

Processos dinâmicos em redes complexas

Projeto 1: Caracterização de redes complexas: Grau, clustering e caminho

Utilize a linguagem de programação que desejar. Os códigos devem ser entregues conjuntamente com o arquivo que descreve os resultados.

Sugiro usar um dos pacotes: NetworX ou igraph

Os resultados deverão ser entregues como um relatório (ver abaixo). Os códigos deverão ser entregues conjuntamente com o relatório.

Bases de dados:

No site Konect (<http://konect.uni-koblenz.de>) considere as seguintes redes:

1. Social

- **Hamsterster friendships:** <http://konect.uni-koblenz.de/networks/petster-friendships-hamster>
- **Les Misérables:** http://konect.uni-koblenz.de/networks/moreno_lesmis

2. Infra-estrutura:

- **Power grid:** <http://konect.uni-koblenz.de/networks/opsahl-powergrid>

3. Transporte:

- **(Roads) Euroroad:** http://konect.uni-koblenz.de/networks/subelj_euroroad
- **Traffic control (Airport network):** <http://konect.uni-koblenz.de/networks/maayan-faa>

4. Biológica:

- **Human protein (Figeys) (Proteínas):** <http://konect.uni-koblenz.de/networks/maayan-figeys>

Transformem as redes dirigidas (se for o caso) em redes não dirigidas. Veja as descrições no site.

Projeto:

1 - Implemente:

- a) uma rotina que leia as redes dos arquivos.
- b) extraia o maior componente da rede. Você pode usar uma função do NetworX ou igraph.

2 - Testes para a distribuição lei de potencia:

- Obtenha a distribuição do grau de cada rede.
- Obtenha a distribuição de probabilidade acumulada complementar ($P(K > k)$) de cada rede. Compare com a distribuição do grau. Há uma suavização da curva?
- Verifique se a distribuição do grau segue uma lei de potencia para cada uma das redes. Isto é, determine o expoente da lei de potência e k_{min} e realize o teste de hipóteses para verificar se os dados são do tipo lei de potência realmente. Use o teste de Kolmogorov-Smirnov. Ajuste a melhor curva aos dados a partir dessa análise. Os valores do expoente crítico da lei de potencia e k_{min} podem ser indicados nos gráficos. Use as funções disponíveis no site:
<http://tuvalu.santafe.edu/~aaronc/powerlaws/>

2 – Medidas globais:

Usando um das bibliotecas, calcule para cada rede:

- Grau médio.
- Segundo momento da distribuição do grau.
- Entropia de Shannon da distribuição do grau ($H = - \sum P(k) \log P(k)$).
- Média do coeficiente de aglomeração local.
- Coeficiente de aglomeração da rede toda pela fórmula da transitividade
- Média dos menores caminhos.
- Eficiência.
- Diâmetro
- Coloque esses dados em uma tabela e compare as redes reais. Monte um tabela como a tabela II nesse artigo:
<http://arxiv.org/abs/cond-mat/0303516> (página 10).

3 – Clustering

Determine:

- Distribuição de probabilidade acumulada ($F(x) = P(X < x)$) do coeficiente de aglomeração local ($cc(i)$).
- Faça um gráfico de $k(i)$ X $cc(i)$. Verifique se há correlação entre essas duas medidas, ou seja, calcule o coeficiente de correlação de Pearson entre essa duas medidas.
- Faça um gráfico de $cc(k)$, ou seja, do coeficiente de aglomeração em função do grau. Algumas redes apresentam estrutura hierárquica? Veja:
<http://arxiv.org/abs/cond-mat/0206130>
Monte um gráfico como o da figura 3 desse artigo.

Apresentação dos resultados:

Apresente os resultados na forma de um relatório (em português ou inglês).

Sugiro que vocês escrevam o relatório segundo as seções:

a) Introdução:

Escreva a introdução, motivando a área de redes e porque é importante caracterizar uma rede real.

b) Bases

Descreva as bases que foram usadas. Veja <http://konect.uni-koblenz.de>.

c) Conceitos e métodos:

Descreva cada medida usada para caracterizar uma rede.

d) Resultados:

Discuta os resultados obtidos.

Compare as distribuições do grau. Todas as redes são scale-free?

Compare as medidas globais. Qual delas tem menor caminho? Maior clustering?

Eficiência? Compare as redes de rodovias e aeroportos. São parecidas? E as redes sociais? São parecidas?

As redes são apresentadas estrutura hierárquica? Como o clustering se relaciona com o grau nas redes?

As distribuições devem ser apresentadas como histogramas e as médias podem ser colocadas em uma tabela, onde cada linha é uma rede e cada coluna, uma medida.

e) Conclusões:

Conclua com um breve sumário de tudo o que foi feito e comente os resultados principais.

Sugestões:

Escreva o artigo usando o latex.

Os gráficos podem ser feitos usando-se o Matlab, Python ou Gnuplot, por exemplo.

Entrega:

Enviar o texto e os códigos desenvolvidos por email para:

redescomplexas@gmail.com

Paper.pdf + códigos

Colocar no Subject: Redes Complexas: Projeto 1