College of Computer & Information Science
Northeastern University
CS 7880: Network Algorithms and Analysis

Spring 2012
10 January 2012
Scribe: Rajmohan Rajaraman

Lecture Outline:

- Course overview: Topics, results, and techniques
- Graph models and example problems

1 Overview of the course

This is an advanced course in algorithms and theory, with a focus on network algorithms and analysis. The goal of the course is to cover some of the fundamental theoretical techniques that are used in the design of network algorithms and analysis of various network processes. Here is a tentative and partial list of topics and results that we plan to cover.

1. Linear algebra overview and spectral analysis of graphs
   - Cheeger's Inequality that relates a certain eigenvalue associated with the graph's adjacency matrix with its expansion
2. Random walks, Markov chains, and graph clustering
   - Analysis of mixing time of random walks, Lovasz-Somonovits Theorem
   - Local graph clustering algorithm using PageRank vectors
3. Diffusion in networks
   - Analysis of basic branching processes
   - Analysis of gossip protocols for information dissemination
   - The small-world phenomenon: Kleinberg's analysis of decentralized search
4. Threshold phenomena and basics of percolation theory
   - Birth of giant component in Erdos-Renyi graph models
   - The Harris-Kesten Theorem on percolation of the 2-D grid
   - Percolation in finite graphs
5. The multiplicative weights method for optimization
   - Online learning
   - Garg-Konemann and related algorithms for multicommodity flow
6. Linear programming (LP) based approximation algorithms
   - Leighton-Rao algorithm for approximating sparsest cut and multicommodity flow
7. Semidefinite programming (SDP) based approximation algorithms
   - Goemans-Williamson algorithm for maximum cut
   - Arora-Rao-Vazirani algorithm for sparsest cut

2 Graphs models and example problems

2.1 Graph models

Erdos-Renyi Random Graph model. One of the most widely studied random graph models is the classic Erdos-Renyi model [AS91]. In the $G(n,p)$ model, the graph consists of a set V of n vertices, and a random set E of edges; for each pair u and v of vertices, (u,v) is in E with probability p, independent of any other pair of vertices. There is a related model $G(n,m)$ in which the set of edges is drawn uniformly at random from the collection of all sets of edges with cardinality m. Probabilists and graph theorists have extensively studied the $G(n,p)$ model, with a focus on how the parameter p impacts the various properties and measures like connectivity and diameter of the graph. One of the central findings, due to Erdos and Renyi, is that there was a sharp threshold for the emergence of many properties.

We consider one such property – connectivity – as a warm-up. Let $p = \lambda/n$, where λ is a real number. Let $C = (S, V-S)$ denote any cut, where S is an arbitrary subset of say k vertices. What is the probability that there is no edge crossing the cut C? It is exactly $(1-\lambda/n)^{k(n-k)}$. Then, by a straighforward union bound, the probability that the graph G is disconnected is at most:

$$\begin{aligned} & \sum_{k=1}^{\lfloor n/2 \rfloor} \binom{n}{k} (1-\lambda/n)^{k(n-k)} \\ \leq & \sum_{k=1}^{\lfloor n/2 \rfloor} \binom{n}{k} (1-\lambda/n)^{kn/2} \\ \leq & \sum_{k=1}^{\lfloor n/2 \rfloor} \binom{n}{k} e^{-k\lambda/2} \\ \leq & \frac{n}{2} \frac{n^k}{e^{k\lambda/2}}. \end{aligned}$$

If $\lambda \geq 4 \ln n$, the above bound is at most $2/n$, for n sufficiently large, implying that with probability at least $1 - 2/n$, the graph is connected.

What about the other direction? How small should λ be to ensure that the graph is disconnected with "high probability". We now show that if $\lambda < a \ln n$, $a < 1$, then with probability tending to 1, there will be at least n^{1-a} isolated vertices, hence implying that the graph is disconnected. Let X_i denote the indicator random variable for the event that vertex i is isolated. So we have:

$$\Pr[X_i = 1] = \left(1 - \frac{\lambda}{n}\right)^{n-1}.$$

Let X equal $\sum_i X_i$. Then, $E[X]$ equals $n(1-\lambda/n)^{n-1}$. We next calculate the variance of X.

$$\begin{aligned}
\mathrm{var}(X) &= E[(X - E[X])^2] \\
&= E[(\sum_i X_i - E[X])^2] \\
&= E[(\sum_i X_i)^2] + E[X]^2 - 2E[X]^2 \\
&= E[(\sum_i X_i)^2] - E[X]^2 \\
&= \sum_i E[X_i^2] + \sum_{i \neq j} E[X_i X_j] - E[X]^2 \\
&= E[X] + \sum_{i \neq j} E[X_i X_j] - E[X]^2.
\end{aligned}$$

We need to calculate $E[X_i X_j]$ for $i \neq j$. This is simply the probability that both i and j are isolated, which equals $(1-\lambda/n)^{2n-3}$. We thus have

$$\begin{aligned}
\mathrm{var}(X) &= E[X] + n(n-1)(1-\lambda/n)^{2n-3} - n^2(1-\lambda/n)^{2n-2} \\
&= E[X] + n^2(1-\lambda/n)^{2n-3}(1 - 1/n - 1 + \lambda/n) \\
&\leq E[X] + n(1-\lambda/n)^{2n-3}(a \ln n - 1) \\
&= E[X]\left(1 + (1-\lambda/n)^{n-2}(a \ln n - 1)\right) \\
&= E[X](1 + o(1))
\end{aligned}$$

We now use Chebyschev's inequality on the distribution of the tail of X to obtain the following.

$$\Pr[|X - E[X]| > c\sqrt{E[X]}] \leq \frac{\mathrm{var}(X)}{c^2 E[X]} \leq \frac{1}{c^2}.$$

Thus, with probability at least $1 - 1/c^2$, the number of isolated vertices is at least $E[X] + c\sqrt{E[X](1+o(1))}$. For $\lambda < a \ln n$, $E[X] \approx n^{1-a}$. Thus, the number of isolated vertices is $n^{1-a}(1-o(1)$ with probability $1 - o(1)$.

Power laws, preferential attachment, and small-world graph models. While the $G(n,p)$ model is arguably most natural, and it lends itself to elegant and rigorous mathematical study, most "real" networks do not fit that model. For instance, the degree distribution observed in real networks is much different from the Poisson distribution for the $G(n,p)$ model. Many real networks have been shown to have a *power law* degree distribution: the fraction of vertices with degree k tends to $Ck^{-\beta}$ as $k \to \infty$, where C and β are positive constants. One way to address this is to develop graph models with specified degree distributions. One such model is the Molloy-Reed model [MR95]. Another approach is to develop a generative model that eventually yields a graph satisfying power law degree distributions. A highly-cited example of this approach is the preferential attachment model of Barabasi-Albert [BA99].

Other graph topologies. We will also study regular interconnection networks such as the d-dimensional grid, hypercube, and families of graphs such as expander graphs and planar graphs. Each of these graph families has numerous applications. For instance, grid and hypercube structures are commonly found in network architectures, and also form the basis of several overlay networks

that are designed on top of physical networks. Planar graphs capture several physical interconnections, such as transportation systems and circuits. Expander graphs are one of the most widely used graph objects in computer science, with applications in coding theory, complexity theory, and randomized algorithms.

2.2 Network optimization problems and approximation algorithms

Many algorithmic problems require the computation of optimal and near-optimal graph structures satisfying certain properties. For instance, path selection for routing in communication and transportation networks can be modeled as the multicommodity flow problem. Many divide-and-conquer algorithms on graphs work by dividing the graph into two parts with few edges crossing between them: a well-studied problem modeling such graph cuts is the sparsest cut problem.

Many natural graph optimization problems are NP-hard, and it is unlikely that we will be able to give polynomial-time algorithms for these problems. This is true of many graph cutting problems, covering problems on graphs (such as vertex cover, minimum dominating set), and network design problems (such as finding the minimum subgraph connecting desired sets of vertices). We will address this intractability by designing *approximation algorithms*. That is, rather than solving optimally for every instance, we will demand *polynomial-time* algorithms that solve *near-optimally* for *every instance*. What does "near-optimal" mean? One notion of approximation is that of an *absolute performance guarantee*, in which the value of the solution returned by the approximation algorithm differs from the optimal value by an absolute constant. While this notion is useful for certain problems such as the minimum-degree spanning tree problem, it turns out to be inappropriate for most NP-complete problems. The notion of approximation that is most widely used is that of a *relative performance guarantee*, which we now define. Let Π be an optimization problem and let I be an instance of Π. Given an algorithm $\mathcal{A}$ for Π, let $\mathcal{A}(I)$ denote both the solution as well as the value of the solution returned by $\mathcal{A}$ on instance I. Also, let $\mathrm{OPT}(I)$ denote the optimal value for instance I.

Definition 1. An approximation algorithm $\mathcal{A}$ for a minimization problem Π has an **approximation ratio** of r if the following condition holds for all instances I of Π:

$$\frac{\mathcal{A}(I)}{\mathrm{OPT}I} \leq r.$$

2.3 Random walks, percolation, and spectral analysis

Random walks and their generalizations such as Markov chains model several natural processes (e.g., Brownian motion, web surfing). They also have wide applications in algorithms; for example, random walks in expander graphs form the basis for some pseudorandom generators. Another related process is that of percolation, in which one selects vertices and/or edges of a graph according to a probability distribution, and studies the properties of the selected subgraph. Percolation has numerous applications in condensed matter physics, and is also a useful model for studying the spread of information and diseases in physical networks.

When studying random walks, percolation, and related processes on finite graphs, the spectral profile of the underlying graph plays a prominent role. For instance, the mixing time of a random

walk on a graph is closely related to the conductance of the graph, which can be bounded in terms of the second eigenvalue of the Laplacian. Spectral methods also appear prominently in the study of massive data sets. We will begin this course with a study of basic spectral graph theory. The course website has links to excellent notes by Fan Chung, Jon Kelner, Dan Spielman, and others, from which we will draw material for this portion of the course.

References

[AS91] N. Alon and J. H. Spencer. *The Probabilistic Method*. Wiley, New York, NY, 1991.

[BA99] A. L. Barabási and R. Albert. Emergence of Scaling in Random Networks. *Science*, 286:509, 1999.

[MR95] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, 6:161–179, 1995.