

13º Congresso de Inovação, Ciência e Tecnologia do IFSP - 2022

CRIAÇÃO DE UM MINICURSO INTRODUTÓRIO SOBRE A CIÊNCIA DAS REDES: TEORIA E APLICAÇÃO COMPUTACIONAL EM PYTHON

Felipe A. S. M. T. Fernandes¹, Alexandre S. Martinez², Olavo H. Menin³

¹ Instituto Federal de Educação, Ciência e Tecnologia de São Paulo.

² Universidade de São Paulo.

³ Instituto Federal de Educação, Ciência e Tecnologia de São Paulo.

Área de conhecimento (Tabela CNPq): 1.03.02.02-6 Modelos analíticos e de simulação.

RESUMO: O mundo vem se tornando cada vez mais conectado e as relações entre nações, pessoas e empresas, mais complexas. Com isso, o tradicional cartesianismo tem perdido espaço para um novo paradigma, a ciência das redes, que busca estudar não só as partes de um sistema separadamente, mas principalmente as interações entre elas. Apesar de seu amadurecimento como uma das mais importantes áreas de pesquisa nas últimas décadas e de seu caráter multidisciplinar, a ciência das redes é pouco abordada em cursos de graduação, mesmo na área de exatas. Nesse contexto, apresentamos os resultados parciais de um projeto de iniciação científica cujo objetivo é criar um minicurso introdutório sobre ciência das redes, com aplicação em Python, que atenda estudantes de graduação da área de exatas.

PALAVRAS-CHAVE: ensino e aprendizagem; teoria dos grafos; modelagem; NetworkX.

CREATION OF AN INTRODUCTORY MINI-COURSE ON NETWORK SCIENCE: THEORY AND COMPUTATIONAL APPLICATION IN PYTHON

ABSTRACT: The world is becoming more and more connected and the relationships between nations, people and companies more complex. As a result, traditional Cartesianism has lost ground to a new paradigm, the science of networks, which seeks to study not only the parts of a system separately, but mainly the interactions between them. Despite its maturation as one of the most important areas of research in recent decades and its multidisciplinary nature, network science is rarely addressed in undergraduate courses, even in the exact sciences area. In this context, we present the partial results of a undergraduate research project whose objective is to create an introductory mini-course on network science, with application in Python, for undergraduate students in the exact sciences area.

KEYWORDS: teaching and learning; graph theory; modelling; NetworkX.

INTRODUÇÃO

O mundo moderno, globalizado e altamente interconectado, é uma das maiores testemunhas de que nada ocorre de forma isolada. De fato, inúmeros fenômenos estudados pela ciência moderna, comu-

mente classificados como “complexos”, resistem em serem descritos pelo tradicional método cartesiano dissecador/integrador (LADYMAN; LAMBERT; WIESNER, 2013). Nesse sentido, a ciência das redes surge como um novo paradigma para o estudo de diversos fenômenos, tais como propagação de doenças infecciosas em uma população ou de rumores e notícias falsas em redes sociais, transações comerciais entre empresas, estrutura de hiperlinks entre páginas da web e conexões entre reações bioquímicas em uma célula (BARABÁSI, 2009).

Historicamente, a origem da ciência das redes remonta ao final do século 18, quando Leonard Euler solucionou o problema das pontes de Königsberg a partir de uma abstração que representava a situação como um grafo, conjunto de vértices e arestas (PAOLETTI, 2011). Já em meados do século 20, os matemáticos Paul Erdős e Alfréd Rényi apresentaram o primeiro modelo de rede, formalizando e sistematizando as propostas de Euler (ERDŐS; RÉNYI, 1960). Já em 1999, Albert-László Barabási e Réka Albert mostram que as redes do tipo Erdős-Rényi não são adequadas para descrever a maior parte das redes observadas no mundo real e, então, propõem um novo modelo baseado em dois mecanismos: crescimento e ligação preferencial (BARABÁSI; ALBERT, 1999; ALBERT; BARABÁSI, 2002). A partir de então, inúmeros grupos de pesquisa ao redor do mundo têm se dedicado à ciência das redes, fazendo dela uma das que mais cresceram em termos de publicações nas últimas décadas. (BARABÁSI; PÓSFAL, 2016).

A despeito de seu amadurecimento e de seu caráter destacadamente multidisciplinar, a ciência das redes ainda é pouco abordada em cursos de graduação na área de exatas, principalmente de engenharia. Com isso, apresentamos aqui os resultados parciais de um projeto de iniciação científica (IC) cujo objetivo é criar um minicurso introdutório sobre ciência das redes que atenda estudantes de graduação da área de exatas. A proposta é que o curso adote uma abordagem didática dinâmica e ativa, integrando a parte conceitual com aplicações computacionais por meio da linguagem Python. Como resultados, espera-se que o material produzido proporcione aos estudantes (e ao público geral) um primeiro contato com essa tão relevante área da ciência e abra portas para estudos futuros.

MATERIAIS E MÉTODOS

A orientação da IC ocorreu por meio de reuniões entre orientador e orientando, nos quais foram tratados os temas específicos do projeto. Além disso, o orientando foi incluído no Grupo de Estudo de Modelagem e Simulação Computacional (GEMSC), participando de reuniões periódicas com os demais membros nas quais, entre outras atividades, os estudantes orientandos e docentes orientadores apresentaram seminários sobre o que estavam desenvolvendo ou outros assuntos pertinentes ao grupo.

Para o desenvolvimento do trabalho, inicialmente foi realizada uma revisão sobre conceitos básicos de cálculo (THOMAS et al., 2009), matrizes (FRANCO, 2016) e probabilidade/estatística (MORETTIN, 2010), bem como escrita em Latex e linguagem Python. Em seguida, deu-se início aos estudos sobre a ciência das redes utilizando o livro base *Network Science* (BARABÁSI; PÓSFAL, 2016) e, quando pertinente, implementado códigos em Python de forma a aplicar os conceitos teóricos que estão sendo trabalhados. À medida que os estudos avançaram, foi redigido o texto principal do minicurso utilizando-se um *template* apropriado em Latex a fim de que o produto final tenha a melhor qualidade possível e siga os padrões de um bom texto acadêmico. Foi estudado também o pacote NetworkX para realização de aplicações computacionais com a linguagem Python. Com o auxílio desse pacote pode-se criar, manipular e estudar a estrutura e dinâmica das redes (HAGBERG; SCHULT; SWART, 2008).

RESULTADOS E DISCUSSÃO

Apresentamos agora os resultados parciais obtidos até o momento da submissão desse resumo. Primeiro descrevemos brevemente alguns dos principais conceitos sobre redes que são abordados no minicurso e, em seguida, a ideia geral de como o material está sendo desenvolvido.

Conceitos básicos sobre redes

Uma rede pode ser entendida como um conjunto de nós que podem ou não estarem conectados entre si por links. Matematicamente, ela pode ser denotada por $G = (V, E)$, onde $V = \{V_1, V_2, \dots, V_N\}$ são os vértices e $E = \{E_1, E_2, \dots, E_L\}$ os links. No presente minicurso, focaremos em redes não direcionada (links sem direção preferencial), sem nós com auto-ligações (links que conectam um nó a si mesmo) e cujos links tem o mesmo peso. Neste caso, a rede pode ser descrita pela matriz de adjacência $\mathbf{A}$, cujos elementos a_{ij} assumem os valores

$$a_{ij} = \begin{cases} 1, & \text{se há um link entre os nós } V_i \text{ e } V_j, \\ 0, & \text{caso contrário.} \end{cases} \quad (1)$$

Cada nó é caracterizado pelo grau k , que corresponde ao número de links que ele possui. Com isso, o grau médio da rede é dado por

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i. \quad (2)$$

Já a rede, como um todo, é caracterizada pela distribuição de graus, que corresponde às probabilidades p_k de um nó ter grau k , com $k = k_{\min}, \dots, k_{\max}$, onde $k_{\min}$ e $k_{\max}$ são, respectivamente, o menor e o maior grau presente na rede e garantindo que $\sum p_k = 1$.

Além desse e outros conceitos gerais sobre redes, o minicurso aborda os dois tipos mais conhecidos de redes, as *redes randômicas* do tipo Erdős-Rényi e as *redes sem escala* do tipo Barabási-Albert. No primeiro, tem-se um número N fixo de nós conectados dois a dois com probabilidade p . Para esse modelo, a distribuição de graus p_k pode ser aproximada por uma distribuição de Poisson,

$$p_k \approx e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}. \quad (3)$$

Já no modelo de rede sem escala, o número de nós da rede não fixo, crescendo continuamente com o tempo. Além disso, utiliza o conceito de conexão preferencial, segundo o qual os novos nós adicionados tendem a se conectarem aos nós existentes mais conectados, ou seja, a probabilidade de um novo nó ser conectado a um nó pré-existente i é

$$\Pi(k_i) = \frac{k_i}{\sum_{j=1}^N k_j}. \quad (4)$$

Neste modelo, a distribuição de graus segue uma lei de potência,

$$p_k = C k^{-\gamma}, \quad (5)$$

onde $2 < \gamma < 3$ é o expoente de grau e C é uma constante de normalização. A característica marcante das redes sem escala é a coexistência entre nós com poucos links e nós altamente conectados (*hubs*), o que não ocorre nas redes randômicas.

Material produzido

O minicurso, que está inicialmente no formato de um texto, é dividido em quatro capítulos. O primeiro, introdutório, contextualiza o surgimento e o desenvolvimento da ciência das redes, destacando algumas de suas possíveis aplicações. No segundo, são abordados os principais conceitos e métricas que caracterizam as redes. O terceiro e quarto capítulos apresentam, respectivamente, os modelos de Erdős-Rényi e de Barabási-Albert, mostrando suas particularidades e diferenças.

Até o momento, estão finalizados o capítulo dois, “Conceitos básicos sobre grafos”, e o capítulo três, “Redes randômicas e modelo de Erdős-Rényi”. O quarto capítulo, “Redes sem escala e modelo de Barabási-Albert”, está parcialmente escrito e o primeiro capítulo, “Introdução e motivação”, será escrito por último.

Com exceção do primeiro, cada capítulo está estruturado com uma breve introdução sobre o tema abordado, o conceito a ser tratado, a aplicação computacional em Python e, por fim, um resumo em tópicos sobre o que foi abordado. Para ilustrar, a Figura 1 mostra essa estruturação, com as seções de um capítulo já finalizado.

3	Redes randômicas e modelo de Erdős-Rényi	17
3.1	Introdução	17
3.2	Rede aleatória de Erdős-Rényi	17
3.3	Características da rede aleatória	18
3.3.1	Número de links e grau médio	18
3.3.2	Distribuição de grau	18
3.3.3	Caminho médio e diâmetro	19
3.3.4	Coefficiente de agrupamento	20
3.4	Evolução das redes aleatória	20
3.5	Aplicações em Python	22
3.6	Resumo	25

Figura 1: Trecho do sumário mostrando o capítulo 3 e suas seções. (Fonte: autores)

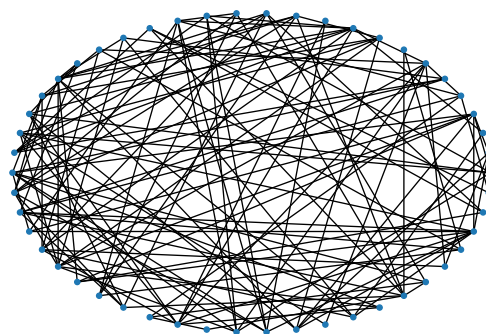
Quanto à aplicação computacional, os códigos são exibidos em um ambiente específico com comentários e, caso esse gere uma saída em formato de imagem, ela é exibida no material, como mostrado na Figura ??.

```
# Instanciando grafo e adicionando os nós
graph = nx.Graph()
list_of_nodes = np.arange(N)
graph.add_nodes_from(list_of_nodes)

# Passando por todas as ligações possíveis
# e verificando se há ou não ligação
for node_i in list_of_nodes:
    for node_j in list(graph.nodes):
        if node_i < node_j:
            random = rd.random()
            if random < p:
                graph.add_edge(node_i, node_j)

nx.draw_circular(graph, node_size=15)
```

(a)



(b)

Figura 2: (a) Código que gera rede do tipo Erdős-Rényi com $N = 50$ e $p = 0.1$ e (b) a imagem da rede gerada por ele. (Fonte: autores)

CONCLUSÕES

Nesse resumo, inicialmente foram discutidos o desenvolvimento e a relevância da ciência das redes, alguns conceitos e métricas fundamentais na análise de redes, os modelos de Erdős-Rényi e Barabási-Albert, bem como foi apontada a necessidade de maior divulgação e disponibilidade de materiais sobre o assunto. A partir disso, foi apresentada uma proposta de minicurso introdutório sobre ciência das redes voltado para alunos de graduação na área de exatas. Espera-se que, após a conclusão desse projeto, o minicurso e o material produzido sejam disponibilizados virtualmente de forma a fomentar estudantes e até o público geral a estudarem e se aprofundarem no assunto.

AGRADECIMENTOS

Os autores agradecem ao Programa Institucional de Bolsa de Iniciação Científica e Tecnológica (PIBIFSP) do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo.

REFERÊNCIAS

- ALBERT, R.; BARABÁSI, A. L. Statistical mechanics of complex networks. *Reviews of Modern Physics*, v. 74, p. 47–97, 2002.
- BARABÁSI, A. L. *Linked: a nova ciência dos networks*. [S.l.]: Leopardo Editora, 2009.
- BARABÁSI, A. L.; ALBERT, R. Emergence of scaling in random networks. *Science*, v. 286, p. 509–512, 1999.
- BARABÁSI, A.-L.; PÓSFAL, M. *Network Science*. Cambridge: Cambridge University Press, 2016.
- ERDŐS, P.; RÉNYI, A. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 1960.
- FRANCO, N. *Álgebra linear*. São Paulo: Pearson Education do Brasil, 2016.
- HAGBERG, A. A.; SCHULT, D. A.; SWART, P. J. Exploring network structure, dynamics, and function using networkx. In: VAROQUAUX, G.; VAUGHT, T.; MILLMAN, J. (Ed.). *Proceedings of the 7th Python in Science Conference*. Pasadena, CA USA: [s.n.], 2008. p. 11 – 15.
- LADYMAN, J.; LAMBERT, J.; WIESNER, K. What is a complex system? *European Journal for Philosophy of Science*, v. 3, p. 33–67, 2013.
- MORETTIN, L. G. *Estatística básica: probabilidade e inferência*. [S.l.: s.n.], 2010.
- PAOLETTI, T. *Leonard Euler's Solution to the Königsberg Bridge Problem*. 2011. Disponível em: <<https://www.maa.org/press/periodicals/convergence/leonard-eulers-solution-to-the-konigsberg-bridge-problem>>.
- THOMAS, G. B. et al. *Cálculo - vol. 1*. São Paulo: Pearson - Addison Wesley, 2009.