



UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO

CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA

BACHARELADO EM SISTEMAS DE INFORMAÇÃO

DISCIPLINA INTRODUÇÃO À CIÊNCIA DE REDES

CALCULANDO CLUSTERIZAÇÃO

VIVIAN RIQUE GIL FERRARO

RIO DE JANEIRO

2022

Para obter as métricas, escrevi um pequeno código na linguagem Python, utilizando a IDE PyCharm e com o auxílio do pacote (biblioteca) Networkx.

Resultados:

Rede: malha elétrica do Oeste dos EUA

```
import networkx as nx
import numpy as np

Grafo = nx.Graph()

with open("C:\\Users\\vifer\\OneDrive\\Documents\\BSI\\7(5)Periodo\\ICR\\networks\\powergrid.edgelist.txt") as file:
    for linha in file:
        Grafo.add_edge(int(str(linha).split(" ")[0]), int(str(linha).split(" ")[1]))

graus = []

for no in Grafo.nodes():
    graus.append(nx.degree(Grafo, no))

print(f"N° de vértices: {Grafo.number_of_nodes()}")
print(f"N° de Arestas: {Grafo.number_of_edges()}")
print(f"Grau médio: {np.mean(graus)}")
print(f"Densidade: {nx.density(Grafo)}")
print(f"Media Menor Caminho: {nx.average_shortest_path_length(Grafo)}")
print(f"Clusterização global: {nx.transitivity(Grafo)}")
print(f"Clusterização local média: {nx.average_clustering(Grafo)}")
```

```
C:\Users\vifer\AppData\Local\Programs\Python\Python311\python.exe "C:\Users\vifer\OneDrive\Documents\BSI\7(5)Periodo\ICR\Codigos\MetricasPowerGrid.py"
N° de vértices: 4941
N° de Arestas: 6594
Grau médio: 2.66909532483303
Densidade: 0.0005403026973346214
Media Menor Caminho: 18.989185424445708
Clusterização global: 0.10315322452860086
Clusterização local média: 0.08010361108159712

Process finished with exit code 0
```

- 1) Podemos perceber que o grafo é esparso, pela densidade que é muito pequena ($\cong 0,00054$), e a média do menor caminho entre os vértices ser bem grande ($\cong 19$). Apesar disso, o coeficiente de clusterização é bem alto ($\cong 0,10$), o que significa que há no geral, em torno de 10% da vizinhança dos vértices é conectada entre si. A média da clusterização local, ou seja, que é medida em cada nó com grau maior que um, está próxima dessa medida global ($\cong 0,08$), os clusters formados tem em torno de 8% da sua vizinhança conectada.
- 2) Coeficiente de clusterização global: 0,10315322452860086.
Média dos Coeficientes de clusterização locais: 0,08010361108159712.

- 3) Em uma rede aleatória, o coeficiente de clusterização esperado seria da mesma ordem da medida de densidade do grafo, o que não acontece nessa rede; mostrando que essa é uma rede clusterizada.

Rede: Metabolismo (*E. Coli*)

```
import networkx as nx
import numpy as np

Grafo = nx.DiGraph()

with open("C:\\Users\\vifer\\OneDrive\\Documents\\BSI\\7(5)Período\\ICR\\networks\\metabolic.edgelist.txt") as file:
    for linha in file:
        Grafo.add_edge(int(str(linha).split(" ")[0]), int(str(linha).split(" ")[1]))

    graus = []

    for no in Grafo.nodes():
        graus.append(nx.degree(Grafo, no)/2)

print(f"N° de nós: {Grafo.number_of_nodes()}")
print(f"N° de Arestas: {Grafo.number_of_edges()}")
print(f"Grau médio: {np.mean(graus)}")
print(f"Densidade: {nx.density(Grafo)}")
print(f"Media Menor Caminho: {nx.average_shortest_path_length(Grafo)}")
print(f"Clusterização global: {nx.transitivity(Grafo)}")
print(f"Clusterização local média: {nx.average_clustering(Grafo)}")
```

```
C:\Users\vifer\AppData\Local\Programs\Python\Python311\python.exe "C:\Users\vifer\OneDrive\Documents\BSI\7(5)Período\ICR\Codigos\MetricasMetabolic.py"
N° de nós: 1039
N° de Arestas: 5802
Grau médio: 5.584215591915303
Densidade: 0.005379783807240177
Media Menor Caminho: 2.571989147709466
Clusterização global: 0.032475074064231654
Clusterização local média: 0.28056141848444327

Process finished with exit code 0
```

- 1) Também é um grafo esparsos, menos que o anterior, mas com densidade $\cong 0,0054$. Nesse grafo, o índice de clusterização global também é maior do que o esperado (considerando a densidade). Uma coisa que chama bastante atenção nesse caso, é a média da clusterização local, que é muito alta, o que indica que os clusters formados são muito conectados entre si.
- 2) Coeficiente de clusterização global: 0,032475074064231654.
Média dos Coeficientes de clusterização locais: 0,28056141848444327.
- 3) Da mesma forma que a rede anterior, essa também é uma rede clusterizada, em relação ao que seria esperado para uma rede aleatória.

Para testar essa hipótese, criei um grafo não-direcionado, de forma aleatória.

Resultados obtidos (Rede aleatória):

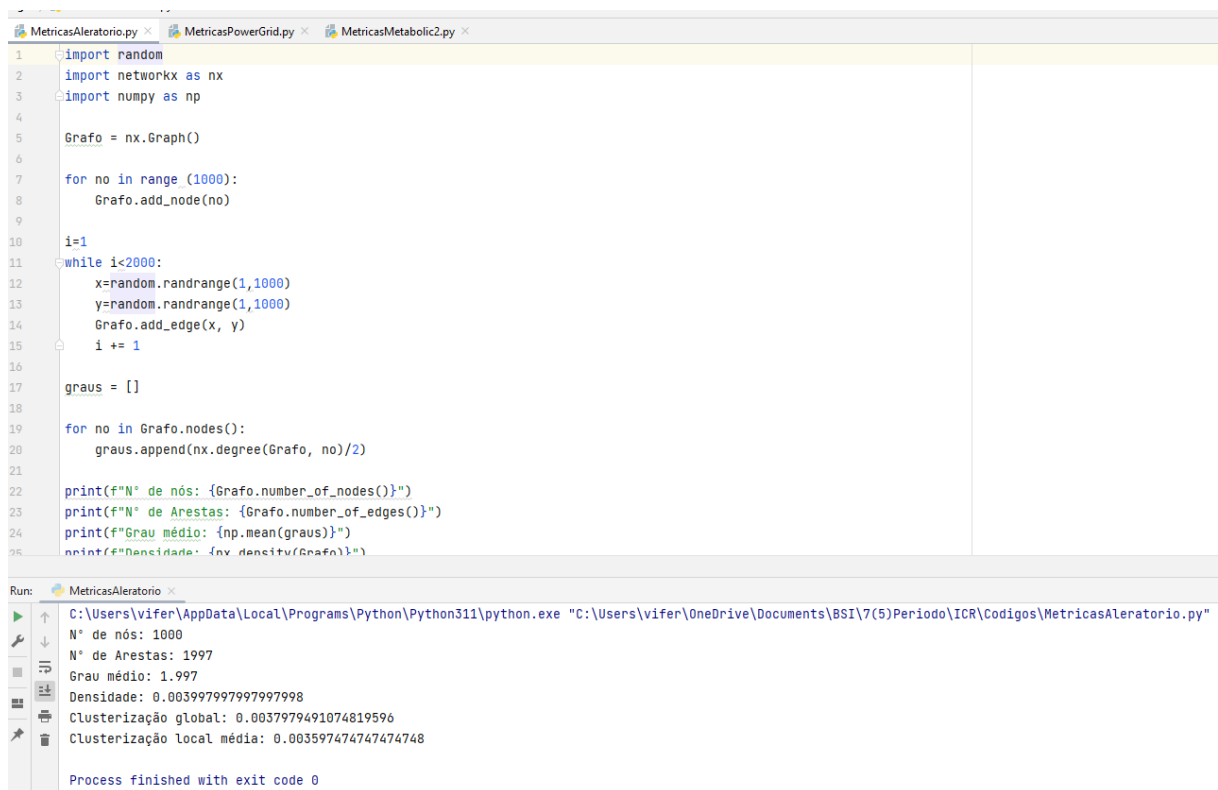
1) Densidade: 0,003997997997997998.

Coeficiente de clusterização global: 0,0037979491074819596.

Média dos Coeficientes de clusterização locais: 0,003597474747474748.

Os resultados obtidos sugerem que essa rede aleatória não é clusterizada, pois os agrupamentos (coeficiente de clusterização), obedecem a uma proporção segundo a densidade do grafo.

Repeti o teste algumas vezes, obtendo resultados semelhantes em todas elas.



```
1 import random
2 import networkx as nx
3 import numpy as np
4
5 Grafo = nx.Graph()
6
7 for no in range(1000):
8     Grafo.add_node(no)
9
10 i=1
11 while i<2000:
12     x=random.randrange(1,1000)
13     y=random.randrange(1,1000)
14     Grafo.add_edge(x, y)
15     i += 1
16
17 graus = []
18
19 for no in Grafo.nodes():
20     graus.append(nx.degree(Grafo, no)/2)
21
22 print(f"N° de nós: {Grafo.number_of_nodes()}")
23 print(f"N° de Arestas: {Grafo.number_of_edges()}")
24 print(f"Grau médio: {np.mean(graus)}")
25 print(f"Densidade: {nx.density(Grafo)}")
```

Run: MetricsAleratorio

```
C:\Users\viifer\AppData\Local\Programs\Python\Python311\python.exe "C:\Users\viifer\OneDrive\Documents\BSI\7(5)Periodo\ICR\Codigos\MetricasAleratorio.py"
N° de nós: 1000
N° de Arestas: 1997
Grau médio: 1.997
Densidade: 0.003997997997997998
Clusterização global: 0.0037979491074819596
Clusterização local média: 0.003597474747474748

Process finished with exit code 0
```