Network Science

Barabási: Ch. 2 — Graph Theory — Lecture 1

Joao Meidanis

University of Campinas, Brazil

September 20, 2020

Summary

Origin of Graph Theory

Walk across all seven bridges and never cross the same one twice

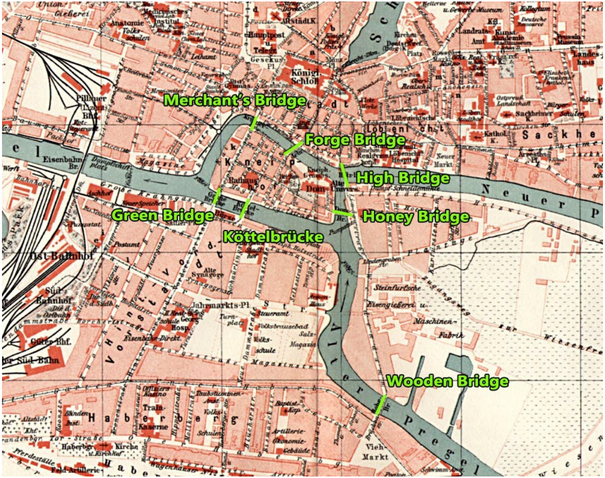


Figure source:

Amusing Planet

(amusingplanet.com)

Author:

Kaushik Patowary

The Solution

Impossible: 4 nodes of odd degree

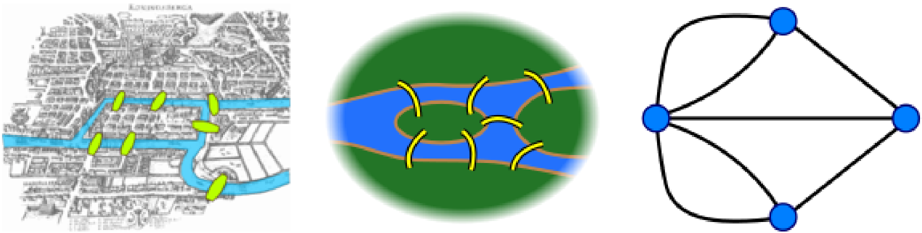


Figure source: Wikipedia, *Seven Bridges of Königsberg*; authors: Bogdan Giuşcă, Chris-martin, Riojajar commonswiki; license: Creative Commons Attribution-Share Alike 3.0 Unported

Networks and Graphs

Networks and Graphs

Parameters

- N = number of nodes
- L = number of links

Type

- Directed or undirected (or mixed)

Terminology

Network Science	**Graph Theory**
Network	Graph
Node	Vertex
Link	Edge

Ten Basic Networks Used in Book

Network	Nodes	Links	Type	N	L
Internet	Routers	Connections	Undir.	192,244	609,066
WWW	Web pages	Links	Dir.	325,729	1,497,134
Power Grid	Plants, stations	Cables	Undir.	4,961	6,594
Mobile	Suscribers	Calls	Dir.	36,595	91,826
Email	Addresses	Messages	Dir.	57,194	103,731
Science Collab.	Scientists	Co-authors	Undir.	23,133	93,437
Hollywood	Actors	Co-acting	Undir.	702,388	29,397,908
Citation	Papers	Citations	Dir.	449,673	4,689,479
E. coli	Metabolites	Reactions	Dir.	1,039	5,802
Cell	Proteins	Interactions	Undir.	2,018	2,930

Degrees

Degree, Average Degree

Degree of a node: number of links to other nodes

$$k_i = \text{ degree of node } i, \text{ for } i = 1 \ldots N$$

Links and degree (undirected network)

$$L = \frac{1}{2} \sum_{i=1}^{N} k_i$$

Average degree

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i = \frac{2L}{N}$$

Directed Networks

Total degree is made of incoming degree and outgoing degree

$$k_i = k_i^{in} + k_i^{out}$$

Links and degree (directed network)

$$L = \sum_{i=1}^{N} k_i^{in} = \sum_{i=1}^{N} k_i^{out}$$

Average degree

$$\langle k^{in} \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i^{in} = \langle k^{out} \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i^{out} = \frac{L}{N}$$

Degree Distribution

Probabilistic distribution of degrees

$$p_k = \text{ probability of a random node having degree } k$$

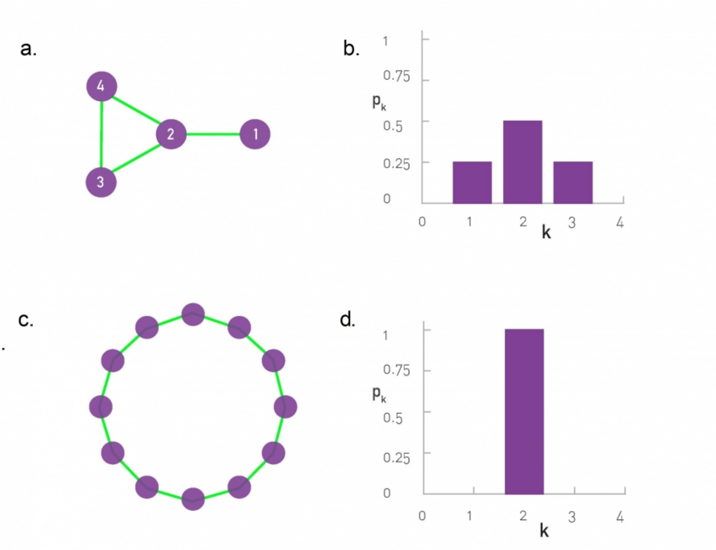
If $N_k =$ number of nodes with degree k then

$$p_k = \frac{N_k}{N}$$

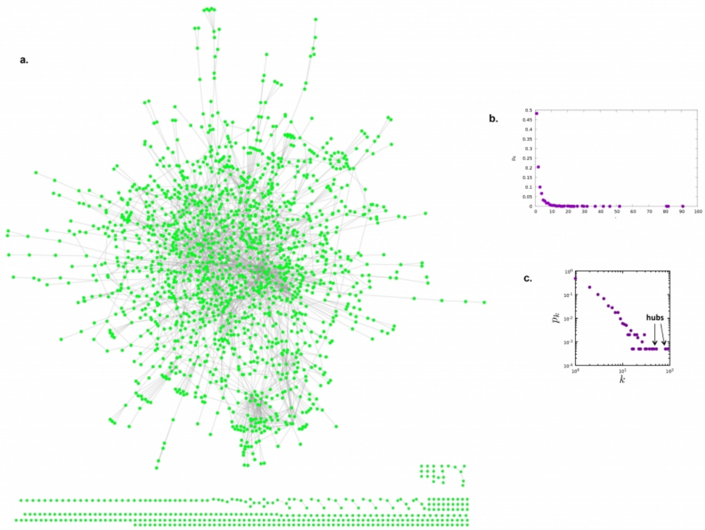
Average degree in terms of degree distribution

$$\langle k \rangle = \sum_{k=0}^{\infty} k p_k$$

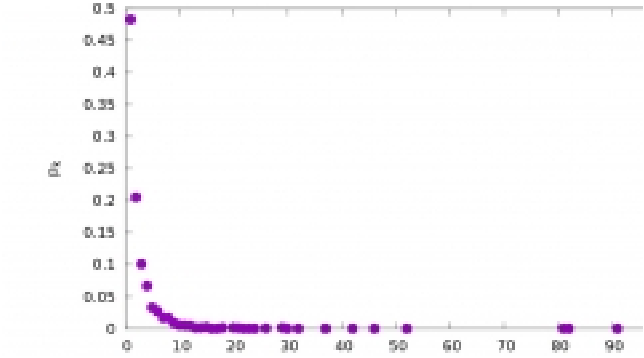
Degree Distribution — Examples



Real Network (Protein Interaction in Yeast)



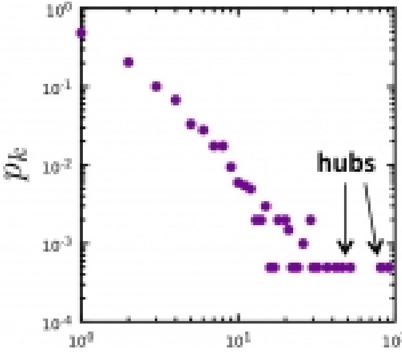
Zoom in — Degree Distribution



$$p_1 = 0.48$$

$$p_{92} = 1/N = 0.0005$$

Log-Log Plot — Degree Distribution



power law becomes a straight line

Adjacency Matrix

Adjacency Matrix

N rows, N columns; elements:

$$A_{ij} = \begin{cases} 1 & \text{if } j \to i \\ 0 & \text{otherwise} \end{cases}$$

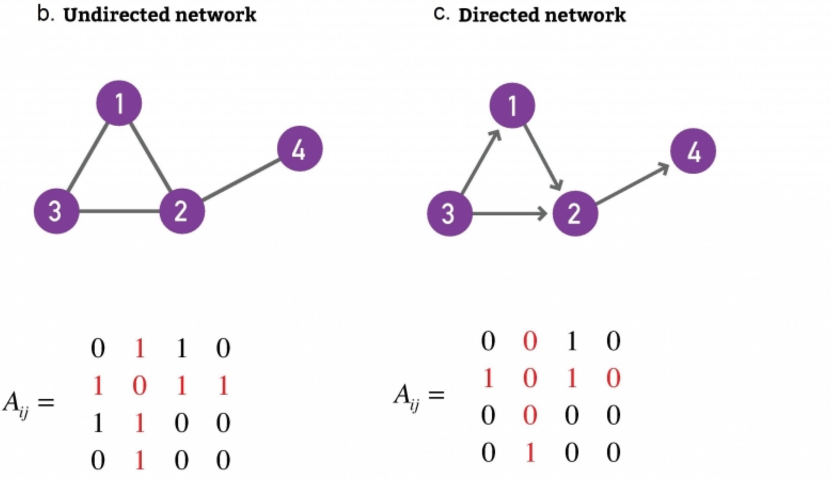
Degrees from adjacency matrix (undirected network)

$$k_i = \sum_{j=1}^{N} A_{ji} = \sum_{j=1}^{N} A_{ij}$$

Degrees from adjacency matrix (directed network)

$$k_i^{in} = \sum_{j=1}^{N} A_{ij}, \qquad k^{out} = \sum_{j=1}^{N} A_{ji}$$

Adjacency Matrix



Real Networks are Sparse

Number of nodes N varies wildly

- *C. elegans*: $N \sim 10^2$ neurons
- Human cell: $N \sim 10^4$ genes
- Social network: $N \sim 10^8$ people
- Human brain: $N \sim 10^{11}$ neurons
- WWW: $N > 10^{12}$ documents

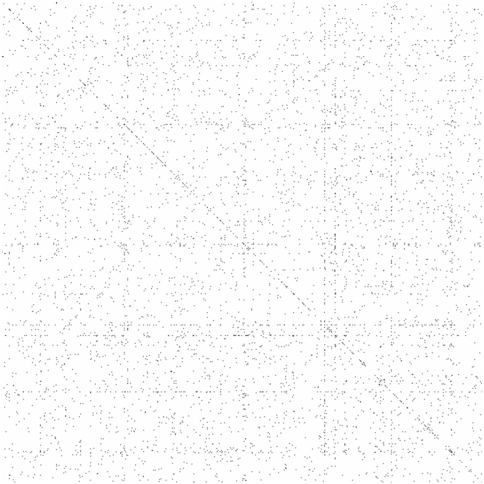
$$0 \leq L \leq L_{max} = \frac{N}{2} k_{max} = \frac{N(N-1)}{2}$$

Number of links L also varies wildly

Usually much closer to N than to L_{max}

Network	N	L	$\langle k \rangle$
Internet	192,244	609,066	6.34
WWW	325,729	1,497,134	4.60
Power Grid	4,961	6,594	2.67
Mobile calls	36,595	91,826	2.51
Email	57,194	103,731	1.81
Science Collaboration	23,133	93,437	8.08
Hollywood	702,388	29,397,908	83.71
Citation network	449,673	4,689,479	10.43
E. coli metabolism	1,039	5,802	5.58
Cell network	2,018	2,930	2.90

Adjacency Matrix of Sparse Network



- Yeast protein-protein interaction network
- Dots in positions where $A_{ij} = 1$
- For efficiency, store just list of 1-positions

Weighted Networks

Weighted Networks

$$A_{ij} = w_{ij}$$

Metcalfe's Law (Used in the late 1990's to evaluate internet companies)

- The value of a network is proportional to the square of the number of its nodes, i.e., N^2

Limitations:

- The value is in fact proportional to the links created
- Most real networks are sparse, so L does not grow like N^2
- Links have different weights
- Some links are used heavily but the vast majority are rarely utilized

Bipartite Networks

Bipartite Networks

All links have one end in U and the other in V