competição, para cada padrão apresentado na camada de entrada, os neurônios da camada de saída vão competir entre si (KOHONEN, 1990), com o propósito de encontrar o neurônio vencedor. O neurônio que melhor responder ao estímulo do dado de entrada é ativado (neurônio vencedor), dada pela Eq. 2.12, onde  $D(t)$  é o padrão (exemplo) de entrada apresentado na rede e  $w_x$  é o vetor de peso do neurônio  $x$  da matriz de pesos  $w$ , enquanto  $y$  é o índice do neurônio vencedor e  $w_y$  corresponde ao seu vetor de pesos e  $\| \cdot \|$  é a função de dissimilaridade.

Ou seja, o neurônio vencedor é aquele que apresentar a menor distância, é o BMU (do inglês *Best Matching Unit*, valores variam de 0 a 1, sendo 0 menos próximo a BMU e 1 mais próximo), e para tal compara-se a distância do exemplo com a distância do vetor de pesos de um neurônio.

$$(\|D(t) - w_y\|) = \operatorname{argmin}(\|D(t) - w_x\|) \quad (2.12)$$

2. **Cooperação:** Nesta etapa, o neurônio escolhido como vencedor compartilha parte do seu estímulo recebido com neurônios da sua vizinhança, isso influencia sua região para que os neurônios próximos respondam a estímulos parecidos (KOHONEN, 1990). Uma das formas de propagar o estímulo para vizinhança, pode ser pelo método da Função Gaussiana.
3. **Adaptação Sináptica:** Por fim, na última etapa, os neurônios excitados (vencedor e sua vizinhança) ajustam seus pesos, dada pela Eq. 2.13, em direção ao estímulo recebido, com relação às taxas de aprendizado estabelecidas (KOHONEN, 1990). No qual  $w_j(t+1)$  representa o novo vetor de peso atualizado para o neurônio  $j$ , enquanto  $w_j(t)$  representa o vetor de peso atual do neurônio  $j$ , e,  $H_{i,j}(t)$  representa a função de vizinhança, que define com que força o neurônio  $i$  pertencente à vizinhança do neurônio  $j$  (vencedor) será atualizada,  $\eta(t)$  é a taxa de aprendizado e  $[D(t) - w_j(t)]$  é a diferença entre o dado de entrada e o peso atual de um neurônio  $j$  na iteração  $t$ .

$$w_j(t+1) = w_j(t) + H_{i,j}(t)\eta(t)(D(t) - w_j(t)) \quad (2.13)$$

Após todos os padrões de treinamento serem passados e os pesos dos neurônios ajustados, é atualizada a taxa de aprendizado, que decresce lentamente com o tempo.

Tabela 2 – Algoritmo de Treinamento da rede SOM

<b>Algoritmo SOM</b>	
01:	Inicialização dos pesos dos neurônios e dos parâmetros de aprendizagem
02:	<b>Repita</b>
03:	<b>para</b> todos os padrões de entrada <b>faça</b>
04:	Etapa de Competição (Eq. 2.12)
05:	Etapa de Cooperação e Adaptação Sináptica (Eq. 2.13)
06:	<b>fim-para</b>
07:	Atualizar a taxa de aprendizagem $\eta$ e a função de vizinhança $H$
08:	<b>até que</b> não existam modificações significativas nos pesos da rede

Assim ao final do processo, é gerado um mapa de *clusters* (agrupamento de dados), na qual, cada um desses *clusters* reflete em grupo de neurônios que corresponde a um padrão, que compartilham alguma característica similar entre si.

## 2.3 Aspectos de Neurociência

Com a neurociência é possível abordar novos temas com perspectivas baseado em sistemas complexos (SPORNS, 2013; RUBINOV; SPORNS, 2009). Ferramentas eficazes tem surgido na área de neurociência, para obter uma melhor compreensão da estrutura de sistemas complexos, e com isso, se tornaram fundamentais para a modelagem de redes (SPORNS, 2013; SPORNS et al., 2004; BULLMORE; SPORNS, 2009). Redes cerebrais são derivadas de observações anatômicas ou fisiológicas, resultando em redes de conectividade estrutural e funcional, respectivamente (SPORNS, 2013; BULLMORE; SPORNS, 2009; HEUVEL; POL, 2010).

### 2.3.1 Rede Cerebrais Funcionais

O cérebro humano é uma rede complexa que descreve as regiões (funcionais e estruturais) do cérebro e suas interações (HEUVEL; POL, 2010; BULLMORE; SPORNS, 2009; SPORNS, 2013; RUBINOV; SPORNS, 2009). A conectividade funcional, aponta importantes tarefas em processos cognitivos, agregando informações entre as regiões do cérebro e pode oferecer importantes compreensões da sua organização (HEUVEL; POL, 2010). As observações são descritas através de séries temporais, procurando um padrão de dependência estatística entre as atividades de regiões anatomicamente diferentes (HEUVEL; POL, 2010; FRISTON et al., 1993; SPORNS, 2013). Esse tipo de conexão é continuamente modulado por estímulos sensoriais e de contexto de tarefas, por isso é dependente do tempo (STAM; REIJNEVELD, 2007; BULLMORE; SPORNS, 2009; SPORNS et al., 2004; STAM et al., 2009). Para a neuroimagem funcional, a conectividade funcional retrata a relação entre os padrões de ativação neural de regiões cerebrais anatomicamente separadas, refletindo o nível de comunicação funcional entre as regiões (HEUVEL; POL, 2010; FRISTON et al., 1993; SPORNS, 2013). O estudo de rede cerebral funcional, demonstra que o desempenho dessas redes é como fenômenos de segregação e integração (SPORNS, 2013; RUBINOV; SPORNS, 2009; BULLMORE; SPORNS, 2009). A primeira (segregação) se refere a capacidade da rede para processamento em um grupo densamente interconectado de vértices, enquanto a integração caracteriza a capacidade de troca de informações via conexões presentes da rede (SPORNS, 2013; RUBINOV; SPORNS, 2009).

### 2.3.2 Imagem por Ressonância Magnética Funcional

A imagem por ressonância magnética funcional (fMRI, do inglês *Functional Magnetic Resonance Imaging*) é uma técnica não invasiva que possibilita medir indiretamente a atividade cerebral, por meio da detecção de alterações associadas ao fluxo sanguíneo

(SALMON; LEONI, 2019; ROCHA, 2017; LINDQUIST; WAGER, 2014). É processada na observação de respostas a uma atividade neural, que levam a um aumento do fluxo sanguíneo cerebral em regiões específicas (SALMON; LEONI, 2019; LINDQUIST; WAGER, 2014). Essa alteração é reflexo do aumento relativo de oxihemoglobina presente no sangue (é diamagnética e, com isso, o campo magnético local é menos afetado), comparado à desoxihemoglobina (é paramagnética, por isso, o campo magnético local é mais afetado significativamente), levando à modulação do sinal da ressonância magnética em uma forma de contraste conhecido como *BOLD* (do inglês *blood oxygenation level-dependent*) (SALMON; LEONI, 2019; HUETTEL; SONG; MCCARTHY, 2009; ROCHA, 2017; LINDQUIST; WAGER, 2014), veja Fig. 7 (a). No caso do contraste *BOLD*, a imagem adquirida é ponderada em  $T2^*$ , bastante suscetível às diferenças locais de campo magnético. Como a oxihemoglobina e a desoxihemoglobina afetam o campo magnético local de maneira diferente, quando há mudanças em suas concentrações, há mudanças no campo magnético local sentido pelos prótons de hidrogênio (H) (SALMON; LEONI, 2019). Assim, o sinal muda.

As imagens de *fMRI* são geralmente representadas como um bloco de dados em quatro dimensões (4D), sendo 3 dimensões espaciais e uma dimensão de tempo (ROCHA, 2017; LINDQUIST; WAGER, 2014). As imagens de *fMRI* podem ser adquiridas durante um estímulo ou tarefa, que pode ser motora, sensorial, auditiva ou cognitiva, revelando redes funcionais específicas. Porém, uma análise global das redes cerebrais também pode ser feita com a pessoa em repouso. Dessa forma, a utilização do *fMRI* em cérebros se resume à atividade intrínseca (ROCHA, 2017), ou seja, são quaisquer atividades neurais e metabólicas, que estejam acontecendo no cérebro que não estão diretamente relacionadas a uma tarefa específica (ROCHA, 2017).

### 2.3.3 Pré-Processamento em *fMRI*

O método de *fMRI* é útil para extrair informações funcionais, e identificar o sinal *BOLD* (PAMPLONA, 2014), porém precisa passar por um pré-processamento para tratar a variabilidade dos dados, e, além disso, preparar para análise estatística (SALMON; LEONI, 2019; PAMPLONA, 2014), veja a Fig. 7 (b). Abaixo são listados alguns métodos para o pré-processamento das imagens funcionais.

1. **Correção de tempo entre fatias.** Faz a correção da diferença temporal de quando as imagens são adquiridas entre fatias de um certo volume (SALMON; LEONI, 2019; PAMPLONA, 2014). A correção garante considerar o tempo correto de aquisição para posterior análise (SALMON; LEONI, 2019; PAMPLONA, 2014).
2. **Correção do movimento tridimensional.** Esse pré-processamento é importante,

pois, procura corrigir o impacto de movimentos involuntários durante o processo de aquisição (SALMON; LEONI, 2019; PAMPLONA, 2014). O cálculo do movimento é pautado com base em um volume de referência, onde parâmetros são passados para alinhar as séries temporais para as imagens de um mesmo indivíduo para o volume de referência (SALMON; LEONI, 2019; PAMPLONA, 2014).

3. **Co-Registro Anatômico-Funcional.** O objetivo é registrar as imagens de *fMRI* de baixa resolução espacial, com as imagens estruturais de alta resolução espacial (SALMON; LEONI, 2019; PAMPLONA, 2014). Assim, temos a informação estrutural e a funcional no mesmo espaço (SALMON; LEONI, 2019; PAMPLONA, 2014).
4. **Segmentação.** Segmenta as imagens em distintas classes de tecidos do cérebro, tais com a substância cinzenta, substância branca e/ou fluído espinhal, utilizando as imagens estruturais de alta resolução espacial (SALMON; LEONI, 2019; PAMPLONA, 2014).
5. **Normalização.** Permite análises de grupo e comparação direta entre sujeitos. A intenção é colocar todos os volumes a um espaço geométrico comum (SALMON; LEONI, 2019; PAMPLONA, 2014). Faz com que a localização de uma região no cérebro de um indivíduo, corresponda a mesma localidade no cérebro de outro (SALMON; LEONI, 2019; PAMPLONA, 2014).
6. **Suavização espacial.** Com a aplicação de um filtro, suaviza volumes de imagens, a fim de reduzir ruídos durante o processo de pró-mediação entre sujeitos. Segundo (PAMPLONA, 2014), a vantagem dessa aplicação é a melhora da relação sinal-ruído, na validade dos testes estatísticos. Com isso, torna mais normal as mudanças anatômicas e funcionais entre os indivíduos, além dos erros que são distribuídos (PAMPLONA, 2014; SALMON; LEONI, 2019).
7. **Filtragem Temporal.** Este passo é importante no pré-processamento, pois, dados de *fMRI* podem exibir sinais em diferentes frequências causados por ruídos, sendo de ordem física ou fisiológica (SALMON; LEONI, 2019; PAMPLONA, 2014). Dessa forma, esses erros podem induzir a falsas correlações e confundir na interpretação dos resultados de conectividade funcional (SALMON; LEONI, 2019; PAMPLONA, 2014).

### 2.3.4 Análise de Conectividade Funcional

As redes cerebrais são grafos que representam o cérebro de uma maneira matemática, constituindo-se de vértices (regiões do cérebro) interligados através de arestas (conexões funcionais entre essas regiões) (BULLMORE; SPORNS, 2009; SPORNS, 2013; RUBINOV; SPORNS, 2009).

Após a etapa de pré-processamento dos dados, dividimos o cérebro em regiões de interesse (ROIs do inglês *region of interest*) (ROCHA, 2017). Dessa forma, a imagem funcional é registrada a um atlas, que contém as divisões em diferentes ROIs do cérebro, considerando a disposição espacial das regiões, determinadas por *voxels*, representando um valor em uma grade tridimensional (ROCHA, 2017). Com isso, os vértices são representados como regiões corticais (uma pequena região do cérebro); as ROIs. Como uma imagem funcional contém vários volumes, ao final há uma série temporal para cada ROI (ROCHA, 2017) (Fig. 7c).

Já que cada ROI tem uma série temporal, para estabelecer uma conexão entre duas regiões, basta utilizar alguma medida de similaridade. Alguns métodos foram formulados para identificar a conexão funcional entre as ROIs do cérebro (HEUVEL; POL, 2010), incluindo métodos de sementes (BISWAL et al., 1995; CORDES et al., 2000), análise de componentes independentes (ICA) (BECKMANN et al., 2005; CALHOUN et al., 2001; VEN et al., 2004) e agrupamento (CORDES et al., 2002; HEUVEL; MANDL; POL, 2008). De forma geral, pode-se colocar esses métodos em dois grupos, sendo o primeiro dependente de modelo e o segundo sem modelo.

1. **Métodos dependentes de modelo: método de semente.** Para investigar conexões funcionais de determinada região do cérebro com outras regiões, basta correlacionar a série temporal da região de interesse (semente) com as outras séries temporais das demais regiões, resultando num mapa de conectividade funcional (fcMap do inglês *functional connectivity map*) que define as conexões funcionais da região predefinida do cérebro (HEUVEL; POL, 2010; BISWAL et al., 1995; CORDES et al., 2000); fornecendo informações sobre com quais regiões a região da semente selecionada está funcionalmente ligada e em que extensão (HEUVEL; POL, 2010). Nesse método, a primeira etapa é a definição da semente, que pode ser selecionada através de um mapa de ativação dependente de tarefa tradicional, adquirido em um experimento de *fMRI* separado, identificando uma região específica de interesse (HEUVEL; POL, 2010). Muitos estudos também utilizam como sementes regiões anatomicamente definidas (por exemplo, modelos com diferentes parcelamentos) ou todos os *voxels* da imagem (análise *voxelwise*). Neste caso, é usada alguma medida de correlação para caracterizar a similaridade entre as regiões (ROIs) (ROCHA, 2017; HEUVEL; POL, 2010). Assim, a imagem de ressonância magnética funcional dá origem a uma rede, onde cada vértice é representado por uma ROI e a conexão de cada arestas entre as ROIs é dada pelo valor da similaridade entre as correlações (ROCHA, 2017). Calcular o nível de correlação entre as séries temporais de duas regiões (ROIs) do cérebro, resulta na matriz de correlação (ROCHA, 2017; BULLMORE; SPORNS, 2009; HEUVEL; POL, 2010; SPORNS, 2013) (Fig. 7d e 7e), formando uma rede ponderada. Há alguns estudos que utilizam rede não ponderada, sendo redes binárias.

Para isso, um valor limiar é atribuído para abstrair algumas conexões entre os pares de vértices (ROIs), i.e., caso o valor da correlação for maior ou igual ao limiar, aceita a ligação das arestas entre ambas regiões (ou seja, entre os vértices  $i$  e  $j$ ), caso contrário, não há ligação entre eles; formando uma matriz binária (Fig. 7f).

2. **Métodos sem modelo.** Essa técnica foi programada para encontrar padrões gerais de conectividade nas regiões do cérebro (HEUVEL; POL, 2010). Um método entre os mais utilizados é a análise de componentes independentes (ICA do inglês *Independent Component Analysis*) (BECKMANN et al., 2005; CALHOUN et al., 2001; VEN et al., 2004), que consegue decompor uma matriz de dados (tempo x *voxels*) em um conjunto de componentes espacialmente independentes com séries temporais e com distribuições espaciais associadas (HEUVEL; POL, 2010). A ICA é uma ferramenta muito útil para aplicações ao *fMRI*, pois pode ser usada para exames com maior profundidade de diferenças topológicas entre grupos de controle e pacientes, quando executado a dados de *voxels* de todo o cérebro e com os sinais temporais dos componentes independentes (HEUVEL; POL, 2010).

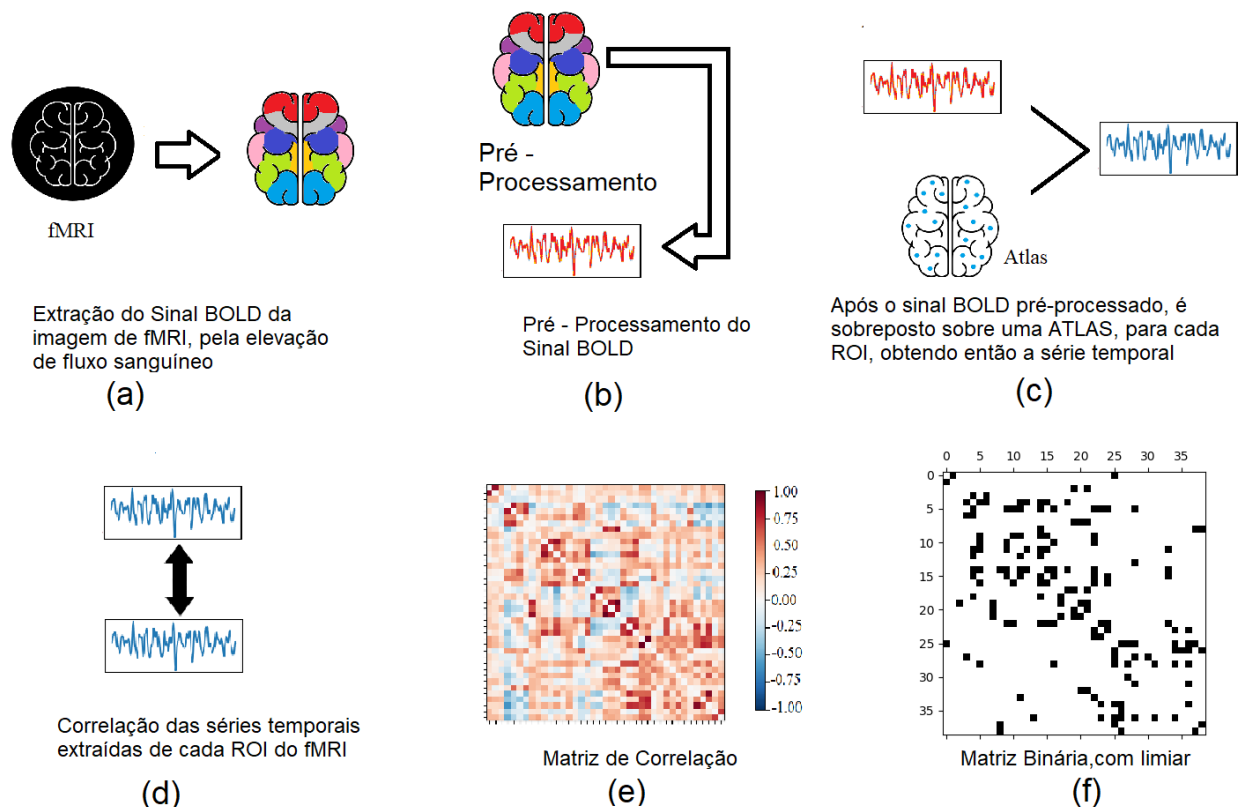


Figura 7 – (a) Extraindo o sinal *BOLD*. (b) Pré-Processamento do Sinal *BOLD*. (c) Sobrepondo com algum Atlas para extração das séries temporais. (d) Correlação entre as séries temporais. (e) Matriz de Correlação. (f) Matriz binária com limiar. Imagem feita pelo autor.

## 2.4 Detecção de Comunidades em Redes Cerebrais

Conforme apresentado na seção 1.1, além de Sporns e Betzel (SPORNS; BETZEL, 2015), outros estudos também destacaram a importância da detecção de módulos (comunidades) em redes cerebrais. Betzel e colaboradores (BETZEL et al., 2019) apresentaram um método que usa redes cerebrais de multicamadas, utilizando a conectividade funcional, no qual procurou-se descobrir sua estrutura modular. Nesse trabalho, a metodologia utilizada foi baseada na maximização de modularidade, que forneceu um meio para estudar as diferenças nas comunidades de redes cerebrais entre os sujeitos (BETZEL et al., 2019).

Sanchez e colaboradores apresentaram em sua pesquisa (SANCHEZ-RODRIGUEZ et al., 2019), estendida em (SANCHEZ-RODRIGUEZ et al., 2020), uma nova abordagem de detecção de comunidades em redes cerebrais. Sua metodologia consiste em duas etapas. A primeira foi soltar caminhantes na rede e deixá-los andar livremente e ir agrupando os vértices por sua ocorrência simultânea em cada escala de tempo, e após isso executar o *K-Modes* (HUANG, 1998), extensão do *K-Means* (MACQUEEN, 1967) para variáveis categóricas de meios binários, ou seja, de variáveis com os valores 0 ou 1. Com isso, retornando as partições desejadas nas redes cerebrais de multicamadas com conectividade estrutural e funcional. Essa técnica se mostrou capaz de identificar as comunidades em redes cerebrais.

No artigo (GARCIA et al., 2018), Garcia e colaboradores apresentaram aplicações das técnicas de detecção da comunidade às redes cerebrais, destacando as implicações algorítmicas para a função neural. Esse estudo, oferece recursos de medidas estatísticas que podem ser usadas para caracterizar a estrutura das comunidades dessas redes, que inclui medidas topológicas, espaciais (físicas) e para avaliação das forças e importâncias que as comunidades possuem na rede, com isso conseguimos extrair características das comunidades para validações e previsão. Também, discute métodos e estatísticas para redes temporais, pois, no estudo da função neural e sua relação com a cognição, geralmente temos um conjunto ordenado de redes, em que a ordem é baseada no tempo, e por fim, apresenta formas de medidas estatísticas que podem ser usadas para caracterizar a estrutura dinâmica da comunidade desses tipos de redes.



## 2.4.1 Análise Topológica das Comunidades em Rede Cerebrais

Metodologias de análises de comunidades foram formuladas para identificar o quão eficaz é um algoritmo para encontrar as comunidades (FORTUNATO, 2010). Garcia, em seu trabalho (GARCIA et al., 2018), também apresentou alguns conceitos específicos para análise topológica de módulos em redes cerebrais, citando exemplos de estudos.

**Modularidade:** A medida de modularidade, já apresentada na subseção 2.1.3, proposta por (NEWMAN; GIRVAN, 2004), indica a quão boa é a divisão da rede em sub-redes, ou seja, quantifica em quanto o algoritmo identifica uma comunidade na rede. Dessa forma, serve como uma medida útil da qualidade da partição de nós em comunidades.

**Conectividade Dentro e Entre os Módulos:** É importante calcular a resistência das arestas, tanto a resistência entre as comunidades como a dentro das comunidades também, ou seja, a força entre as comunidades (GARCIA et al., 2018). A fórmula é apresentada na Eq. 2.14:

$$I_{k1,k2} = \frac{\sum_{i \in C_{k1}, j \in C_{k2}} A_{i,j}}{N_{C_{k1}} N_{C_{k2}}} \quad (2.14)$$

onde  $A_{i,j}$  é o elemento na linha  $i$  e coluna  $j$  da matriz de adjacência da rede,  $C_{k1}$  é a primeira comunidade e  $C_{k2}$  é a segunda comunidade, e  $N_{C_{k1}}$  é o número de vértices contidos na primeira comunidade e  $N_{C_{k2}}$  na segunda.

**Força Intra-Modular:** Para quantificar o quão forte um vértice está conectado a outro vértice da mesma comunidade, usamos a Força Intra-Modular (GARCIA et al., 2018), dada pela Eq. 2.15.

$$z_i = \frac{S_{C_i} - S_{C_i}^-}{\sigma S_{C_i}^-} \quad (2.15)$$

onde  $S_{C_i}$  denota a resistência (isto é, o peso total das arestas) do vértice  $i$  e a resistência de outros vértices na mesma comunidade  $C_i$ , a quantidade  $S_{C_i}^-$  é a média de  $S_{C_i}$  sobre todos os vértices em  $C_i$ , e  $\sigma S_{C_i}^-$  é o desvio padrão de  $S_{C_i}$  dentro de  $C_i$ . O resultado do módulo é dado pela média do valor de  $z_i$ .

**Força Inter-Modular:** A Força Inter-Modular buscar saber como as conexões

---

de um vértice são compartilhadas entre os vértices de diferentes comunidades, também conhecida como coeficiente de participação (GARCIA et al., 2018).

$$P_i = 1 - \sum_{k=1}^K \left( \frac{S_{iC_k}}{S_i} \right)^2 \quad (2.16)$$

onde  $S_{iC_k}$  representa o número de arestas conectadas do vértice  $i$  para outros vértices na comunidade  $C_k$ . A média do módulo é dada pela média do valor de  $P_i$ .

Além disso, foi realizado uma Análise de Variância (*ANOVA*), que é uma análise que permite verificar se há diferenças entre as medianas (ou médias) de grupos diferentes, e testes pareados de Tukey, para comparar os valores dos conceitos das análises topológicas dos módulos, sendo Intra-Modular, Inter-Modular, conectividade e modularidade entre os métodos.

# Modelo do Competição de Partículas Paralela

As estratégias e o plano de desenvolvimento do método proposto serão descritas abaixo.

## 3.1 Considerações Iniciais

A Tabela 3 apresenta as definições dos termos das equações que serão descritas da metodologia do procedimento do método proposto.

Tabela 3 – Termos das equações do Método Proposto.

<b>Termos</b>	<b>Descrição dos Termos das Fórmulas</b>
$i$ ou $j$	Vértices na rede
$k$	Partícula na rede
$ V $	Número total de vértices na rede
$E$	Matriz Energética
$E_{min}$	Energia Mínima Estabelecida
$B$	Matriz que estabelece os valores para o cálculo da transição da matriz $P_{ale}$
$P_{ale}$	Matriz que calcula a transição aleatória da partícula
$P_{pref}$	Matriz que calcula a transição preferencial da partícula
$P$	Matriz que calcula a transição da partícula da rede, levando e conta as matrizes $P_{ale}$ e $P_{pref}$
$\Delta$	Energia propagada para o vértice visitado
$\alpha$	Energia que é propagada para os vértices vizinhos, do vértice visitado sendo metade do valor de $\Delta$
$\Upsilon$	Informa se algum vértice é vizinho do outro
$\Xi$	Informa se alguma partícula está visitando algum vértice

## 3.2 Metodologia do Procedimento

Dada uma rede  $G = \langle V, F \rangle$  com  $|V|$  vértices e  $|F|$  arestas,  $K$  partículas  $K_1, K_2, \dots, K_n$  são colocadas na rede de acordo com a centralidade de proximidade dos vértices, ou seja, cada uma das partículas  $K$  é colocada em um dos vértices  $V$  com os maiores valores de centralidade de proximidade da rede. A centralidade de proximidade (FREEMAN, 1978) de um vértice  $i$  é definida como o inverso da soma das distâncias do vértice  $i$  a todos os outros vértices  $j$  da rede, conforme mostrado pela Eq. 3.1, onde  $|V|$  é número de vértices na rede, e  $d(i, j)$  é a distância do caminho mais curto entre os vértices  $i$  e  $j$ . Portanto, se a centralidade de proximidade de um vértice for maior, significa que a propagação do sinal pode chegar a outros vértices a partir do vértice atual em menos etapas.

$$C_p(i) = \frac{|V| - 1}{\sum_{j \in |V|} d(i, j)} \quad (3.1)$$

O modelo possui uma matriz energética  $E$ , que indica o nível de dominação de cada partícula em cada vértice. Para começar o processo, todas as partículas são iniciadas com valor de energia no nível mínimo, conforme mostrado na Eq. 3.2.

$$E_{k,i}(0) = E_{min} \quad (3.2)$$

onde  $E$  é a matriz energética e  $E_{k,i}(0)$  é o nível de energia da partícula  $k$ ,  $k = 1, 2, \dots, K$  no vértice  $i$ , com  $i = 1, 2, \dots, |V|$ , na iteração 0, e  $E_{min}$  é o nível mínimo de energia permitido pelo modelo.

Para a caminhada da partícula na rede, usa-se a combinação de duas matrizes, caminhadas aleatória e preferencial. Para a caminhada aleatória, um vértice que ainda esteja com o limite igual de energia ao de  $E_{min}$ , é aleatoriamente selecionado para ser visitado pela partícula  $k$  na próxima iteração. Com isso, é garantido, que todos os vértices recebam modificação na sua matriz de energia. As Eqs 3.3 e 3.4 resumem o procedimento.

$$B_{k,i}(t) = \begin{cases} 1, & \text{se } E_{q,i}(t) = E_{min} \text{ para todo } q = 1, 2, \dots, Q \text{ e } q \neq k \\ 0, & \text{caso contrário} \end{cases} \quad (3.3)$$

onde  $i$  é o vértice e  $Q$  é o número total de partículas na rede. Para calcular a probabilidade é a Eq. 3.4.

$$P_{ale}^{(k,i,j)} = \frac{B_{k,i}}{\sum_{i=1}^{|V|} B_{k,i}} \quad (3.4)$$

$P_{ale}^{(k,i,j)}$  é a probabilidade de que a partícula  $k$  selecione o nó  $i$ , na iteração  $t$  definida pela Eq. 3.4, para visitar a partir do nível de energia  $E_{min}$ . Enquanto isso, a caminhada

preferencial, é a probabilidade de a partícula viajar para um vértice que tenha alto nível de sua energia. Dada pela Eq. 3.5.

$$P_{pref}^{(k,i,j)}(t) = \frac{E_{k,i}(t)}{\sum_{i=1}^{|V|} E_{k,i}(t)} \quad (3.5)$$

onde  $P_{pref}^{(k,i,j)}(t)$  é a probabilidade da partícula  $k$  selecionar o vértice  $i$  para visitar a partir do vértice  $j$  na iteração  $t$ . Vemos que é mais provável que uma partícula visite um vértice, cujo nível de dominação (de energia) é alto nesse vértice. Portanto, a Eq. 3.5 define um papel de andar preferencial em contraste ao andar aleatório. Para a caminhada da partícula, é dada uma probabilidade  $P^{(k,i,j)}(t)$ , na qual suponha que a partícula  $k$  esteja no vértice  $i$  na iteração  $t - 1$ , ela selecionará um vértice vizinho  $j$ , de  $i$ , para visitar na iteração  $t$  conforme a Eq. 3.6, que é a combinação das duas caminhadas anteriores:

$$P^{(k,i,j)}(t) = \lambda P_{ale}^{(k,i,j)} + (1 - \lambda) P_{pref}^{(k,i,j)}(t) \quad (3.6)$$

onde  $P_{ale}$  é a matriz de probabilidade de visitar um vértice que ainda não tenha recebido o sinal de propagação. O  $P_{pref}$  é a matriz de probabilidade de a partícula visitar um vértice que tenha um alto nível de dominação, garantindo que a determinada partícula  $k$  seja a dona de determinada região na rede. E conforme os modelos anteriores da competição de partícula (SILVA; ZHAO, 2016; SILVA; ZHAO, 2012b; SILVA; ZHAO, 2012a; QUILES et al., 2008; GAO et al., 2019; MARTINS; ZHAO, 2020), também temos um padrão de equilíbrio  $\lambda$ , onde  $\lambda \in [0, 1]$ .

Depois que uma partícula  $k$  seleciona um vértice de destino  $i$  para visitar, introduzimos duas funções para atualizar a matriz energética, definidas pela Eq. 3.7 e Eq. 3.8. O primeiro papel é um processo sequencial, como apareceu no modelo original de competição de partículas, ou seja, quando uma partícula  $k$  visita um vértice  $i$ , ela aumenta sua própria energia em  $i$  e diminui a energia de todas as outras partículas no mesmo vértice. O segundo papel, inspirado em SOM, propaga sua energia para todos os vértices vizinhos de  $i$ .

$$E_{k,i}(t+1) = \begin{cases} E_{k,i}(t) + \Delta, & \text{se } \Xi(k,i)(t) = 1 \\ E_{q,i}(t) - \Delta, & \text{se } \Xi(q,i)(t) = 0 \end{cases} \quad (3.7)$$

Na equação acima, novamente  $E$  é a matriz energética. A energia da partícula  $k$  sobre o vértice  $i$  é aumentada em um valor  $\Delta$  se a partícula  $k$  estiver visitando  $i$  na iteração  $t$ . Em simultâneo, as energias de outras partículas (rivais)  $q$  depositadas no vértice  $i$  são diminuídas. A função  $\Xi(k,i)(t)$  informa se a partícula  $k$  está visitando o vértice  $i$ , com valor 1, caso não o valor será 0, na iteração  $t$ .

$$E_{k,j}(t+1) = \begin{cases} E_{k,j}(t) + \alpha, & \text{se } \Upsilon(i,j) = 1 \wedge \Xi(k,i)(t) = 1 \\ E_{q,j}(t) - \alpha, & \text{se } \Upsilon(i,j) = 1 \wedge \Xi(q,i)(t) = 0 \end{cases} \quad (3.8)$$

Na Eq. 3.8,  $i$  é o vértice visitado pela partícula  $k$  na iteração  $t$  e  $j$  é um vértice vizinho de  $i$ , com  $\alpha \in [0, 1]$ . Se o vértice  $j$  é vizinho de  $i$ , a função  $\Upsilon(i, j)$  retorna 1; caso contrário, retorna 0. Eq. 3.8 propaga a energia de uma quantidade  $\alpha$  (sendo metade do valor de energia de  $\Delta$ ) simultaneamente para todos os vizinhos  $j$  do vértice  $i$  visitado. No caso de redes ponderadas, consideramos apenas o peso da aresta multiplicado por  $\alpha$  como a quantidade de propagação.

A cada iteração, também precisamos verificar se um elemento da matriz energética é reduzido ao nível abaixo de  $E_{min}$ . Se isso acontecer, o nível de energia retorna ao seu valor inicial, ou seja,

$$E_{k,i}(t) = E_{min}, \quad \text{se } E_{k,i}(t) < E_{min} \quad (3.9)$$

A cada iteração, cada partícula seleciona um vértice vizinho para visitar segundo a probabilidade de caminhada. Então, a partícula propaga sua energia de duas maneiras: 1) apenas para o vértice visitado constituindo um processamento sequencial e 2) para todos os vizinhos do vértice visitado para realizar um processamento paralelo. Em ambos os casos, o nível de energia da partícula atual nos vértices é aumentado e o nível de energia das partículas rivais é diminuído. Este processo é repetido até que o sistema convirja. Finalmente, as comunidades são determinadas pelo nível de energia (nível de dominação) de cada partícula em cada vértice, ou seja, cada partícula domina uma comunidade. O algoritmo apresentado na seção 3.4 resume as etapas para iterar o sistema.

### 3.3 Descrição do Processo

O modelo proposto é inspirado em duas técnicas de aprendizagem não supervisionadas. A primeira é a competição de partículas, na qual as partículas viajam pela rede tentando dominar o maior número de vértices possível, enquanto repelem as partículas opostas, descrito em (SILVA; ZHAO, 2016; SILVA; ZHAO, 2012b; SILVA; ZHAO, 2012a; QUILES et al., 2008; GAO et al., 2019; MARTINS; ZHAO, 2020). O segundo mecanismo é a propagação de sinal paralelo inspirada em SOM (KOHONEN, 1990), também conhecido como Mapa de Kohonen ou Mapas Auto Organizáveis, ou seja, são as etapas de Cooperação e Adaptação Sináptica conforme abordado na subseção 2.2.2.

No modelo original de competição de partículas, não é possível detectar comunidades desbalanceadas (comunidades têm tamanhos ou densidades de conexão diferentes), visto

que todas as partículas seguem as mesmas regras de movimento. Por outro lado, a propagação paralela introduzida no modelo proposto, permite que cada partícula capture a estrutura local dos vértices visitados e gere um comportamento particular para cada partícula, adaptando-se assim a estrutura da rede. Como consequência, o modelo proposto pode ser capaz de detectar comunidades desbalanceadas.

Outra vantagem do modelo proposto é sua alta eficiência de processamento em comparação ao modelo original de competição de partículas. Especificamente, no novo modelo, as partículas não precisam andar pela rede devido à introdução da propagação paralela, elas só precisam andar em algumas regiões emitindo seu sinal de propagação para os vértices próximos. Então, quem anda na rede é a energia das partículas para dominar a região.

Dessa forma, a inspiração nas etapas de cooperação e adaptação sináptica dos mapas de Kohonen (KOHONEN, 1990), proporcionará esse mecanismo de propagação paralela. Com isso, é emitido um sinal de energia para a vizinhança dos vértices na rede (etapa de cooperação dos Mapas Auto Organizáveis (KOHONEN, 1990)), fazendo-os tornar, ou ficar parecidos, com o vértice que recebeu a visita da partícula e emitiu a energia. Após essa propagação os vértices ajustam os níveis de energia que contém de cada partícula (etapa de adaptação sináptica dos Mapas Auto Organizáveis (KOHONEN, 1990)). Então as partículas que conter mais energia nos vértices indicam que esses vértices são mais parecidos e conseqüentemente podem estar na mesma comunidade, a Fig. 8 ilustra o processo de propagação inspirados nos mapas auto-organizáveis. A nova metodologia funcionará da seguinte forma:

1. Todos os elementos da matriz energética são inicializados com o valor mínimo estabelecido;
2. As posições iniciais das partículas (número estabelecido) são definidas nos vértices com a maior medida de centralidade de proximidade;
3. A cada iteração, as partículas propagam sua energia de duas maneiras:
  - Apenas para o vértice visitado constituindo um processamento sequencial;
  - Para todos os vizinhos do vértice visitado para realizar um processamento paralelo. Em ambos os casos, o nível de energia da partícula atual nos vértices é aumentado e o nível de energia das partículas rivais é diminuído;
4. Verifica-se cada elemento da matriz energética. Se algum deles tiver um valor de energia abaixo do mínimo estabelecido, é definido de volta para esse valor;
5. Após, cada partícula vai escolher um novo vértice para visitar, ou seja, um novo destino, por meio da probabilidade de caminhada  $P^{(k,i,j)}(t)$ , dada pela Eq. 3.6;

6. Este processo é repetido até que o sistema convirja ou repete até o limiar de repetições estabelecido.

Finalmente, as comunidades são determinadas pelo nível de energia (nível de dominação) de cada partícula em cada vértice, ou seja, cada partícula domina uma comunidade. A Fig. 8, ilustra a descrição do processo de detecção do método. Também ilustramos todo o processo de propagação de energia (nível de dominação) do modelo proposto por uma rede clássica: uma rede agrupada aleatoriamente com  $|V| = 60$  vértices e 4 comunidades (Fig. 9).

### 3.4 Algoritmo Proposto

Tabela 4 – Algoritmo Proposto

<b>Algoritmo de Competição Paralela</b>	
01:	<b>procedimento</b> PARALLELCOMPETITION( $G, K, T, weight$ )
02:	$E(0) = \text{inicia\_Energia}(G, K)$ : pela Eq. 3.2
03:	$\text{inicia\_Particula}()$ : Seta as posições iniciais das $K$ partículas
04:	$t = 1$
05:	<b>enquanto</b> $t < T$ ou $\ E(t) - E(t - 1)\  \leq \epsilon$
06:	<b>para</b> $k = 1$ to $K$ <b>faça</b>
07:	$\text{propaga\_Sinal}(G, E(k, i))$ pelas Eqs. 3.7 e 3.8
08:	$\text{checa\_Energia}(G, E(k, i))$ pela Eq. 3.9
09:	$P_{ale} = \text{calculaAleatória}(K)$ pelas Eqs. 3.3 e 3.4
10:	$P_{pref} = \text{calculaPreferencial}(K)$ pela Eq. 3.5
11:	$\text{seleciona\_No}(G, K, P_{ale}, P_{pref})$ pela Eq. 3.6
12:	<b>fim-para</b> $k$
13:	$t = t + 1$
14:	<b>fim-enquanto</b>
15:	<b>fim-procedimento</b>

O procedimento de iteração do algoritmo é explicado por meio das etapas descritas a seguir.

1. Passo 01: Entre com os parâmetros de entrada. No caso de redes e/ou grafos ponderados, defina  $weight = TRUE$ ; caso contrário,  $weight = FALSE$ .
2. Passo 02: Todos os elementos da matriz energética são inicializados com o valor  $E_{min}$ .
3. Passo 03: As posições iniciais das partículas  $K$  são definidas nos vértices  $V$  com a maior medida de centralidade de proximidade.
4. Passo 04: Inicializa a iteração 1 do procedimento.



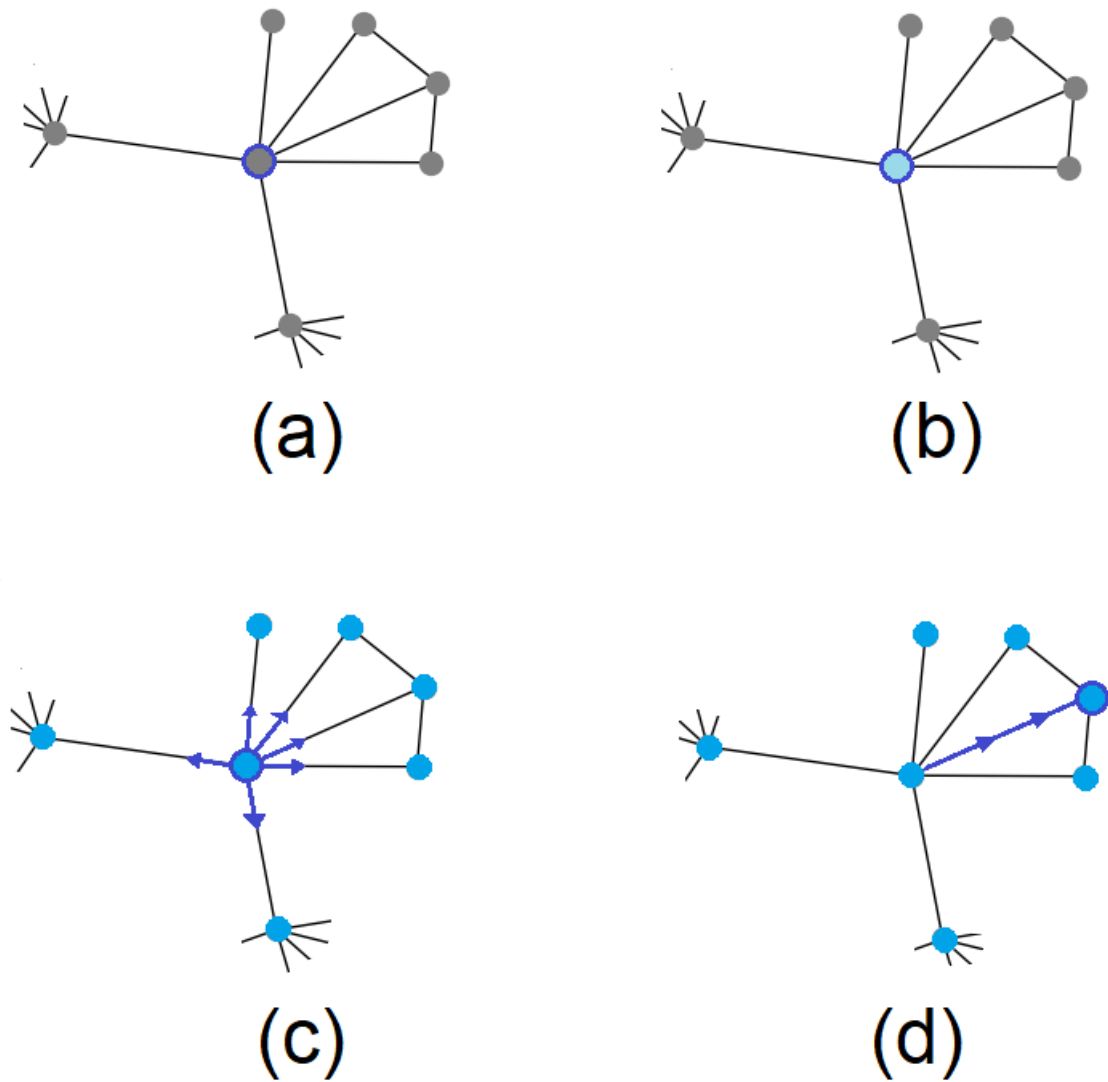


Figura 8 – Ilustra a descrição do processo do método proposto. (a) - Posição inicial da Partícula em um vértice com maior medida de centralidade de proximidade. (b) - Propagação de energia no vértice visitado. (c) - Propagação de energia para os vizinhos do vértice visitado. (d) - Escolha um novo vértice para visitar e fazer o processo novamente. Imagem feita pelo autor.

5. Passo 05: Executa o processo até um valor limite predefinido ou checa o padrão de convergência apresentado em (SILVA; ZHAO, 2012b).
6. Passo 06: Para cada partícula execute o processo.
7. Passo 07: As  $K$  partículas propagam o sinal e atualizam a matriz energética.
8. Passo 08: Verifica cada elemento da matriz energética. Se algum deles tiver valor de energia abaixo de  $E_{min}$ , defina novamente para  $E_{min}$ .

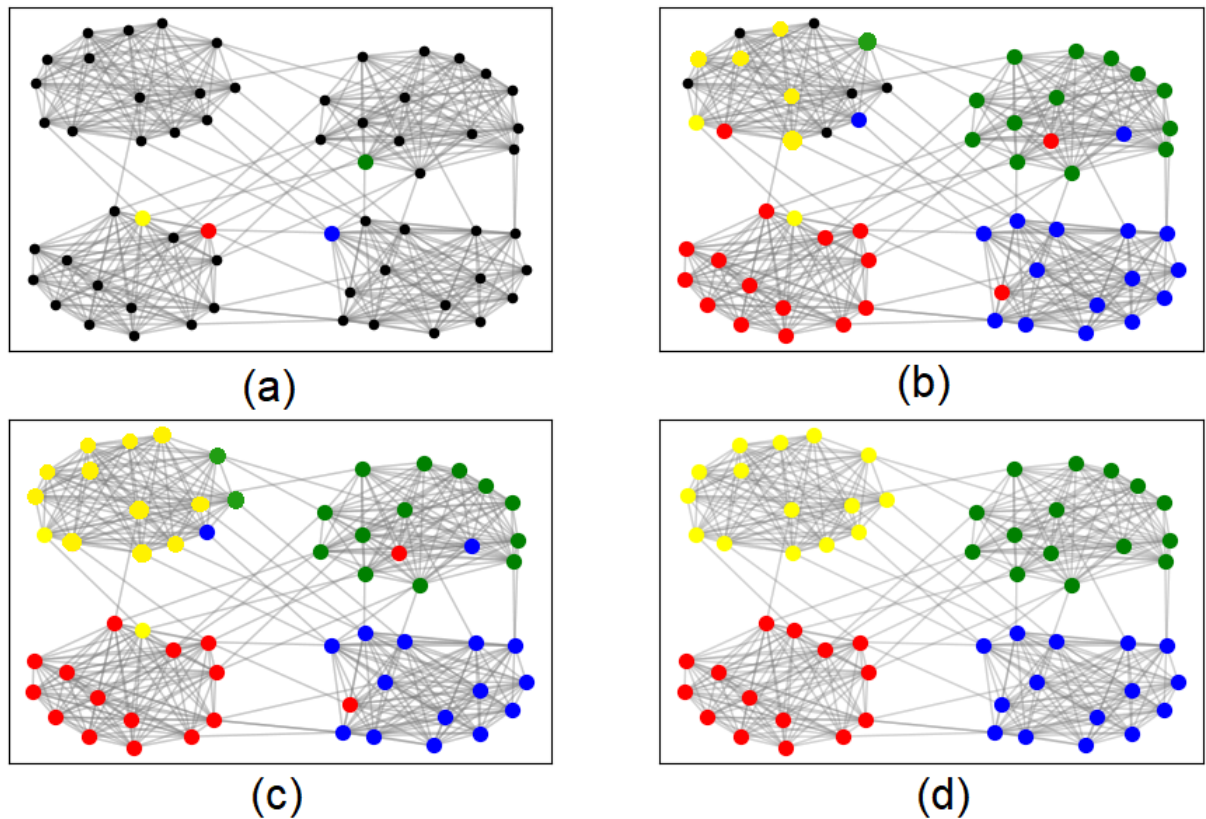


Figura 9 – Ilustrando o processo de detecção de comunidades com a técnica proposta. A construção da rede teve  $|V| = 60$ ,  $M = 4$  e 15 vértices para cada comunidade. Cada cor corresponde a uma partícula. Se um vértice for marcado com uma cor específica, significa que a partícula correspondente tem o nível de dominação mais alto no vértice na iteração atual. (a) Configuração inicial.  $t = 0$ . (b)  $t = 1$ . (c)  $t = 10$ . (d)  $t = 100$ . Rede gerada pelo autor através do módulo *NetworkX* na linguagem de programação *Python*.

9. Passo 09: Cálculo da caminhada aleatória para as partículas.
10. Passo 10: Cálculo da caminhada preferencial para as partículas.
11. Passo 11: A partícula escolhe um novo vértice para visitar, com base na combinação das matrizes  $P_{ale}$  e  $P_{pref}$ , resultando na matriz  $P(t)$ . Com relação à matriz gerada,  $P(t)$ , a partícula seleciona qual será o seu próximo destino.
12. Passo 12: Fim da repetição para cada partícula.
13. Passo 13: Atualiza o tempo de iteração.
14. Passo 14: Fim repetição.
15. Passo 15: Fim do procedimento.

## 3.5 Condições do Sistema

### 3.5.1 Critérios de Inicialização

Para iniciar a execução do algoritmo, é necessário saber onde as partículas vão começar. Para isso, calcula-se a centralidade de proximidade, conforme aborda na seção 3.2 na Eq. 3.1. Com isso, mesmo que se tenha diversas execuções do modelo, as partículas irão sempre iniciar nos mesmos vértices. Também é necessário atribuir o valor de energia inicial aos valores da matriz energética, que também foi abordado na seção 3.2 na Eq 3.2.

### 3.5.2 Critérios de Finalização

Para finalizar a execução do algoritmo, pode executar até um limite de execuções que são passadas no parâmetro  $T$ . Também pode ser por um critério de convergência proposto por (SILVA; ZHAO, 2012b), que checa se a maior mudança nos níveis de dominação (da matriz energética) está abaixo de um certo limite  $\epsilon$ , como apresentado no algoritmo da seção anterior em 3.4. No algoritmo proposto de competição de partículas paralelas, foi usado como critério de finalização a execução de um limite de iterações pelo parâmetro  $T$ .

## 3.6 Complexidade Computacional

A complexidade computacional do algoritmo pode ser facilmente determinada. A etapa 07 envolve a ordem de cálculo  $O(\langle k \rangle)$ , onde  $\langle k \rangle$  é o grau médio da rede. Isso ocorre, porque a propagação é feita para apenas um vértice vizinho pelo processamento sequencial e para todos os vizinhos ( $\langle k \rangle$  em média) pelo processamento paralelo; o passo 08 é apenas  $O(1)$ ; as etapas 09 e 10 são  $O(|V|)$  porque só precisamos atualizar os elementos  $|V|$  para obter novos  $P_{ale}$  e  $P_{pref}$  para cada partícula e cada vértice; e a etapa 11 é novamente  $O(\langle k \rangle)$ , porque selecionamos um vértice de todos os seus vizinhos. Portanto, a ordem de complexidade do algoritmo é  $O(TK|V| \langle k \rangle)$ , onde  $T$  é o número máximo de iterações e  $|V|$  é o número de vértices na rede. Como o número de partículas é muito menor que o número de vértices, ou seja,  $K \ll |V|$ , a ordem de complexidade é geralmente  $O(T|V| \langle k \rangle)$ .

Em redes esparsas, ou seja,  $\langle k \rangle \ll |V|$ , ele se reduz a  $O(T|V|)$ . Devido à propagação paralela, um número muito menor de iterações é necessário para terminar o processo de dominação das partículas, especialmente em grandes redes. Portanto, a ordem de complexidade pode até mesmo atingir  $O(|V|)$  (linear) em redes grandes e esparsas.

### 3.7 Parâmetros do Algoritmo

Tabela 5 – Descrição dos parâmetros iniciais do Algoritmo.

Parâmetros	Descrição dos parâmetros do Algoritmo
$K$	Número de comunidades a ser identificadas pelo procedimento.
$E_{min}$	Energia mínima a ser iniciada na matriz energética do processo.
$\Delta$	Nível de energia que é incrementado no vértice da rede, pela partícula que visita.
$\alpha$	Nível de energia que a partícula propaga para os outros vértices vizinhos, ao vértice que está visitando.
$\lambda$	Esse valor é variado de 0 à 1. Define a ênfase da caminhada da partícula. Caso ela vá fazer uma caminhada aleatória ou preferencial.
$T$	O número mínimo de iterações que deseja realizar durante a execução do processo de identificar as comunidades na rede.
<i>weight</i>	Valor booleano, <i>True</i> ou <i>False</i> , que determina a ponderação das arestas na rede. Com isso quando o sinal $\alpha$ for propagado, ele será multiplicado pelo peso da conexão entre os vértices, conforme foi explicado na seção 3.2 na Eq. 3.8.

---

## Análise Empírica do Modelo

Os resultados obtidos das análises do modelo proposto, serão descritos nessa seção.

### 4.1 Considerações Iniciais

A Tabela 6 apresenta os valores dos parâmetros usados nas simulações aplicando a técnica proposta. Usamos o valor  $\lambda$  com 0.4, como demonstrando na Eq. 3.6. Assim, quando a partícula andar pela rede, ela não ficará viciada em um método de caminhada, e também pode dar saltos com a taxa de aprendizado.  $E_{min}$  foi estabelecida em zero, para melhor aprendizagem da matriz energética, dessa forma, por estar tudo zerado, as partículas não vão ter influência de energia, e terão que competir entre si. Enquanto ao número de iterações estabelecemos um valor padrão (mínimo), com 10000 iterações.

Tabela 6 – Parâmetros usados pelo Método Proposto.

Parâmetros	Valor
$\lambda$	0.4
$E_{min}$	0
$T$	$\leq 10000$

A Fig. 10 (a) mostra os resultados da detecção de comunidade em uma rede artificial com 4 comunidades balanceadas, enquanto a Fig. 10 (b) mostra o resultado em outra rede artificial com duas comunidades desbalanceadas. Em ambos os casos, a técnica proposta pode detectar corretamente as comunidades.

### 4.2 Propagação de Energia

Conforme já discutido, o método proposto busca encontrar comunidades com sinais de propagação de energia das partículas que são aumentadas e/ou diminuídas em relação aos

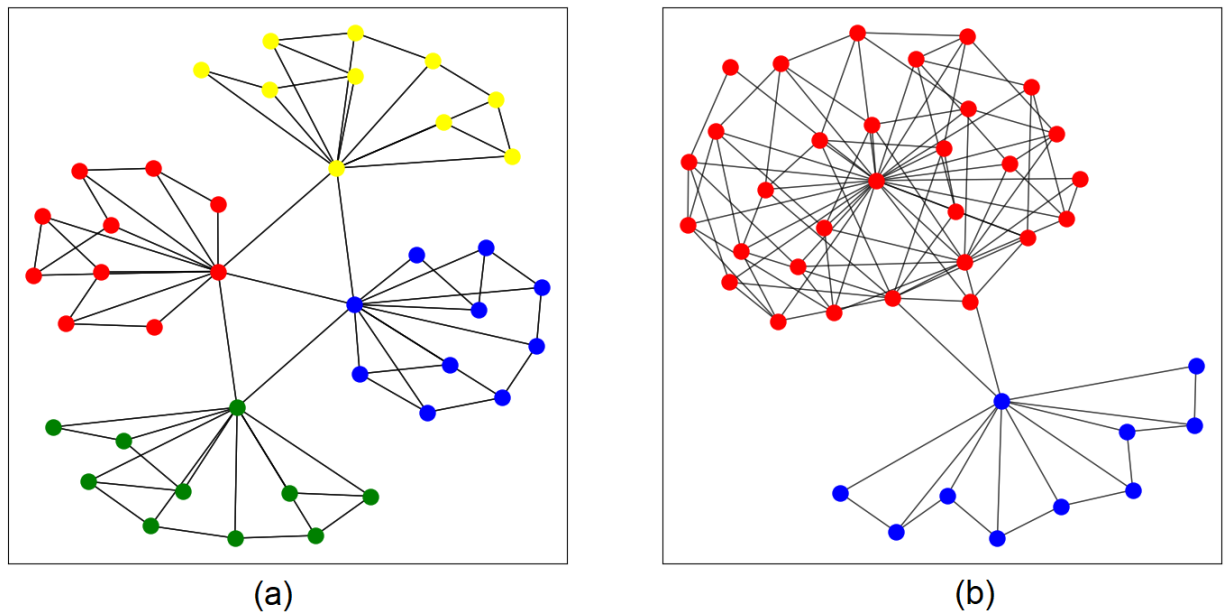


Figura 10 – Resultados de detecção de comunidade obtidos em redes artificiais. (a) Uma rede com 4 comunidades equilibradas. (b) Uma rede com 2 comunidades desbalanceadas. Redes geradas pelo autor através do módulo *NetworkX* na linguagem de programação *Python*.

vértices. Portanto, uma análise do nível de energia de cada partícula em cada comunidade ao longo do tempo é realizada e os resultados são mostrados na Fig. 11, para rede da Fig. 10 (a). Vemos, nesta figura, que cada partícula domina uma comunidade após algumas iterações, ou seja, uma partícula tem alto nível de energia sobre uma comunidade, enquanto outras partículas têm baixos níveis de energia na mesma comunidade. Nesta figura, os níveis de energia são normalizados em  $[-1, 1]$  pela função Tangente Hiperbólica. Também analisamos a propagação de energia em redes com comunidades desbalanceadas, como mostrado na Fig. 12, para rede da Fig. 10 (b).

### 4.3 Análise da Complexidade do Algoritmo

Conforme discutido, o método proposto consegue ser mais rápido na iteração e na velocidade de detecção das comunidades na rede, em relação aos métodos desenvolvidos anteriormente (SILVA; ZHAO, 2016; SILVA; ZHAO, 2012b; SILVA; ZHAO, 2012a; QUILES et al., 2008; GAO et al., 2019; MARTINS; ZHAO, 2020). Na Fig. 13, apresenta-se o resultado da iteração do processo, ou seja, quanto tempo levou para executar todo o processo em ambas as técnicas para as 1000 iterações. A Fig. 14, ilustra o resultado de quanto tempo levou para identificar comunidades de rede. Conforme foi feito em (SILVA; ZHAO, 2012b), foi comparado com duas técnicas de detecção de comunidades da metodologia de

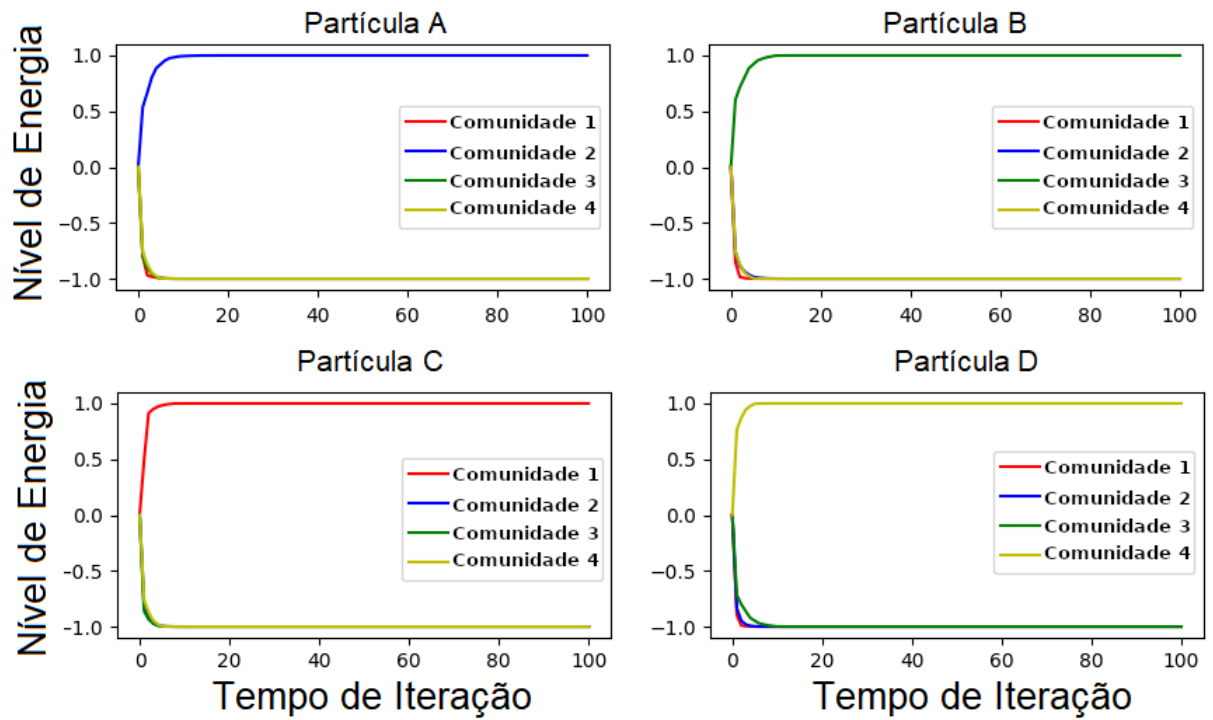


Figura 11 – Nível médio de energia ao longo do tempo para a rede da Fig. 10 (a)

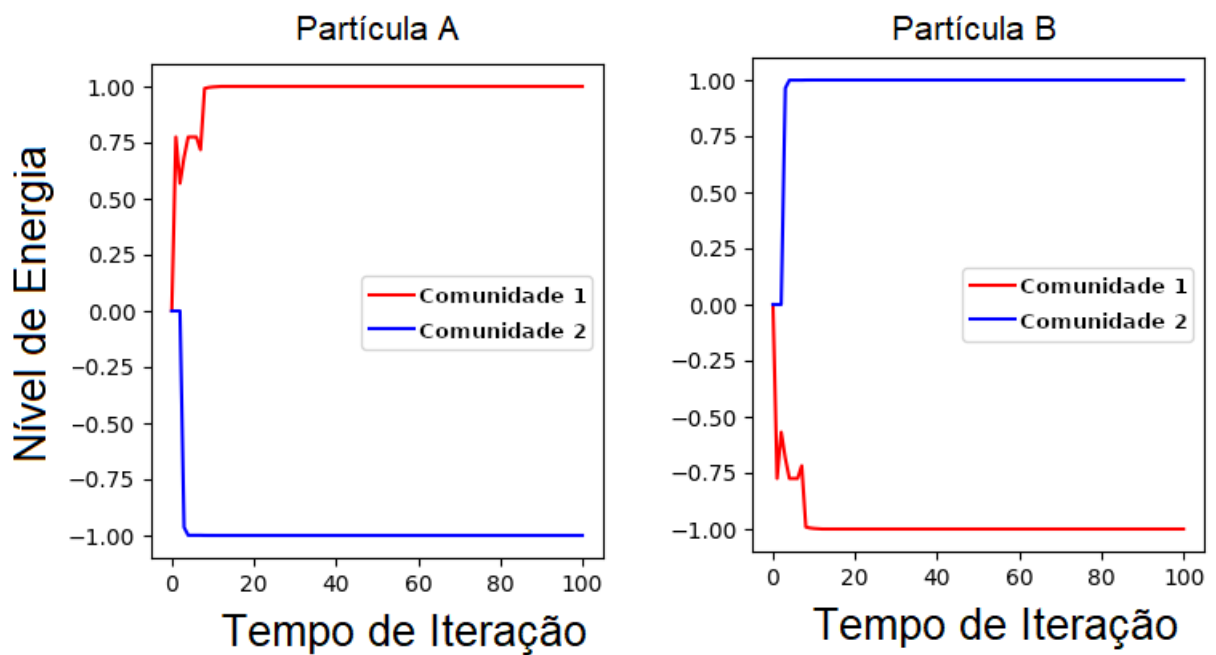


Figura 12 – Nível médio de energia ao longo do tempo para a rede da Fig. 10 (b)

modularidade, *Louvain* (BLONDEL et al., 2008) e *Fast Greedy* (CLAUSET; NEWMAN; MOORE, 2004). O método proposto mostrou-se mais rápido do que o método *Fast Greedy* (CLAUSET; NEWMAN; MOORE, 2004) para identificação das comunidades, contudo foi mais lento do que a modularidade de *Louvain* (BLONDEL et al., 2008).

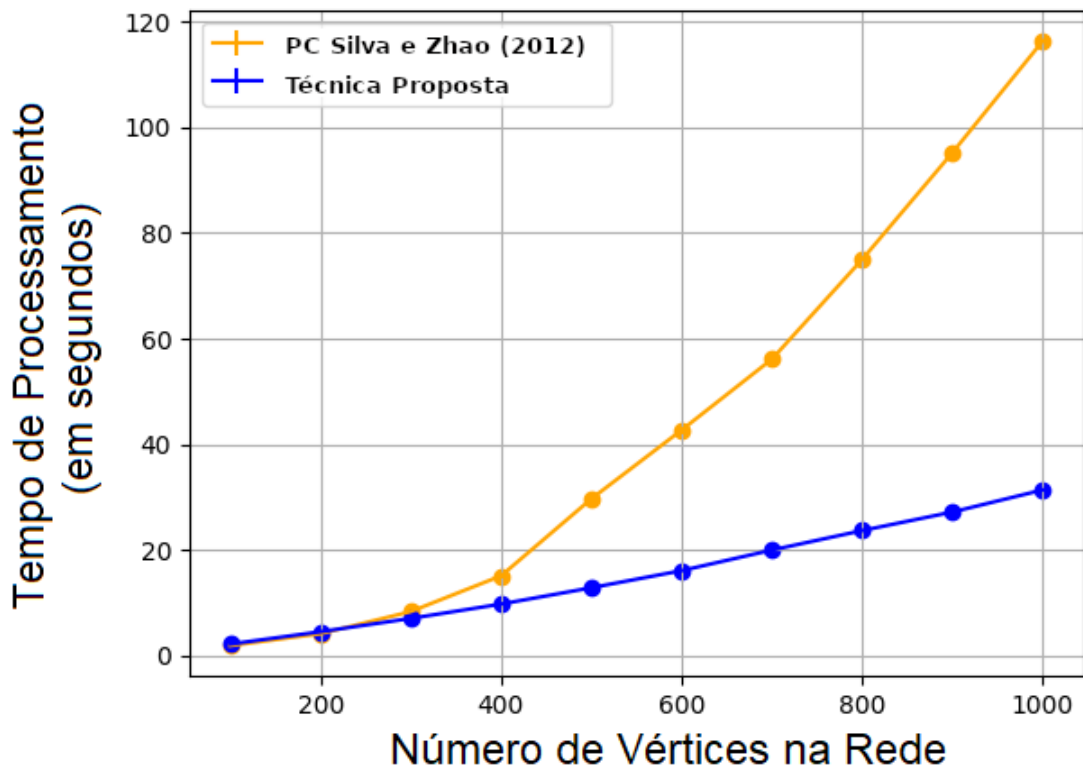


Figura 13 – Resultado obtido para a complexidade do método proposto em relação à competição de partículas implementada por (SILVA; ZHAO, 2012b), por tempo de iteração. Cada ponto no traço tem uma média de 30 de execuções.

## 4.4 Impacto do Parâmetro do Sinal de Propagação $\alpha$

O Impacto do Parâmetro do Sinal de Propagação  $\alpha$ , é a forma que o algoritmo encontra para identificar as comunidades. Então, avaliamos o seu impacto sendo menor, igual e maior que o sinal  $\Delta$  em três categorias de redes. Veja a Fig. 15, todas com padrão de mistura médio. Nota-se que o impacto do parâmetro de energia não interferiu para encontrar as comunidades.

## 4.5 Redes de Referências

A Fig. 16 mostra as precisões de detecção de comunidades em relação aos níveis de mistura nas redes de Benchmark Girvan e Newman (NEWMAN; GIRVAN, 2004; SILVA; ZHAO, 2012b; FORTUNATO, 2010), utilizando a técnica proposta. Para cada ponto da trajetória, 100 iterações independentes foram realizadas, e pegamos a média delas.



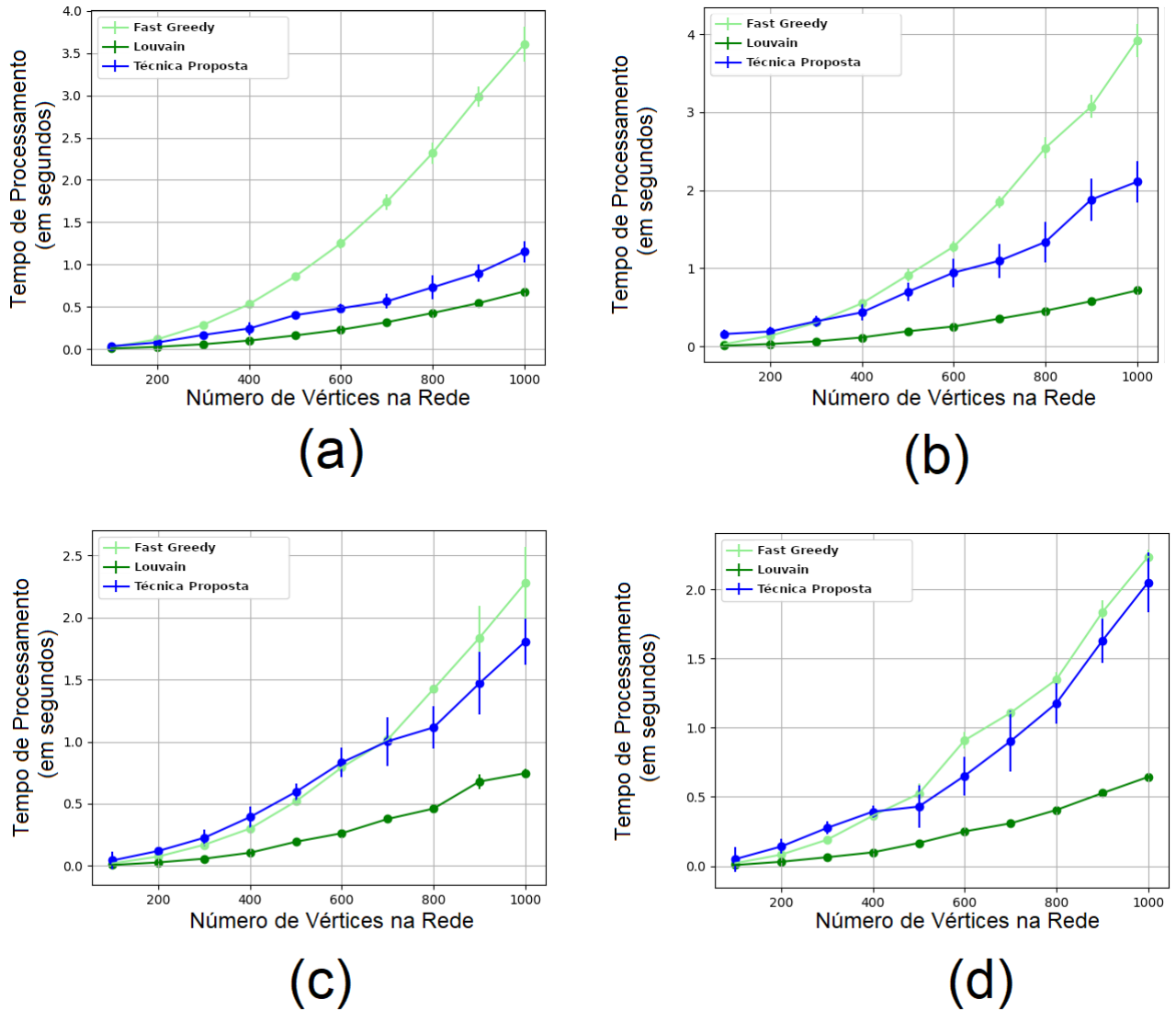


Figura 14 – Resultado obtido para a complexidade do método proposto em relação às modularidades. (a) - Rede com  $P_{in} = 16$ ,  $P_{out} = 0.1$ ,  $M = 4$ . (b) - Rede com  $P_{in} = 16$ ,  $P_{out} = 0.3$ ,  $M = 4$ . (c) - Rede com  $P_{in} = 0.15$ ,  $P_{out} = 0.1$ ,  $M = 4$ . (d) - Rede com  $P_{in} = 0.15$ ,  $P_{out} = 0.3$ ,  $M = 4$ . Cada ponto no traço tem uma média de 30 execuções.

Também analisamos a rede de Benchmark Lancichietti (LANCICHINETTI; FORTUNATO; RADICCHI, 2008; FORTUNATO, 2010) (Fig. 17). Essas redes seguem uma função de lei de potência para sua construção, nas distribuições de grau e também ao tamanho das comunidades (SILVA; ZHAO, 2012b). É utilizado um padrão de mistura  $\mu$ , que faz a conexão das comunidades (SILVA; ZHAO, 2012b). Assim, através das arestas, cada vértice emite  $\mu$  a outro vértice de outra comunidade, e uma fração  $1 - \mu$  para os vértices pertencentes à mesma comunidade (SILVA; ZHAO, 2012b; LANCICHINETTI; FORTUNATO; RADICCHI, 2008; FORTUNATO, 2010). Com isso, analisamos o impacto do método proposto para esse tipo de rede, comparando-o com outros algoritmos da literatura e, dessa forma, o método proposto consegue ser tão eficiente quanto a moduari-

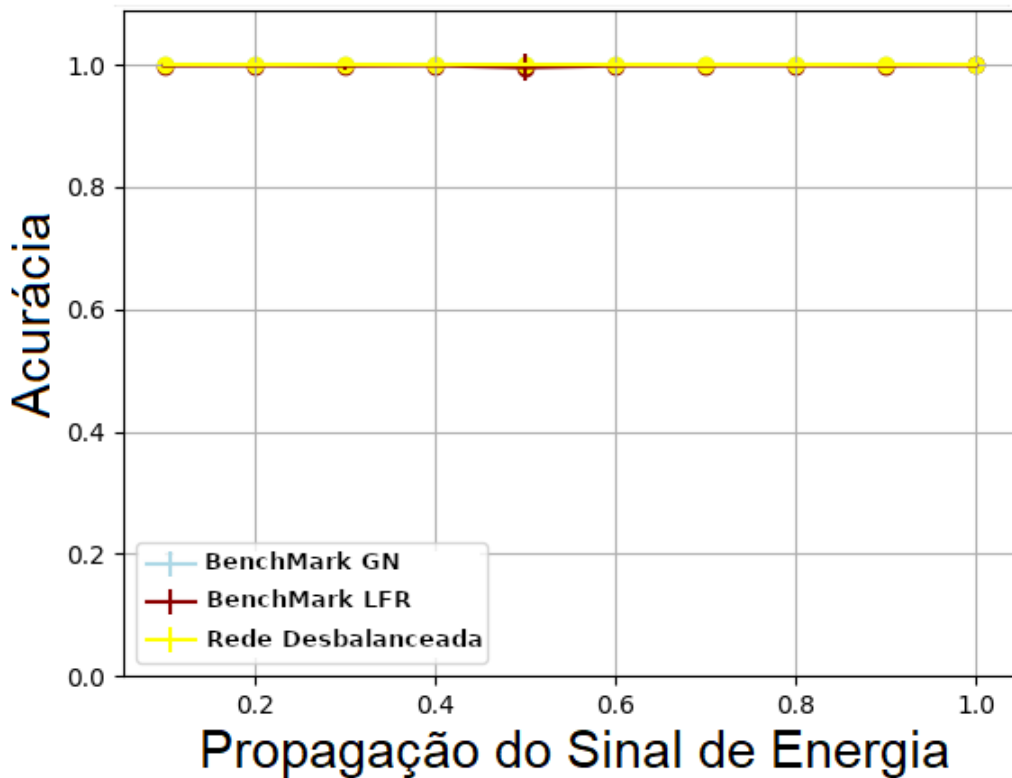


Figura 15 – Avaliando o impacto do sinal de propagação  $\Delta = 0.5$ ,  $\alpha$  variando de 0.1 à 1.0. BenchMark de Girvan e Newman com  $V = 128$ ,  $M = 4$  e  $P_{out} = 0.2$ . BenchMark de Lancichinetti com  $V = 500$ ,  $M = 4$  e  $P_{out} = 0.15$ . Rede com comunidades Desbalanceadas com  $V = 110$ ,  $M = 2$  e  $P_{out} = 0.25$ . Cada ponto no traço tem uma média de 50 execuções.

dade, em outros casos, a acurácia fica próxima. E apresentou melhor desempenho que a metodologia de propagação de rótulos (RAGHAVAN; ALBERT; KUMARA, 2007).

Aplicamos a técnica de rede real conhecida como clube de caratê de Zachary (ZACHARY, 1976), que consiste em 34 vértices divididos em duas comunidades. A Fig. 18 mostra o resultado da detecção de comunidade. As cores em vermelho e azul denotam as comunidades detectadas pelo algoritmo proposto. Foram feitas 100 iterações para essa rede, e foi analisado em qual comunidade os vértices estavam identificados. Alguns algoritmos de detecção de comunidades, classificam os vértices 3 (GIRVAN; NEWMAN, 2002) e 10 (NEWMAN, 2004) incorretamente (SILVA; ZHAO, 2012b).

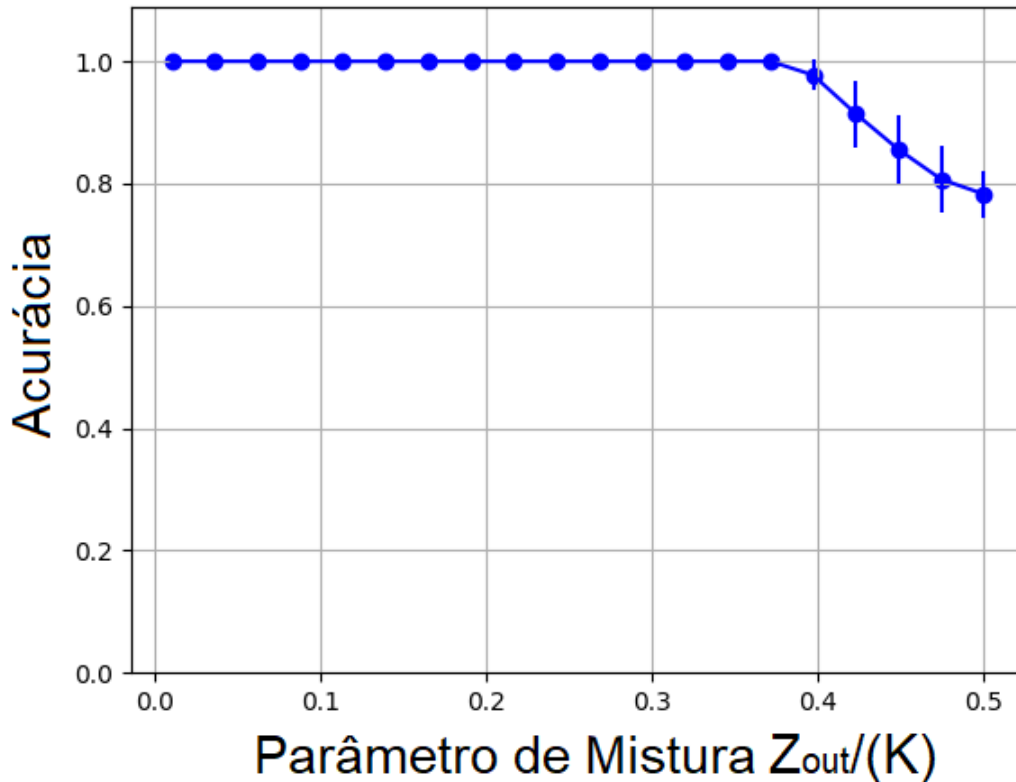


Figura 16 – Resultado obtido para detecção de comunidades usando Benchmark Girvan e Newman. A construção da rede contou com  $V = 128$ ,  $M = 4$ , totalizando 32 vértices para cada comunidade. Cada ponto no traço tem uma média de 100 execuções.

## 4.6 Comunidades Desbalanceadas

Outra tarefa em que a técnica proposta funciona bem é a detecção de comunidades desbalanceadas (comunidades de diferentes tamanhos, densidades ou distribuições de graus). Na Fig. 19, é ilustrado o processo de identificação de duas comunidades desbalanceadas. Observa-se que, a menos que as partículas comecem no mesmo módulo, o procedimento pode expulsar uma partícula para outro módulo na rede. No final, as comunidades foram identificadas corretamente.

Conforme foi feito em (MARTINS; ZHAO, 2020), usamos o índice de NMI (do inglês *Normalized Mutual Information*) ou Informação Mútua Normalizada, para apresentar a taxa de acerto em redes com comunidades desbalanceadas. A Fig. 20 mostra o resultado da avaliação em questão do desbalanceamento. Em cada rede, uma comunidade é fixada com 10 vértices e o número de vértices na segunda comunidade é aumentado de uma rede para a próxima começando em 100, o que conseqüentemente aumenta o desequilíbrio da

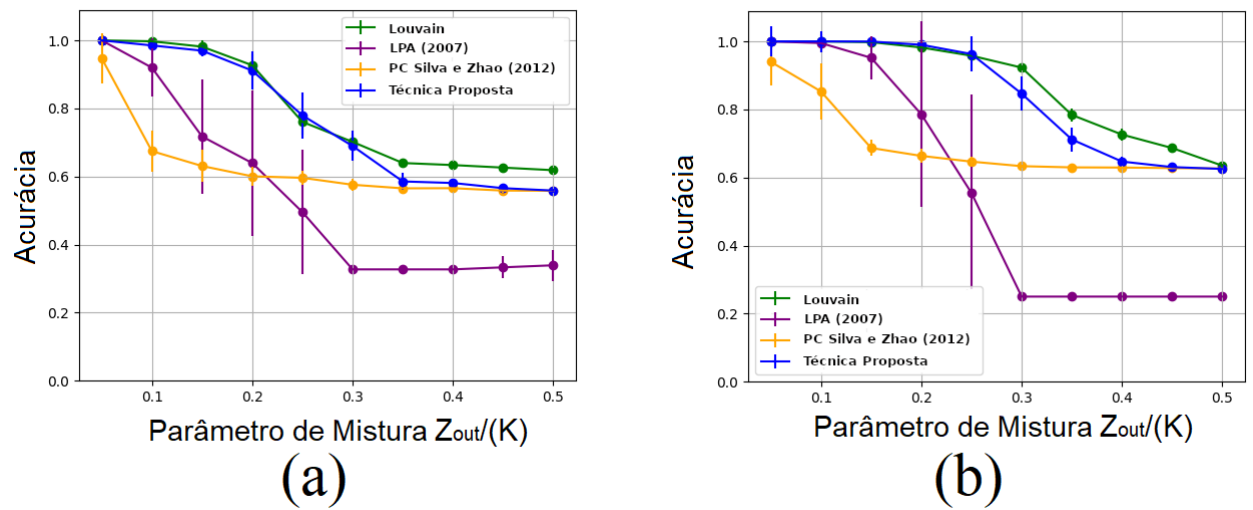


Figura 17 – Resultado obtidos para Benchmark Lancichinetti et al. A construção da rede contou com  $\gamma=3$  e  $\beta = 1.5$ . Cada ponto no traço tem uma média de 30 execuções. (a) - Rede com  $|V| = 100$ ,  $M = 3$ . (b) - Rede com  $|V| = 500$ ,  $M = 4$ .

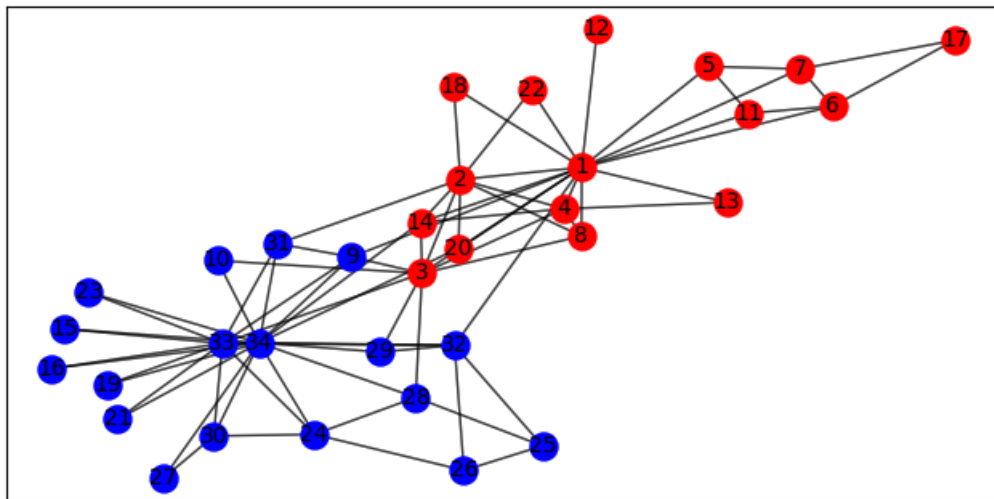


Figura 18 – Resultado de detecção da comunidade na rede do clube de caratê Zachary pelo método proposto. As cores em vermelho e azul denotam às duas comunidades. Média de 100 execuções. Rede gerada através do módulo *NetworkX* na linguagem de programação *Python*.

rede. Vemos que a técnica proposta obtém uma precisão de detecção em 100%.

Ainda na questão do desbalanceamento, avaliamos as densidades da rede. O resultado é mostrado na Fig. 21. Assim, a densidade em ambas as comunidades aumentam. Para baixas densidades, o método proposto não obteve acurácia satisfatória, com uma limitação. Observe que quanto mais densa a rede, melhor será a detecção de suas comunidades.

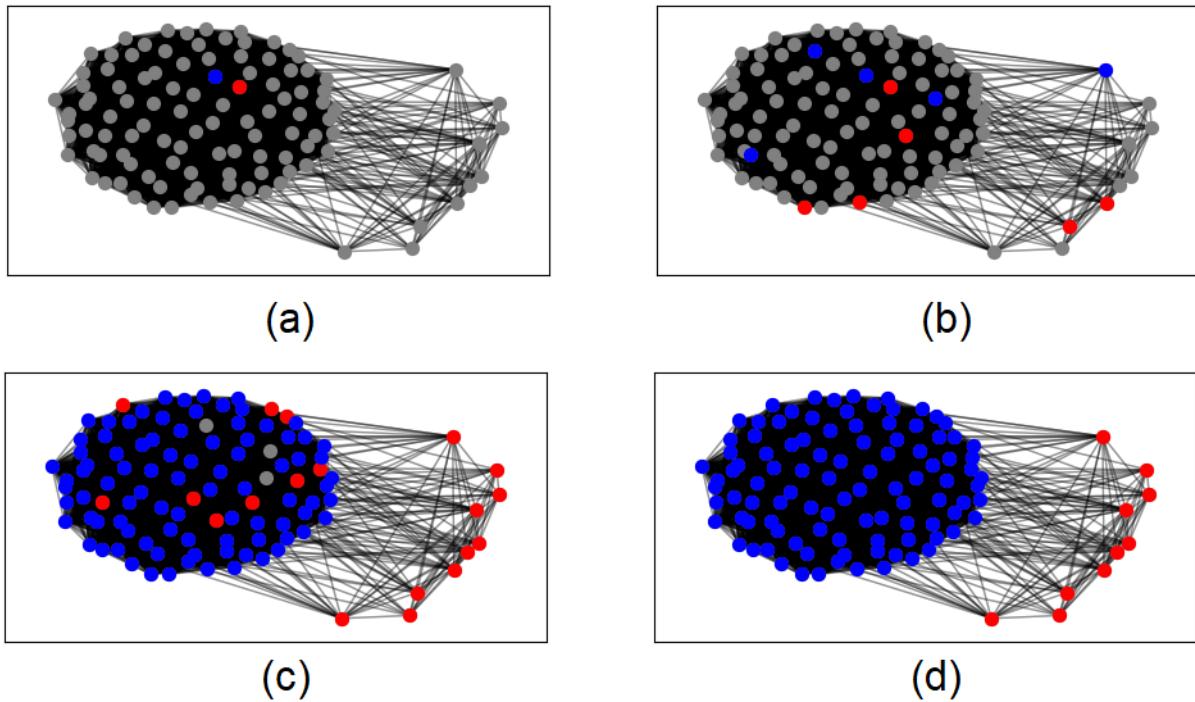


Figura 19 – Rede simples com comunidades desbalanceada. Duas comunidades, uma com 10 vértices e outra com 100.  $P_{out} = 0.1$  e  $P_{in} = 1.0$ . (a) - Configuração Inicial  $t = 0$ . (b) -  $t = 1$ . (c) -  $t = 10$ . (d) -  $t = 100$ . Rede gerada pelo autor através do módulo *NetworkX* na linguagem de programação *Python*

## 4.7 Considerações Finais

Neste capítulo, foram apresentados os resultados de análise do modelo proposto para a detecção de comunidades em redes da literatura e na detecção de comunidades desbalanceadas. O Método proposto consegue uma acurácia aceitável, embora haja limitação quando as densidades das comunidades nas redes são baixas.

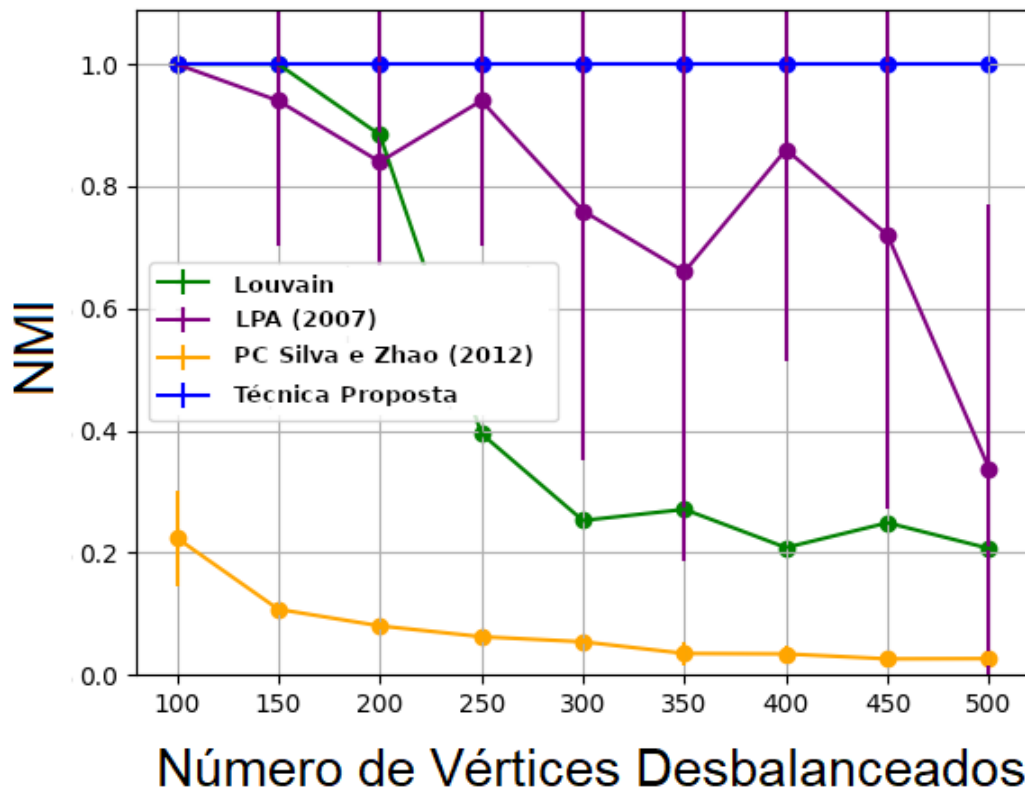


Figura 20 – Precisão de detecção de comunidade em relação ao nível de desbalanceamento em redes agrupadas aleatórias. Cada rede contém 2 comunidades, com  $P_{out} = 0.01$  e  $P_{in} = 1.0$ . A primeira comunidade teve um número fixo de vértices (10 vértices) e a segunda comunidade teve um número crescente de vértices de 100 para 500. Cada simulação tem uma média de 50 execuções para cada ponto.

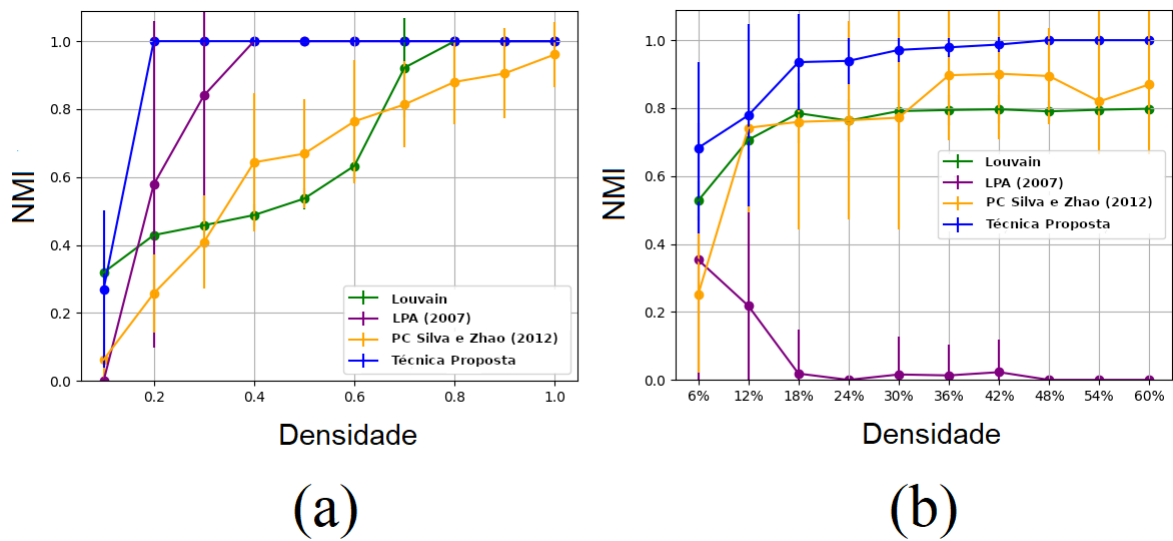


Figura 21 – Precisão de detecção de comunidade em relação ao nível de desbalanceamento em redes agrupadas aleatórias. (a) - Usamos uma rede simples com duas comunidades, uma com 50 vértices e outra com 200, com  $P_{out} = 0.01$  fixo e  $P_{in}$  variável, aumentando a densidade intra-modular, iniciando em 0.1 até o valor máximo de 1.0. (b) - Foi usado uma rede com duas comunidades, ambas com 100 vértices cada. Uma comunidade ficou com a densidade fixada em 6% enquanto na outra, a densidade foi variada de 6% até 60% em relação ao número de vértices. Cada simulação tem uma média de 50 execuções para cada ponto.

# Detecção de Comunidades em Redes Cerebrais Funcionais

Nessa seção são descritos os resultados aplicados a detecção de comunidades em redes cerebrais funcionais.

## 5.1 Considerações Iniciais

Na construção das redes cerebrais, as correlações foi feita pelo  $zScore$ . As redes cerebrais são formadas por uma visão tridimensional. Portanto, ilustramos a rede e as comunidades com 3 visões diferentes em relação aos eixos,  $x$ ,  $y$  e  $z$  (Fig. 22). À primeira vista, temos os eixos  $x$  e  $y$  com a vista posterior do cérebro, ou seja, as costas. Na segunda, temos a vista lateral dos eixos  $y$  e  $z$ , sendo o lado direito. E, por fim, a vista superior, em relação aos eixos  $x$  e  $z$ .

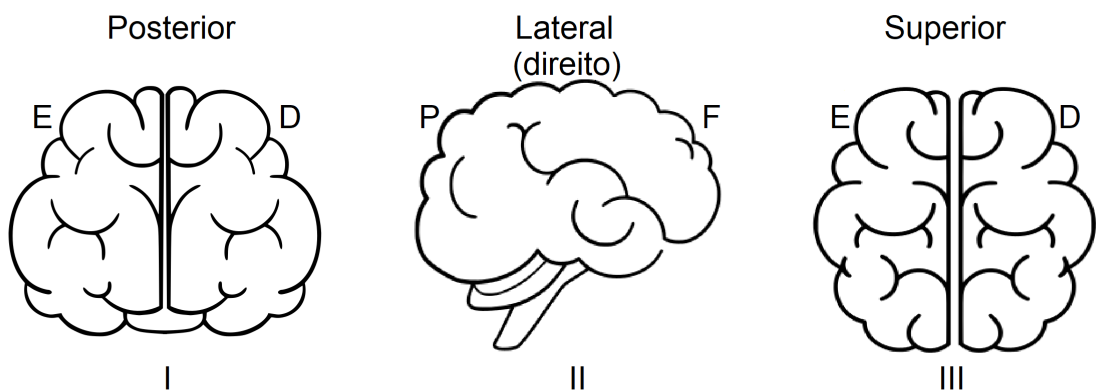


Figura 22 – Ilustrando as posições coordenadas  $(x, y, z)$  do cérebro. A letra  $E$  representa o lado esquerdo,  $D$  o lado direito do cérebro. A letra  $F$  é a região frontal e  $P$  é a parte de trás (posterior) do cérebro. Imagem feita pelo autor.

Neste estudo, são realizadas simulações em duas etapas. A primeira é a detecção de comunidades com o método proposto. Na segunda, analisamos as diferenças topológicas



ao nível de comunidade aplicadas em uma rede binária e rede ponderada usando a técnica proposta. Para isso, duas redes cerebrais funcionais são construídas. Foi utilizado o módulo da *Nilearn* <https://nilearn.github.io/stable/index.html> para a construção da Rede. Essa biblioteca fornece dados de imagens médicas e, além disso, atlas de mapeamento do cérebro que permitem efetuar a formulação da rede, ou seja, transformar as imagens médicas em um grafo com o qual podemos trabalhar.

## 5.2 Conjunto de Dados

Selecionamos imagens de 54 indivíduos cognitivamente normais (idade:  $80 \pm 6$  anos, 26 homens, 28 mulheres) do conjunto de dados *OASIS-3 Brain* (LAMONTAGNE et al., 2019). É um conjunto de dados de neuroimagem longitudinal, clínico, cognitivo e biomarcador gratuito para envelhecimento normal e doença de Alzheimer. O acesso é feito através do site [www.oasis-brains.org](http://www.oasis-brains.org).

As imagens foram obtidas no *BioGraph mMR PET-MR 3T* pelo *Knight Alzheimer Research Imaging Program*. Imagens estruturais ponderadas em *T1* e *BOLD-fMRI* em estado de repouso foram usadas. Durante a aquisição de *T1* e *BOLD-fMRI* em estado de repouso de 6 minutos, os participantes foram solicitados a deitar em silêncio, com os olhos abertos. Usamos o módulo *Nilearn* (uma biblioteca da linguagem de programação Python) para baixar as imagens através do link <https://nilearn.github.io/modules/reference.html#module-nilearn.datasets>.

## 5.3 Processamento de Imagens e Análise da Comunidades

Os dados funcionais foram processados e analisados no *Connectivity Toolbox (CONN)* (WHITFIELD-GABRIELI; NIETO-CASTANON, 2012), que faz parte do kit de ferramentas SPM (do inglês *Statistical Parametric Mapping, University College London*) (FRISTON et al., 1994). O pré-processamento incluiu reorientação da imagem, utilizando a comissura anterior como ponto de referência da origem; correção de tempo de corte; realinhamento de imagens funcionais; registro entre imagens funcionais e anatômicas; segmentação de imagens anatômicas; normalização para espaço *MNI*; e suavização espacial de imagens funcionais (filtro gaussiano,  $FWHM = 6\text{mm}$ ). As flutuações indesejadas no sinal de *fMRI* foram removidas usando os seis parâmetros de correção de movimento residual e suas derivadas de primeira vez, como também, o sinal global de substância branca e cinzenta como regressores.

A conectividade funcional (FC) baseada em regiões de interesse (ROI) foi realizada em *CONN* usando o atlas AAL-3v1 (ROLLS et al., 2019) (versão 3), consistindo em 166 regiões corticais para a construção do mapeamento de rede <https://www.gin.cnrs.fr/en/tools/aal/>. O atlas AAL (TZOURIO-MAZOYER et al., 2002; ROLLS et al., 2019), é bastante referenciado na literatura e disponibilizado pelo grupo de pesquisa *Neurofunctional Imaging Group (GIN-IMN) de Montreal Neurological Institute (MNI)*. Para a análise *ROI-to-ROI*, foi obtida uma matriz de correlação a partir da correlação bivariada entre as séries temporais de cada *ROI* (regiões do atlas).

Em seguida, realizamos a análise da comunidade. Combinamos o valor  $R$  (Eq. 5.1) e o valor de modularidade (NEWMAN; GIRVAN, 2004) para determinar o número ideal de partículas (comunidades) a serem detectadas pelo modelo proposto. O  $R$ , proposto e utilizado em (SILVA; ZHAO, 2012b), é a média do valor máximo de dominância de todos os nós. Se o número de partículas estiver correto,  $R$  tende a um valor alto. O experimento foi realizado 30 vezes para cada valor de partícula, e assim calculado o  $R$  e a modularidade, para evitar impacto negativo da inicialização incorreta. Comparamos as comunidades obtidas para o número ideal de partículas com os resultados obtidos pelo método da Modularidade Maximizada de *Louvain* (BLONDEL et al., 2008).

$$R(k) = \frac{1}{|V|} \sum_{u=1}^{|V|} \text{MAX}_{m \in K} (E(m, u)(T)) \quad (5.1)$$

Analizamos as diferenças topológicas ao nível de comunidade aplicadas em uma rede cerebral binária, e rede ponderada usando a técnica proposta. Para tanto, foram construídas duas redes cerebrais funcionais a partir da matriz de correlação média entre os 54 indivíduos, com limiar de 0.2, os valores dos limiares foram testados, e quando estava abaixo do estabelecido ( $< 0.2$ ) as redes ficaram muito densas e, por outro lado, ficaram esparsas quando o valor do limiar foi maior que 0.2, tanto para a matriz binária quanto para a ponderada. Usamos o número ideal de partículas obtido na etapa anterior. As seguintes medidas das duas redes foram calculadas: conectividade dentro e entre os módulos (Eq. 2.14), e forças Intra-Modulares (Eq. 2.15) e Inter-Modulares (Eq. 2.16), conforme descrito na subseção 2.4.1.

Como vários estudos são baseados no método de modularidade maximizada para identificar comunidades em redes cerebrais (SPORNS; BETZEL, 2015; GARCIA et al., 2018; SANCHEZ-RODRIGUEZ et al., 2019; SANCHEZ-RODRIGUEZ et al., 2020), aqui, comparamos as medidas topológicas em combinação ao número ideal de partículas em relação ao método de modularidade maximizada de *Louvain* (BLONDEL et al., 2008), e outro algoritmo de modularidade conhecido como *Fast Greedy* (CLAUSET; NEWMAN; MOORE, 2004).

A comparação das medidas topológicas entre as diferentes técnicas foi realizada

por Análise de Variância (*ANOVA*) e teste *pos-hoc* de *Tukey*. Consideramos diferenças significantes para  $p < 0.05$ .

## 5.4 Resultado Experimentais

As Fig.s 23 (a) e (b) mostram o cálculo de  $R$  nas redes binária e ponderada variando o número de partículas, respectivamente. As Fig.s 23 (c) e (d) mostram a modularidade calculada para as mesmas redes variando o número de partículas. As Fig.s 23 (e) e (f) mostram a combinação de  $R$  e modularidade, onde o maior valor combinado é obtido com 7 partículas para ambas as redes.

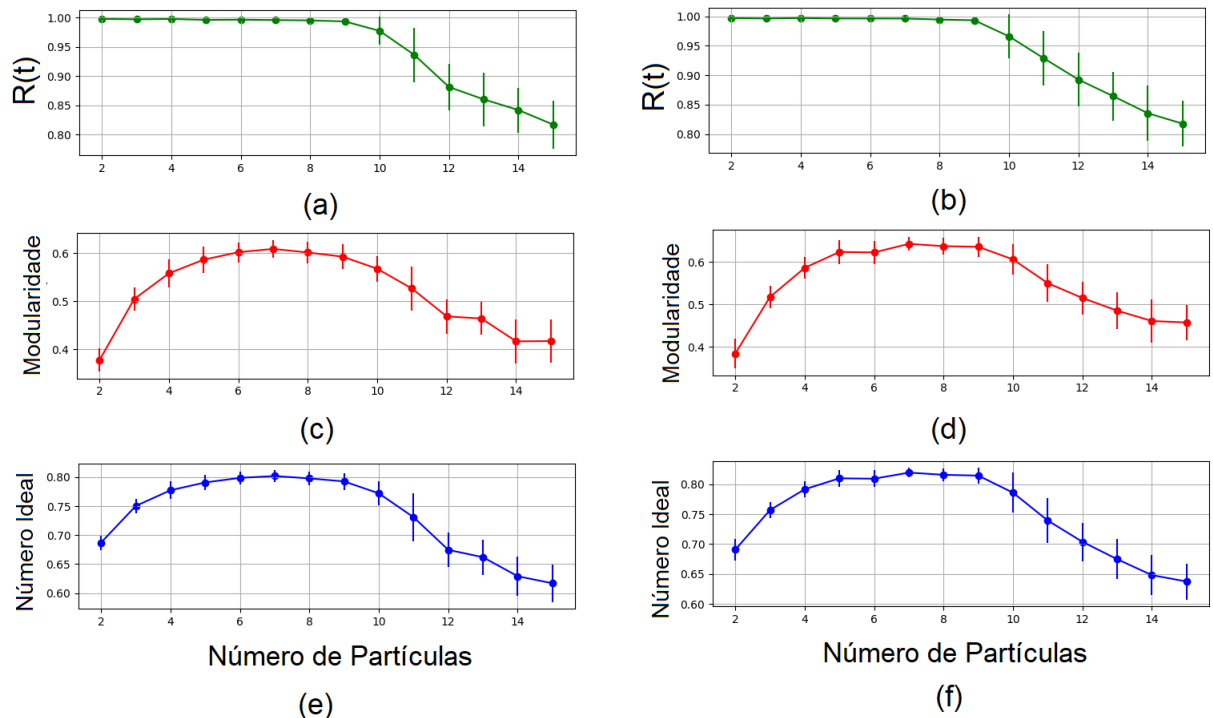


Figura 23 – Avaliação para encontrar o número ideal de partículas (comunidades). (a), (c) e (e) para a Rede de Binária e (b), (d) e (f) para a Rede da Ponderada. (a) e (b)  $R$  contra número de partículas. (c) e (d) Modularidade em relação ao número de partículas. (e) e (f) Combinação de  $R$  e modularidade contra o número de partículas. Cada simulação tem uma média de 30 execuções.

A Fig. 24 ilustra as comunidades detectadas para o melhor caso, ou seja, com o número ideal de acordo com o cálculo e análise apresentados na Fig. 23 (7 partículas), podemos observar que na rede binária (Fig. 24 (a)), em uma região central da rede cerebral, o método proposto identificou uma comunidade em que o método da modularidade maximizada identificou várias comunidades com apenas um vértice. Para efeito de comparação, também mostramos os resultados para o método da Modularidade Maximizada de *Louvain* (BLONDEL et al., 2008) (Fig. 25).

Com isso, o método proposto mostra-se eficaz na localização de comunidades em redes cerebrais funcionais, nas quais consegue captar as informações inter e intra vértices na rede cerebral. Assim, os resultados aqui apresentados mostram uma certa robustez na estrutura de comunidades da rede cerebral resultante, além de resultar em um agrupamento de regiões responsáveis por funcionalidades semelhantes. Veja as Figs. 24 e 25.

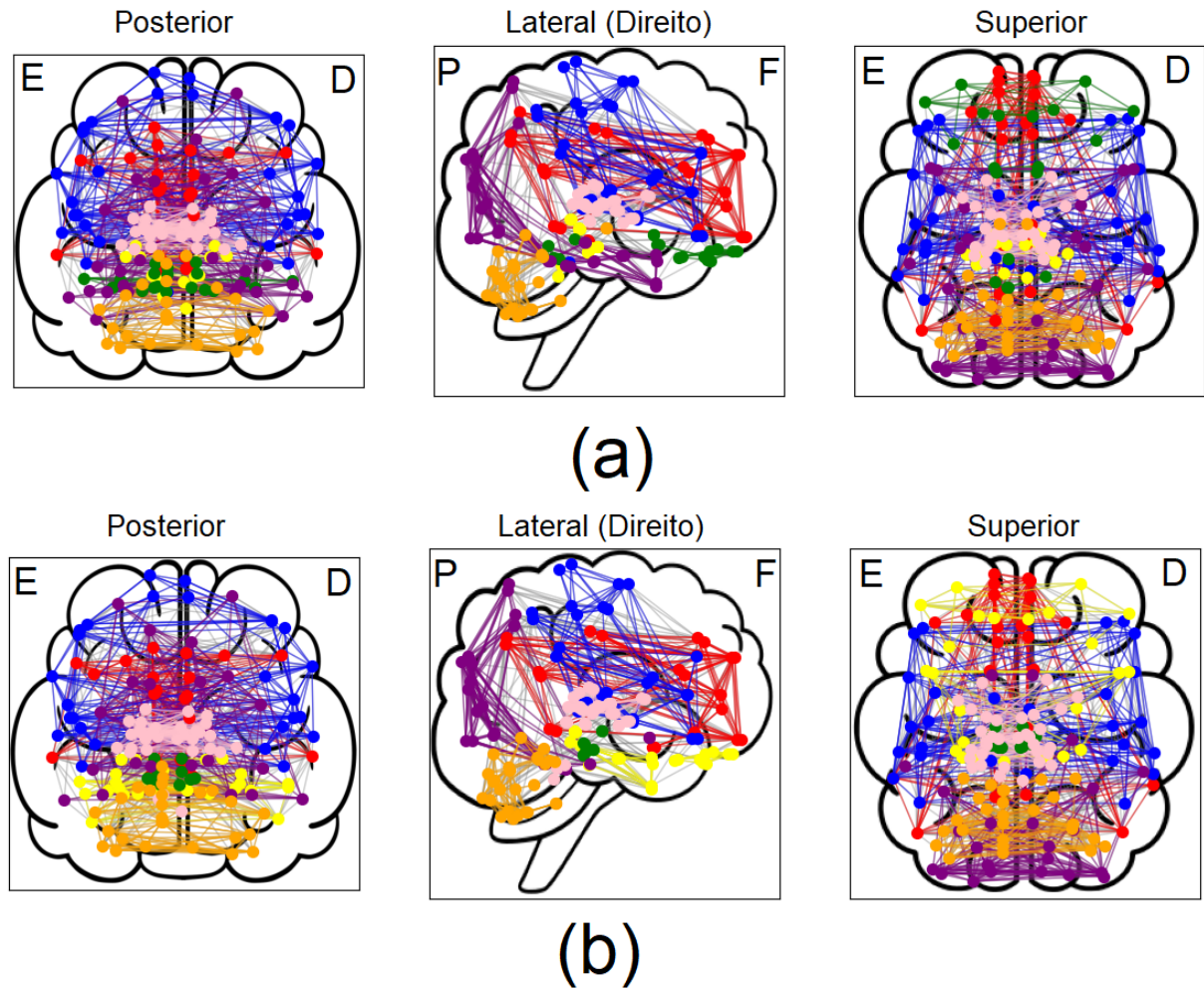


Figura 24 – Ilustração das comunidades identificadas pela técnica proposta. As arestas da mesma cor representam as conexões entre os vértices do mesmo módulo, as arestas com a cor cinza representam as conexões entre os diferentes módulos. (A) - Rede Binária. (B) - Rede Ponderada. E - Esquerda. D - Direita. P - Posterior. F - Frontal.

A tabela 7, mostra as medidas baseadas na comunidade calculadas para analisar a mudança topológica da rede binária para a ponderada. Os valores Inter-Modulares foram significativamente menores para a técnica proposta em relação ao da Modularidade de *Louvain* (BLONDEL et al., 2008) para os casos binário e ponderado ( $p = 0.03$ ). Em contrapartida, os valores Intra-Modulares foram significativamente maiores para a técnica proposta em relação às demais ( $p = 1.552 \times 10^{-5}$ ). Não foram observadas diferenças estatisticamente significativas para conectividade ( $p = 0.86$ ) e modularidade ( $p = 0.101$ )

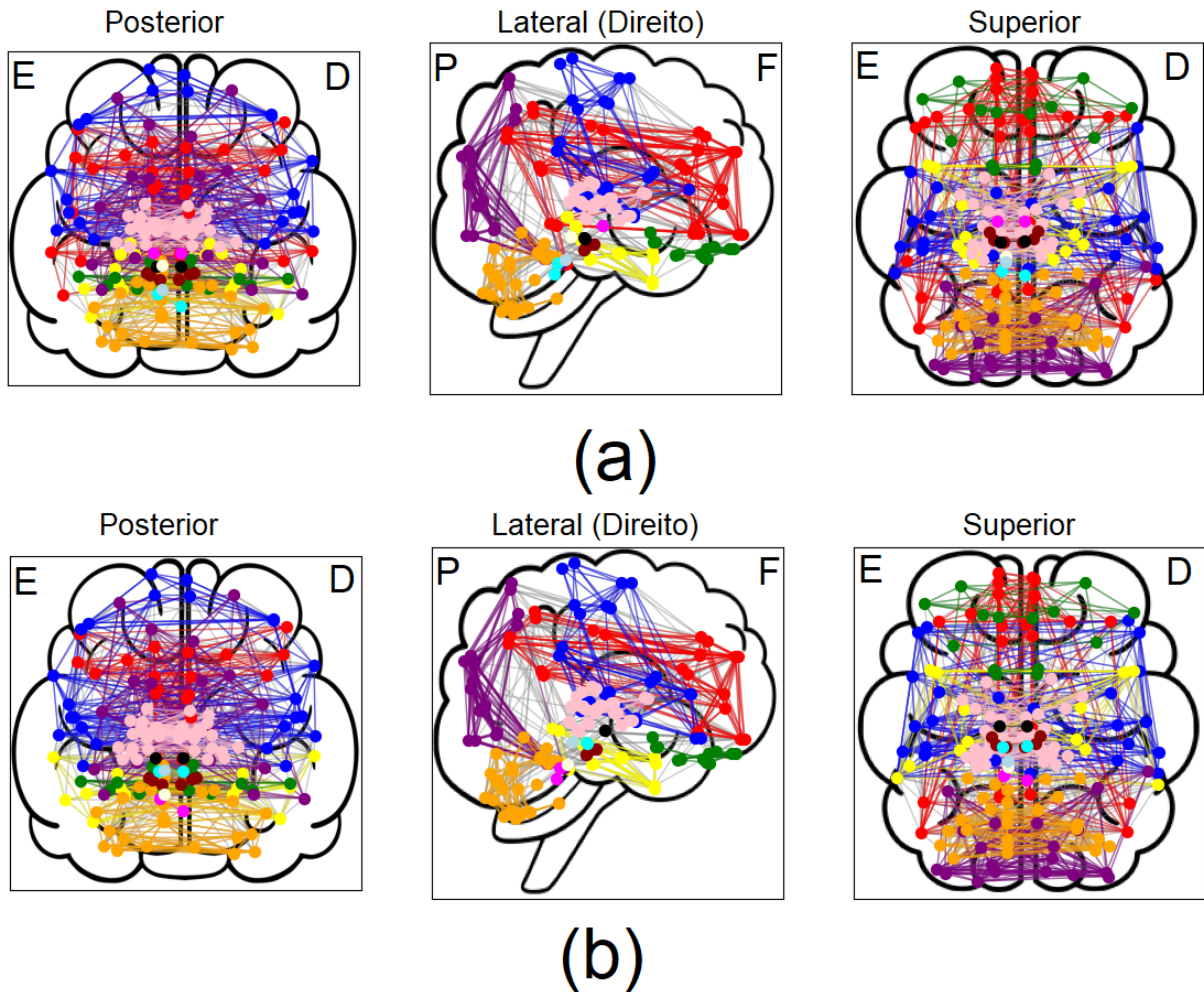


Figura 25 – Ilustração das comunidades identificadas pela Modularidade Maximizada de *Louvain* (BLONDEL et al., 2008). As arestas da mesma cor representam as conexões entre os vértices do mesmo módulo, as arestas com a cor cinza representam as conexões entre os diferentes módulos. (A) - Rede Binária. (B) - Rede Ponderada. E - Esquerda. D - Direita. P - Posterior. F - Frontal.

entre as técnicas.

Para ambas as redes, acreditamos que essas diferenças entre as forças modulares ocorreram porque o método de modularidade maximizada identificou comunidades na rede com apenas um vértice. No entanto, a técnica *Louvain* (BLONDEL et al., 2008) identificou 13 para a rede binária e ponderada, e para ambas as redes 6 comunidades tinham apenas 1 nó <sup>1</sup>. *Fast Greedy* (CLAUSET; NEWMAN; MOORE, 2004) identificou 11 comunidades para rede binária e rede ponderada, e também para ambas as redes, 2 comunidades tinham apenas 1 nó <sup>2</sup>.

Considerando apenas o modelo proposto e o número ideal de partículas, analisamos as regiões que constituem cada comunidade. Para ambas as redes (binárias e ponderadas),

<sup>1</sup> Na tabela 7, (7) estes são os números de comunidades identificadas que tinham mais de um nó.

<sup>2</sup> Na tabela 7, (9) estes são os números de comunidades identificadas que tinham mais de um nó.

Tabela 7 – Análise Topológica Comparando com Maximização de Modularidade

<i>Rede Binária</i>			
	<b>Tec. Proposta (Número Ideal)</b>	<b>Max. Modularidade (Louvain)</b>	<b>Max. Modularidade (Fast Greedy)</b>
<b>Comunidades</b>	7	13 (7)	11 (9)
<b>Inter-Modular</b>	$0.27 \pm 0.11$	$0.31 \pm 0.33$	$0.30 \pm 0.36$
<b>Intra-Modular</b>	$0.81 \pm 0.04$	$0.50 \pm 0.40$	$0.44 \pm 0.40$
<b>Conectividade</b>	$0.24 \pm 0.00$	$0.33 \pm 0.10$	$0.29 \pm 0.00$
<b>Modularidade</b>	$0.63 \pm 0.08$	$0.64 \pm 0.00$	$0.59 \pm 0.00$

<i>Rede Ponderada</i>			
	<b>Tec. Proposta (Número Ideal)</b>	<b>Max. Modularidade (Louvain)</b>	<b>Max. Modularidade (Fast Greedy)</b>
<b>Comunidades</b>	7	13 (7)	11 (9)
<b>Inter-Modular</b>	$0.28 \pm 0.17$	$0.33 \pm 0.34$	$0.30 \pm 0.36$
<b>Intra-Modular</b>	$0.82 \pm 0.04$	$0.51 \pm 0.40$	$0.44 \pm 0.40$
<b>Conectividade</b>	$0.10 \pm 0.00$	$0.14 \pm 0.06$	$0.13 \pm 0.09$
<b>Modularidade</b>	$0.67 \pm 0.08$	$0.68 \pm 0.00$	$0.59 \pm 0.00$

as comunidades não refletiram inteiramente as redes cerebrais funcionais em estado de repouso; no entanto, elas foram formadas por regiões com funções semelhantes. Para a rede binária (Fig. 26 a), as comunidades refletiram os seguintes grupos funcionais: (1) rede de modo padrão; (2) linguagem, frontoparietal e sensório-motor; (3) regiões subcorticais, incluindo o tálamo; (4) regiões com funções pouco conhecidas; (5) visual, visuoespacial e memória; (6) gânglios da base e tálamo; e (7) rede cerebelar. Para a rede ponderada (Fig. 26 b), as comunidades refletiram os seguintes grupos funcionais: (1) rede de modo padrão; (2) linguagem, frontoparietal e sensório-motor; (3) memória; (4) regiões com funções pouco conhecidas; (5) visual e visuoespacial; (6) gânglios da base e tálamo; e (7) rede cerebelar. Em geral, os métodos parecem considerar as informações funcionais e também a localização da região para separar as comunidades.

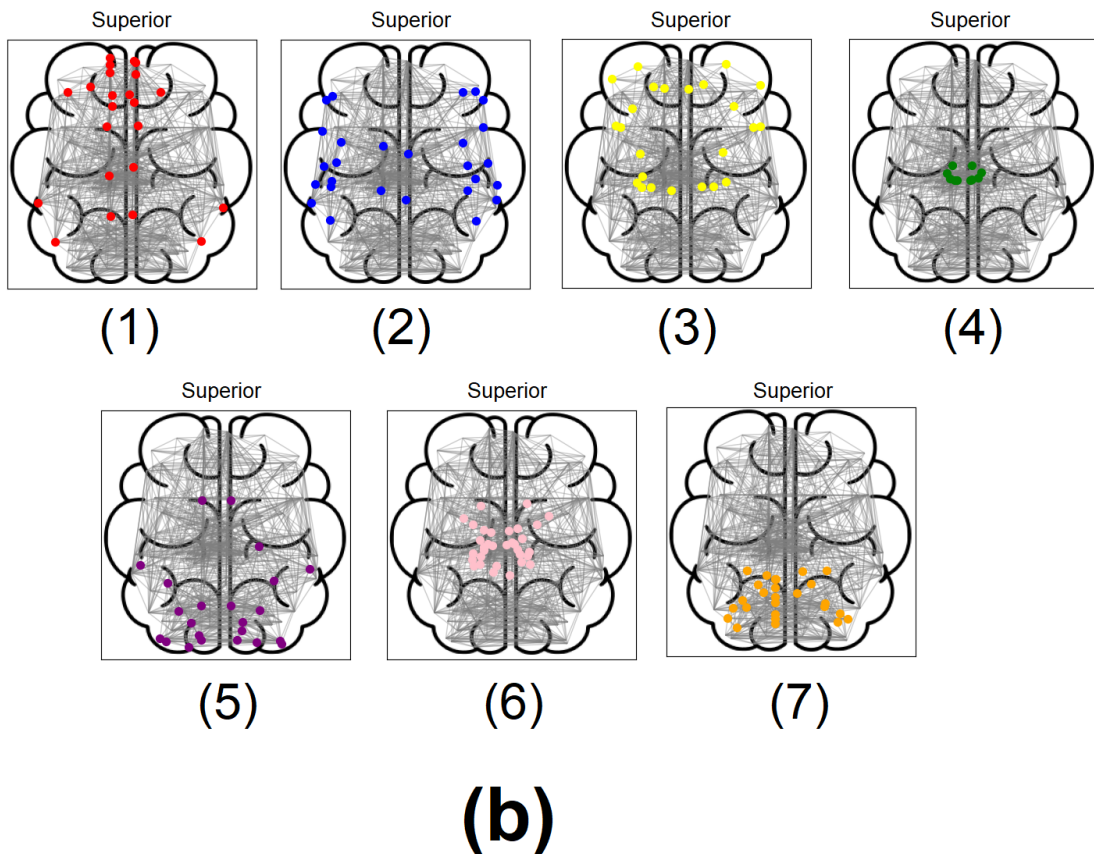
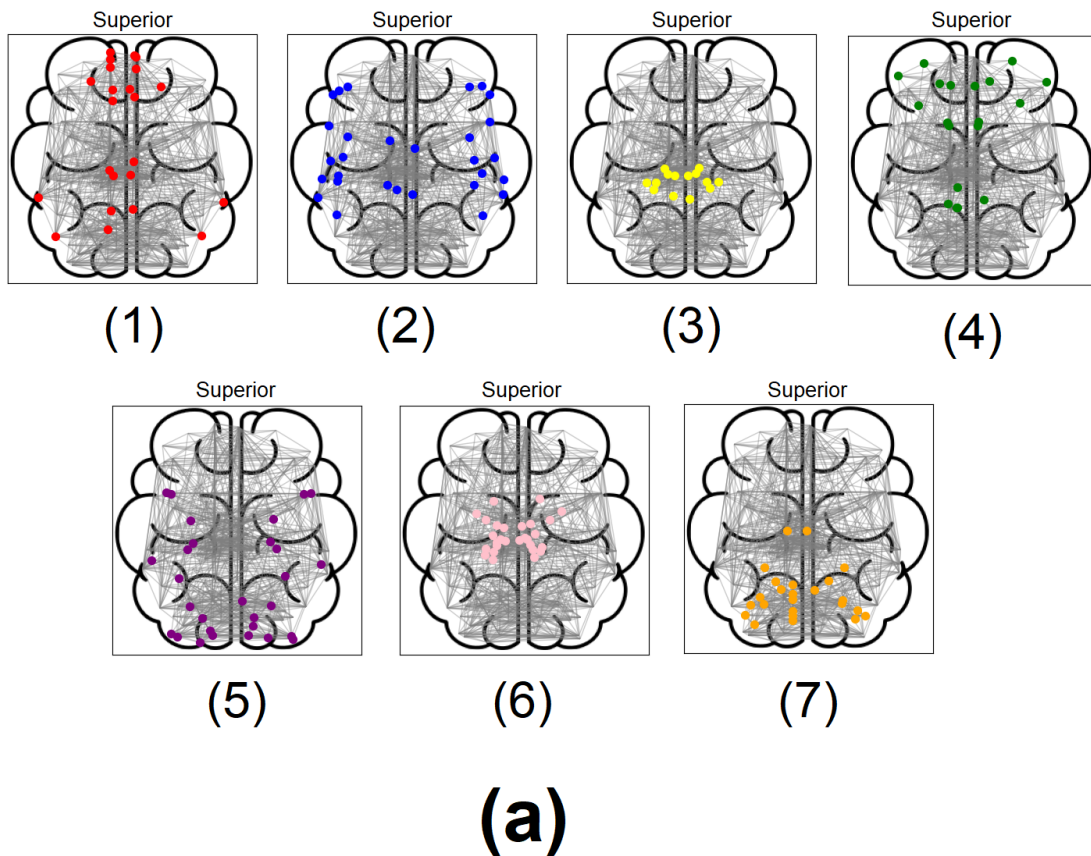


Figura 26 – Análise das regiões que constituem cada comunidade. (a) Rede binária. (b) Rede ponderada.

---

## Conclusão

A nova técnica de detecção de comunidade foi proposta com a combinação do método de competição de partículas e uma propagação paralela inspirada nos Mapas de Kohonen. O primeiro é um processo, sequencial, que pode verificar os vértices visitados de forma estocástica. Por outro lado, o último fornece um mecanismo de propagação de sinal paralelo, com a vantagem de capturar rapidamente a estrutura da rede local em torno do vértice visitado por uma partícula. O modelo apresenta duas características salientes em comparação com o modelo original de competição de partículas:

1. A propagação paralela ajuda a capturar a estrutura local da rede. Como consequência, cada partícula tem uma capacidade diferente de exploração conforme a estrutura da rede. Portanto, o novo modelo pode detectar comunidades desbalanceadas;
2. Novamente, devido à propagação paralela, cada partícula não precisa visitar todos os vértices. Na verdade, ela só precisa visitar uma pequena parte dos vértices para gerar os resultados de detecção da comunidade. Por outro lado, a competição de partículas original só pode detectar comunidades depois que todos os vértices são visitados várias vezes. Portanto, o novo modelo é mais eficiente que os originais.

Além disso, o método proposto se mostrou eficaz na localização de comunidades em redes cerebrais funcionais, nas quais é capaz de capturar informações intra-vértices na rede cerebral. Assim, os resultados aqui apresentados mostram certa robustez na estrutura de comunidades da rede cerebral resultante, além de resultar em um agrupamento de regiões responsáveis por funcionalidades semelhantes. Nosso estudo mostra que existe uma forte associação entre as comunidades obtidas e as redes cerebrais em repouso.

Portanto, a introdução da propagação paralela no modelo original de competição de partículas permitiu a detecção robusta de comunidades desbalanceadas. Sua aplicação em dados de *fMRI* em estado de repouso identifica comunidades que estão relacionadas a padrões de redes funcionais, sugerindo-a como uma ferramenta útil e adicional para realizar mapeamento funcional do cérebro.



## 6.1 Limitações

Uma das limitações apresentada pelo método proposto foi em relação à baixa densidade e quando há mistura alta entre as comunidades. Dessa forma, para baixas densidades nas redes, a acurácia do método será baixa, como apresentados nos resultados exibidos na seção 4.6. A acurácia também foi baixa quando ocorreram altos valores de mistura entre as comunidades na rede.

## 6.2 Publicações

O trabalho aqui apresentado gerou um artigo de conferência intitulado "*A New Particle Competition Model for Community Detection with Application in Functional Brain Networks*", publicado em *2021 The International Conference on Neural Networks (IJCNN)*, em colaboração com o coautor: Zhao Liang. Nessa primeira publicação apresentamos a nova abordagem para a detecção de comunidades em redes cerebrais.

## 6.3 Trabalhos Futuros

Tanto no conceito de detectar comunidades quanto para as análises em neurociência, pretendemos aprimorar a acurácia da técnica proposta quando para casos de alto nível de mistura entre as comunidades. Também queremos nos estender no estudo de redes cerebrais, aplicando para diferentes contextos de outras doenças neurológicas e/ou psiquiátricas. Além das redes cerebrais funcionais, vamos incrementar as abordagens do método proposto para redes cerebrais estruturais.

---

# Referências

- ALBERT, R.; JEONG, H.; BARABASI, A.-L. Error and attack tolerance of complex networks. *Nature*, v. 406, 08 2000.
- ALONSO, R. L. Dinâmica de partículas e aprendizado competitivo para detecção de comunidades em redes complexas. *Tese (Doutorado em Ciências de Computação e Matemática Computacional) - Universidade de São Paulo, São Carlos (São Paulo, Brasil)*, 2008.
- BARABASI, A.-L.; ALBERT, R. Albert, r.: Emergence of scaling in random networks. *science* 286, 509-512. *Science (New York, N.Y.)*, v. 286, p. 509–12, 11 1999.
- BARABÁSI, A. *Network science*. [S.l.]: Cambridge: Cambridge University Press. ISBN: 9781107076266 1107076269, 2016.
- BECKMANN, C. F. et al. Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 360, n. 1457, p. 1001–1013, 2005. Disponível em: <<https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2005.1634>>.
- BETZEL, R. et al. The community structure of functional brain networks exhibits scale-specific patterns of inter- and intra-subject variability. *NeuroImage*, v. 202, 07 2019.
- BISWAL, B. et al. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic Resonance in Medicine*, v. 34, n. 4, p. 537–541, 1995. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.1910340409>>.
- BLONDEL, V. D. et al. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, IOP Publishing, v. 2008, n. 10, p. P10008, Oct 2008. ISSN 1742-5468. Disponível em: <<http://dx.doi.org/10.1088/1742-5468/2008/10/P10008>>.
- BULLMORE, E.; SPORNS, O. Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature reviews. Neuroscience*, v. 10, p. 186–198, 03 2009.
- BULLMORE, E.; SPORNS, O. The economy of brain network organization. *Nature Reviews Neuroscience*, v. 13, p. 336 –349, 2012.
- CALHOUN, V. et al. A method for making group inferences from functional mri data using independent component analysis. *Human Brain Mapping*, v. 14, n. 3, p. 140–151, 2001. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.1048>>.

- CLAUSET, A.; NEWMAN, M. E. J.; MOORE, C. Finding community structure in very large networks. *Physical Review E*, American Physical Society (APS), v. 70, n. 6, Dec 2004. ISSN 1550-2376. Disponível em: <<http://dx.doi.org/10.1103/PhysRevE.70.066111>>.
- CORDASCO, G.; GARGANO, L. Community detection via semi-synchronous label propagation algorithms. *IEEE International Workshop on: Business Applications of Social Network Analysis (BASNA)*, p. 1–8, 2010.
- CORDES, D. et al. Mapping functionally related regions of brain with functional connectivity mri (fcmri). *AJNR. American journal of neuroradiology*, v. 21, p. 1636–44, 11 2000.
- CORDES, D. et al. Hierarchical clustering to measure connectivity in fmri resting-state data. *Magnetic resonance imaging*, v. 20, p. 305–17, 06 2002.
- ERDÖS, P.; RÉNYI, A. Additive properties of random sequences of positive integers. *Acta Arithmetica*, v. 6.1, p. 83–110, 1960.
- FALLANI, F. D. V. et al. Cortical functional connectivity networks in normal and spinal cord injured patients: Evaluation by graph analysis. *Human Brain Mapping*, John Wiley and Sons Inc., v. 28, n. 12, p. 1334–1346, dez. 2007. ISSN 1065-9471.
- FERREIRA, L.; ZHAO, L. Time series clustering via community detection in networks. *Information Sciences*, v. 326, 08 2015.
- FORTUNATO, S. Community detection in graphs. *Physics Reports*, Elsevier BV, v. 486, n. 3-5, p. 75–174, Feb 2010. ISSN 0370-1573. Disponível em: <<http://dx.doi.org/10.1016/j.physrep.2009.11.002>>.
- FREEMAN, L. C. Centrality in social networks conceptual clarification. *Social Networks*, v. 1, n. 3, p. 215–239, 1978. ISSN 0378-8733. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0378873378900217>>.
- FRISTON, K. J. et al. Functional connectivity: The principal-component analysis of large (pet) data sets. *Journal of Cerebral Blood Flow & Metabolism*, v. 13, n. 1, p. 5–14, 1993. PMID: 8417010. Disponível em: <<https://doi.org/10.1038/jcbfm.1993.4>>.
- FRISTON, K. J. et al. Statistical parametric mapping. *Functional neuroimaging: Technical foundations*, p. 79–93, 1994.
- GAO, X. et al. Particle competition for multilayer network community detection. *In Proceedings of the 2019 11th International Conference on Machine Learning and Computing (ICMLC '19) New York, USA*, Association for Computing Machinery, p. 281–297, 2019.
- GARCIA, J. O. et al. Applications of community detection techniques to brain graphs: Algorithmic considerations and implications for neural function. *Proceedings of the IEEE*, v. 106, n. 5, p. 846–867, 2018.
- GIRVAN, M.; NEWMAN, M. E. J. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, Proceedings of the National Academy of Sciences, v. 99, n. 12, p. 7821–7826, Jun 2002. ISSN 1091-6490. Disponível em: <<http://dx.doi.org/10.1073/pnas.122653799>>.

HARTWELL, L. et al. From molecular to modular cell biology. *Nature*, v. 402, p. C47–52, 01 2000.

HEUVEL, M. P. van den; POL, H. E. H. Exploring the brain network: a review on resting-state fmri functional connectivity. *European Neuropsychopharmacology*, v. 20, n. 20, p. 519–534, 2010.

HEUVEL, M. van den; MANDL, R.; POL, H. H. Normalized cut group clustering of resting-state fmri data. *PLOS ONE*, Public Library of Science, v. 3, n. 4, p. 1–11, 04 2008. Disponível em: <<https://doi.org/10.1371/journal.pone.0002001>>.

HUANG, Z. Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data Mining and Knowledge Discovery*, v. 2(3), p. 283 – 304, 1998.

HUETTEL, S.; SONG, A.; MCCARTHY, G. Functional magnetic resonance imaging. *he Yale Journal of Biology and Medicine*, v. 82(4), p. 233, 2009.

KOHONEN, T. The self-organizing map. *Proceedings of the IEEE*, v. 78, n. 9, p. 1464–1480, 1990.

LAMONTAGNE, P. J. et al. Oasis-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease. *medRxiv*, Cold Spring Harbor Laboratory Press, 2019. Disponível em: <<https://www.medrxiv.org/content/early/2019/12/15/2019.12.13.19014902>>.

LANCICHINETTI, A.; FORTUNATO, S.; RADICCHI, F. Benchmark graphs for testing community detection algorithms. *Physical Review E*, American Physical Society (APS), v. 78, n. 4, Oct 2008. ISSN 1550-2376. Disponível em: <<http://dx.doi.org/10.1103/PhysRevE.78.046110>>.

LI, W. et al. Community detection for information propagation relying on particle competition. In: . Public Library of Science, 2008. v. 3, n. 4, p. 1–11. Disponível em: <<https://doi.org/10.1371/journal.pone.0002001>>.

LINDQUIST, M. A.; WAGER, T. D. Principles of functional magnetic resonance imaging. In: *Ombao, H., Lindquist, M., Thompson, W., Aston, J. (Eds.), Handbook of neuroimaging data analysis*. London, England: Chapman and Hall/CRC, 2014. p. 3–48. ISBN 9781315373652. Disponível em: <[https://cpb-us-e1.wpmucdn.com/sites.dartmouth.edu/dist/2/2150/files/2019/12/2014\\_Lindquist\\_Handbook-of-Neuroimaging-Data-Analysis.pdf](https://cpb-us-e1.wpmucdn.com/sites.dartmouth.edu/dist/2/2150/files/2019/12/2014_Lindquist_Handbook-of-Neuroimaging-Data-Analysis.pdf)>.

MACQUEEN, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, p. 281–297, 1967.

MARTINS, L. V. C.; ZHAO, L. Particle competition for unbalanced community detection in complex networks. In: *Intelligent Systems: 9th Brazilian Conference, BRACIS 2020, Rio Grande, Brazil, October 20–23, 2020, Proceedings, Part II*. Springer-Verlag, 2020. p. 322–336. ISBN 978-3-030-61379-2. Disponível em: <[https://doi.org/10.1007/978-3-030-61380-8\\_22](https://doi.org/10.1007/978-3-030-61380-8_22)>.

NEWMAN, M. E. J. Fast algorithm for detecting community structure in networks. *Physical Review E*, American Physical Society (APS), v. 69, n. 6, Jun 2004. ISSN 1550-2376. Disponível em: <<http://dx.doi.org/10.1103/PhysRevE.69.066133>>.

- NEWMAN, M. E. J.; GIRVAN, M. Finding and evaluating community structure in networks. *Physical Review E*, American Physical Society (APS), v. 69, n. 2, Feb 2004. ISSN 1550-2376. Disponível em: <<http://dx.doi.org/10.1103/PhysRevE.69.026113>>.
- PALLA, G. et al. Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, v. 435, p. 814–818, 07 2005.
- PAMPLONA, G. S. P. Conectividade funcional do cérebro: Uma análise das associações com desempenho intelectual a atenção sustentada usando imagens por ressonância magnética. *Dissertação (Mestrado em Física Aplicada a Medicina e Biologia) - Faculdade de Filosofia Ciências e Letras de Ribeirão Preto (FFCLRP) - Universidade de São Paulo, Ribeirão Preto (São Paulo, Brasil)*, 2014.
- QUILES, M. G. et al. Particle competition for complex network community detection. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, AIP, v. 18, n. 3, p. 033107, 2008.
- RAGHAVAN, N.; ALBERT, R.; KUMARA, S. Near linear time algorithm to detect community structures in large-scale networks. *Physical review. E, Statistical, nonlinear, and soft matter physics*, v. 76, p. 036106, 10 2007.
- RAVASZ, E.; BARABÁSI, A.-L. Hierarchical organization in complex networks. *Physical Review E*, American Physical Society (APS), v. 67, n. 2, Feb 2003. ISSN 1095-3787. Disponível em: <<http://dx.doi.org/10.1103/PhysRevE.67.026112>>.
- ROCHA, W. A. A. da. Aspectos de redes complexas com aplicações em neurociência. *Dissertação (Mestrado em Matemática Aplicada) - Universidade Estadual de Campinas, Campinas (São Paulo, Brasil)*, n. 76-91, 2017.
- ROLLS, E. et al. Automated anatomical labelling atlas 3. *NeuroImage*, v. 206, p. 116189, 09 2019.
- RUBINOV, M.; SPORNS, O. Complex network measures of brain connectivity: uses and interpretations. *NeuroImage*, v. 52, p. 1059–69, 10 2009.
- SALMON, C. E. G.; LEONI, R. F. Conectividade funcional cerebral utilizando técnicas de imagens por ressonância magnética. *Revista Brasileira de Física Médica*, v. 13, n. 1, p. 66–75, set. 2019. Disponível em: <<https://rbfm.org.br/rbfm/article/view/518>>.
- SANCHEZ-RODRIGUEZ, L. M. et al. A method for multiscale community detection in brain. *bioRxiv*, Cold Spring Harbor Laboratory, 2019. Disponível em: <<https://www.biorxiv.org/content/early/2019/08/22/743732>>.
- SANCHEZ-RODRIGUEZ, L. M. et al. Detecting brain network communities: considering the role of information flow and its different temporal scales. *bioRxiv*, Cold Spring Harbor Laboratory, 2020. Disponível em: <<https://www.biorxiv.org/content/early/2020/08/28/743732>>.
- SILVA, T. C.; ZHAO, L. Network-based stochastic semi-supervised learning. *IEEE Transactions on Neural Networks and Learning Systems*, v. 23, p. 451–466, 2012.
- SILVA, T. C.; ZHAO, L. Stochastic competitive learning in complex networks. *IEEE Transactions on Neural Networks and Learning Systems*, v. 23, p. 385–398, 2012.

- SILVA, T. C.; ZHAO, L. Uncovering overlapping cluster structures via stochastic competitive learning. *Information Sciences*, v. 247, p. 40–61, 2013. ISSN 0020-0255. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0020025513004507>>.
- SILVA, T. C.; ZHAO, L. Machine learning in complex networks. *Springer*, p. 331, 2016.
- SOARES, I.; OLIVEIRA, C. S.; BRITO, J. A. d. M. Um estudo do problema de detecção de comunidades em redes. *Sistemas Gestão*, v. 9, n. 4, p. 566–574, 2014. Disponível em: <<https://www.revistasg.uff.br/sg/article/view/V9N4A14>>.
- SPORNS, O. Structure and function of complex brain networks. *Dialogues Clin Neurosci*, v. 15, n. 3, p. 247–262, 2013.
- SPORNS, O.; BETZEL, R. Modular brain networks. *Annual review of psychology*, v. 67, 09 2015.
- SPORNS, O. et al. Organization, development and function of complex brain networks. *Trends in cognitive sciences*, v. 8, p. 418–25, 10 2004.
- STAM, C. et al. Graph theoretical analysis of magnetoencephalographic functional connectivity in alzheimer’s disease. *Brain*, v. 123, p. 213–224, 2009.
- STAM, C.; REIJNEVELD, J. Graph theoretical analysis of complex networks in the brain. *Nonlinear biomedical physics*, v. 1, p. 3, 02 2007.
- TZOURIO-MAZOYER, N. et al. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *NeuroImage*, v. 15, p. 273–89, 02 2002.
- VEN, V. et al. Functional connectivity as revealed by spatial independent component analysis of fmri measurements during rest. *Human brain mapping*, v. 22, p. 165–78, 07 2004.
- WATTS, S. S. D. Collective dynamics of ‘small-world’ networks. *Nature*, v. 393, p. 440–44, 1998.
- WHITFIELD-GABRIELI, S.; NIETO-CASTANON, A. Conn: A functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connectivity*, v. 2, n. 3, p. 125–141, 2012. PMID: 22642651. Disponível em: <<https://doi.org/10.1089/brain.2012.0073>>.
- ZACHARY, W. An information flow model for conflict and fission in small groups. *Journal of anthropological research*, v. 33, 11 1976.

# Apêndices



---

# Disponibilidade do Algoritmo de Aprendizado Competitivo Paralelo

O código final do Algoritmo estará disponível no *GitHub*. Pode ser acessado pelo *link* <https://github.com/darkaiba/ParallelCompetitiveLearning>. Contém exemplos de implementação em redes complexas, utilizadas nesse estudo.