Vertex similarity in networks

A. Leicht, Petter Holme, and M. E. J. Newman
Phys. Rev. E 73, 026120 (2006)

Apresentador: Eduardo Hargreaves

Equivalência Estrutural

- Dois vértices são equivalentes se possuem vértices em comum
- Problema: vértices de grau baixo podem ser considerados semelhantes a vértices de grau alto
- Solução: Normalização

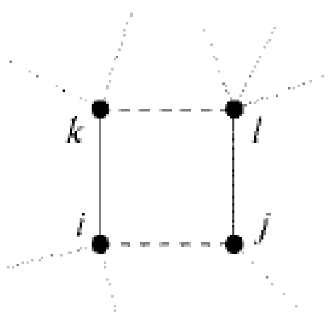
$$\sigma_{\text{Jaccard}} = \frac{|\Gamma_i \cap \Gamma_j|}{|\Gamma_i \cup \Gamma_j|},$$

$$\sigma_{\text{cosine}} = \frac{|\Gamma_i \cap \Gamma_j|}{\sqrt{|\Gamma_i||\Gamma_j|}}, \longrightarrow \cos\theta = \frac{u.v}{\|u\|\|v\|} = \frac{|\tau_i \cap \tau_j|}{\sqrt{|\tau_i||\tau_j|}}$$

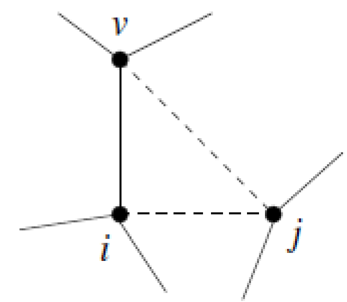
$$\sigma_{\min} = \frac{|\Gamma_i \cap \Gamma_j|}{\min(|\Gamma_i|\,|\Gamma_i|)}.$$

Equivalência Regular

- Para estas definições só existe similaridade, se dois vértices possuem algum vértice em comum
- A equivalência regular extrapola este requisitos quando incorpora papéis



O autor propõe uma equivalência triangular
I é similar a j, se i possui uma aresta com um vértice v que é similar a j
.



Premissas

- As arestas implicam em similaridades entre os vértices a ela conectados
- Implica em relações entre semelhantes
- A similaridade é calculada recursivamente por:

$$S_{ij} = \phi \sum_{v} A_{iv} S_{vj} + \psi \delta_{ij},$$

Similaridade Interpretada como número de caminhos entre ij

- Expandindo-se esta equação em série de potência
$$\mathbf{S} = \mathbf{I} + \phi\mathbf{A} + \phi^2\mathbf{A}^2 + \cdots$$
- $[A^l]_{ij}$ é o número de caminhos entre i e j de comprimento l
- Esta equação diz que um vértice é idêntico a ele mesmo, vértices adjacentes tem similaridade ϕ, vértices simultaneamente adjacentes a i e j tem similaridade ϕ^2, etc...
- Implica em rede conectada

Proposta

- Vértices centrais apresentam mais caminhos de l do que vértices periféricos. Então é preciso normalizar a similaridade
- Proposta: Normalizar pelo número de caminhos esperados em grafo aleatório
- Vértices interligados por mais caminhos do que o esperado são mais similares

- Após algumas manipulações algébricas...

$$\mathrm{DSD} = \frac{\alpha}{\lambda_1}\mathrm{A}(\mathrm{DSD}) + \mathrm{I}.$$

- O único parâmetro independente desta fórmula é α
- $\lambda 1$ é o maior autovalor de A

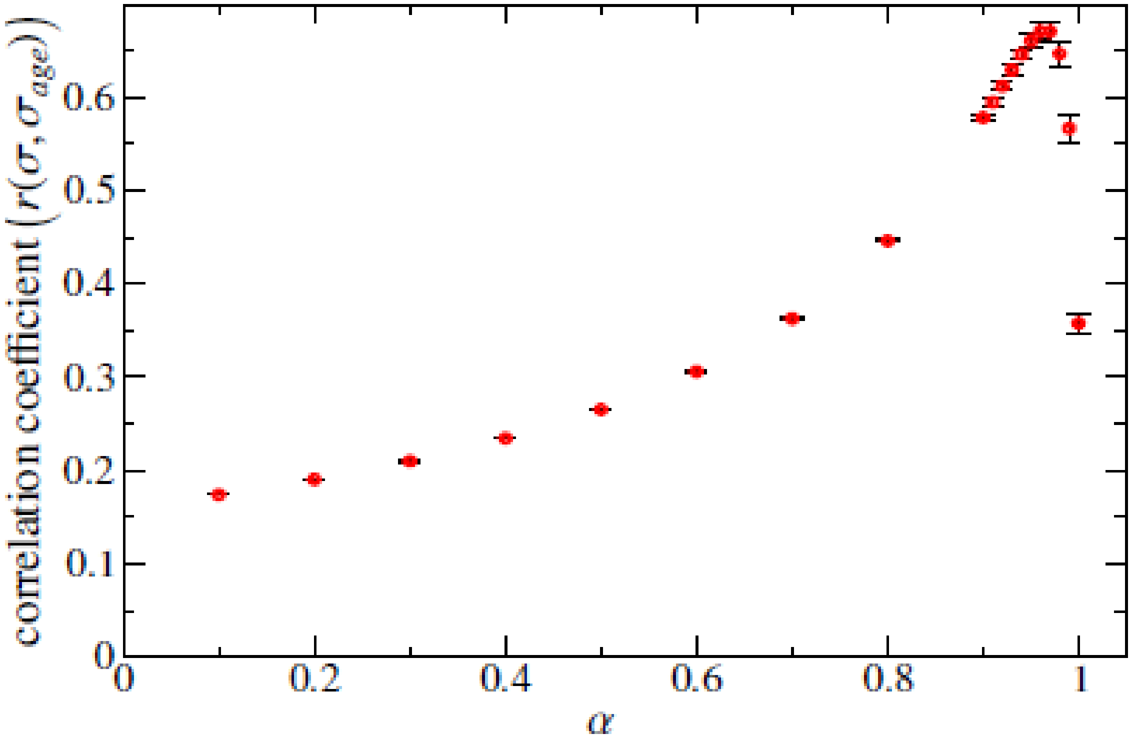
Equivalência Estrutural

- Se for considerados apenas os caminhos de comprimento 2, a medida de similaridade pode ser considerada uma equivalência estrutural

$$\sigma = \frac{|\Gamma_i \cap \Gamma_j|}{k_i k_j} = \frac{|\Gamma_i \cap \Gamma_j|}{|\Gamma_i|\,|\Gamma_j|}.$$

- A normalização é feita pelo número esperado de adjacências em comum. Dois nós são similares se apresentam um número improvável de vizinhos

Escolha de α



Valores entre 0.9 e 0.99

Comentários

- A equivalência necessita de uma rede conectada o que enfraquece o conceito de equivalência regular
- Uma das soluções é permitir que termos fora da diagonal da função de Kronecker tenham valores diferentes de zero. Esta informação viria de fontes além da rede
-