

Análise de Centralidade em Redes Complexas

Rodrigo José Zonzin Esteves¹

¹Departamento de Computação – Universidade Federal de São João del Rei

1. Dados

A rede foi obtida através do portal “Stanford Network Analysis Project” e apresenta uma rede social virtual obtida por meio do Facebook. Os dados foram anonimizados, retirando-se o ID identificador do perfil real, e apresentam atributos como afinidade partidária - com uma máscara numérica para democratas e republicanos. [McAuley and Leskovec 2012].

2. Análise

2.1. Centralidade por grau - Clustering

Também conhecido como “Coeficiente de Clustering”.

Figura 1. Vértices por Clustering

	vertice	clustering
0	0	0.041962
1	1	0.419118
2	2	0.888889
3	3	0.632353
4	4	0.866667
...
4034	4034	1.000000
4035	4035	0.000000
4036	4036	1.000000
4037	4037	0.666667
4038	4038	0.555556

Observou-se $\hat{C} = 0.605547$, $\sigma = 0.214462$, $Q_{0.25} = 0.466667$, $M_{c_v} = 0.600000$, $Q_{0.75} = 0.752381$. Os 10 vértices com maior centralidade estão apresentados na imagem a seguir.

2.2. Closeness

Os vértices com maiores valores de Closeness são apresentados na tabela a seguir.

$Closeness_{avg} = 6.669574e - 04$, $\sigma = 1.164634e - 02$, $Q_{0.25} = 3.997507e - 07$, $M = 2.918300e - 06$ e $Q_{0.75} = 1.515292e - 05$.

Figura 2. 10 vértices com maior Closeness

	vertice	closeness
107	107	0.480518
1684	536	0.337797
3437	2600	0.236115
1912	767	0.229295
1085	1599	0.149015
0	0	0.146306
698	1212	0.115330
567	1081	0.096310
58	58	0.084360
428	942	0.064309

2.3. Betweenness

Figura 3. 10 vértices com maior Betweenness

	vertice	betweenness
107	107	0.459699
58	58	0.397402
428	942	0.394837
563	1077	0.393913
1684	536	0.393606
171	171	0.370493
348	348	0.369916
483	997	0.369848
414	928	0.369543
376	606	0.366558

$Betweenness_{avg} = 0.276168$, $\sigma = 0.036124$, $Q_{0.25} = 0.260348$, $M = 0.282457$ e $Q_{0.75} = 0.315001$.

2.4. Eigenvector

Figura 4. 10 vértices com maior eigencentrality

	vertice	autovetor
1912	767	0.095407
2266	2686	0.086983
2206	2163	0.086052
2233	2903	0.085173
2464	2905	0.084279
2142	2196	0.084193
2218	2377	0.084156
2078	3525	0.084136
2123	3179	0.083671
1993	848	0.083532

3. Discussão

O clustering se dá em função do grau de cada vértice, como na equação a seguir. Da análise realizada, observa-se que os 10 vértices com maior clustering têm coeficiente

igual a 1. Matematicamente,

$$C_{grau}^i = \frac{k_i}{n-1} = 1 \iff k_i = n-1$$

Desse resultado, segue-se que a rede em análise apresenta vértices cuja conectividade é máxima (estão ligados a todos os demais vértices). Topologicamente, essa propriedade é interessante uma vez que esses vértices são capazes de atingir qualquer outro vértice de interesse.

Ademais, no quartil superior, os vértices estão ligados a pelo menos 75% da rede. De forma similar, esse é um valor expressivo e pode ser utilizado, por exemplo, caso não se soubesse a priori da existência de conectividade maximal em alguns vértices da rede – escolhas ótimas, com custo unitário, para se atingir qualquer outro vértice da rede.

A medida de Closeness apresenta a distância média entre um vértice v e todos os demais $u \in V - \{v\}$ através do caminho mais curto. Essa medida é capaz de mensurar o quão rápido uma propriedade pode trafegar na rede.

De acordo com a figura 6(a), os nós com maior closeness pertencem a clusters diferentes. Desse modo, tais vértices podem ser bastante influentes na velocidade de tráfego de propriedades na rede, mas com maior influência ao seu cluster.

A centralidade Betweenness quantifica como um nó age como ponte entre o caminho mínimo de dois outros vértices. Essa métrica é capaz de aferir a influência que cada vértice sobre o fluxo de propriedade de uma rede.

Diferentemente do closeness, a topologia da rede sugere que os vértices têm uma influência similar na propagação de propriedades sobre a rede. De fato, tomando o coeficiente de variação das duas centralidades, observamos $CV_{clos} = 17.4597$ e $CV_{bet} = 0.1307$, indicando que não há valores de betweenness que se destacam.

A centralidade por autovetor (*eigenvector centrality*) considera a conexão de um vértice a "vértices importantes" na rede. Isto é, quanto mais próximo de vértices importantes ele está, mais importante ele se torna – e maior será sua centralidade por *eigenvector*.

A figura 4 apresenta os 10 vértices com maior centralidade por autovetor. Pode-se observar que os vértices estão dispostos um ao outro (entre 1912 e 2464). De maneira similar, a figura 6(d) apresenta que o "vértice mais importante" é cercado por outros vértices importantes (sobrepostos em cor laranja) e seguidos por vértices de cor azul.

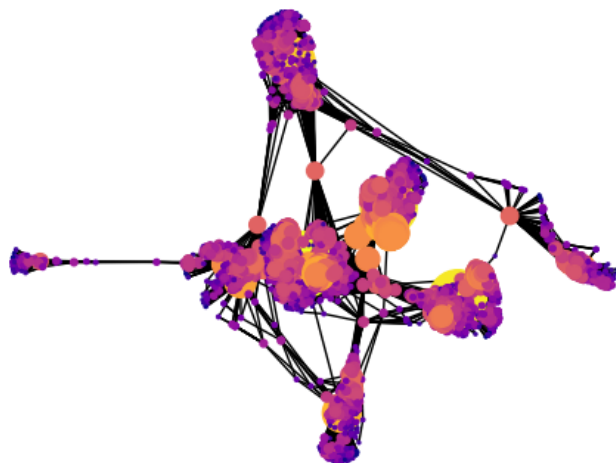
As demais regiões da rede apresentam vértices pequenos, dos quais muitos se quer foram representados.

4. Visualização da Rede

As figuras a seguir apresentam atributos de centralidade da rede discricionados por tamanho e cor do vértice. Para cada figura, quanto maior o nó, maior é o respectivo atributo.

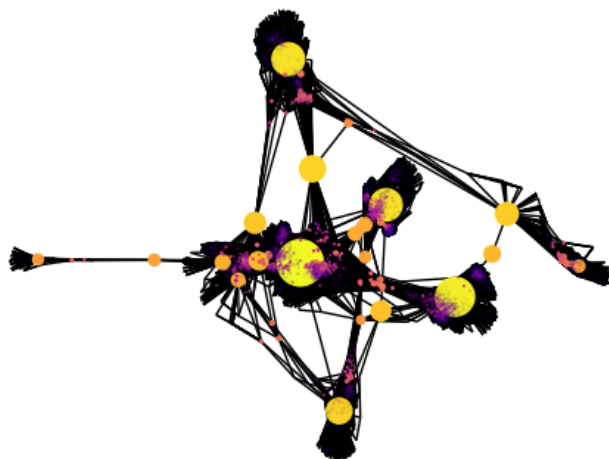
Figura 5. Centralidade por Autovetor

Grau



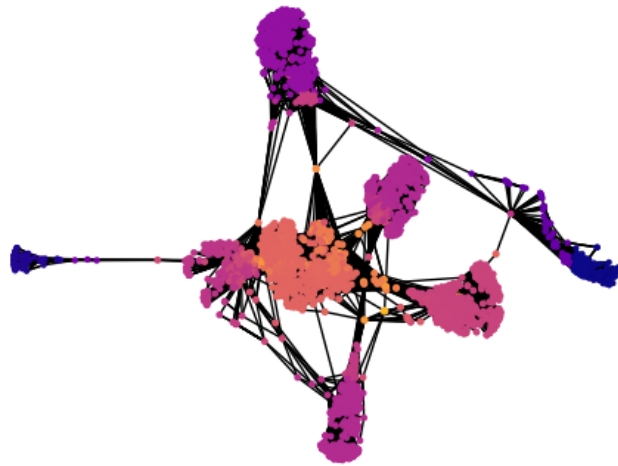
(a) Grau

Closeness



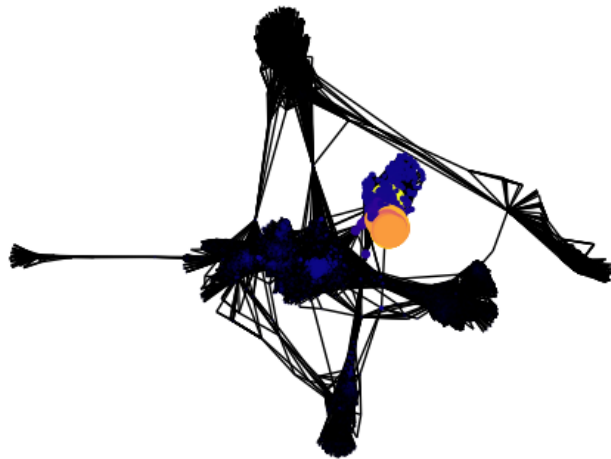
(b) Closeness

Betweenness



(c) Betweenness

Autovector



(d) Autovector

Referências

McAuley, J. and Leskovec, J. (2012). Learning to discover social circles in ego networks. NIPS.