

Universidade Federal do Rio de Janeiro
PESC/COPPE/UFRJ
CPS765 - Redes Complexas 2023/3

Trabalho Prático 1

Filipe Prates - DRE: 116011311

https://github.com/FilipePrates/CPS765-Complex_Networks_Exploratory_Graph_Analysis

Enunciado

Nesta tarefa você deve instalar e se familiarizar com alguma biblioteca de análise de redes, como graph-tool, NetworkX, **igraph**. Você deve escolher ao menos **três redes** disponíveis nos repositórios abaixo (ou outros repositórios), e caracterizá-las utilizando diferentes métricas como grau, distância, tamanho das componentes conexas, e clusterização (você pode escolher as outras). Para cada métrica analisada, calcule as seguintes estatísticas: **máximo, mínimo, média, mediana, desvio padrão, e distribuição empírica (CCDF)**. Faça uma tabela com os dados das diferentes redes, e um gráfico com as diferentes redes.

Dados Escolhidos

1. "Kaggle chess players (2010)" Primeira rede escolhida para o trabalho. Consiste em um grafo com nós representando jogadores de alto nível de xadrez online, e arestas os jogos entre esses jogadores. A direção da aresta indica a cor das peças, e o peso o resultado do jogo. [$Nodes = 7301$; $Edges = 65053$]
 - Graus:
 - (a) Max: 280
 - (b) Min: 0
 - (c) Média: 17.817858121062724
 - (d) Mediana: 7
 - (e) Desvio Padrão: 28.876823622324924
 - (f) CCDF (1):
 - Distâncias (excluindo Infinito):
 - (a) Max: 13.0
 - (b) Min: 0.0
 - (c) Média: 4.009875245912235
 - (d) Mediana: 4.0
 - (e) Desvio Padrão: 1.0689429074208772
 - (f) CCDF (2)
 - $|CC|$:
 - (a) Max: 7115
 - (b) Min: 1

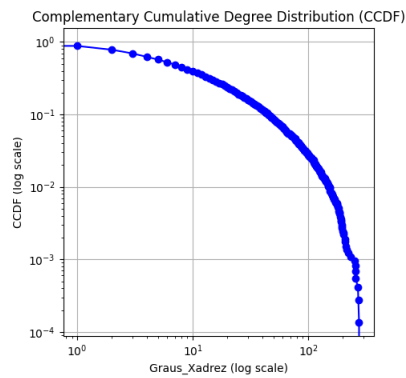


Figure 1: CCDF (1): Graus da Rede de Jogos de Xadrez

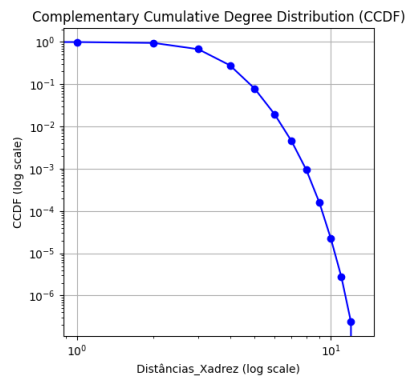


Figure 2: CCDF (2): Distâncias da Rede de Jogos de Xadrez

- (c) Média: 93.61538461538461
- (d) Mediana: 2
- (e) Desvio Padrão: 805.3407156798098
- (f) CCDF (3)
- Betweenness:
 - (a) Max: 263805.1353626446
 - (b) Min: 0.0
 - (c) Média: 10434.023555190315
 - (d) Mediana: 1307.7004230226116
 - (e) Desvio Padrão: 23229.127661514773
 - (f) CCDF (4)

Todos os jogadores possuem jogos, não observamos uma diferença grande de ordens de grandeza nos graus dos nós, nem em suas distâncias. Uma rede bem conectada, onde os

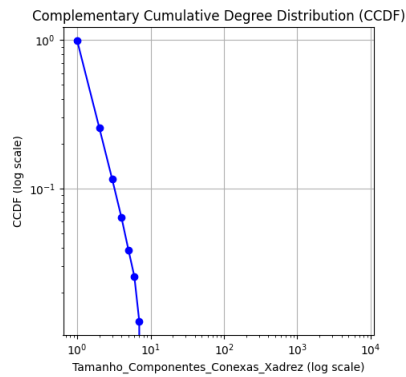


Figure 3: CCDF (3): Tamanho das Componentes Conexas da Rede de Jogos de Xadrez

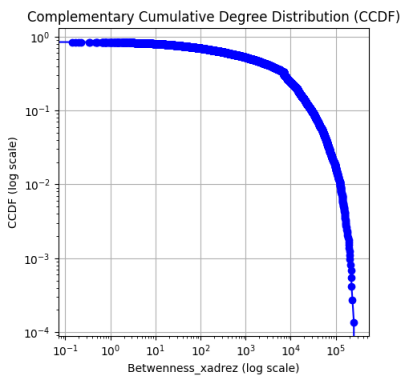


Figure 4: CCDF (4): Betweness dos vértices da Rede de Jogos de Xadrez

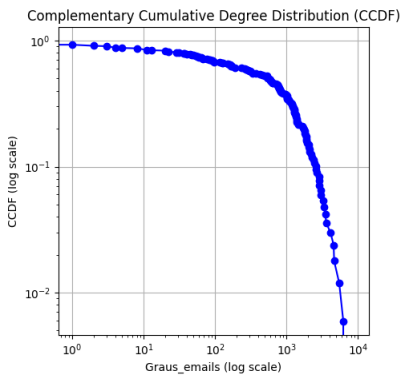


Figure 5: CCDF (5): Graus da Rede de emails corporativos

valores de Betweness são em geral altos e possuindo poucas componentes conexas, sendo uma dominante. Porém possui diâmetro maior do que o esperado.

2. "Manufacturing company emails (2010)" Rede de emails entre colaboradores da empresa Radoslaw de manufatura. Cada nó representa um trabalhador, e cada aresta um email trocado entre colegas. [Nodes = 167 ; Edges = 82967]

- Graus:
 - (a) Max: 9053
 - (b) Min: 0
 - (c) Média: 987.2261904761905
 - (d) Mediana: 575
 - (e) Desvio Padrão: 1311.6600566534498
 - (f) CCDF (5):
- Distâncias:
 - (a) Max: 5.0
 - (b) Min: 0.0
 - (c) Média: 1.9555396199354607
 - (d) Mediana: 2.0
 - (e) Desvio Padrão: 0.7008001432558403
 - (f) CCDF (6)
- |CC|:
 - (a) Max: 167
 - (b) Min: 1
 - (c) Média: 84.0
 - (d) Mediana: 167
 - (e) Desvio Padrão: 117.37972567696688

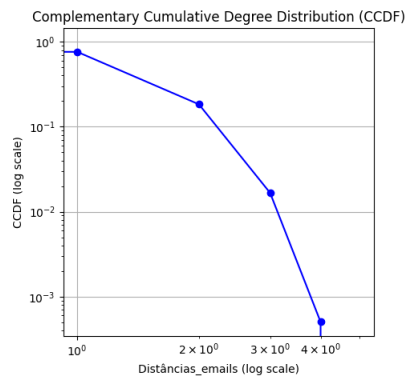


Figure 6: CCDF (6): Distâncias da Rede de emails corporativos

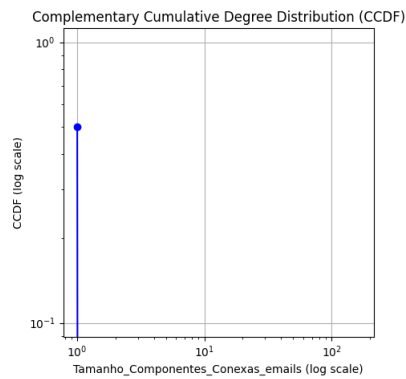


Figure 7: CCDF (7): Tamanho das Componentes Conexas da Rede de emails corporativos

- (f) CCDF (7)
- Betweness:

(a) Max: 986.035562949238

(b) Min: 0.0

(c) Média: 79.81547619047605

(d) Mediana: 13.855632616714734

(e) Desvio Padrão: 161.87530298124815

(f) CCDF (8)

Existe trabalhadores que não enviaram nem receberam emails de ninguém. Tirando estes, a rede é bem conectada, com diâmetro total baixo. Possui apenas duas componentes conexas, sendo a segunda do trabalhador que não enviou emails. A diferença grande entre a betweness máxima e a média/mediana indica que poucos colaboradores são centrais na comunicação da empresa por emails.

3. "Openflights airport network (2010)" A rede de voos entre todos os aeroportos comerciais nos

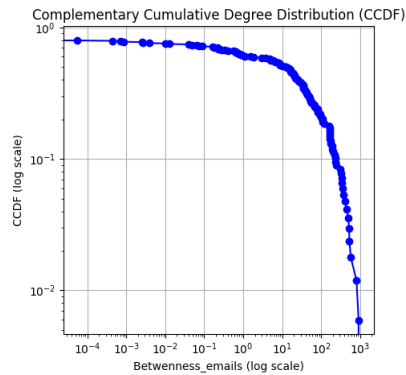


Figure 8: CCDF (8): Betweness dos vértices da Rede de emails corporativos

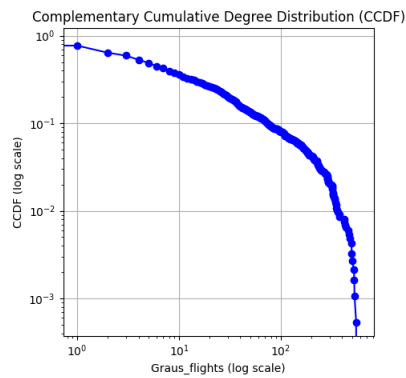


Figure 9: CCDF (9): Graus da Rede de voos entre aeroportos

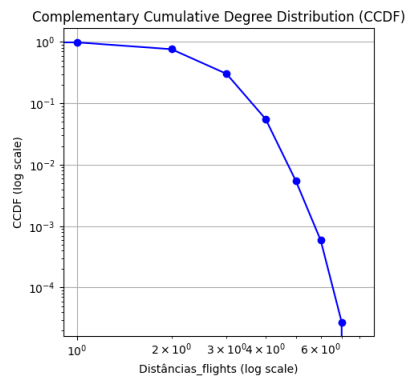


Figure 10: CCDF (10): Distâncias da Rede de voos entre aeroportos

Estados Unidos. Dados derivados do U.S. Bureau of Transportation Statistics (BTS). Pesos representam o número de assentos disponíveis nos aviões entre esses aeroportos. [$Nodes = 1858$; $Edges = 28234$]

- Graus
 - (a) Max: 596
 - (b) Min: 0
 - (c) Média: 30.37547068316299
 - (d) Mediana: 5
 - (e) Desvio Padrão: 71.70021131200969
 - (f) CCDF (9):
- Distâncias
 - (a) Max: 8.0
 - (b) Min: 0.0
 - (c) Média: 3.112818954526309
 - (d) Mediana: 3.0
 - (e) Desvio Padrão: 0.8795860636389855
 - (f) CCDF (10)
- $|CC|$:
 - (a) Max: 1572
 - (b) Min: 1
 - (c) Média: 6.477351916376307
 - (d) Mediana: 1
 - (e) Desvio Padrão: 92.73301029599729
 - (f) CCDF (11)
- Betweenness:

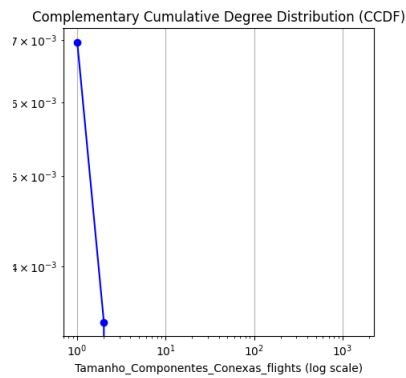


Figure 11: CCDF (11): Tamanho das Componentes Conexas da Rede de voos entre aeroportos

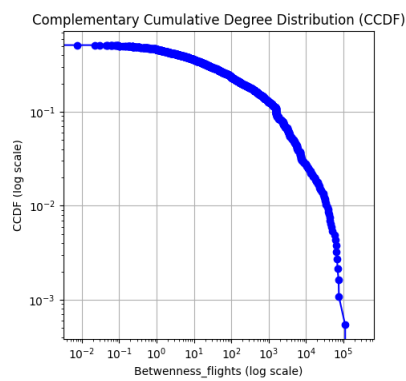


Figure 12: CCDF (12): Betweenness dos vértices da Rede de voos entre aeroportos

- (a) Max: 272795.1319948575
- (b) Min: 0.0
- (c) Média: 1404.9580419580402
- (d) Mediana: 0.18745642964392967
- (e) Desvio Padrão: 9063.772545314147
- (f) CCDF (12)

Apenas uma componente conexa, excluindo os nós que possuem grau 0. Diâmetro da rede pequeno, porém ainda significativo. Os poucos aeroportos principais (de capitais por exemplo), possuem betweenness muitas ordens de grandeza maiores que a média, o que poderia indicar uma lei de potência nessa propriedade, porém tal lei não se mostra de maneira aparente no gráfico de CCDF.