



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
DEPARTAMENTO DE ESTATÍSTICA E INFORMÁTICA
BACHARELADO EM SISTEMAS DE INFORMAÇÃO - 2020.1
MODELOS EM REDES - GE3

Analisando a rede de usuários de gênero feminino do MyAnimeList

Stefany Vitória da Conceição Izidio

Recife
Julho de 2021

Introdução

Este projeto visa modelar uma rede real com os dados de usuários do gênero feminino do MyAnimeList [1] e verificar as propriedades de redes complexas, tais como CC, degree, TMC, distribuição do grau etc. Na rede os usuários são representados pelos nós e as relações entre eles se dão através de 3 ou mais animes favoritos em comum, existe uma relação entre um nó x e um nó y se e somente se x tem pelo menos 3 animes favoritos em comum com y e x é diferente de y .

Para a criação da rede e verificação das propriedades foi criado um notebook no Google colab (apêndice 1) e a linguagem utilizada foi o *Python* na sua versão 3.7.10 e suas bibliotecas *pandas*, *networkx*, *jolib*, *numpy* e *matplotlib*.

Extração e criação da rede

A rede foi extraída do dataset “Anime Dataset with Reviews - MyAnimeList” [2], disponível gratuitamente no Kanggle [3].

Para a criação da rede Inicialmente a base de dados que continha os perfis dos usuários foi carregada e as duplicidades foram removidas, em seguida os registros que continham usuários somente do gênero feminino foram selecionados e então os registros com valores nulos foram removidos, então restaram 9264 registros na base. Por sua vez, o campo que continha os dados sobre os animes favoritos de cada usuário foi transformado em uma lista, pois inicialmente era um campo com strings. Adiante, a base de dados foi percorrida e cada usuário foi adicionado a rede, para cada usuário adicionado a base foi percorrida novamente dele em diante e quando um outro usuário que tinha ao menos 3 animes em comum com o primeiro era encontrado uma aresta foi criada entre eles.

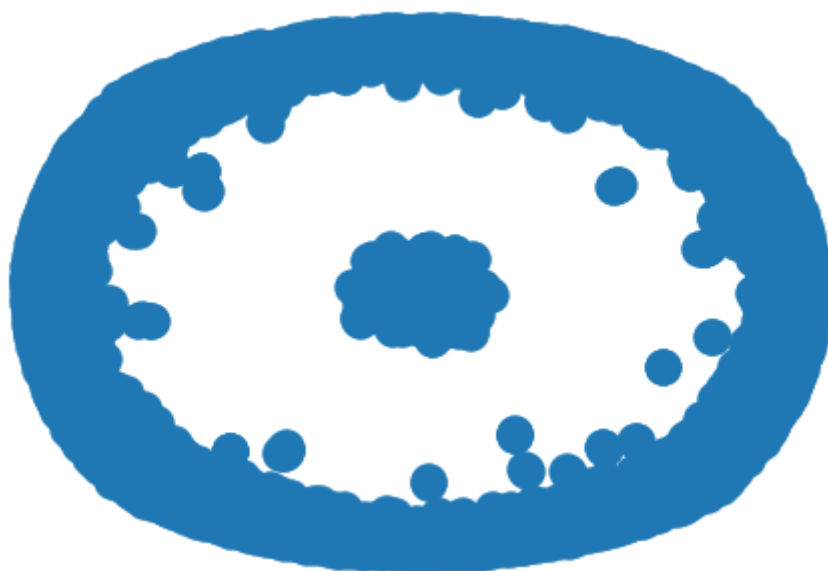


Figura 01. Rede de pessoas de gênero feminino com pelo menos 3 animes em comum.
Fonte: Autor.

Análise das propriedades da rede

A rede tem um total de 9.264 nós — que representam os usuários — e 36.590 enlaces — que representam as relações entre eles —. A rede gerada não é totalmente conectada e no total tem 4.298 componentes, esse número elevado se deu devido ao fato

de que muitos usuários na base não tinham nenhum anime favorito, tinham menos que 3 ou até tinham mais porém eram animes em comum com poucos outros usuários — 8 ou menos — o que faz com que eles fiquem em várias componentes pequenas. A rede tem uma componente conectada que se sobressai em relação ao tamanho com 4.882 nós e 36.491 enlases.

Como este projeto tem fins educacionais para a medida das próximas métricas, foi considerado a maior componente conectada (MCC).

Quanto a densidade que vale aproximadamente 0,3% pode-se observar que a rede segue a tendência das redes complexas de ser Esparsa, um possível fator que ocasiona isso é o fato de que as pessoas são seletivas ao favoritar um anime e não inserirem todos na sua lista.

Outra propriedade de redes complexas que pode ser encontrada nesta rede é que ela tem muitos nós com grau baixo, a tabela abaixo apresenta algumas medidas associadas ao grau dos nós.

Métrica	Valor
Grau médio	14.949201147070873
Hub	396
Porcentagem de nós abaixo do grau médio	71,73%
Porcentagem de nós acima do grau médio	28,26%

Tabela 01. Métricas relacionadas ao grau do nó.
Fonte: Autor.

A distribuição dos graus segue uma lei de potência, por isso a rede é livre de escala.

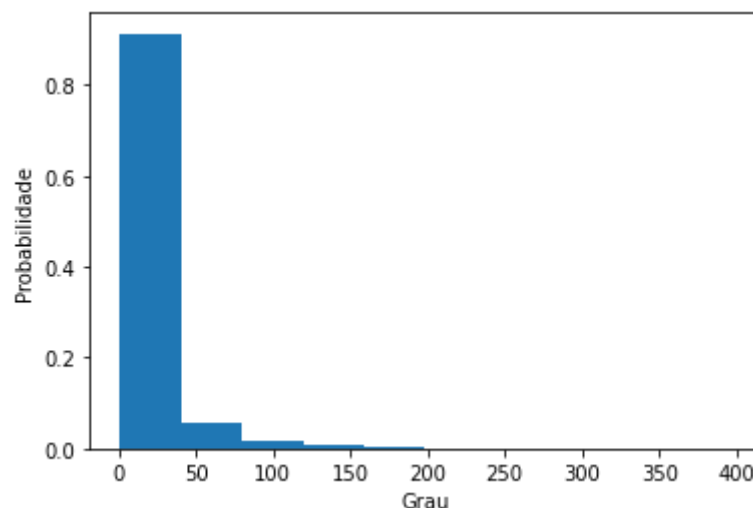


Figura 02. Histograma com a distribuição do grau.
Fonte: Autor.

A rede tem um Coeficiente de Aglomeração (CC) muito elevado de aproximadamente 0,286, uma possível explicação é que as pessoas tendem a favoritar animes semelhantes, por exemplo animes de um gênero x, então seus vizinhos também terão animes desse gênero o que aumenta a chance de estarem conectados o que

fortalece a conexão da rede. Outra possível explicação pode ser a existência de fãs de animes — geralmente os mais populares —, o que também fortaleceria a conectividade na rede, devido ao fato de que os fãs estarão ligados entre eles desde que favoritem o anime. É importante notar que ao favoritar um anime o que acontece na prática é que aquele usuário entra para um grupo daquele anime, formando um enlace de curto alcance com todos os outros usuários que também favoritaram aquele anime, ao favoritar um segundo anime o usuário estará criando um enlace de longo alcance com o segundo grupo, podemos notar que esta rede é criada por agrupamentos o que explica o alto CC.

No tocante aos caminhos a rede tem um baixo Tamanho Médio de Caminhos (TMC), que vale aproximadamente 3,82 e uma eficiência de enlaces de 0,99. Com estes resultados pode-se notar mais uma propriedade das redes complexas, a propriedade de serem eficientes na forma como seus enlaces são distribuídos para encurtar distâncias, essa propriedade é chamada de Efeito Mundo-Pequeno. Vale lembrar que o fato da rede ser livre de escala contribui para os pequenos caminhos na rede [4].

Conclusão

Podemos concluir que a rede criada neste projeto possui propriedades claras das redes complexas, tais como o efeito de mundo-pequeno, a propriedade de caminhos curtos, a formação por grupos e o alto índice de aglomeração, a distribuição de grau que segue uma lei de potência e a baixa densidade.

Para a fundamentação deste projeto foi utilizado em suma o material do blog do Doutor Glaucio Estácio Gonçalves [7].

Referências

[1] **MyAnimeList**: <https://myanimelist.net/>

[2] **Marlesson**. Anime Dataset with Reviews - MyAnimeList. Kaggle, 2020.
<https://www.kaggle.com/marlesson/myanimelist-dataset-animes-profiles-reviews?select=reviews.csv>

[3] **Kanggle**: <https://www.kaggle.com/>

[4] **Gonçalves, Glaucio Estácio**. Redes Redes complexas e seus caminhos curtos, 2020:
[https://redesemexame.blogspot.com/2020/10/redes-complexas-e-seus-caminhos-curtos.htm](https://redesemexame.blogspot.com/2020/10/redes-complexas-e-seus-caminhos-curtos.html)
l

[5] **Gonçalves, Glaucio Estácio**. Redes Distribuição do grau em redes reais, 2020:
<https://redesemexame.blogspot.com/2020/09/distribuicao-do-grau-em-redes-reais.html>

[6] **Gonçalves, Glaucio Estácio**. Redes e grupos de nós, 2020:
<https://redesemexame.blogspot.com/2020/09/redes-e-grupos-de-nos.html>

[7] **GONÇALVES, G. E.**. Ciência de Redes em Exame, 2020. (Blog).

Apêndices

[1] Extração da rede:

https://colab.research.google.com/drive/1PYuyg-nCkRSWGD_0OM0vpHRQVh7aZLWr?usp=sharing

[2] Verificação das Propriedades:

<https://colab.research.google.com/drive/1KIWnT-n34ZrFKL6RivEhCSO2DU72Z0Z?usp=sharing>