

# Grafos

## Trabalho Prático 1

Heitor Lourenço Werneck  
heitorwerneck@hotmail.com

26 de Dezembro de 2019

### Conteúdo

<b>1</b>	<b>Introdução</b>	<b>2</b>
<b>2</b>	<b>Modelagem e Solução</b>	<b>2</b>
2.1	Floyd-Warshall . . . . .	2
<b>3</b>	<b>Análise de Resultados</b>	<b>2</b>
<b>4</b>	<b>Conclusão</b>	<b>4</b>

# 1 Introdução

Algoritmos de obtenção de menor caminho em grafos são muito utilizados em diversas áreas como por exemplo: recomendação de sistemas, mapas geográficos, redes de telefonia e em diversas outras áreas.

Em um problema de menor caminho é dado um grafo ponderado e direcionado  $G = (V, E)$ , a função de ponderamento é definida como  $w : E \rightarrow \mathbb{N}$ .cite:rayward-smith91\_introd\_to\_algor

O caminho com menor peso do vertice  $u$  até o  $v$  é  $\delta(u, v)$ . O mesmo é definido como:

$$\delta(u, v) = \begin{cases} \min\{w(p) : u \rightsquigarrow^p v\} & \text{se existir um caminho de } u \text{ para } v \\ \infty & \text{se não} \end{cases}$$

Esse trabalho tem por objetivo apresentar um algoritmo que solucione o problema de encontrar o caminho mínimo entre todos pares de vértices de um grafo  $G$  e também fazer análise de um grafo de uma rede real. O algoritmo feito foi o Floyd-Warshall, o Bellman-Ford também foi feito porém foi utilizado somente para validação do Floyd-Warshall.

## 2 Modelagem e Solução

Primeiramente é importante definir como será representado o grafo  $G$ . A escolha foi uma matriz de adjacência por sua facilidade de manipulação e também por causa de seu acesso aos elementos ser de complexidade assintótica de tempo  $O(1)$ , essa operação é realizada muitas vezes nos algoritmos logo essa complexidade será benéfica.

O grafo será representado como:

$$\begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1j} \\ v_{21} & v_{22} & \cdots & v_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ v_{i1} & v_{i2} & \cdots & v_{ij} \end{bmatrix} \quad (2.1)$$

Com  $v_{ij} \in \{0(\text{Não tem aresta}), 1(\text{Existe aresta})\}$  para  $i = 1, \dots, |V| \wedge j = 1, \dots, |V|$ .

### 2.1 Floyd-Warshall

O algoritmo ref:alg:floydwarshall descreve o funcionamento detalhadamente. Como é possível ver pelos laços a complexidade assintótica de tempo é  $O(n^3)$ . No final do algoritmo uma matriz de distâncias de todos pares de vértices é obtida.

---

**Algorithm 1** Floyd-Warshall.

---

**Input:**  $G$  **Output:**  $d$

---

```
1: procedure FLOYD_WARSHALL
2:    $d \leftarrow \text{copy}(G)$ 
3:   for  $k = 1$  to  $G.V$  do
4:     for  $i = 1$  to  $G.V$  do
5:       for  $j = 1$  to  $G.V$  do
6:         if  $d[i, j] > d[i, k] + d[k, j]$  then
7:            $d[i, j] \leftarrow d[i, k] + d[k, j]$ 
8:   return  $d$ 
```

---

## 3 Análise de Resultados

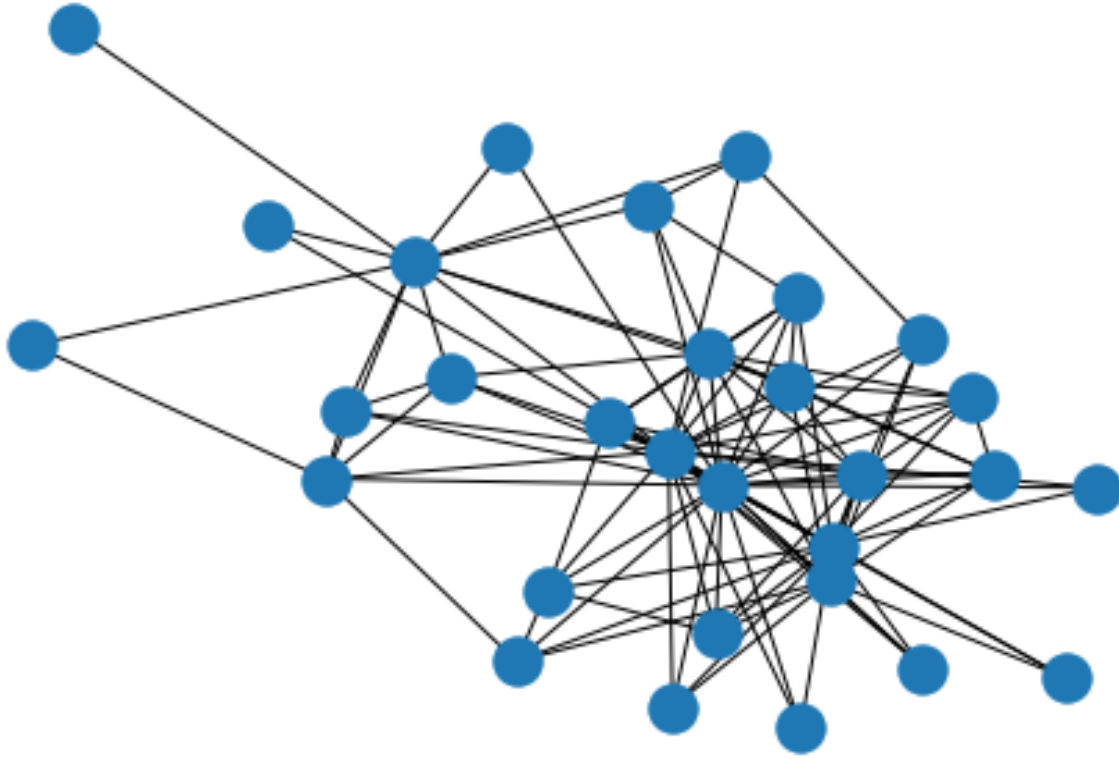
A base de dados escolhida para ser feito a análise foi o *Yelp Dataset Challenge* cite:WinNT. Dessa base de dados foi retirado os dados sobre as amizades dos usuários da cidade de Madison, esse grafo é não direcionado.

O grafo terá seus vértices como usuários e as arestas como relação de amizade entre usuários. Foi obtido um total de 1022 usuários(vértices) e 5078 amizades(arestas).

A densidade do grafo é de 0.0097329, ou seja, bem baixa. Isso mostra que pessoas são amigas de poucas pessoas.

O clique do grafo é 17, isso mostra que existe uma pequena comunidade de amizade no grafo.

Na figura ref:fig:graph mostra uma amostragem de 30 usuários do grafo.



**Figura 1:** Grafo da rede de usuários.

A partir desses dados iniciais foi iniciado o processo de análise de localização dos vértices. Primeiro foi executado o algoritmo ref:alg:floydwarshall no grafo e foi obtido a matriz de distâncias. Com essa matriz foi obtido o número mínimo, máximo e médio de cada linha. Não é necessário fazer para as colunas pois o grafo é não direcionado e o processo seria redundante.

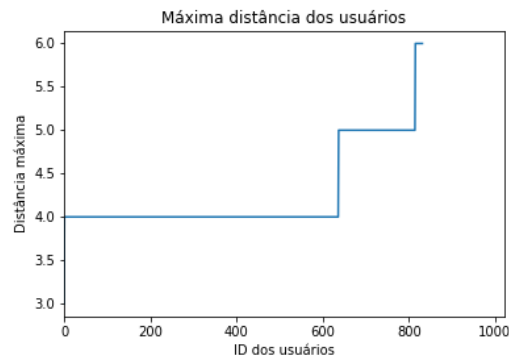
$$\begin{bmatrix} 0 & 2 & 1 & \text{min} = 1, \text{média} = 1.5, \text{max} = 2 \\ 2 & 0 & 1 & \text{min} = 1, \text{média} = 1.5, \text{max} = 2 \\ 1 & 1 & 0 & \text{min} = 1, \text{média} = 1, \text{max} = 1 \end{bmatrix} \quad (3.1)$$

Um problema na análise desse grafo é que se existir um usuário com distância máxima infinito então todos outros também terão essa mesma distância máxima, que é o que aconteceu, logo não poderá se ter uma análise detalhada com base nesse aspecto. Porém para utilizar a informação de distância máxima foi considerado o maior valor sendo um valor menor que infinito se o mesmo existir, se não existir então o valor é infinito mesmo.

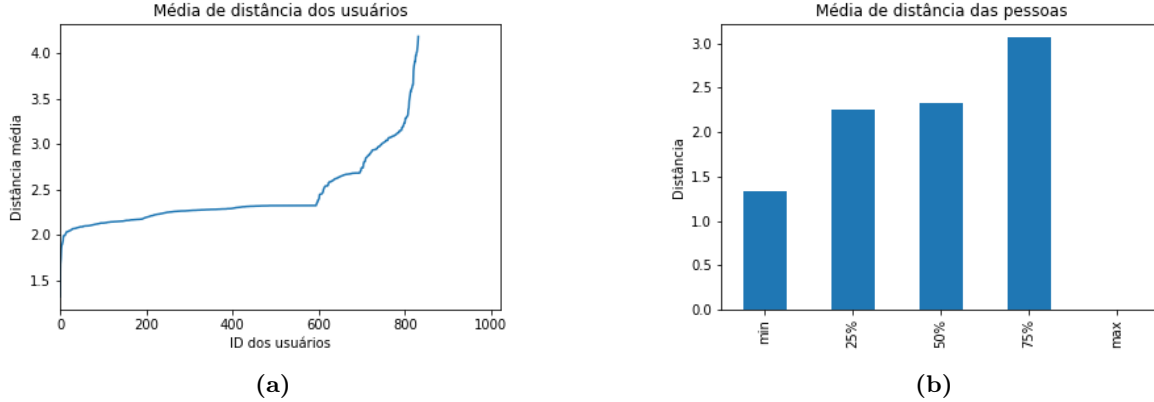
Como é possível ver na figura ref:fig:maxdist aproximadamente 80% dos usuários possuem amizades, já os outros 20% não possuem amizade. E de acordo com essa mesma figura existem mais pessoas com distâncias máximas menores em relação as outras, isso pode indicar que a maioria dos usuários não pertencem a redes grandes de amizade.

O problema no cálculo de distância máxima também se propaga para a distância média, logo para fazer a análise de distância média de um usuário foi desconsiderado os valores infinitos, porém se não existirem valores diferentes de infinito então a distância média é realmente infinito.

De acordo com a figura ref:fig:pmeandist e ref:fig:meandist pelo menos 50% dos usuários tem um valor médio de distância bem constante.



**Figura 2:** Distância máxima dos usuários.



**Figura 3:** Distância média.

Porém pela figura ref:fig:pmeandist do usuário 600 até o 800 há um crescimento considerável em relação a de 0 a 600. Isso pode mostrar uma relação de diferenciação grande de uma pequena parte dos usuários.

A distância mínima pode ser analisada tranquilamente sem os problemas anteriores mencionados.

Foi observado que 832 usuários possuem um valor mínimo menor que infinito, ou seja, eles possuem pelo menos uma amizade. Porém os outros 190 não possuem nenhuma amizade, isso também feito analisado de outra forma nos outros gráficos.

Isso indica que o *dataset* escolhido possui 81.4% de usuários com alguma amizade.

## 4 Conclusão

Com o trabalho realizado foi possível fazer diversas observações sobre a base selecionada. A complexidade de tempo do algoritmo implementado na prática fez com que os resultados fossem produzidos com necessidade de muito tempo, o que corresponde a complexidade de  $O(n^3)$ .

Foi possível obter diversas informações importantes sobre o grafo utilizando informação de distância, isso mostra sua importância na análise.

O *dataset* analisado apresentou muitos usuários sem vértices ligando-os, isso mostra algo que pode ser recorrente em grafos de amizades em uma cidade.

Apesar dos problemas para fazer a análise as medidas tomadas para se solucionar os problemas se mostraram eficientes e foi possível prosseguir com a análise.

Então pode-se concluir que o trabalho foi um sucesso pois um algoritmo foi implementado para solucionar o problema de distância mínima e uma análise de uma base real foi feita.

bibliographystyle:plain bibliography:doc.bib