

REDES SOCIAIS

Entrega: Medidas de Centralidade

Aluno(s): Eduardo Tirta e Filipe Borba

Data: 02/10

A atividade em questão propôs formas de medidas de centralidade numa rede, sendo elas closeness e betweenness. Closeness indica que indivíduos serão mais centrais se estiverem, simultaneamente, próximos dos outros indivíduos na rede. Enquanto isso, betweenness indica que indivíduos serão mais centrais se estiverem no “menor caminho” (geodésicas) das díades. A hipótese inicial da atividade foi a seguinte: Quando consideramos outros tipos de trajetória e outros tipos de difusão, os nós com maior closeness simulado e betweenness simulado não são necessariamente os nós com maior closeness e betweenness segundo as fórmulas clássicas.

Com isso, foi realizado um teste-t para cada trajetória e difusão, utilizando como parâmetro a geodésica e transferência e como hipótese nula a igualdade das médias. Este teste comparou a média de closeness e betweenness entre cada uma das trajetória e difusões e a geodésica com transferência. No final, foi obtida a **tabela 1** com os resultados dos testes-t para cada comparação de médias. Com os resultados apresentados na **tabela 1** fica claro que o p-value para cada teste é extremamente baixo. Isso significa que num mundo onde a hipótese nula é verdadeira, extremamente raramente as médias seriam iguais. Então pode-se concluir que a hipótese nula é improvável e, por isso, a hipótese inicial tem uma forte justificativa.

Além disso, verificando as simulações de closeness e betweenness das outras trajetórias e difusões, pode-se perceber que quando realizado uma

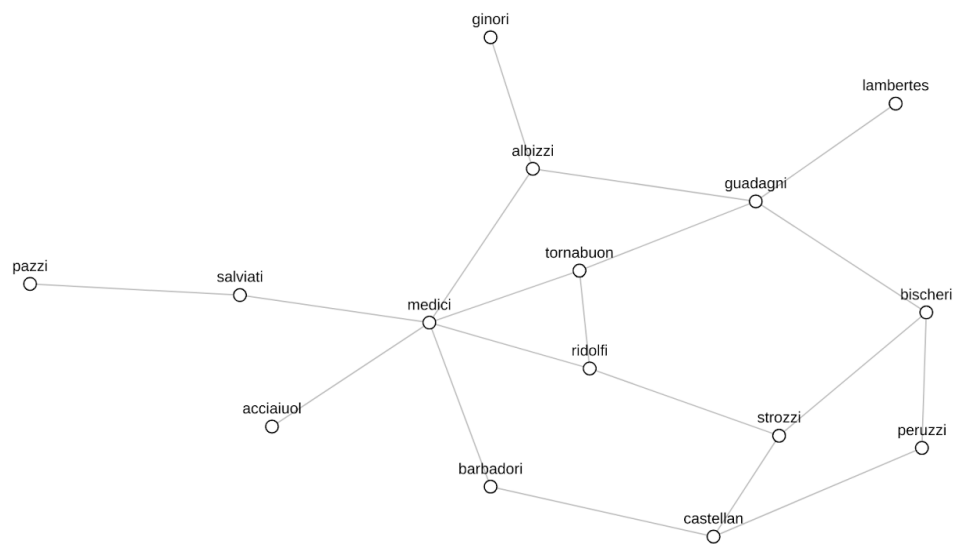
ordenação decrescente utilizando as duas variáveis, a ordem das famílias mais “centrais” muda drasticamente para cada um. É possível observar casos como a família Medici oscilando entre os mais influentes, a família Peruzzi oscilando entre o meio e o final da tabela, até um caso extremo como a família Pazzi que se apresenta na base da **tabela 10**, sendo uma trajetória geodésica, indo para o topo da **tabela 3**, quando se trata de um passeio com transferência. Essa comparação ganha mais expressão quando vista em conjunto com o **grafo 1** que mostra visualmente a rede das famílias. Nota-se que novos “centros” começam a se formar, assim como diferentes formas de enxergar a relação entre as famílias.

De modo geral, é possível verificar como as medidas de closeness e betweenness mudam para cada trajetória e difusão. Tanto as tabelas geradas pela simulação, quanto os testes-t demonstram que a ordem de influência das famílias muda conforme a trajetória ou difusão é alterada. Dessa forma, dadas as evidências, pode-se concluir que a hipótese inicial foi corroborada.

Referências:

<https://www.vox.com/science-and-health/2017/7/31/16021654/p-values-statistical-significance-redefine-0005>

Gráficos



Grafo 1. Representação das famílias em forma de um grafo

Tabelas

	type	closeness tstatistic	closeness pvalue	betweenness tstatistic	betweenness pvalue
0	geo_duplic	nan	nan	3.3331679399	0.0049262132
1	path_duplic	-10.44607919 91	0.0000000543	15.153802257 5	0.0000000004
2	path_transf	-12.36303404 10	0.0000000064	4.1595921170	0.0009633944
3	trail_duplic	-11.15845394 56	0.0000000236	10.792092301 0	0.0000000360
4	trail_transf	-12.30451627 47	0.0000000068	4.6287886814	0.0003904952
5	walk_duplic	-15.69815880 15	0.0000000003	5.6895793785	0.0000558738
6	walk_transf	-19.30255698 10	0.0000000000	7.2147012940	0.0000044617

Tabela 1. Tabela com o resultado dos testes-t para cada trajetória e difusão em comparação à geodésica com transferência.

	closeness tstatistic	closeness pvalue	betweenness tstatistic	betweenness pvalue
count	6.0000000000	6.0000000000	7.0000000000	7.0000000000
mean	-13.54546654 05	0.0000000152	7.2816748528	0.0009057821
std	3.3472469138	0.0000000210	4.2663135235	0.0018077723
min	-19.30255698 10	0.0000000000	3.3331679399	0.0000000004
25%	-14.86437761 14	0.0000000018	4.3941903992	0.0000022489
50%	-12.33377515 79	0.0000000066	5.6895793785	0.0000558738
75%	-11.44496952 79	0.0000000194	9.0033967975	0.0006769448
max	-10.44607919 91	0.0000000543	15.153802257 5	0.0049262132

Tabela 2. Tabela gerada pela função *describe* do pandas, trazendo algumas informações relevantes como a média de p-value dos testes-t.

	família	closeness simulado	closeness analítico	betweenness simulado	betweenness analítico
5	salviati	0.040369	0.388889	1.536758	0.142857
4	pazzi	0.039047	0.285714	0.687582	0.000000
12	peruzzi	0.038499	0.350000	1.589231	0.021978
0	ginori	0.037620	0.333333	0.770330	0.000000
10	acciaiuol	0.036487	0.368421	0.760989	0.000000
13	barbadori	0.036468	0.437500	1.632582	0.115385
1	lambertes	0.036308	0.325581	0.773407	0.000000
8	bischeri	0.036138	0.400000	2.481703	0.120879
2	albizzi	0.036038	0.482759	2.510055	0.212454
7	tornabuon	0.035388	0.482759	2.507692	0.091575
9	ridolfi	0.034698	0.482759	2.502418	0.086081
11	strozzi	0.034680	0.424242	2.464011	0.075092
6	medici	0.034437	0.560000	5.056044	0.521978
14	castellan	0.034088	0.388889	2.438462	0.087912
3	guadagni	0.034060	0.466667	3.363462	0.260073

Tabela 3: Comparação entre o analítico e o simulado, usando *passeio* como trajetória e *transferência* como difusão da simulação.

	família	closeness simulado	closeness analítico	betweenness simulado	betweenness analítico
6	medici	0.560000	0.560000	1.390714	0.521978
2	albizzi	0.482759	0.482759	0.570275	0.212454
9	ridolfi	0.482759	0.482759	0.251648	0.086081
7	tornabuon	0.482759	0.482759	0.245385	0.091575
3	guadagni	0.466667	0.466667	0.687692	0.260073
13	barbadori	0.437500	0.437500	0.317198	0.115385
11	strozzi	0.424242	0.424242	0.207143	0.075092
8	bischeri	0.400000	0.400000	0.311648	0.120879
5	salviati	0.388889	0.388889	0.406593	0.142857
14	castellan	0.388889	0.388889	0.230879	0.087912
10	acciaiuol	0.368421	0.368421	0.000000	0.000000
12	peruzzi	0.350000	0.350000	0.055330	0.021978
0	ginori	0.333333	0.333333	0.000000	0.000000
1	lambertes	0.325581	0.325581	0.000000	0.000000
4	pazzi	0.285714	0.285714	0.000000	0.000000

Tabela 4: Comparação entre o analítico e o simulado, considerando *trajetória geodésica* e *duplicação* como difusão da simulação.

	família	closeness simulado	closeness analítico	betweenness simulado	betweenness analítico
7	tornabuon	0.278030	0.482759	0.681099	0.091575
9	ridolfi	0.275726	0.482759	0.702088	0.086081
6	medici	0.265837	0.560000	0.820495	0.521978
3	guadagni	0.265041	0.466667	0.743901	0.260073
13	barbadori	0.263984	0.437500	0.580879	0.115385
11	strozzi	0.263603	0.424242	0.680440	0.075092
8	bischeri	0.262429	0.400000	0.650769	0.120879
2	albizzi	0.254752	0.482759	0.616978	0.212454
14	castellan	0.253489	0.388889	0.617637	0.087912
12	peruzzi	0.253409	0.350000	0.526923	0.021978
10	acciaiuol	0.236083	0.368421	0.396868	0.000000
1	lambertes	0.234127	0.325581	0.402418	0.000000
0	ginori	0.228952	0.333333	0.393516	0.000000
5	salviati	0.225178	0.388889	0.440440	0.142857
4	pazzi	0.203899	0.285714	0.264560	0.000000

Tabela 5: Comparação entre o analítico e o simulado, usando *caminho* e *duplicação* como difusão da simulação.

	família	closeness simulado	closeness analítico	betweenness simulado	betweenness analítico
6	medici	0.305926	0.560000	0.629066	0.521978
3	guadagni	0.269255	0.466667	0.402747	0.260073
7	tornabuon	0.261050	0.482759	0.334780	0.091575
9	ridolfi	0.258842	0.482759	0.364341	0.086081
2	albizzi	0.257231	0.482759	0.228736	0.212454
14	castellan	0.253950	0.388889	0.383791	0.087912
8	bischeri	0.246580	0.400000	0.365824	0.120879
13	barbadori	0.242199	0.437500	0.240769	0.115385
5	salviati	0.242110	0.388889	0.142857	0.142857
10	acciaiuol	0.238060	0.368421	0.000000	0.000000
11	strozzi	0.237827	0.424242	0.373736	0.075092
12	peruzzi	0.221625	0.350000	0.247253	0.021978
1	lambertes	0.211719	0.325581	0.000000	0.000000
0	ginori	0.208298	0.333333	0.000000	0.000000
4	pazzi	0.198528	0.285714	0.000000	0.000000

Tabela 6: Comparação entre o analítico e o simulado, considerando *caminho* e *transfêrencia* como difusão da simulação.

	família	closeness simulado	closeness analítico	betweenness simulado	betweenness analítico
7	tornabuon	0.261941	0.482759	0.923791	0.091575
9	ridolfi	0.261346	0.482759	0.929780	0.086081
6	medici	0.256561	0.560000	1.253901	0.521978
3	guadagni	0.253668	0.466667	0.900934	0.260073
13	barbadori	0.253498	0.437500	0.602637	0.115385
8	bischeri	0.249716	0.400000	0.780110	0.120879
11	strozzi	0.249646	0.424242	0.883132	0.075092
2	albizzi	0.248451	0.482759	0.688242	0.212454
14	castellan	0.242129	0.388889	0.827692	0.087912
12	peruzzi	0.238096	0.350000	0.592747	0.021978
10	acciaiuol	0.227774	0.368421	0.367033	0.000000
1	lambertes	0.226908	0.325581	0.400055	0.000000
0	ginori	0.225156	0.333333	0.372582	0.000000
5	salviati	0.221475	0.388889	0.420604	0.142857
4	pazzi	0.200382	0.285714	0.262473	0.000000

Tabela 7: Comparação entre o analítico e o simulado, considerando *trilha* e *duplicação* como difusão da simulação.

	família	closeness simulado	closeness analítico	betweenness simulado	betweenness analítico
6	medici	0.291128	0.560000	0.701923	0.521978
2	albizzi	0.276296	0.482759	0.267912	0.212454
13	barbadori	0.259748	0.437500	0.269945	0.115385
3	guadagni	0.257665	0.466667	0.410110	0.260073
7	tornabuon	0.246191	0.482759	0.314505	0.091575
9	ridolfi	0.243966	0.482759	0.332253	0.086081
12	peruzzi	0.240704	0.350000	0.226923	0.021978
5	salviati	0.238170	0.388889	0.142857	0.142857
11	strozzi	0.233742	0.424242	0.327527	0.075092
8	bischeri	0.229909	0.400000	0.341703	0.120879
14	castellan	0.227674	0.388889	0.376484	0.087912
10	acciaiuol	0.227460	0.368421	0.000000	0.000000
0	ginori	0.221905	0.333333	0.000000	0.000000
1	lambertes	0.206122	0.325581	0.000000	0.000000
4	pazzi	0.196370	0.285714	0.000000	0.000000

Tabela 8: Comparação entre o analítico e o simulado, usando *trilha* como trajetória e *transferência* como difusão da simulação

	família	closeness simulado	closeness analítico	betweenness simulado	betweenness analítico
6	medici	0.173774	0.560000	6.930824	0.521978
7	tornabuon	0.168356	0.482759	2.396593	0.091575
9	ridolfi	0.166838	0.482759	2.528956	0.086081
13	barbadori	0.162888	0.437500	1.473022	0.115385
3	guadagni	0.162499	0.466667	4.509670	0.260073
2	albizzi	0.160570	0.482759	3.040110	0.212454
11	strozzi	0.157669	0.424242	2.707473	0.075092
8	bischeri	0.153606	0.400000	2.760604	0.120879
14	castellan	0.152636	0.388889	3.110275	0.087912
10	acciaiuol	0.152590	0.368421	0.580220	0.000000
1	lambertes	0.146232	0.325581	0.701429	0.000000
5	salviati	0.145602	0.388889	1.817692	0.142857
0	ginori	0.145554	0.333333	0.787912	0.000000
12	peruzzi	0.144725	0.350000	1.674725	0.021978
4	pazzi	0.129184	0.285714	0.773516	0.000000

Tabela 9: Comparação entre o analítico e o simulado, usando *passeio* como trajetória e *duplicação* como difusão da simulação

	família	closeness	closeness	betweenness	betweenness
--	----------------	------------------	------------------	--------------------	--------------------

		simulado	analítico	simulado	analítico
6	medici	0.560000	0.560000	0.521484	0.521978
2	albizzi	0.482759	0.482759	0.210879	0.212454
9	ridolfi	0.482759	0.482759	0.089835	0.086081
7	tornabuon	0.482759	0.482759	0.089670	0.091575
3	guadagni	0.466667	0.466667	0.256813	0.260073
13	barbadori	0.437500	0.437500	0.115110	0.115385
11	strozzi	0.424242	0.424242	0.079176	0.075092
8	bischeri	0.400000	0.400000	0.120549	0.120879
5	salviati	0.388889	0.388889	0.142857	0.142857
14	castellan	0.388889	0.388889	0.088242	0.087912
10	acciaiuol	0.368421	0.368421	0.000000	0.000000
12	peruzzi	0.350000	0.350000	0.021648	0.021978
0	ginori	0.333333	0.333333	0.000000	0.000000
1	lambertes	0.325581	0.325581	0.000000	0.000000
4	pazzi	0.285714	0.285714	0.000000	0.000000

Tabela 10: Comparação entre o analítico e o simulado, considerando *trajetória geodésica* e *transferencia* como difusão da simulação.