

# OTIMIZAÇÃO DE INCLUSÕES DE NÓS EM REDES COMPLEXAS

Carlos Miguel Moreira Gonçalves<sup>1\*</sup>

\*Corresponding author

<sup>1</sup>Mestrado em Modelagem e Métodos Quantitativos (first author)

E-mails: carlosmiguel@fisica.ufc.br

**ABSTRACT.** Nos últimos anos o estudo sobre intervenções em redes tem ganhado mais interesse de pesquisadores. Remover ou adicionar sítios ou arestas tem várias aplicações epidemiológicas e sociais. Nesse sentido pretende-se otimizar qual a melhor posição na rede deve ser retirado ou colocado o nó, fazer essa ação tem custos na vida real e isso é um fator limitante na qual limita quantas vezes se deve fazer esse processo. Existem várias métricas a serem otimizadas, uma delas é a chamada *betweenness centrality* que é a métrica que quantifica quanto o nó é importante sobre as informações transmitidas entre cada par de nós na rede. Neste trabalho será discutido a problemática de inserção de nós em redes com o objetivo de reunir o máximo de informação possível da rede. Para resolver utilizaremos diferentes meta-heurísticas para comparar a sua eficiência, pois além de ser um problema combinatório a *betweenness centrality* tem um alto custo computacional e se deve ter cuidado em qual algoritmo utilizar para otimização.

**Keywords:** Fluxo Máximo, Grafos, Simplex, Algoritmo Ford-Fulkerson

## 1 Introdução

A inserção de nós em redes é um tema de grande relevância e interesse em diversos campos, incluindo ciência da computação, engenharia de redes e teoria dos grafos. Essa prática consiste em adicionar novos nós a uma rede existente, alterando sua estrutura e dinâmica. A principal motivação para realizar a inserção de nós em redes é a capacidade de melhorar o desempenho, a eficiência e a resiliência da rede, além de permitir a adaptação a novas demandas e cenários.

Existem várias aplicações práticas em que a inserção de nós se mostra benéfica. Por exemplo, em redes de infraestrutura, como redes elétricas ou sistemas de distribuição de água, a adição estratégica de nós pode ajudar a otimizar a distribuição de recursos e melhorar a confiabilidade do sistema. Em redes de transporte, como redes rodoviárias ou sistemas de transporte público, a inserção de nós pode levar a rotas mais eficientes, redução de congestionamentos e melhorias no tempo de viagem. Além disso, a inserção de nós pode ser utilizada em redes sociais e colaborativas para facilitar a formação de comunidades, conectar grupos de interesse semelhantes e promover interações mais efetivas.

Em sistemas de comunicação, a inserção de nós pode ampliar a cobertura da rede, estender sua capacidade e melhorar a conectividade. Esses são apenas alguns exemplos das várias aplicações em que a inserção de nós desempenha um papel crucial na otimização e aprimoramento de redes existentes.

Na Seção 2 definiremos o que é uma rede e qual o critério será utilizado para ser a nossa função a ser minimizada e na Seção 3 iremos formalizar o problema e explicar quais serão os algoritmos que serão utilizados para encontrar soluções viáveis para o problema.

## 2 Teoria de Grafos

Grafos (Goldbarg and , Auth.) é o campo da matemática que estuda objetos (vértices ou nó) e as interações entre si (arestas ou conexões). Apesar de ter uma formulação bem simples é um campo com bastante interdisciplinaridade, é de interesse de vários campos da ciência (como Física, Sociologia e Biologia) podem ser representados como Grafos. Ao ser estudado o campo de redes sociais e como uma *fake news* se propaga no *Twitter* é possível representar cada conta no site como um nó e o número de seguidores como as suas conexões, ou ao ser estudado uma cadeia alimentar em alguma região é possível utilizar grafos também sendo cada animal um vértice e quem se alimenta de quem podem ser representadas pelas arestas. A terminologia de Grafos/Redes, Vértices/Nós e Arestas/Conexões variam do campo, mas na literatura, em geral, tem o mesmo significado.

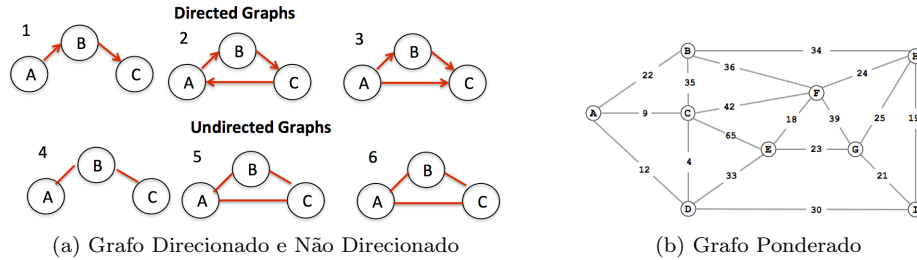


Figure 1: Tipos de Grafos

Dado  $N = 0, 1, 2, \dots$  um conjunto não vazio de nós e  $L = (i, j)$  um número de pares ordenados entre os  $N$  nós que chamaremos de ligações, podemos definir um Grafo como uma estrutura que contém nós e ligações  $G = (N, L)$ . Caso não exista alguma ligação do tipo  $(i, i)$  o grafo é chamado **grafo simples**, caso para cada conexão  $(i, j)$  exista também uma ligação  $(j, i)$  então esse grafo é chamado de **não-direcionado**, caso haja uma ligação  $(i, j)$  e não existir nenhuma  $(j, i)$  então ele é chamado de **grafo direcionado**. Caso o Grafo se de qualquer vértice  $i$  seja possível chegar em qualquer vértice  $j$  esse grafo é chamado de *grafo conectado* e caso tenha todas as possíveis ligações entre todos os nós caracterizamos esse grafo como **grafo completo**. Em nossa modelagem podemos atribuir pesos aos vértices ou arestas no nosso grafo, esse tipo de grafo chamaremos de

**grafo ponderado.** Por fim no caso de uma rede não direcionada definimos que  $j$  é vizinho de  $i$  se  $i$  tem uma conexão com  $j$ . No caso de redes direcionadas aparece uma ambiguidade, pois no caso podemos ter um vizinho que tem uma conexão de entrada e uma de saída, portanto definiremos para essa rede dois tipos de vizinhos, o de entrada e o de saída

A depender do nosso problema é necessário saber qual tipo de Grafo utilizar, por exemplo, ao estudar a rede de amizade *Facebook* sabe-se que se um usuário adiciona outro necessariamente o outro adiciona o primeiro, caracterizando-se como um grafo não direcionado. Porém ao tratar da rede de seguidores do *Instagram* se um usuário segue o outro não necessariamente o segundo segue o primeiro, caracterizando-se como um grafo direcionado. Portanto, ao modelar um problema é necessário estar atento a que características da nossa rede.

Estudar métricas de centralidade é de fundamental importância para entender e analisar a estrutura e o funcionamento de redes complexas. Essas métricas nos fornecem informações valiosas sobre a importância e influência de cada nó em uma rede, permitindo identificar os atores-chave e compreender os padrões de interação entre eles. Entre as diversas medidas de centralidade, destaca-se a Betweenness Centrality (BC), que quantifica a capacidade de um nó intermediar ou controlar o fluxo de informação dentro de uma rede. Ela revela quais nós atuam como "pontes" ou "conectores" entre diferentes regiões, desempenhando um papel crucial na comunicação e no fluxo de recursos. Compreender a *betweenness centrality* é essencial para identificar os atores que possuem maior poder de influência e que podem exercer um controle significativo sobre o fluxo de informações e a disseminação de ideias dentro de uma rede, sendo uma medida valiosa para análise de redes sociais, transporte, infraestruturas e muitos outros domínios. Ela é definida como:

$$C_B(G, v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (1)$$

$C_B(G, v)$  representa a centralidade de intermediação do nó  $v$  na rede. A centralidade de intermediação é calculada somando-se as frações  $\frac{\sigma_{st}(v)}{\sigma_{st}}$  para todos os pares de nós distintos  $(s, t)$ , onde  $s$  e  $t$  são nós diferentes de  $v$ . Aqui,  $\sigma_{st}$  denota o número total de caminhos mais curtos entre  $s$  e  $t$ , e  $\sigma_{st}(v)$  é o número de caminhos mais curtos que passam pelo nó  $v$ .

### 3 Definição do Problema

O Problema  $k - l$  foi introduzido por Lozano and Trujillo (2019), nele dado uma quantidade  $l$  e  $k$  de novos nós a serem adicionados a um grafo  $G(N, L)$  esses novos sítios devem ser colocados estrategicamente para que o novo grafo  $G'(N', L')$  tenha uma BC maximizada. Formalmente dado um conjunto  $S = s_1, s_2, \dots, s_i, \dots, s_k$  de novos nós cada com com  $l$  ligações  $L' = (s_i, k) s_i \in S k \in N$  na qual maximizamos:

$$\max \min C_B(G', v) \quad (2)$$

Achar a BC de um grafo é um problema  $O(N \cdot L + N^2 \cdot \log(N))$  e dado que encontrar quais as ligações os sítios novos devem fazer deixa o problema ainda mais complicado pela quantidade possíveis de combinações. Portanto urge a necessidade de algoritmos para tentar encontrar soluções aproximadas dado o critério de maximin  $C_B(G', v)$ , para isso usaremos diferentes meta-heurísticas (Antônio Gaspar-Cunha, 2012) e iremos comparar seus resultados, pois pela demora computacional que é o cálculo da *betweenness centrality* é necessário analisar com mais cautela o gasto computacional. Serão analisados os seguintes algoritmos

- **Busca local** é uma técnica de otimização que busca melhorar iterativamente uma solução inicial explorando seu espaço de busca local. Ele é aplicado em problemas nos quais a qualidade da solução depende de um conjunto limitado de vizinhos próximos da solução atual. A ideia central do algoritmo é iniciar com uma solução inicial arbitrária ou obtida por meio de heurísticas específicas. Em seguida, a solução é iterativamente aprimorada por meio de movimentos locais que modificam gradualmente a solução atual.
- **Algoritmo Genético (AG)** é uma técnica de otimização baseada na teoria da evolução biológica. Ele busca encontrar soluções aproximadas para problemas complexos, inspirando-se no processo de seleção natural e reprodução dos seres vivos. O AG opera com uma população de soluções candidatas, que são representadas por cromossomos ou indivíduos. Cada cromossomo é composto por genes, que representam partes da solução. Por exemplo, em um problema de otimização, um gene pode ser uma variável ou parâmetro. A evolução da população ocorre em várias gerações, em que os indivíduos mais adaptados têm uma maior probabilidade de reprodução e de transmitir seus genes para a próxima geração. O processo de evolução é guiado por operadores genéticos, que são aplicados aos cromossomos para criar novas soluções.
- **Otimização por Enxame de Partículas (PSO)** é uma metaheurística inspirada no comportamento de um enxame de partículas, como pássaros ou cardumes de peixes, em busca de comida. Cada partícula representa uma solução candidata em um espaço de busca multidimensional. O algoritmo começa gerando uma população inicial de partículas, distribuídas aleatoriamente no espaço de busca. Cada partícula possui uma posição e uma velocidade, que são atualizadas em cada iteração do algoritmo. Durante a busca, cada partícula é influenciada pela sua melhor posição encontrada até o momento, que é a posição em que a partícula obteve o melhor desempenho, e pela melhor posição encontrada por todo o enxame. A atualização das posições e velocidades das partículas é feita através de fórmulas que combinam a direção da velocidade atual, a distância entre a posição atual e a melhor, e a distância entre a posição atual e a melhor. Essas fórmulas permitem que as partículas se movam em direção às melhores soluções encontradas até o momento.

## References

- António Gaspar-Cunha, Ricardo Takahashi, C. H. A. (2012). *Manual de computação evolutiva e metaheurística*. Imprensa da Universidade de Coimbra, 1 edition.
- Goldbarg, M. and (Auth.), E. G. (2012). *Grafos*.
- Lozano, M. and Trujillo, H. M. (2019). Optimizing node infiltrations in complex networks by a local search based heuristic. *Computers & Operations Research*, 111:197–213.