LECTURES ON RANDOM GRAPHS

History

In the 1950's Pál Erdős realized the power of the probabilistic method in graph theory and started to exploit it systematically. In these proofs one often chooses a graph uniformly at random among all $2^{\binom{n}{2}}$ graphs on n vertices. (Recall Erdős' proof of his lower bound for the diagonal Ramsey numbers $R(k,k)$.)

Then the question became natural: what do random graphs "look like"?

In a series of papers Pál Erdős and Alfréd Rényi initiated the systematic study of random graphs (between 1959 and 1961). These truly remarkable papers reveal many surprising and beautiful features and establish the main technical tools.

A flurry of activities followed and by the mid 1980's random graph theory became a well-established independent field. The excellent book of Béla Bollobás (1985) summarized the state of the art.

Today the study of random graphs became widely spread among many areas of research such as (just to name a few), statistical physics, epidemiology, study of social networks, biological and chemical networks, communication networks, etc.

BOOKS

Janson, Łuczak, & Ruciński : RANDOM GRAPHS, 2000
Penrose : RANDOM GEOMETRIC GRAPHS, 2003
Durrett : RANDOM GRAPH DYNAMICS, 2010
Grimmett : PROBABILITY ON GRAPHS, 2010
Lu & Chung : COMPLEX GRAPHS AND NETWORKS, 2006
van der Hofstad : RANDOM GRAPHS AND COMPLEX NETWORKS, 2014+

MODELS OF RANDOM GRAPHS

- **Uniform:** In the simplest model one chooses a graph uniformly at random among all $2^{\binom{n}{2}}$ graphs on n (labeled) vertices. Thus, the probability that the random graph G_n equals any particular graph g is $\mathbb{P}(G_n = g) = 2^{-\binom{n}{2}}$.

  Equivalently, every one of the $\binom{n}{2}$ edges is present with probability $\frac{1}{2}$, independently. It is this independence that makes the model attractive and tractable. Note that the number of edges of G_n has binomial distribution $\mathrm{Bin}(\binom{n}{2}, \frac{1}{2})$ and it is concentrated around $\binom{n}{2}/2 \pm \Theta_p(n) \approx \frac{n^2}{4} \pm \Theta_p(n)$

- $G(n,p)$ — Erdős-Rényi model:

This is a generalization of the uniform distribution. Each one of the $\binom{n}{2}$ edges is present with probability p for some $p \in [0,1]$, independently. The number of edges is $\mathrm{Bin}\left(\binom{n}{2}, p\right)$ and the degree of any given vertex is $\mathrm{Bin}(n-1, p)$. If p is a constant, the graph is dense as it has $\sim \frac{n^2 p}{2} \pm \Theta\left(\sqrt{n^2 p(1-p)}\right)$ edges. One often considers $p = p(n) \to 0$ to get sparser graphs. One may define the entire process $G(n,p) : p \in [0,1]$ on a single probability space by assigning an independent $\mathrm{Unif}([0,1])$ random variable U_e to each edge e and defining $X_e = \mathbb{1}_{U_e < p}$ as the indicator of edge e.

- $G(n,M)$: also defined by Erdős and Rényi, this is a uniformly distributed random graph among all $\binom{\binom{n}{2}}{M}$ graphs on n vertices with exactly M edges.

This model is quite similar to $G(n,p)$ with $p = \frac{2M}{n^2}$ because in $G(n,p)$ the number of edges is concentrated around $\frac{n^2 p}{2}$. There are no major differences in the behavior of $G(n,M)$ and $G(n,p)$.

This model also has a "process" version: start with the empty graph and at each time step select an edge at random and add it to the graph. After M steps we arrive at $G(n,M)$ (Why?)

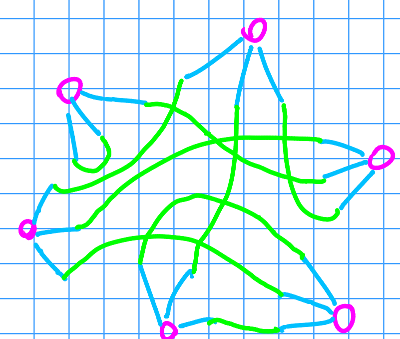
- **Random regular graphs**

A graph is called d-regular if every vertex has degree d. (We need that $n \cdot d$ is even.)

Exercise: How many 2-regular graphs are there?

(The number for general d is not known except for asymptotic formulas.)

The analysis of random regular graphs crucially uses the so-called **configuration model**: generate a random graph by creating d "half edges" for each vertex and match the nd half edges at random:



Note that what we obtain is, in general, a **multigraph** that may contain loops and multiple edges. However this doesn't happen too often:

The probability that the multigraph is **simple** (i.e., no loops, no multiple edges) converges to

$\exp\left(-\frac{(d-1)^2}{4}\right)$ as $n \to \infty$. Crucially, given that the multigraph is simple, the conditional distribution of the obtained graph is uniform among all d-regular graphs. As a consequence, if one proves that a certain property holds with probability $1-o(1)$ in the configuration model, it also holds for RRG's.

• General degree sequences

For $d_1 \le d_2 \le \dots \le d_n$ positive integers, we may consider graphs chosen uniformly at random from all graphs such that the vertices have degrees $d_1, \dots, d_n$. Of course, not all degree sequences are "graphic". Erdős & Gallai (1960) proved that there exist graphs with degrees $d_1 \le d_2 \le \dots \le d_n$ if and only if $\sum_{i=1}^{n} d_i$ is even and for all $k = 1, \dots, n$,

$$\sum_{i=1}^{k} d_i \le \underbrace{k(k-1)}_{\text{max total degree of subgraph induced by the first } k \text{ vertices}} + \underbrace{\sum_{i=k+1}^{n} \min(k, d_i)}_{\text{upper bound on total degree of vertices } 1,\dots,k \text{ coming from edges to } v_{k+1},\dots,v_n}$$

(Necessity is easy: max total degree of subgraph induced by the first k vertices; upper bound on total degree of vertices $1, \dots, k$ coming from edges to $v_{k+1}, \dots, v_n$)

The configuration model generalizes to this case.

Inhomogeneous random graphs

One may generalize the Erdős–Rényi model to obtain graphs with uneven degree sequences. Each vertex receives a **weight** $w_i > 0$, $i = 1, \dots, n$.

Let $\ell_n = \sum_{i=1}^{n} w_i$. Connect vertex i and j at random, independently, with probability

$$\frac{w_i w_j}{\ell_n + w_i w_j} \qquad \left(\text{or } \min\left(\frac{w_i w_j}{\ell_n}, 1\right) \text{ in a variant}\right)$$

The expected degree of vertex i is

$$\sum_{j : j \neq i} \frac{w_i w_j}{\ell_n + w_i w_j} \le w_i \sum_j \frac{w_j}{\ell_n} \le w_i$$

If $w_i = w$ for all i, we get the Erdős–Rényi $G(n,p)$ model with $p = \frac{w}{n+w}$.

Another special case is the stochastic block model or planted partition: partition the vertex set into k "clusters" $C_1, \dots, C_k$: $C_i \cap C_j = \emptyset$, $\bigcup_{i=1}^{k} C_i = \{v_1, \dots, v_n\}$ given a symmetric matrix of nonnegative numbers $[q_{ij}]_{k\times k}$, $\mathbb{P}(v_m \sim v_\ell) = q_{ij}$ if $v_m \in C_i$, $v_\ell \in C_j$ $(m \neq \ell)$.

In random bipartite graphs $k=2$ and $q_{11} = q_{22} = 0$.

- W-random graphs

 Let $w : [0,1]\times[0,1] \to [0,1]$ be a measurable symmetric function, i.e., $w(x,y) = w(y,x)$ (a "graphon"). Let $X_1, \dots, X_n$ be independent, uniformly distributed on $[0,1]$. Vertex i and vertex j are connected with probability $w(X_i, X_j)$. (Note the two levels of randomness.) If $w(x,y) \equiv p$, then we recover $G(n,p)$.

 W-random graphs are important in the study of graph limits, see Lovász & Szegedy.

- Random geometric graphs

 Let $X_1, \dots, X_n$ be independent, identically distributed random vectors distributed according to a density $f(x)$ on $\mathbb{R}^d$. In the simplest model, f is uniform on $[0,1]^d$.

 In a neighborhood graph vertices i and j are joined by an edge if and only if $\|X_i - X_j\| < r_n$ for some $r_n > 0$.

 In the k-nearest neighbor graph $i \sim j$ if X_i is among the $k = k_n$ nearest neighbors of X_j (or vice versa).

- **Dynamic models; preferential attachment**

Such random graphs are built by a recursive process. For example, in *preferential attachment* vertices are added one-by-one. When a new vertex is added, it connects to a fixed number of previously present vertices such that the probability of connection is an (increasing) function of the degrees of the "older" vertices.

- **Percolation**

Given a fixed graph G on n vertices, delete each edge with probability $1-p$ at random, independently. For $G = K_n$, this is just $G(n,p)$. One often considers infinite ground graphs such as the integer grid $\mathbb{Z}^d$ or the hexagonal lattice. Typical questions are about the existence of infinite components.

This is *bond percolation*. In *site percolation* one deletes vertices instead of edges.

- **Random trees**: a huge variety of models such as Galton-Watson trees, uniform trees, random spanning trees, binary search trees, etc.