

Trabalho Prático 1

Redes Complexas
Departamento de Ciência da Computação - UFMG

Lucas Pereira Monteiro
lucasmonteiro@dcc.ufmg.br

1 - Introdução

Este trabalho visa coletar a rede de relações no Facebook do usuário Lucas Monteiro (<https://www.facebook.com/lucasmonteiro001>). O trabalho foi dividido em duas partes, sendo a primeira o desenvolvimento de um Crawler para realizar a coleta de dados do usuário acima. A segunda parte se refere à análise do grafo da rede social que foi coletado.

2 - Crawler

Para o desenvolvimento do crawler, a linguagem utilizada foi Python, devido à sua grande gama de recursos e relativa facilidade de manipulação. A biblioteca de manipulação de grafos utilizada foi a [Networkx](#). O grafo coletado foi extraído do Facebook utilizando-se para tal a [Graph API](#).

Algoritmo e Limitações

Devido à limitação imposta pela API, a qual impõe que não é possível que seja realizada a coleta de dados de usuários que não forneceram a permissão explícita para tal, restringiu consideravelmente o universo de busca. Logo, a rede coletada foi uma rede de relações da qual o usuário Lucas Monteiro, já descrito acima, faz parte. Entende-se como relação qualquer foto que o usuário postou ou foi marcado (*tagged*) e posts que foram feitos pelo próprio usuário ou por terceiros que o marcaram.

Cada relação explicita um relacionamento entre todos os usuários que foram marcados ou que sejam autores dela. Exemplo: userX postou uma foto e marcou o usuário Lucas Monteiro e o userY. Assim, é criada a relação no grafo entre userX e userY, userX e Lucas Monteiro, Lucas Monteiro e userY. Para todos os casos, todas as interações que ocorreram implicam o

surgimento de relações entre todos os usuários que fizeram parte dela de uma forma todos-para-todos.

3 - Análise da rede coletada

A. Rede analisada

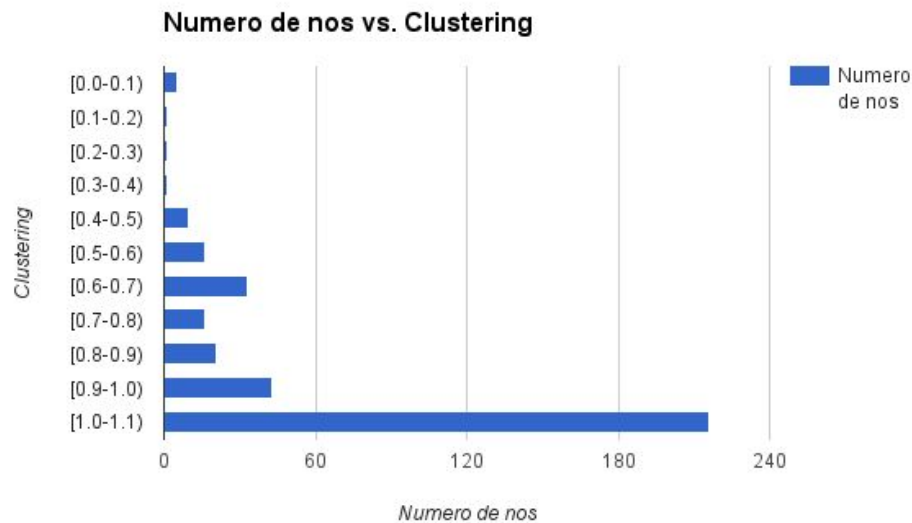
A rede coletada possui algumas características peculiares, já que todos os seus nodos foram extraídos a partir do nodo principal (usuário Lucas Monteiro). Assim, algumas métricas como a distância máxima ficam comprometidas, já que qualquer nodo está à uma distância máxima de 2 de qualquer outro nodo, visto que todos os nodos compartilham o nodo do usuário Lucas Monteiro como relação em comum.

B. Grau dos nodos



A figura acima mostra a distribuição do número de graus do grafo. Observa-se que há um número razoável de usuários que possuem entre 20 - 30 conexões, o que é plausível, dado que esse não é um número muito elevado para uma rede social que caracteriza ciclos de amizade. A partir de 30 conexões por usuários, percebe-se que há um declínio no número de conexões, visto que são poucos os usuários que tiveram interação com muitos outros na rede. Nota-se que para números muito elevados de conexões +60, o valor de apenas um usuário para aquele grau é comum.

C. Coeficiente de clusterização



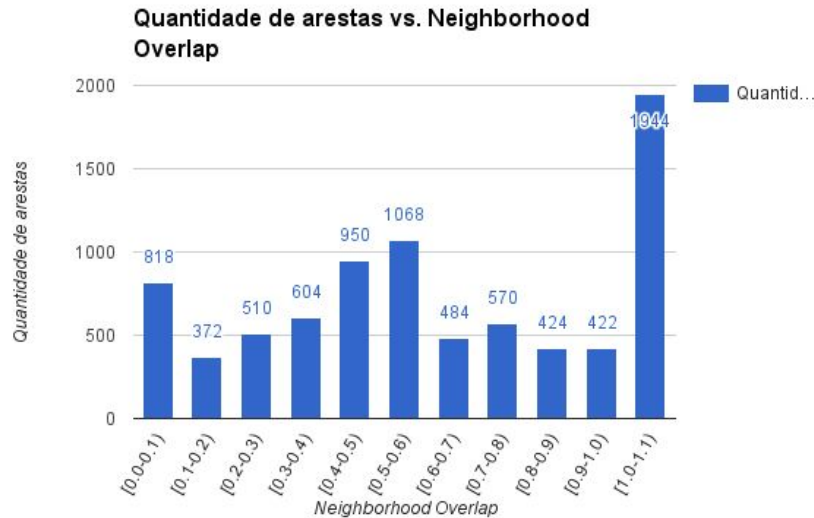
Nota-se um alto valor de clusterização para a grande maioria dos nós. Isso é bem caracterizado para esta rede, a qual possui comunidades bem definidas, as quais podem ser visualizadas na letra I dessa seção. Tais comunidades tendem a se fechar, o que faz com que seus nós aumentem o seu coeficiente de clusterização.

O **coeficiente de clusterização global** desse grafo é (0.872), um número que indica que esse grafo tende a ficar mais perto de ser um grafo com muitos triângulos fechados.

D. Componentes

Dada a especificidade desse grafo, a qual faz com que todos os nodos sejam conectados ao usuário Lucas Monteiro, o grafo possui apenas um componente.

E. Overlap da vizinhança



O gráfico acima reforça a análise do coeficiente de clusterização. Podemos ver que há uma grande quantidade de nós que são extremamente conectados dentro de sua própria comunidade, aproximadamente 27% dos nodos. Então, nota-se uma tendência de aproximadamente 28% dos usuários terem um overlap entre 0.4 e 0.6. Por fim, percebe-se que há aproximadamente 11% dos usuários que não possuem muita interação com os vizinhos de seus vizinhos.

F. Distância

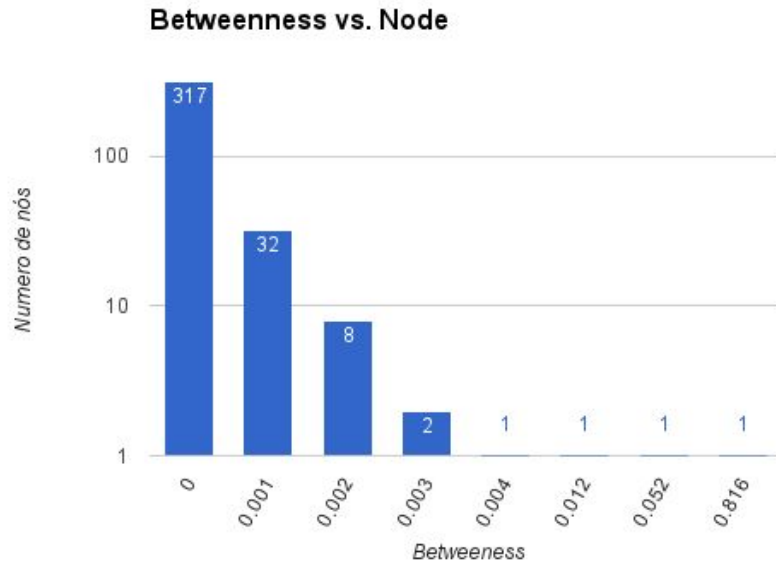


Como todos os nodos se conectam ao usuário utilizado como base da coleta, é esperado que a maior distância entre dois nodos que não possuem conexão direta se dê pela conexão com o usuário Lucas Monteiro. **A distância média é 1.94.**

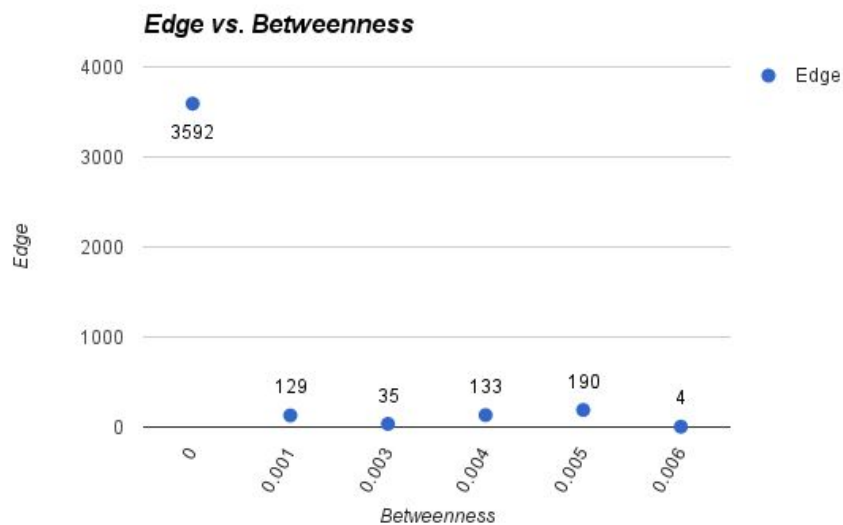
G. Betweenness

Observando-se tanto os gráfico de Betweenness em Nodos, quanto Arestas, é possível notar que eles são correlatos, visto que apenas uma pequena parcela de nodos das redes faz parte dos nodos que ligam grupos distintos dentro dessa própria rede.

a. Nodos



b. Arestas



H. Bridge span

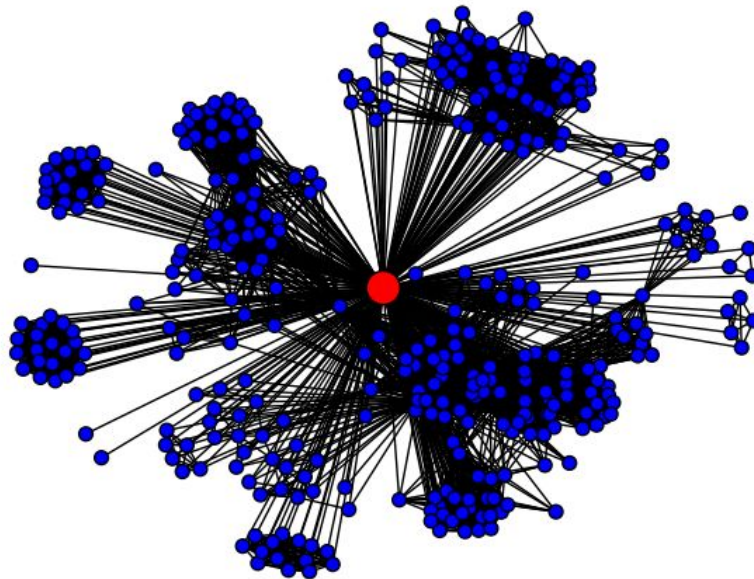
A abordagem utilizada para calcular o Bridge span foi a seguinte:

```
Para cada aresta no grafo, faça:  
  Retire a aresta do grafo  
  Calcule a distancia entre os nodos  
  Insira novamente a aresta retirada
```

Com a execução do algoritmo acima, é possível obter-se todos os nodos que estão acima de uma determinada distância quando a aresta que os conecta é retirada. Exemplo: Pegue a aresta que conecta X e Y e a retire do grafo, calcule a distância entre X e Y. Se for 2, indica que eles têm um amigo em comum, se for maior que 2, essa pode ser uma indicação de que a aresta pode ser uma ponte.

Novamente, devido à particularidade desse grafo, foram achadas apenas **4** arestas que se caracterizam como ponte, o restante está todo conectado a uma distância de **2**.

I. Imagem do grafo



A imagem acima mostra o grafo que foi obtido a partir da rede social. Nota-se que há algumas comunidades muito bem definidas e outras com nós que poderiam ser ponte, se não houvesse a conexão de todos com o nó principal. Também é possível visualizar os nós que estão isolados (nesse caso são 4), o que caracteriza a aresta deles com o nodo principal como ponte.

4 - Curiosidades : Remoção do nó Lucas Monteiro

Por motivo de curiosidade, a remoção do nó Lucas Monteiro da rede trouxe algumas características interessantes, as quais são mostradas a seguir:

- O número de componentes passou a ser 8;
- Houve casos de distâncias entre nós iguais a 1, 2, 3, 4 e 5, além de distâncias “infinitas”, visto que os nós faziam parte de uma ponte com um único nó.