

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO  
DEPARTAMENTO DE ESTATÍSTICA E INFORMÁTICA  
SISTEMAS DE INFORMAÇÕES - 2020.4  
MODELOS EM REDES



THAYS CONCEIÇÃO DA SILVA

Para esse projeto foi feita uma coleta de dados da rede Instagram para a modelagem de uma rede para o estudo. O ponto de partida para modelar esse conjunto de dados foi selecionar 5 perfis como raízes da busca e obter os seguidores e quem ele segue de cada um desses perfis, bem como as ligações de amizade entre alguns de seus seguidores e ele, para captar bem como as ligações da rede Instagram são compostas. Os dados foram anonimizados.

Para recolher os dados de maneira mais prática e eficiente, foi feito um código na linguagem Python e o uso da biblioteca Instaloader, que permitiu o acesso ao Instagram e aos seguidores dos perfis escolhidos para a amostra. Devido à limitação ao acesso dos dados do Instagram por APIs, no código em questão foi possível recolher apenas os seguidores dos perfis raízes. Já para a segunda parte, os seguidores em comum entre cada uma das raízes e quem ele segue, foi feita uma coleta manual desses dados. Com esse código e essa coleta manual foi possível selecionar todos os nós e enlaces necessários para a construção dessa rede.

Após a coleta de dados, foi estudado algumas propriedades de redes complexas, para que pudéssemos analisar bem a estrutura dessa rede, como a sequência do grau, grau médio, tamanho médio dos caminhos, eficiência dos enlaces, coeficiente de aglomeração, componente conectada, densidade, diâmetro, grau do hub e também o histograma da rede. Nesse caso do Instagram, em que analisaremos quem as raízes seguem e quem segue eles, iremos lidar com um grafo direcionado. Por isso, algumas propriedades serão estudadas como nós de saída e outras como nós de entradas, dependendo da direção do seu enlace.

Com os resultados dessas análises foi possível concluir primeiro que, por ter um alto coeficiente de aglomeração, é uma rede bem estruturada e coesa, o que podemos ver por conta da sua maior e única componente conectada, que foi analisada como a componente mais fraca conectada, uma vez que, por se tratar de um grafo direcionado, é preciso analisar os caminhos de ida e volta de cada nó para outros nós. O que nos faz ver que estamos tratando de uma rede complexa.

Apesar de os nós estarem bastante conectados, há uma certa distância entre eles, sabendo que redes complexas tendem a ser esparsas, o que podemos observar pela baixa densidade dessa rede. No entanto, é uma rede que possui um

curto tamanho médio dos caminhos e uma alta eficiência dos enlaces, uma vez que é preciso de poucos saltos para alcançar algum outro nó da rede.

A respeito do grau médio e máximo do grafo, podemos ver que o grau máximo se conecta com pouco menos da metade de todos os nós, o que vem de uma das raízes da busca. Isso indica a assimetria na distribuição do grau dos seus nós, o que faz com que haja muitos nós de entrada com grau baixo, e muitos nós de saída com grau elevado, por se tratar de um grafo direcionado. O que pode ser explicado por ser uma rede cuja distribuição do grau segue uma lei de potência.

A partir dessas observações podemos concluir então que, assim como as redes estudadas ao longo do curso (Facebook, Erdos, Energia e Proteína), a rede Instagram montada e estudada segue o mesmo padrão das demais redes complexas, apresentando característica semelhantes, como ser esparsa, possuir uma única componente conectada, alto coeficiente de aglomeração, tamanho de caminho curto, uma alta eficiência dos enlaces e uma distribuição de grau que segue uma lei de potência.