

1 Redes e Pacote Utilizado

Analizamos neste trabalho 4 redes:

1. **Zachary's karate club**¹: Rede social de amizade entre os membros de um clube de karatê em uma universidade americana nos anos 70. Cada vértice representa um membro. Dois membros são conectados se foram observados algum evento externo ao clube. Destacam-se dois membros influentes: o gestor do clube e o instrutor.
2. **Metabolic**²: Rede que representa as reações metabólicas na bactéria *E. coli*. Cada vértice é um metabolismo. Cada conexão $A \rightarrow B$ (rede direcionada) significa que existe uma reação tal que A é a entrada e B o produto.
3. **Powergrid**²: Representação da rede elétrica da Western States dos EUA. Cada vértice é uma unidade planta (unidade geradora, transformadora ou consumidora). Dois vértices são conectados se existe conexão física via cabos entre as plantas.
4. **Protein**²: Rede representando uma interação proteína-proteína na levedura. Cada vértice representa uma proteína e elas estão conectadas se interagem fisicamente dentro da célula.

As características básicas destas redes são mostradas na tabela 1.

Tabela 1: Características básicas

	Karate	Metabolic	PowerGrid	Protein
Vértices	34	1.039	4.941	2.018
Arestas	78	5.802	6.594	2.930
Direcionada?	Não	Sim	Não	Não

Para a caracterização das redes, utilizamos a biblioteca **networkx**³. Os arquivos de código-fonte deste trabalho estão disponibilizados no GitHub⁴.

2 Caracterização das Redes

Para cada rede, realizamos a caracterização através de 12 métricas: Grau; Distância; Tamanho das componentes conexas; Clusterização (local e global); Centralidade de grau; *Betweenness*; *Closeness*; Centralidade de auto-vetor; Page-rank; Similaridade de Jaccard; Similaridade de Adamic/Adar.

¹<http://www-personal.umich.edu/~mejn/netdata>

²<http://networksciencebook.com/translations/en/resources/data.html>

³<https://networkx.org/>

⁴<https://github.com/cleitonmoya/CPS765-Trabalho1>

As métricas computadas para cada redes são mostradas nas tabelas 2, 3 4 e 5.

Tabela 2: Métricas - Rede Karate

	Nom.	Máx.	Mín.	Média	Mediana	Desv. Pad.
Grau	-	17	1	4,6	3	3,8
Distância	-	5	1	2,4	2	0,9
Tam. comp. conex.	34	-	-	-	-	0
Clust. local	-	1	0	0,57	0,5	0,34
Clust. global	0,25	-	-	-	-	-
Centr. de grau	-	0,51	0,03	0,14	0,09	0,12
Betweenness	-	0,43	0	0,04	0,0025	0,09
Closeness	-	0,57	0,28	0,43	0,38	0,07
Centr. auto-vetor	-	0,37	0,02	0,14	0,10	0,09
Page Rank	-	0,10	0,0085	0,029	0,021	0,02
Jaccard	-	1	0	0,15	0,09	0,20
Adamic/Adar	-	4,71	0	0,35	0,35	0,46

Tabela 3: Métricas - Rede Metabolic

	Nom.	Máx.	Mín.	Média	Mediana	Des. Pad.
Grau Entr.	-	576	0	5,84	3	22,46
Grau Saída	-	399	0	5,58	3	19,11
Distância	-	8	1	2,98	3	0,82
Tam. comp. cox.	-	(a)	(a)	(a)	(a)	(a)
Clust. local	-	1	0	0,28	0,25	0,21
Clust. global	0,032	-	-	-	-	-
Centr. de grau	-	0,87	0,00096	0,11	0,0057	0,036
Betweenness	-	0,50	0	0,0016	6,8e-5	0,018
Closeness	-	0,64	0	0,30	0,31	0,10
Centr. auto-vetor	-	0,52	0	0,016	0,0087	0,026
Page Rank	-	0,07	0,0001	0,0009	0,0004	0,0032
Jaccard	-	(a)	(a)	(a)	(a)	(a)
Adamic/Adar	-	(a)	(a)	(a)	(a)	(a)

(a) - Função não disponível na biblioteca para grafos direcionados

Tabela 4: Métricas - Rede Powergrid

	Nom.	Máx.	Mín.	Média	Mediana	Desv. Pad.
Grau	-	19	1	2,7	2	1,8
Distância	-	46	1	19	19	6,5
Tam. comp. cox.	4941	-	-	-	-	-
Clust. local	-	1	0	0,08	0	0,22
Clust. global	0,10	-	-	-	-	-
Centr. de grau	-	0,004	0,0002	0,0005	0,0004	0,0003
Betweeness	-	0,28	0	0,003	0,0004	0,017
Closeness	-	0,08	0,03	0,05	0,05	0,007
Centr. auto-vetor	-	0,28	0	0,001	e-09	0,014
Page Rank	-	0,001	5e-5	0,0002	0,0002	0,0001
Jaccard	-	1	0	0,0003	0	0,010
Adamic/Adar	-	3,8	0	0,001	0	0,03

Tabela 5: Métricas - Rede Protein

	Nom.	Máx.	Mín.	Média	Mediana	Desv. Pad.
Grau	-	91	1	2,9	2	5
Distância	-	14	1	5,6	6	1,6
Tam. comp. cox.	-	1647	1	11	2	120
Clust. local	-	1	0	0,05	0	0,18
Clust. global	0,024	-	-	-	-	-
Centr. de grau	-	0,05	0,0005	0,001	0,001	0,002
Betweeness	-	0,18	0	0,0015	0	0,007
Closeness	-	0,24	0	0,12	0,14	0,06
Centr. auto-vetor	-	0,46	0	0,007	0,0005	0,02
Page Rank	-	0,01	0,0002	0,0005	0,0004	0,0006
Jaccard	-	1	0	0,0024	0	0,033
Adamic/Adar	-	14,3	0	0,004	0	0,06

O gráfico da figura mostra a distribuição empírica de grau da forma CCDF (*Complementary Cumulative Distribution Function*) das 4 redes analisadas. As distribuições das demais métricas são disponibilizadas no

O gráfico das figuras 1 e 2 mostra as distribuições empíricas do grau e da distância de cada vértice.

3 Discussão dos Resultados

Com relação à distribuição de graus, observamos na figura 1 que as redes Metabolic e Protein possivelmente podem ser modeladas através de uma lei de potência. Já na rede Powergrid, a distribuição claramente não segue uma lei de potência.

Com relação à distância, é interessante observar na figura 2 que as redes Metabolic e Protein possuem valores de distância esperada próximos, apesar de possuírem caracterís-

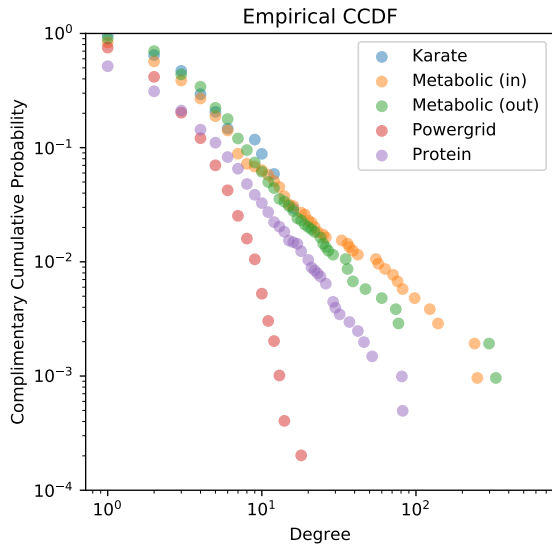


Figura 1: Graus - CCDF

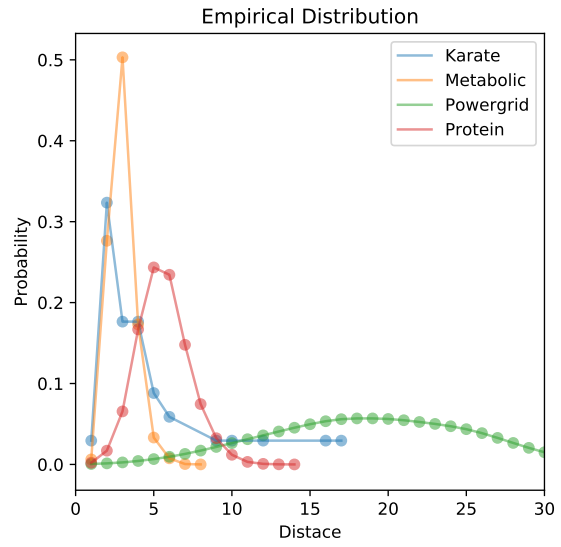


Figura 2: Distância - PDF

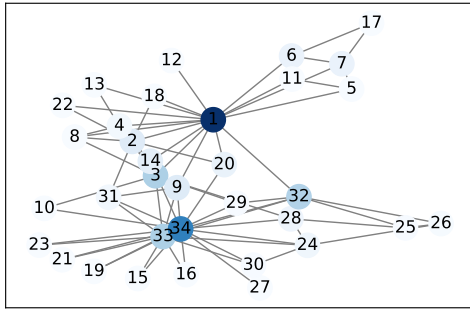


Figura 3: Karate - Betwensess

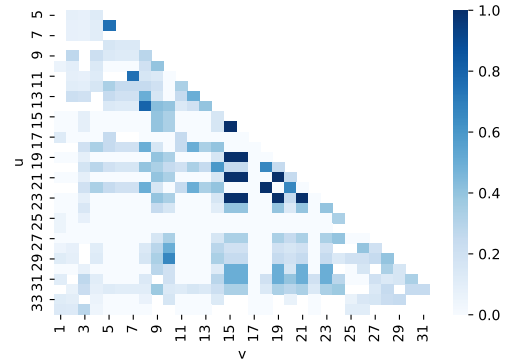


Figura 4: Karate - Jaccard

tivas estruturais diferentes. A rede Powergrid, por outro lado, apresenta distância média bem superior às demais redes.

Para a rede Protein, observamos que a maior componente conexa corresponde a 82% dos vértices da rede, o que corrobora para a característica de “tudo conectado”.

Com relação à centralidade *Closeness*, verificamos que as redes Karatê e Metabolic apresentam centralidade média maior que as outras duas. Isso é intuitivo, uma vez que a métrica é baseada na distância média de um vértice com o restante do grafo, e as redes PowerGrid e Protein possuem distância média maior.

Para a Karatê, é possível visualizar a distribuição de *betweeness*, por exemplo, com a coloração dos próprios vértices (3). A métrica identifica os vértices mais centrais da rede (1, 34). De fato, sabe-se que os membros 1 e 34 são o gerente e o instrutor do clube. Ainda, a distribuição do índice de similaridade de Jaccard pode ser visualizada através de um mapa de calor (figura 4), o qual poderia ser utilizado, por exemplo, para identificar possíveis grupos entre os membros do clube.