



## Projeto 1

### Grupo:

Gustavo Choueiri - 232014010

Giovanni Daldegan - 232002520

Este projeto visa analisar uma estrutura de rede social e avaliar a influência de seus usuários, utilizando das métricas de grafos e algoritmos a eles aplicados.

Para a representação do grafo, plotagem de gráficos e demais requisitos, foi utilizada a linguagem de programação *Python* e a plataforma *Jupyter Notebook*, com auxílio dos seguintes pacotes: *NetworkX*, *Pandas*, *NumPy*, *Matplotlib*, *seaborn* e *random*.

### 1. Coleta de dados

O conjunto de dados utilizado é um recorte da rede social Facebook, realizado por meio de uma pesquisa em 2012. Cada nó representa um usuário e as arestas, relacionamentos entre usuários. Todos os participantes foram anonimizados. Esses dados representam um grafo não direcionado de 4.039 nós e 170.174 arestas totais.

Para selecionar 2.000 nós de forma aleatória foi utilizada a amostragem do pacote *random* do Python, definindo a *seed* como 42 (ou seja, é um método pseudo-aleatório), o que garante que, em todas as execuções do código produzido, o subconjunto de nós selecionados é sempre o mesmo. O subconjunto de 2.000 nós será chamado de *V*.

### 2. Construção do grafo

Para construir um grafo dos nós selecionados, foram filtradas as arestas que tinham ambos os nós terminais pertencentes a *V*. Esse subconjunto de arestas, que somam 22.066 no total, será chamado de *E*. Com esses conjuntos, é possível criar um grafo  $G(V, E)$  utilizando o pacote *NetworkX*.

O grafo resultante *G* é um subgrafo induzido do grafo total. Sua densidade é de 0,0110 e foram detectados 79 componentes conectados no total.

### 3. Extração de métricas do grafo

#### a. Mapeamento de comunidades

Como passo inicial da análise realizada no projeto, foi feita a detecção de comunidades pelo algoritmo de Louvain, método que busca maximizar a medida de modularidade no grafo. Isso é feito agrupando os vértices em comunidades e comparando qual arranjo de comunidades tem a maior modularidade. Essa medida, que varia no intervalo  $[-1, 1]$  indica o quanto os vértices de uma mesma comunidade estão mais conectados entre si que em relação aos demais vértices do grafo.

Em código, foi utilizado o algoritmo de Louvain implementado pelo pacote *NetworkX*, que detectou 100 comunidades para a amostra de dados selecionada.

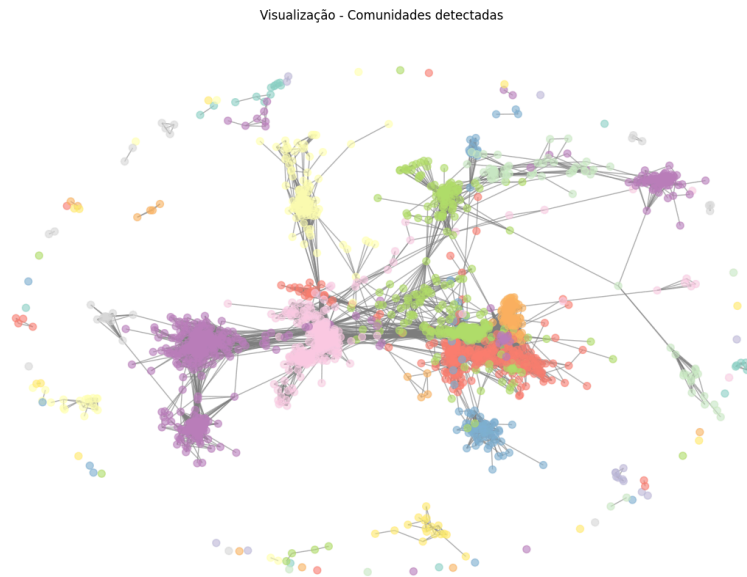


Figura 1 – Visualização das comunidades do grafo G.

#### b. Medidas de centralidade

Para cada medida de centralidade requisitada, serão apresentados sua média e desvio médio; visualização dos nós organizados de acordo com suas medidas, com os 10 nós de maior medida destacados em vermelho; e análise de influência pertinente.

##### i. Centralidade de grau

Na amostra coletada do grafo, a média da centralidade de grau foi de 0.0110, com um desvio padrão de 0.0123.

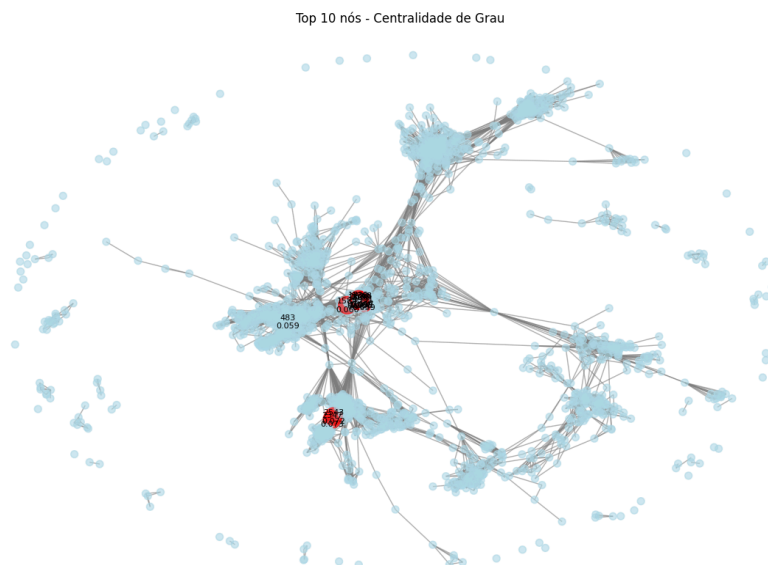


Figura 2 – Visualização de nós organizados segundo grau.

Podemos identificar os nós de maior centralidade de grau como usuários com muitos contatos e mais populares, atingindo diretamente mais pessoas. Esses podem ser *influencers*

que se comunicam com um público na rede que compartilha de gostos e interesses. Como notável na visualização, os nós de maior centralidade de grau (valores normalizados variando entre 0.0730 e 0.0565) se concentram em poucas comunidades, numeradas como 54, 58, 27 e 61 na detecção de comunidades.

ii. Centralidade de intermediação

A centralidade de intermediação teve média de 0.0018 e desvio padrão de 0.0097.

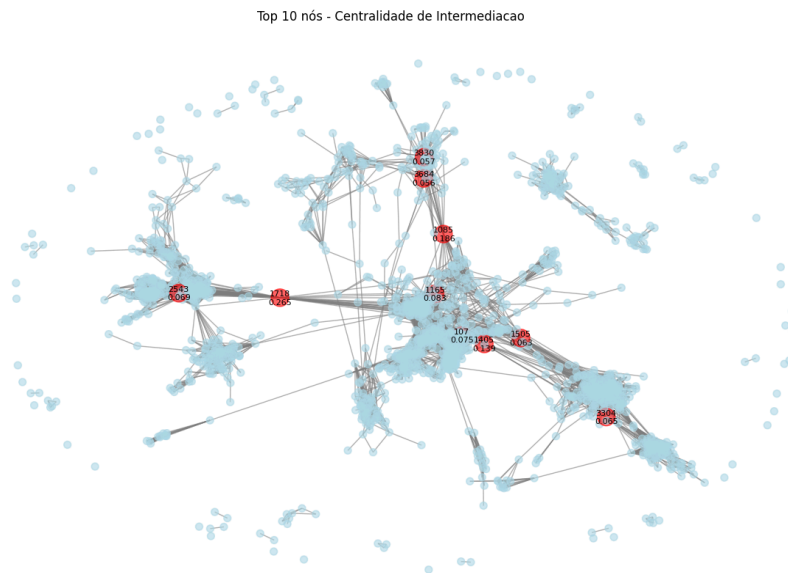


Figura 3 – Visualização de nós organizados segundo intermediação.

Os nós com maior centralidade de intermediação podem atuar como “pontes” entre diferentes partes da rede, mediando grande parte do fluxo de informações e trazendo uma cobertura maior sobre a rede. Os nós mais significativos nessa métrica (variam entre 0.2652 e 0.0562) se encontram mais distantes, pertencendo a um maior número de comunidades: 50, 61, 54, 27, 78, 79.

iii. Centralidade de proximidade

A centralidade de proximidade teve média de 0.1433, com um desvio padrão de 0,0576.

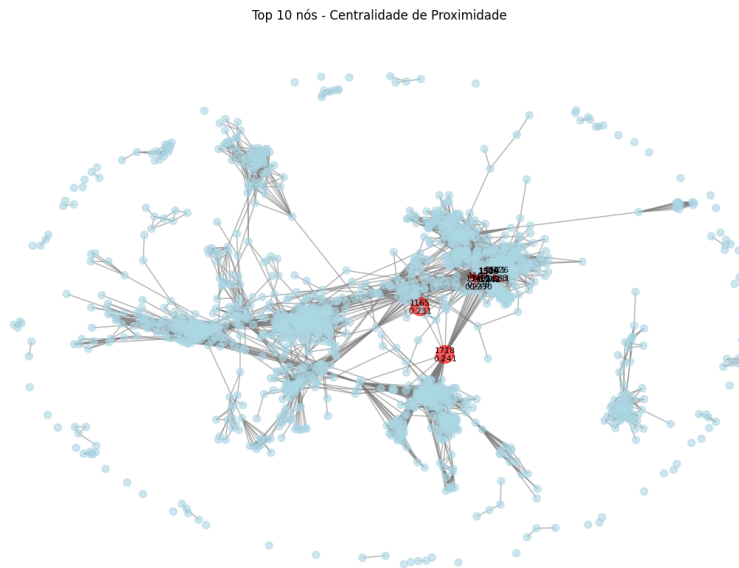


Figura 4 – Visualização de nós organizados segundo proximidade.

Os usuários com maior centralidade de proximidade estão mais conectados com toda a rede, indicando que podem propagar informações mais rapidamente. Seus nós mais influentes (com valores entre 0.2417 e 0.2302) se concentram nas comunidades 54, 50 e 51.

#### iv. Centralidade de autovetor

A média da centralidade de autovetor foi de 0.0055, com um desvio padrão de 0.0217.

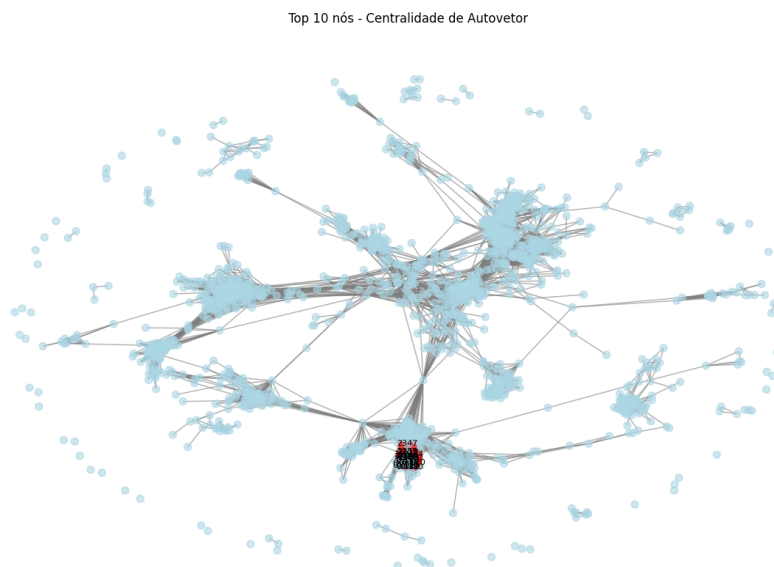


Figura 5 – Visualização de nós organizados segundo autovetor.

Nós com maior centralidade de autovetor têm mais vizinhos influentes, significando que estão no centro de comunidades mais significativas. Se esses nós forem fontes de informações para os nós vizinhos, pode haver um efeito cascata em que as informações trazidas pelo nó de maior autovetor são propagadas pelos seus vizinhos, cujos vizinhos

também podem propagá-las, assim sucessivamente. Todos os 10 nós de maior centralidade de autovetor identificados (seus valores vão de 0.1221 a 0.1177) pertencem à comunidade 58.

#### 4. Análise final

Consideradas todas as comunidades detectadas e medidas de centralidades calculadas, é perceptível que algumas comunidades são mais significativas em várias métricas, demonstrando uma forte conexão entre seus vértices e também com outras comunidades, possivelmente representando um núcleo da rede amostrada.

Porém, ao realizar uma análise similar para os vértices em particular, pouquíssimos se destacam em todas as métricas, indicando que seus usuários correspondentes desempenham diferentes papéis na rede. A avaliação da correlação entre as métricas de todos os vértices está exposta no seguinte mapa de calor.

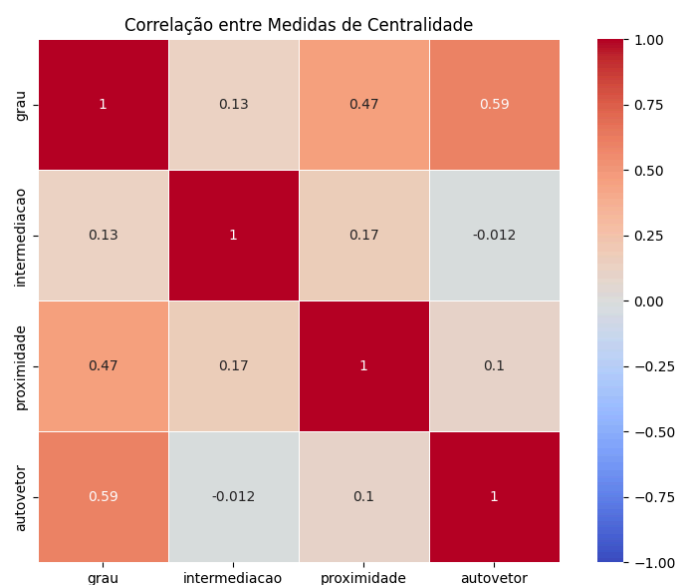


Figura 6 – Mapa de calor da correlação entre medidas de centralidade.

Apenas a correlação entre as centralidades de grau e autovetor são mais significativas, o que significa que usuários com muitos contatos acabam tendo vizinhos mais influentes. As demais métricas não apresentam fortes correlações entre si.

#### 5. Referências bibliográficas

WIKIPEDIA. *In: Wikipedia: Louvain method*. [S. l.], 29 set. 2025. Disponível em: [https://en.wikipedia.org/wiki/Louvain\\_method](https://en.wikipedia.org/wiki/Louvain_method). Acesso em: 20 out. 2025.

DÍBIO Leandro Borges. *In: Teoria e Aplicação de Grafos: Aula 7*. [S. l.], 4 set. 2025. Disponível em: <https://aprender3.unb.br>. Acesso em: 19 out. 2025.