



Departamento de
Ciência da Computação



Universidade Federal
de São João del-Rei

Trabalho Redes Complexas

Análise mais elaborada de Redes Complexas (Redes Reais e Sintáticas)

Pedro Antônio Machado Costa

São João Del Rei

2023

1. Introdução

Redes Reais e Redes Sintéticas são altamente utilizadas em diversos conceitos no mundo atual. Tendo em vista essa definição, foi realizado um estudo e uma análise dessas redes e seus tipos em forma de uma atividade, com o intuito de colocar em prática sobre como essas redes funcionam e como elas caracterizam suas próprias redes. Ao longo da atividade, será demonstrado e explicado sobre as escolhas da metodologia e também será aplicado alguns conceitos aprendidos em sala de aula.

2. Metodologia

De início, é importante decidir sobre qual linguagem de programação será utilizada ao longo de todo o processo de análise dessas Redes. Python foi escolhida devido a sua ampla diversidade e simplicidade, auxiliando fortemente com suas bibliotecas e ajuda dinâmica ao longo do desenvolvimento.

As bibliotecas utilizadas foram a NetworkX, para geração das redes e análise das mesmas, e a Matplotlib, para geração de gráficos e tabelas a partir dos resultados obtidos, sendo de grande importância para uma alta análise posterior de todos os resultados.

Ao longo de toda atividade, foi realizado uma linha de raciocínio para todas as redes e suas análises:

1. Geração da rede;
2. Realização das análises e conceitos de cada rede;
3. Plotagem da rede e geração de gráficos e tabelas para os resultados;

A partir disso, pode-se analisar as diferentes redes de maneira prática e intuitiva, não sendo de difícil realização e não possuindo dificuldades. Códigos e algoritmos estão na pasta correspondente deste relatório. Com isso, resultados foram obtidos e analisados separadamente, podendo ser visualizados no tópico seguinte.

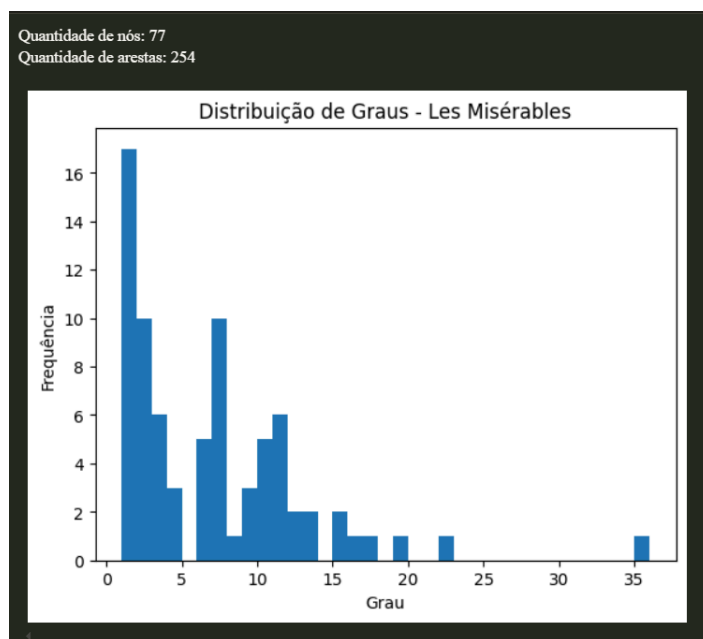
3. Resultados

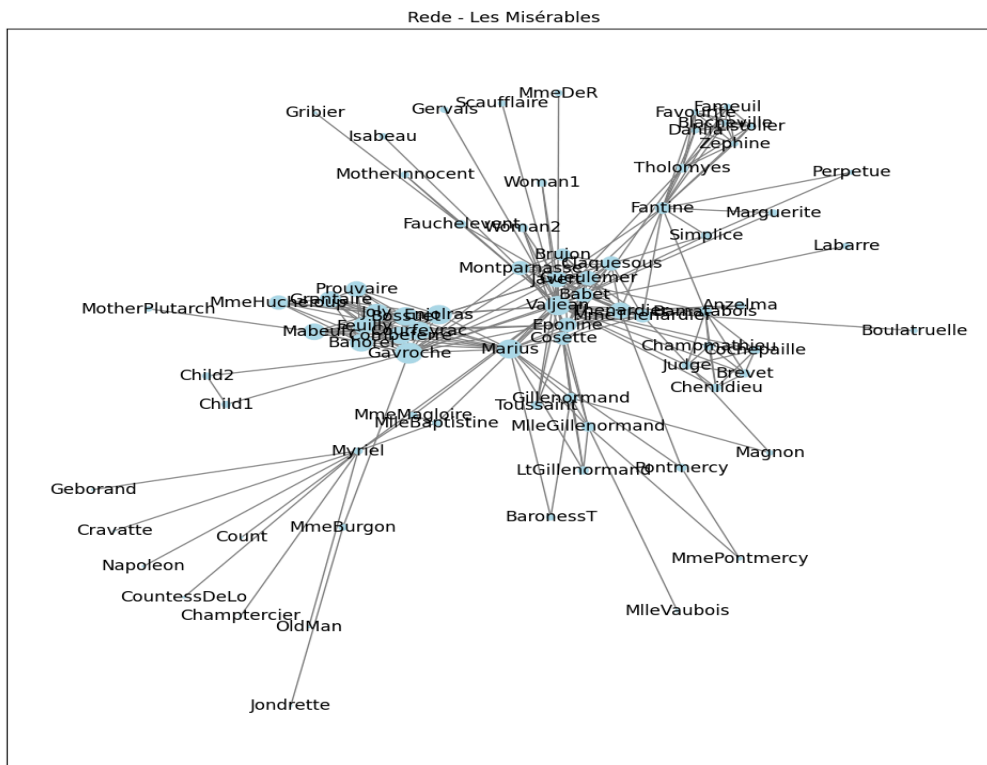
A partir dos resultados obtidos, entre redes plotadas, gráficos e tabelas, foi possível realizar uma ampla análise, caracterizadas da seguinte maneira.

3.1 Redes Reais

Foi realizado uma análise em duas redes Reais (Les Misérables e Karate Club), sendo gerado os seguintes resultados:

Rede Les Misérables





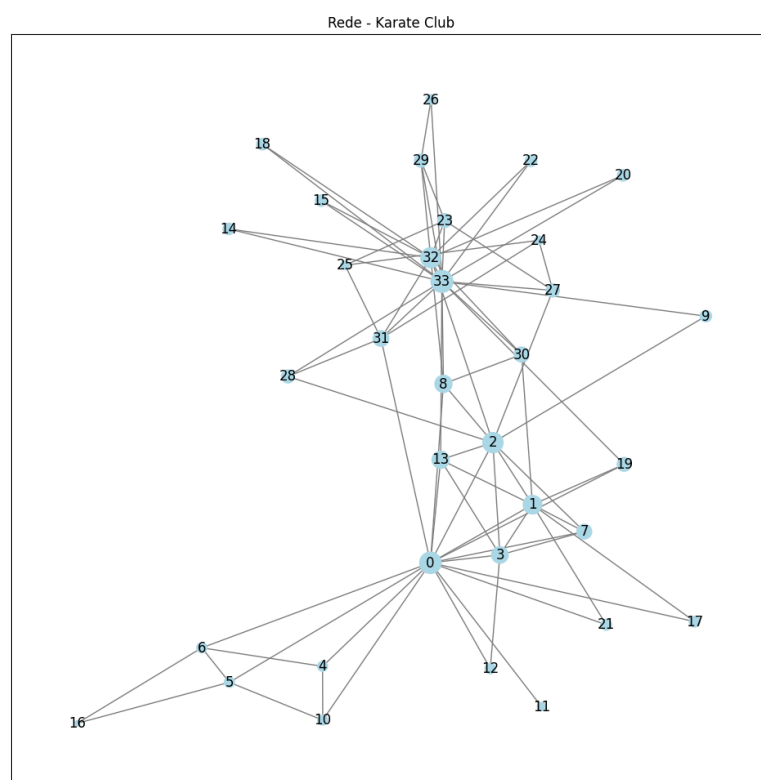
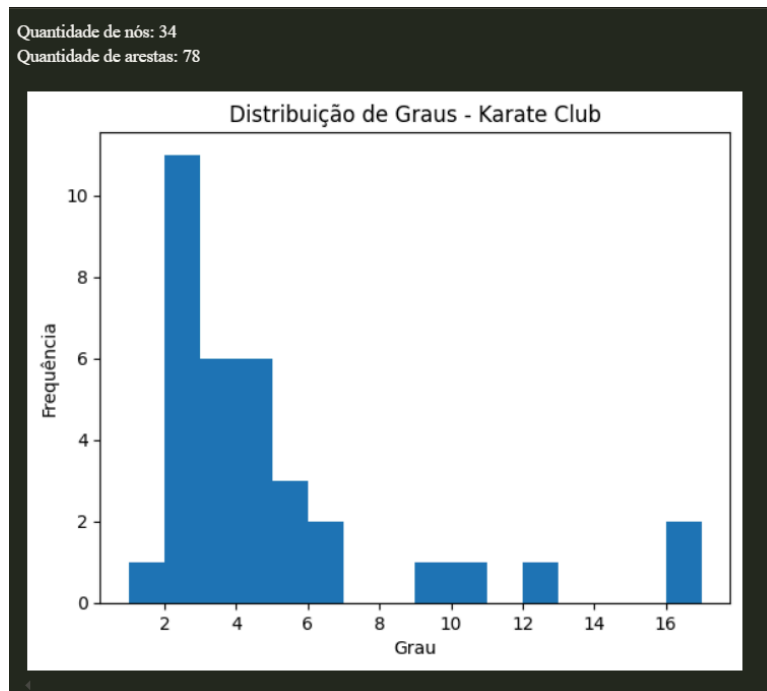
Grau Médio (Les Misérables): 6.597402597402597
Coeficiente de Clustering Médio (Les Misérables): 0.5731367499320135
Distância Média (Les Misérables): 2.6411483253588517

Análise:

- **Grau Médio:** O resultado obtido indica que, em média, cada nó está conectado a aproximadamente 6,6 outros nós. Isso sugere uma rede moderadamente densa, onde os personagens tendem a interagir com vários outros personagens.
- **Coeficiente de Clustering Médio:** O resultado indica uma alta presença de triângulos na rede, ou seja, se dois personagens estão conectados a um terceiro, há uma alta probabilidade de que eles também estejam conectados entre si. Isso é típico em redes sociais ou de interação, onde grupos ou cliques de personagens estão fortemente interconectados.
- **Distância Média:** O resultado indica que a rede é relativamente "pequena", no sentido de que qualquer nó pode ser alcançado a partir de outro em cerca de 2,6 passos em média. Isso é indicativo de uma

rede bem conectada, com poucos "saltos" necessários para atravessá-la.

Rede Karate Club



Grau Médio (Karate Club): 4.588235294117647

Coeficiente de Clustering Médio (Karate Club): 0.5706384782076823

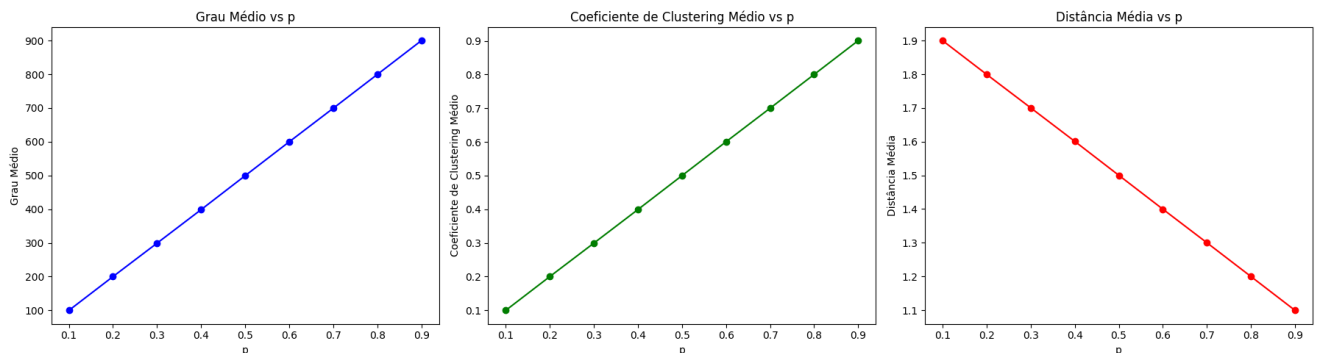
Distância Média (Karate Club): 2.408199643493761

Análise:

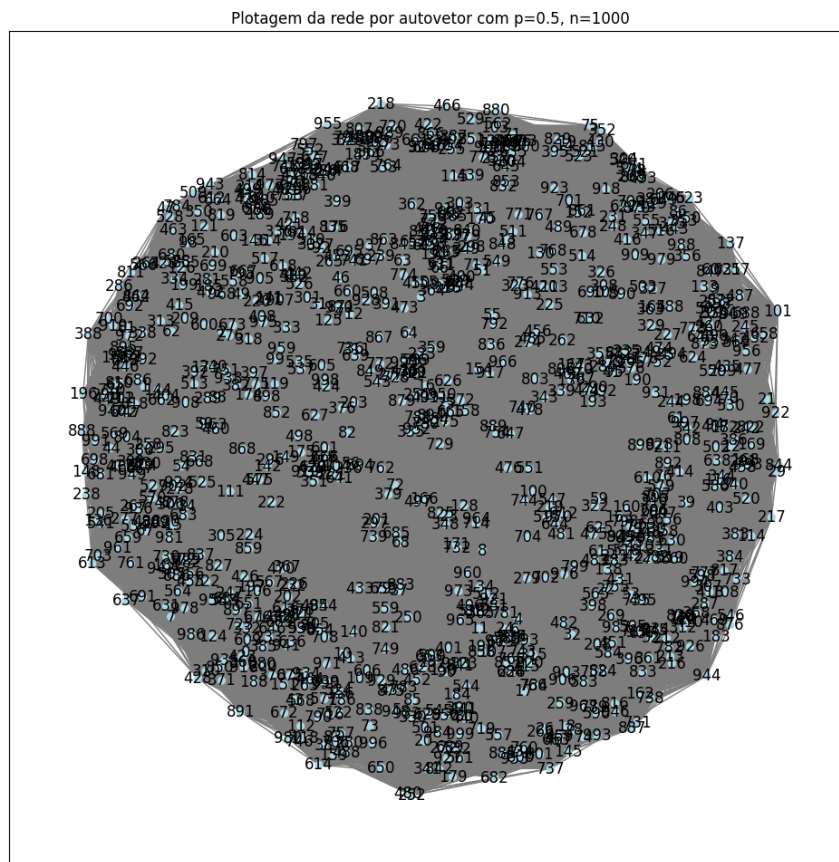
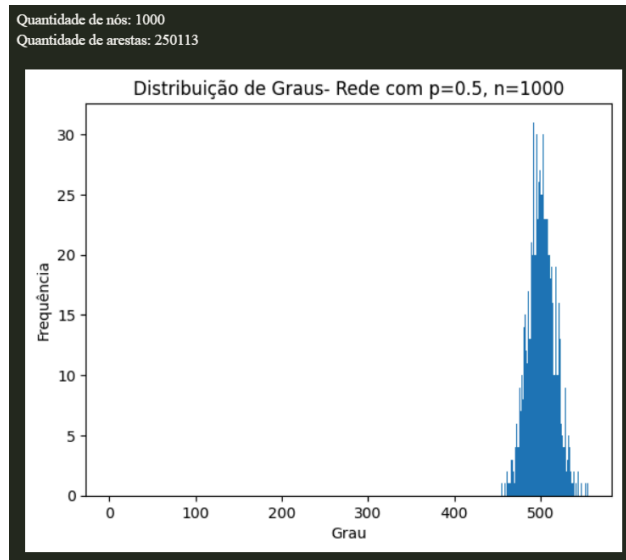
- **Grau Médio:** Esse grau médio obtido indica que a rede não é extremamente densa, mas ainda assim apresenta um nível razoável de conectividade. Em termos sociais, isso pode significar que os membros têm um círculo social relativamente pequeno dentro do clube.
- **Coeficiente de Clustering Médio:** O resultado obtido é bastante alto, o que significa que, se dois membros do clube têm um amigo em comum, há uma probabilidade de 57.1% de que esses dois membros também sejam amigos.
- **Distância Média:** O valor obtido indica que, em média, quaisquer dois membros do clube podem ser conectados em cerca de dois passos ou menos. Essa métrica sugere que a rede é bastante compacta.

3.2 Rede Erdos-Rényi

Foi gerado redes a partir da biblioteca NetworkX do Python, sendo gerada pra diferentes parâmetros: parâmetro $n = 1000$ e variando o parâmetro p de 0.1 a 0.9, de 0.1 em 0.1 (ou seja, $p = 0.1, 0.2, 0.3, \dots, 0.9$). Com isso, foi obtido os seguintes resultados:



(p = 0.1: Grau Médio = 99.932, Coeficiente de Clustering Médio = 0.09972913825173606, Distância Média = 1.900016016016016)
 (p = 0.2: Grau Médio = 200.082, Coeficiente de Clustering Médio = 0.2000944527468203, Distância Média = 1.7997177177177177)
 (p = 0.3: Grau Médio = 299.16, Coeficiente de Clustering Médio = 0.2993762948620784, Distância Média = 1.7005405405405405)
 (p = 0.4: Grau Médio = 398.574, Coeficiente de Clustering Médio = 0.39908888799548664, Distância Média = 1.601027027027027)
 (p = 0.5: Grau Médio = 499.252, Coeficiente de Clustering Médio = 0.4997761642398445, Distância Média = 1.5002482482482482)
 (p = 0.6: Grau Médio = 599.696, Coeficiente de Clustering Médio = 0.6003266613011751, Distância Média = 1.3997037037037037)
 (p = 0.7: Grau Médio = 699.256, Coeficiente de Clustering Médio = 0.6999954687642072, Distância Média = 1.300044044044044)
 (p = 0.8: Grau Médio = 799.448, Coeficiente de Clustering Médio = 0.8002226944625008, Distância Média = 1.1997517517517517)
 (p = 0.9: Grau Médio = 899.812, Coeficiente de Clustering Médio = 0.9007148949253496, Distância Média = 1.0992872872872872)

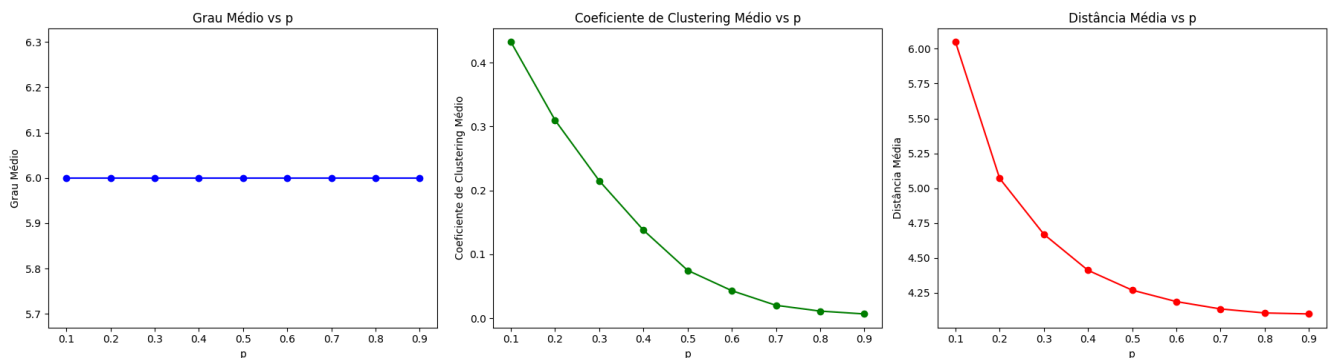


Análise:

- **Grau Médio:** O grau médio aumenta linearmente com p porque, conforme a probabilidade de conexão aumenta, cada nó se conecta a mais vizinhos.
- **Coeficiente de Clustering Médio:** Em redes Erdős-Rényi, o coeficiente de clustering é aproximadamente p , já que a probabilidade de um triângulo existir entre três nós conectados é p .
- **Distância Média:** Inicialmente, a distância média diminui rapidamente à medida que p aumenta, porque a rede se torna mais conectada e "menor". No entanto, à medida que p se aproxima de 1, a rede se torna quase completamente conectada, e a distância média atinge um valor mínimo.

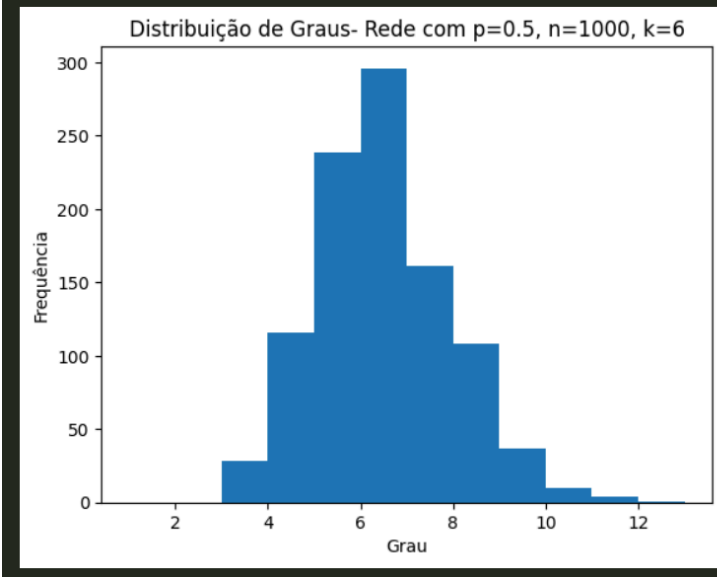
3.3 Rede Watts-Strogatz

Foi gerado redes a partir da biblioteca NetworkX do Python, sendo gerada pra diferentes parâmetros: parâmetro $n = 1000$, $k = 6$ e variando o parâmetro p de 0.1 a 0.9, de 0.1 em 0.1 (ou seja, $p = 0.1, 0.2, 0.3, \dots, 0.9$). Com isso, foi obtido os seguintes resultados:

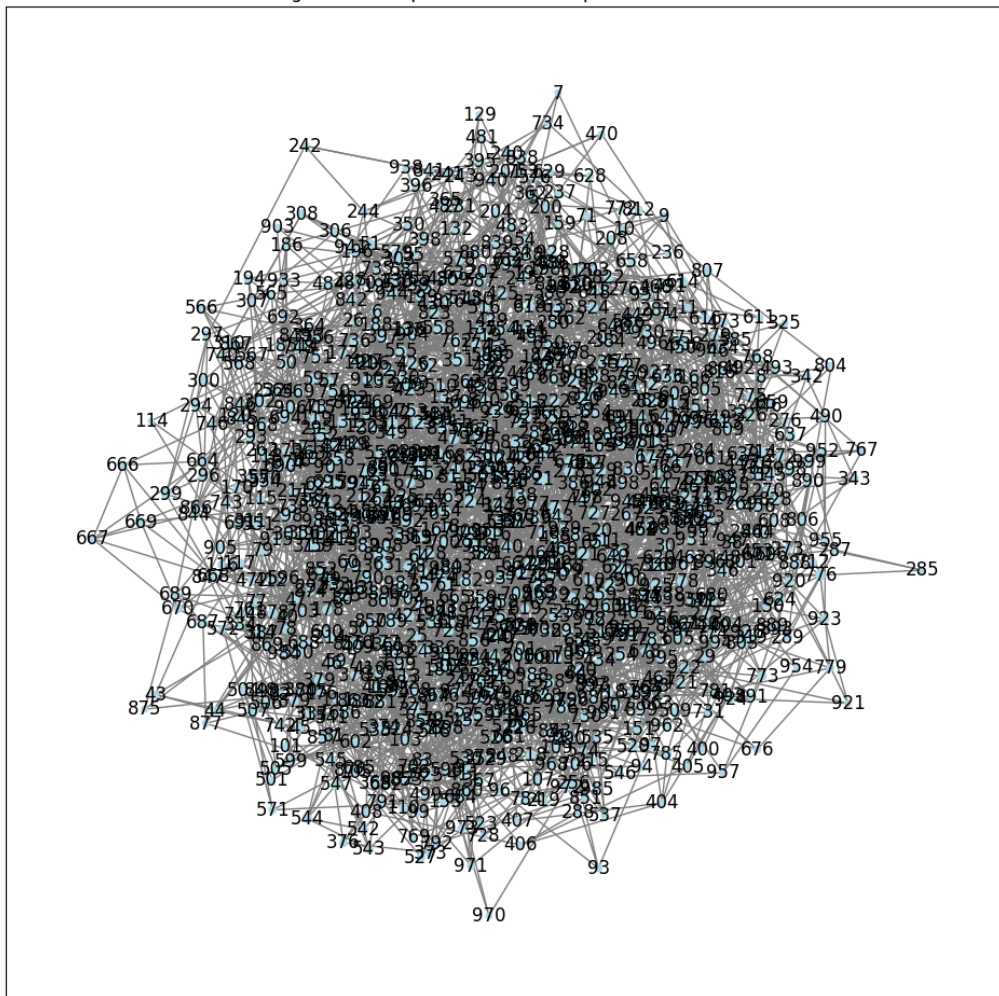


```
(p = 0.1: Grau Médio = 6.0, Coeficiente de Clustering Médio = 0.43263730158730157, Distância Média = 6.048624624624624)
(p = 0.2: Grau Médio = 6.0, Coeficiente de Clustering Médio = 0.31003838383838384, Distância Média = 5.070362362362363)
(p = 0.3: Grau Médio = 6.0, Coeficiente de Clustering Médio = 0.21514877344877345, Distância Média = 4.670318318318318)
(p = 0.4: Grau Médio = 6.0, Coeficiente de Clustering Médio = 0.13778080808080806, Distância Média = 4.411473473473474)
(p = 0.5: Grau Médio = 6.0, Coeficiente de Clustering Médio = 0.07483621933621933, Distância Média = 4.2692912912912915)
(p = 0.6: Grau Médio = 6.0, Coeficiente de Clustering Médio = 0.042987734487734484, Distância Média = 4.187245245245245)
(p = 0.7: Grau Médio = 6.0, Coeficiente de Clustering Médio = 0.020046681096681094, Distância Média = 4.134484484484484)
(p = 0.8: Grau Médio = 6.0, Coeficiente de Clustering Médio = 0.011003896103896103, Distância Média = 4.105821821821822)
(p = 0.9: Grau Médio = 6.0, Coeficiente de Clustering Médio = 0.006551637251637252, Distância Média = 4.098736736736737)
```


Quantidade de nós: 1000
Quantidade de arestas: 3000



Plotagem da rede por autovetor com $p=0.5$, $n=1000$, $k=6$



Análise:

- **Grau Médio:**

- O grau médio é constante e igual a 6 para todas as redes geradas. Isso é esperado, pois na rede Watts-Strogatz, o número de conexões (k) de cada nó com seus vizinhos mais próximos é fixo, independentemente do valor de p .

- **Coefficiente de Clustering Médio:**

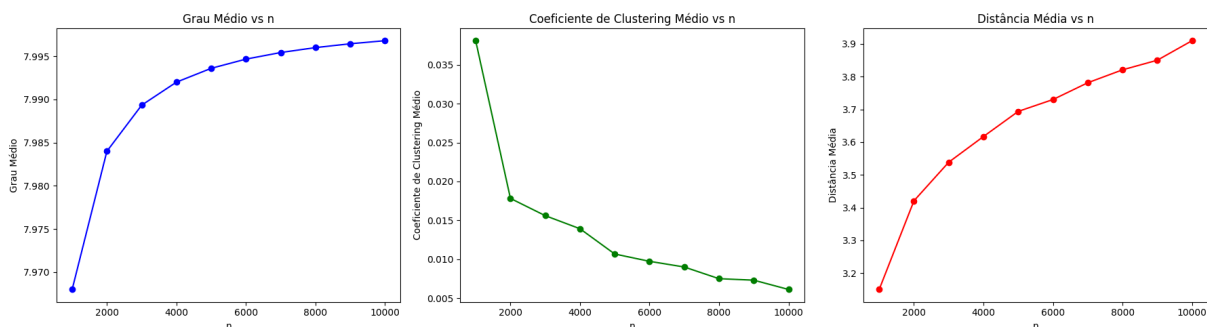
- Em $p=0$, a rede é regular e altamente clusterizada, o que significa que os nós têm muitos triângulos (alta conectividade local).
- Conforme p aumenta, mais arestas são reconfiguradas aleatoriamente, destruindo triângulos e, conseqüentemente, reduzindo o coeficiente de clustering.

- **Distância Média:**

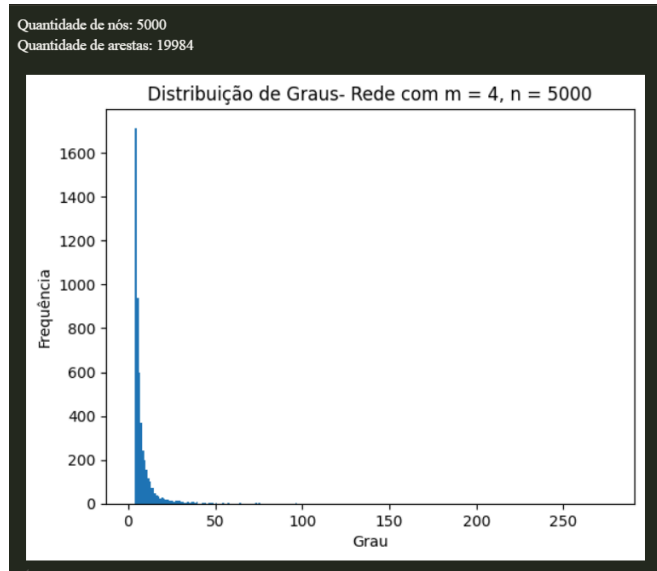
- Com p pequeno, a rede ainda mantém grande parte de sua estrutura regular, resultando em distâncias médias maiores entre os nós.
- Quando p aumenta, a presença de atalhos (conexões aleatórias) reduz significativamente a distância média, aproximando a rede das propriedades de "small-world".
- Para p alto, a distância média se estabiliza, indicando que a maioria dos nós está conectada por caminhos curtos, uma característica das redes "small-world".

3.4 Rede Barabási-Albert

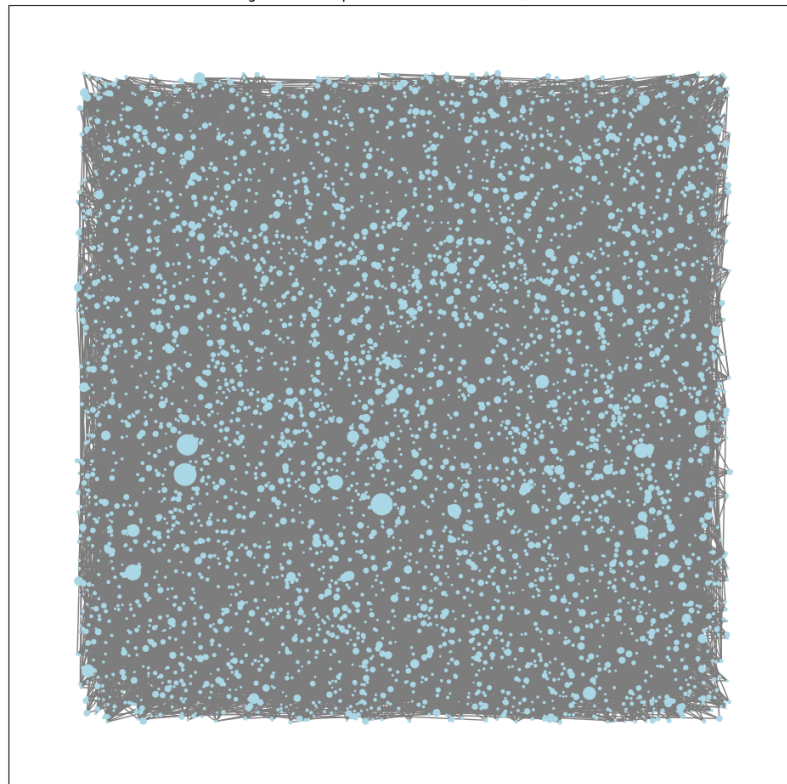
Foi gerado redes a partir da biblioteca NetworkX do Python, sendo gerada pra diferentes parâmetros: parâmetro $m = 4$ e variando o parâmetro n de 1000 a 10000, de 1000 em 1000 (ou seja, $m = 1000, 2000, 3000, \dots, 10000$). Com isso, foi obtido os seguintes resultados:



(p = 1000.0: Grau Médio = 7.968, Coeficiente de Clustering Médio = 0.038111800499664616, Distância Média = 3.1502522522522525)
 (p = 2000.0: Grau Médio = 7.984, Coeficiente de Clustering Médio = 0.017820006443823193, Distância Média = 3.420911955977989)
 (p = 3000.0: Grau Médio = 7.989333333333334, Coeficiente de Clustering Médio = 0.015602645981210275, Distância Média = 3.5387355785261754)
 (p = 4000.0: Grau Médio = 7.992, Coeficiente de Clustering Médio = 0.013920343554640568, Distância Média = 3.617215803950988)
 (p = 5000.0: Grau Médio = 7.9936, Coeficiente de Clustering Médio = 0.010666936962585078, Distância Média = 3.693937827565513)
 (p = 6000.0: Grau Médio = 7.994666666666666, Coeficiente de Clustering Médio = 0.009726843792543025, Distância Média = 3.7301505806523307)
 (p = 7000.0: Grau Médio = 7.9954285714285716, Coeficiente de Clustering Médio = 0.008995920035410802, Distância Média = 3.78137276753822)
 (p = 8000.0: Grau Médio = 7.996, Coeficiente de Clustering Médio = 0.007491928269626778, Distância Média = 3.8208348856107013)
 (p = 9000.0: Grau Médio = 7.996444444444444, Coeficiente de Clustering Médio = 0.00730096597837282, Distância Média = 3.850149942586213)
 (p = 10000.0: Grau Médio = 7.9968, Coeficiente de Clustering Médio = 0.0061053357477139965, Distância Média = 3.9098123412341232)



Plotagem da rede por autovetor com m = 4, n = 5000



Análise:

- **Grau Médio:**

- No modelo BA, o grau médio tende a ser $2 \times m$, onde $m=4$ é o número de arestas adicionadas a cada novo nó. Portanto, o grau médio esperado é aproximadamente 8, que corresponde bem aos valores obtidos;
- Pequenas variações são normais e podem ocorrer devido à aleatoriedade no processo de crescimento da rede.

- **Coeficiente de Clustering Médio:**

- Redes BA tendem a ter coeficientes de clustering baixos, especialmente em redes maiores, porque novos nós preferencialmente se conectam a nós de alto grau, o que cria hubs mas não necessariamente triângulos;
- A diminuição do coeficiente de clustering com o aumento de n reflete o fato de que, em redes maiores, a probabilidade de formação de triângulos diminui;

- **Distância Média:**

- No modelo BA, a distância média tende a crescer de forma logarítmica em relação ao número de nós, o que é consistente com as propriedades de redes de "scale-free";
- A distância média aumenta à medida que a rede cresce, mas de forma lenta, o que indica que mesmo em redes grandes, a distância entre dois nós ainda é relativamente curta.

4. Conclusão

Após realizar uma análise comparativa entre redes reais e sintéticas, os resultados obtidos demonstraram diferenças significativas em suas estruturas e propriedades. Através de gráficos e tabelas, foi possível identificar padrões distintos de conectividade, distribuição de graus e outras características entre as redes analisadas. Os grafos gerados permitiram uma visualização clara dessas diferenças, evidenciando, por exemplo, como redes reais tendem a apresentar características como assortatividade e clusters mais evidentes,

enquanto redes sintéticas, dependendo do modelo utilizado, podem seguir distribuições mais homogêneas ou aleatórias.

Além disso, as visualizações gráficas e análises quantitativas fornecem uma compreensão mais profunda sobre a complexidade e as particularidades de cada rede analisada, o que é fundamental para aplicações práticas em áreas como ciência de dados, engenharia e ciências sociais.

Por fim, a realização desta atividade foi de constante aprendizado e evolução, sendo de bastante proveito para o decorrer do curso e para a vida de um cientista de computação e analista de dados.