CS5016: Computational Methods and Applications

Networks, Random Graphs and Percolation

Albert Sunny

Department of Computer Science and Engineering
Indian Institute of Technology Palakkad

25 January, 2024

What are networks?

A network (or graph, as it is often referred to in mathematics) is a data structure in which nodes are connected by edges.

They provide a very general concept that plays a role in many scientific problems.

Random graphs

- **Random graph** is the general term to refer to probability distributions over graphs. Random graphs may be described simply by a probability distribution, or by a random process which generates them[1].
- The theory of random graphs lies at the intersection between graph theory and probability theory.
- Its practical applications are found in all areas in which complex networks need to be modeled.

[1] https://en.wikipedia.org/wiki/Random_graph

Erdős-Rényi random graphs

- Simplest and well-studied class of random graphs; named after Hungarian mathematicians Paul Erdős and Alfréd Rényi.

- In the $G(n, p)$ model, a graph is constructed by connecting labeled nodes randomly. Each edge is included in the graph with probability p, independently from every other edge.

- $G(n, p)$ can be thought of sampling a graph with n vertices and M edges with probability
$$p^M (1-p)^{\binom{n}{2}-M}$$

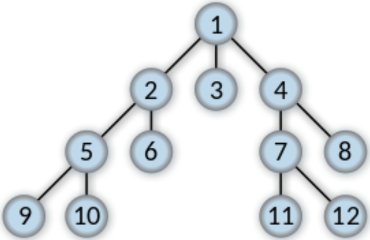
A few properties of Erdős-Rényi random graphs

- If $p < 1/n$, then a graph in $G(n, p)$ will almost surely have **no connected components** of size larger than $O(log(n))$.
- If $p > 1/n$, then a graph in $G(n, p)$ will almost surely have **a unique giant connected component** containing a positive fraction of the vertices, and no other component will contain more than $O(log(n))$ vertices.
- If $p < \frac{(1-\epsilon)\ln n}{n}$, then a graph in $G(n, p)$ will almost surely **contain isolated vertices**, and thus be disconnected.
- If $p > \frac{(1-\epsilon)\ln n}{n}$, then a graph in $G(n, p)$ will almost surely be **connected**.

Breadth First Search/Traversal

Breadth-first search (BFS) is an algorithm for traversing or searching tree or graph data structures. It starts at an arbitrary node of a graph and explores all of the neighbor nodes at the present depth prior to moving on to the nodes at the next depth level [2]. A few applications of BFS are:

- Shortest path and minimum spanning tree
- Cycle detection in undirected graph
- Ford–Fulkerson algorithm
- Finding all nodes within one connected component



A YouTube video on BFS is available at
https://www.youtube.com/watch?v=oDqjPvD54Ss

[2] https://en.wikipedia.org/wiki/Breadth-first_search

Percolation

- In statistical physics and mathematics, percolation theory describes the behavior of a network when nodes or links are removed.
- This is a geometric type of phase transition, since at a critical fraction of removal the network breaks into significantly smaller connected clusters.
- Percolation theory finds applications in materials science and in many other disciplines.

Percolation

- In statistical physics and mathematics, percolation theory describes the behavior of a network when nodes or links are removed[3].
- This is a geometric type of phase transition, since at a critical fraction of removal the network breaks into significantly smaller connected clusters.
- Percolation theory finds applications in materials science and in many other disciplines.

[3] `https://en.wikipedia.org/wiki/Percolation_theory`

Bond percolation

- Consider a large 2D sheet made up of a porous material. Assume that some liquid is poured on top of it.
- A graph of $n \times n$ vertices (n is large), usually called *"sites"*, in which the edge or *"bonds"* between two neighbors may be open (allowing the liquid through) with probability p, or closed with probability $1-p$ (they are assumed to be independent).
- For a given p, what is the probability that an open path (meaning a path, each of whose links is an "open" bond) exists from the top to the bottom?
- The square lattice $\mathbb{Z}^2$ in two dimensions exhibits a sharp phase transition at $p = 1/2$.

Python's NetworkX module

NetworkX is a Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks.

To learn more, visit
`https://networkx.org/documentation/stable/index.html`

Thank You