

## UMA INTRODUÇÃO ÀS REDES COMPLEXAS

Alexandre de Ramos Alabora (PIBIC/CNPq), Haroldo Valentin Ribeiro (Orientador),  
e-mail: [alexandre.alabora@gmail.com](mailto:alexandre.alabora@gmail.com)

Universidade Estadual de Maringá / Centro de Ciências Exatas/ Maringá, PR

### FÍSICA/FÍSICA ESTATÍSTICA E TERMODINÂMICA

**Palavras-chave:** Sistemas Complexos, Redes Complexas, Redes de Dependências.

#### Resumo:

Esse trabalho consiste num estudo teórico de conceitos envolvendo as Redes Complexas e a subsequente construção e caracterização de uma rede real. O processo envolveu o estudo qualitativo de modelos, aspectos probabilísticos e estatísticos da teoria de grafos aplicada às redes complexas assim como a simulação computacional desses modelos. A construção da rede real baseada nas dependências entre os pacotes do Debian (um sistema operacional Linux) permitiu a aplicação de conceitos investigados e comparação entre os resultados dos modelos estudados e as características empíricas da rede real, como o alto coeficiente de aglomeração, o baixo caminho médio e a falta de escala.

#### Introdução

O estudo de sistemas complexos consiste em uma empreitada naturalmente multidisciplinar devido à variedade de fenômenos que possuem a característica de: por meio de muitas partes interagentes produzir um comportamento emergente a princípio não previsível somente conhecendo cada parte. Dada essa aplicabilidade variada o tema foge da Física usual, mas ainda sim a busca de leis gerais que regem esses sistemas se mantém, utilizando principalmente métodos da Física Estatística.

Uma rede complexa é a representação das relações existentes em um sistema complexo por meio de grafos. Trata-se de um conceito utilizado para modelar diversos sistemas sociais, biológicos, tecnológicos e físicos. Em geral, o caráter real dessas redes garante que elas tenham características topológicas não triviais que não ocorrem em grafos puramente aleatórios. Dentre essas características, pode-se citar alta aglomeração, caminho médio pequeno, e falta de escala, características que podem ser capturadas por vários modelos teóricos propostos.

O objeto matemático utilizado para o entendimento e representação de redes complexas são os grafos. Trata-se de um par de conjuntos  $G=\{V,E\}$ , no qual  $V$  é o conjunto de  $N$  vértices e  $E$  é o conjunto de arestas que conectam dois elementos não ordenados (ordenados, quando o grafo é direcionado) de  $V$ . É usual representar um grafo como um conjunto de pontos, no qual dois deles são conectados por uma

linha caso haja uma aresta entre eles. Dadas essas características e o conhecimento do ferramental teórico, a rede de dependência entre os pacotes do sistema operacional *Debian* mostra um bom campo para se explorar, buscando caracterizar uma rede complexa real.

## Revisão Bibliográfica

Aqui apresenta-se uma breve revisão de alguns modelos para representar certas características de redes complexas [1].

*Modelo de Erdős-Rényi (ER)*: Corresponde a uma rede aleatória e foi uns dos primeiros modelos da teoria. Nesse modelo, os vértices se conectam aleatoriamente com uma probabilidade  $\square$ . O modelo ER possui (entre outras propriedades) caminho médio baixo e um coeficiente de aglomeração baixo.

*Modelo de Watts-Strogatz (WS)*: Essa rede permite a descrição do chamado “fenômeno de mundo pequeno”, apresentando um caminho médio muito baixo e um alto coeficiente de aglomeração, características que são observadas em redes reais. O modelo é similar ao ER, mas além do parâmetro  $\square$ , há um parâmetro  $k$  que indica com quantos vizinhos predeterminados um vértice será conectado. Isso permite mapear entre redes completamente regulares e completamente aleatórias.

*Modelo de Barabási-Albert (BA)*: Esse modelo gera as chamadas redes “livres de escala”, nas quais a distribuição de grau dos vértices segue uma lei de potência, ou seja, alguns poucos vértices da rede apresentam uma quantidade muito grande de conexões quando comparados aos demais. O modelo produz esse resultado usando a ideia de conexão preferencial, na qual a probabilidade de uma conexão existir é proporcional à quantidade de conexões já existentes no vértice. Dessa forma, vértices com mais conexões “atrairão” cada vez mais conexões.

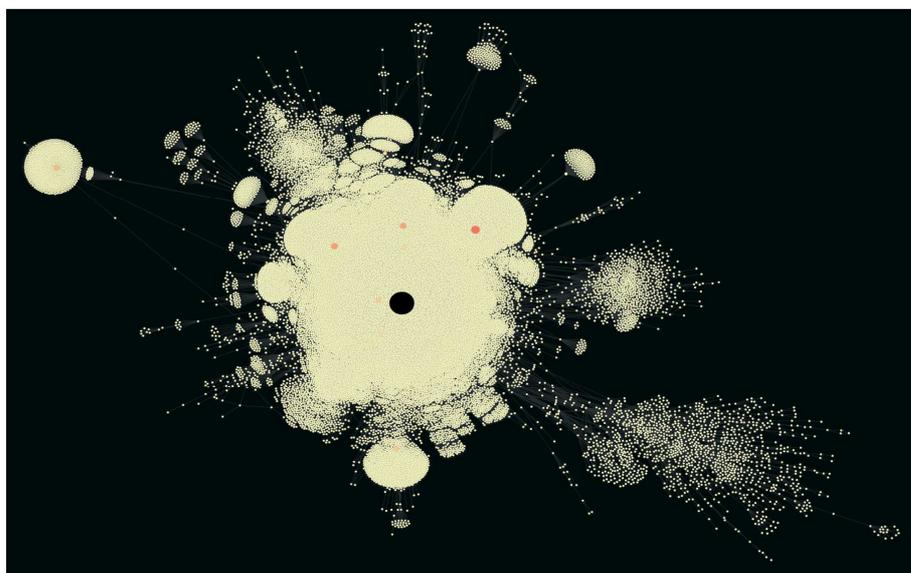
## Materiais e métodos

Todos os aspectos de análise de rede presentes nesse trabalho (grafos, gráficos e métricas) foram construídos usando a linguagem de programação *Python* na distribuição feita para computação científica *Anaconda* junto com o pacote *NetworkX* no ambiente *Jupyter Lab*. O *NetworkX* é uma ferramenta de livre acesso para a análise de redes [2]. Todas as métricas, tanto dos modelos quanto da rede real, foram obtidas usando esse módulo. A única exceção é representação da rede de dependência dos pacotes do *Debian* (Figura 1), na qual foi utilizado módulo *graph\_tool*, capaz de lidar com uma quantidade maior de dados.

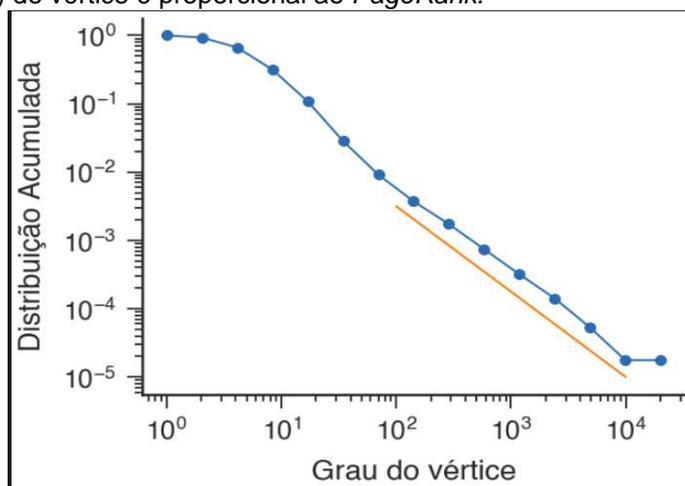
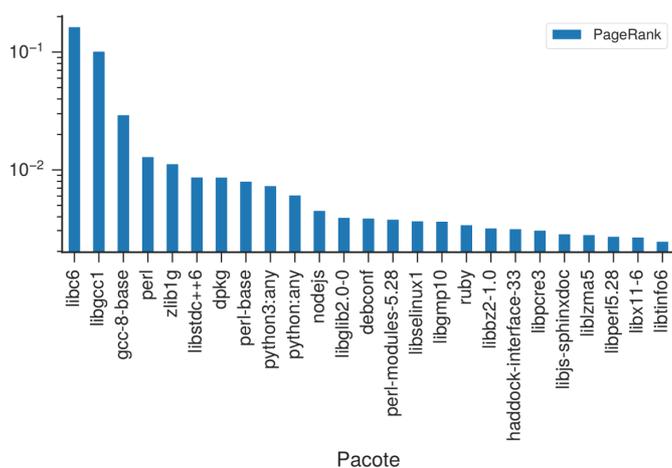
Os dados para criação da rede de dependência entre os pacotes do *Debian* foram obtidos a partir da documentação do *Debian* [3].

## Resultados e Discussão

O Debian é um sistema operacional Linux *open source* cujo o objetivo é criar um sistema universal. Ele possui uma coletânea de pacotes bem estabelecida e, normalmente, para a instalação de um determinado pacote é necessário instalar outros pacotes necessários (as chamadas dependências). Isso cria uma intrincada rede de dependência que já foi alvo de estudos, por exemplo, na referência [4] observou-se a estabilidade da rede frente a ataques aleatórios e dirigidos. Usando esses dados, foi possível construir a rede de dependência, que possui 56594 vértices e 312006 arestas, como representado na Figura 1.



**Figura 1** – Rede das dependências entre os pacotes do Debian. O tamanho e a cor (do branco, passando pelo vermelho até o preto) do vértice é proporcional ao *PageRank*.



**Figura 2** - Ao lado esquerdo, os pacotes com os 25 maiores valores de *PageRank*. Ao lado direito, a distribuição acumulada do grau da rede.

Na Figura 1 é possível notar visualmente a grande conectividade da rede, indicativo de um valor alto para o coeficiente de aglomeração, que de fato é de  $0,16$ , e baixo para o caminho médio. O gráfico no lado direito da Figura 2, mostra que a rede é do tipo livre de escala. O gráfico no lado esquerdo da Figura 2, faz uma classificação da influência dos pacotes usando o *PageRank* [5] (medida da centralidade) e mostra que os principais pacotes são linguagens de programação, suas respectivas bibliotecas, bem como os próprios administradores de pacotes necessários para o funcionamento do sistema.

## Conclusões

A partir do estudo de vários aspectos sobre a teoria de redes complexas foi possível adquirir um panorama razoável de como se comportam as redes reais e quais as limitações dos modelos atuais em descrevê-las. Além disso, utilizando-se de métodos computacionais obteve-se sucesso em construir, representar, e caracterizar a rede de dependência entre os pacotes do Debian.

## Agradecimentos

Agradeço ao CNPq pelo apoio financeiro.

## Referências

- [1] R. Albert, A. L. Barabasi. **Statistical mechanics of complex networks**. Reviews of Modern Physics 74, 47 (2002).
- [2] A. A. Hagberg, D. A. Schult, P. J. Swart. **Exploring network structure, dynamics, and function using NetworkX**, in Proceedings of the 7th Python in Science Conference (SciPy2008), Gael Varoquaux, Travis Vaught, and Jarrod Millman (Eds), (Pasadena, CA USA), 11–15 (2008).
- [3] DependencyHell - Debian Wiki. disponíveis em: <https://wiki.debian.org/DependencyHell>. (Acessado: 28/7/2019).
- [4] O. F. de Sousa, M. A. de Menezes, T. J. P. Penna. **Analysis of the package dependency on Debian GNU/Linux**. Journal of Computational Interdisciplinary Sciences 1, 127 (2009).
- [5] L. Page, S. Brin, R. Motwani, T. Winograd. **The Pagerank Citation Ranking: Bringing Order to the web**, Stanford InfoLab, Stanford - CA, 1998.