Random Graphs II

Dmitri Loguinov
Texas A&M University

March 23, 2017
Agenda

• Midterm solutions
• Random graph models
  - $G(n,p)$
  - $G(n,M)$
  - $G(n,k_{out})$
  - $G(n,r)$
Random Graph Models 4

- Many of the early random-graph models stemmed from the work of Paul Erdos in the 1950s
  - One family of graphs is the famous $G(n,p)$ model
  - But we also study several other models with similar characteristics

- Definition: $G(n,p)$ is a random undirected graph with $n$ nodes, where each edge $(i, j)$ exists with an independent probability $p$
  - Goal of $G(n,p)$ is not necessarily to explain the Internet, but it serves as a good starting point

- Total edges in an undirected graph whose degree sequence is $d_1, \ldots, d_n$?
  \[ \frac{1}{2} \sum_{i=1}^{n} d_i \]
Graphs 2

• Task: derive the distribution of degree $d_i$ in $G(n,p)$
  - Assume that self-loops are not allowed, but this is not essential if the graph is large enough

• Possibly a simpler starting question would be
  - What is the average degree $E[d_i]$ in the graph?

• Assume that $L_{ij}$ is an indicator variable that signals the existence of link $(i,j)$
  - Then:

$$L_{ij} = \begin{cases} 
1 & (i,j) \in E \\
0 & \text{otherwise}
\end{cases}$$
• Thus:

\[ d_i = \sum_{j=1}^{n} L_{ij} \]

• We can further write that \( L_{ij} \) is a Bernoulli variable:

\[ P(L_{ij} = 1) = \begin{cases} p & i \neq j \\ 0 & \text{otherwise} \end{cases} \]

• Thus, we have a binomial distribution for \( d_i \):

\[ d_i = B(n - 1, p) \quad E[d_i] = (n - 1)p \]
While the $G(n,p)$ model provides plenty of non-isomorphic graphs, there are certain families of graphs it cannot build

- Example?

Notice that every instance of $G(n,p)$ may have a different number of edges

- To overcome this limitation, we have $G(n,M)$

Definition: random graph $G(n,M)$ is a uniformly random instance among all undirected graphs with $n$ nodes and exactly $M$ edges
Random Graph Models

- How to construct a $G(n,M)$ graph?
- There is a total of $N = n(n-1)/2$ possible edges in an undirected graph
  - $G(n,M)$ uniformly selects $M$ edges out of $N$
  - Each graph is built with an equal probability:
    \[ p_M = \frac{1}{\binom{N}{M}} \]

- Algorithm:
  - Create an array of $N$ elements representing edges
  - Throw a uniform random point into this array, swap the selected edge with the last element, and shrink array by 1
  - Repeat the last step $M$ times
Random Graph Models 9

- The next model is $G(n, k_{\text{out}})$
  - Each node randomly selects $k$ nodes from the remaining $n - 1$ vertices to be its neighbors
  - This is a directed graph
  - The out-degree is $k$, but what is the in-degree?

- The probability that node $v$ is selected by $u$ is:
  
  $$p_{uv} = 1 - \prod_{i=1}^{k} \left(1 - \frac{1}{n - i}\right) = \frac{k}{n - 1}$$

- Since there are $n - 1$ nodes $u$, the in-degree of $v$ is:
  
  $$d_{v}^{\text{in}} \sim B(n - 1, p_{uv})$$
Random Graph Models 10

- Easy to notice that $E[d_v^{in}] = k$

- Graphs $G(n, k_{out})$ are often found in user-driven networks with freedom to choose neighbors
  - A prime example would be a P2P system, where each joining user randomly selects $k$ other peers to be his/her neighbors

- Another version of this model is the social network
  - Each person randomly selects $k$ other people in the world to be their friends

- **Q:** What is the probability that each person in an $n$-node network is connected to every other one through a chain of acquaintances?
Random Graph Models 7

- The final model is $G(n,r)$
  - A fairly recent development that models wireless ad-hoc networks (i.e., users form a graph using each other’s computers as routers)
  - Two nodes $u$ and $v$ are linked if and only if $||u - v|| \leq r$
  - Parameter $r$ is the communication radius and $|| \cdot ||$ is some norm (usually Euclidean) in the 2D space
Random Graph Models 7

• Formalization
  – Coordinates \((x,y)\) of each node in \(G(n,r)\) are iid uniform random variables in \([0,A]\)
  – Nodes scattered in a 2D square whose area is \(A^2\)
  – Edge \((u,v)\) exists iff distance between \(u\) and \(v\) is less than \(r\)

• Define node density \(\rho\):

\[
\rho = \frac{n}{A^2}
\]

• Number of nodes in an area of size \(S\)?
  – Binomial with parameters \((n, S/A^2)\)
  – The mean is simply \(\rho S\)
• **Q:** What is the expected number of neighbors (degree) of each node?

\[ E[d_i] = \rho \pi r^2 - 1 \]

• **Q:** How many wireless sensors with a communication radius of 200 feet does one need to scatter in a zone 50x50 miles to obtain a connected network *with high probability*?

• This relatively simple question requires rather complex mathematical tools, which we omit here.