

Lista de Exercícios

Programação para Bioinformática com Python
PPGBM0102 - Tópicos Avançados III - T02

MINI-TESTE 03 - Redes complexas

Data de entrega: 05/09/2023

BioGRID representa um repositório de interações biomédicas que reúne informações provenientes de processos de curadoria. A versão 4.4.224 abrange um espectro de 82.750 publicações, abrigando um conjunto de 2.636.898 interações entre proteínas e genes, bem como 30.725 interações químicas e 1.128.339 modificações pós-tradução. Estes dados abrangem diversas espécies de organismos modelo e estão livremente acessíveis, além de serem passíveis de serem baixados em diversos formatos padronizados.

Para realizar o download do arquivo BioGrid, acesse o link:

<https://downloads.thebiogrid.org/File/BioGRID/Release-Archive/BIOGRID-4.4.224/BIOGRID-ALL-4.4.224.tab.zip>.

Após descomprimir o arquivo .zip, utilize o arquivo de interações para realizar as análises abaixo. A Tabela 1 lista estatísticas das interações referente ao *Homo Sapiens*.

Organismo	Tipo de experimento	Interações	Interações unicas	Genes	Publicações
Homo sapiens	Physical	1,235,664	935,746	27,640	37,267
	Genetic Interactions	18,058	17,591	7,063	391

Tabela 1: Estatísticas da base de dados BioGRID. Interações entre proteína-proteína.

Neste mini-teste considere o arquivo de interações depositado na pasta /data/ e crie um código para carregar os dados da rede, construir uma rede de interações proteína-proteína e realizar análises básicas da rede.

- Verifique a quantidade de nós na rede
- Verifique a quantidade de arestas
- Calcule o número de componentes da rede;
- Calcule o grau de centralidade de cada gene;
- Calcule o coeficiente closeness de cada gene;

- Calcule o clustering de cada gene;
- Represente num gráfico de dispersão a relação entre grau, closeness e clustering;
- Identifique os genes hubs da rede da rede em geral;
- Identifique módulos na rede e seus respectivos hubs;
- Utilize a função *write_gml()* para imprimir em arquivos cada módulo como um subgrafo.