Spring 2016
22 January 2016
Scribe: Rajmohan Rajaraman

- Different measures of graph expansion
- Random bipartite graphs expand
- Rumor spreading: Analysis for general graphs

1 Different measures of graph expansion

We formally state several notions of graph expansion.

- Spectral: For a regular graph, $1 - \lambda_2$, where λ_2 is the second largest eigenvalue of the random walk matrix. Alternatively, the second smallest eigenvalue of the Laplacian.
- Vertex expansion: Let $N(S)$ denote the set of nodes in $V - S$ that have at least one neighbor in S. The vertex expansion is defined as:
$$\min_{S:0<|S|\leq|V|/2} \frac{|N(S)|}{|S|}.$$
- Edge expansion: Let $E(S, V - S)$ denote the set of edges with one endpoint in S and the other endpoint in $V - S$. The edge expansion is defined as:
$$\min_{S:0<|S|\leq|V|/2} \frac{|E(S, V - S)|}{|S|}.$$
- Conductance: Let $\mathrm{vol}(S)$ denote the sum of the degrees of all the vertices in S. The conductance Φ is defined as:
$$\min_{\emptyset\neq S\neq V} \frac{|E(S, V - S)|}{\min\{\mathrm{vol}(S), \mathrm{vol}V - S\}}.$$

A well-known result – Cheeger's Inequality – relates conductance with the spectral grap.

Theorem 1. *Cheeger For any d-regular graph, we have*
$$\frac{\Phi^2}{8} \leq 1 - \lambda_2 \leq \Phi.$$

2 Random bipartite graphs expand

Consider a random bipartite graph $G = (L \cup R, E)$, where L and R are two disjoint sets of n vertices, and E is a collection of edges between L and R drawn as follows: each vertex in L selects d vertices in R uniformly at random, and independent of choices of any other vertex in L. Clearly, every vertex in L has exactly d incident edges. We ask the following question: how well does an arbitrary set in L expand?

Theorem 2. *With probability at least* $1/2$*, every subset* S *of vertices in* L *with* $|S| \leq n/d$ *has at least* $d|S|/4$ *neighbors in* R*.*

Proof: Consider any set S of size $s \leq n/d$ vertices. The probability that it has fewer than $ds/4$ neighbors is at most

$$\binom{n}{ds/4}\left(\frac{\binom{ds/4}{d}}{\binom{n}{d}}\right)^s.$$

Simplifying, using the approximation $\binom{ds/4}{d}/\binom{n}{d} \approx (ds/4n)^d$, and using the inequality $\binom{a}{b} \leq (ea/b)^b$ for $a \geq b > 0$, we obtain the bound:

$$\left(\frac{4en}{ds}\right)^{3ds/4} \cdot \left(\frac{ds}{4n}\right)^{ds} \leq \left(\frac{eds}{4n}\right)^{3ds/4}.$$

Since $ds \leq n$, we obtain that the probability that there exists any set S of size at most n/d that has fewer than $d|S|/4$ neighbors is at most

$$\sum_{s\geq 1} \frac{e}{4}^{3ds/4} \leq 1/2$$

for $d \geq 4$. □

3 Analysis of rumor spreading on general graphs

The paradigm of rumor spreading or gossiping is considered as a robust mechanism for spreading information in a distributed network, or influence in a social network. Suppose we have an undirected connected network G with n nodes. A node, say r, has a piece of information M that it wants to broadcast to the entire network. Consider the following gossiping protocol.

In each step, each node that has a copy of M, sends a copy of M to a neighbor chosen uniformly at random. Assume that all the nodes are sychronized in their steps. This is called the Push protocol.

Theorem 3. *The* Push *protocol completes in* $O(n \log n)$ *steps with probability at least* $1 - 1/n$ *for any* n*-vertex graph.*

Proof: Our proof follows the following steps.

(a) Suppose a node u has a copy of M and degree d. What is the expected number of steps, in terms of d, before u sends a copy of M to a specific neighbor v?

The probability that u sends a copy of M to v in any given step is $1/d$. Thus, the expected number of steps it takes before u sends a copy of M to v equals:

$$\frac{1}{d} + 2 \cdot \frac{1}{d} \cdot \frac{d-1}{d} + 3 \cdot \frac{1}{d} \cdot \left(\frac{d-1}{d}\right)^2 + \ldots \sum_{i=1}^{\infty} i \cdot \frac{1}{d}\left(\frac{d-1}{d}\right)^{i-1}.$$

Using elementary algebra/calculus, we simplify the above to obtain the expectation to be d.

(b) Let P be a shortest path from u to v. We now show that the sum of the degrees of all the nodes on P is at most $3n$. We argue that a node x can be a neighbor of at most 3 nodes on a shortest path. Note that this is sufficient to establish the desired claim.

Suppose otherwise; let x be a neighbor of distinct nodes u_1, u_2, u_3, and u_4. Without loss of generality, assume that P first visits u_1, then u_2, then u_3, and then u_4. It follows that the subpath of P from u_1 to u_4 has at least three edges. However, replacing this subpath by the two-hop path $u_1 \to x \to u_4$ contradicts the fact that P is a shortest path from u to v.

(c) Using parts (a) and (b), we now derive an upper bound, in terms of n, on the expected number of steps it takes for an arbitrary node v to receive a copy of M.

By part (a) and linearity of expectation, the expected number of steps it takes for an arbitrary node v to receive a copy of M is at most the sum of the degrees of the nodes along the shortest path from r to v, which is at most $3n$ by part (b).

Unfortunately, part (c) does not give us a bound on the expected completion time, since it only bounds the time taken for an arbitrary node v – not *all nodes* – to receive M.

(d) Let us revisit part (b). Again, suppose a node u has a copy of M and degree d. We find an upper bound, in terms of d, on the number of steps it takes for a specific neighbor v of u to receive a copy of M from u with probability at least $1 - 1/n^3$.

Let t be the number of steps it takes for v to receive a copy of M from u with probability at least $1 - 1/n^3$. The probability that v has not received a copy of M from u in t steps is $(1 - 1/d)^t$. So t is the first step at which this probability is at most $1/n^3$; in other words

$$t \le \ln(1/n^3)/\ln(1 - 1/d) \le 3d \ln n,$$

where we use the inequality $(1 - 1/d)^d \le 1/e$ for $d \ge 1$.

(e) Using parts (b) and (d), we derive an upper bound, in terms of n, on the number of steps it takes for an arbitrary node v to receive a copy of M with probability at least $1 - 1/n^2$. We argue that the same bound yields an upper bound on the number of steps it takes for *all nodes* to receive a copy of M with probability at least $1 - 1/n$.

Consider a shortest path from r to v. In at most $3d_r \ln n$ steps, where d_r is the degree of r, the message crosses the first hop (to, say node u) with probability at least $1 - 1/n^3$. Conditioned on the fact that M has reached u, in at most $3d_u \ln n$ additional steps, where d_u is the degree of u, the message crosses the second hop with probability at least $1 - 1/n^3$. Thus, in at most $3(d_r + d_u) \ln n$ steps, M has reached u with probability at least $1 - 2/n^3$ (using Boole's inequality)'. Continuing with this argument and invoking part (b), we obtain that M reaches an arbitrary node v in at most $3n \log n$ steps with probability at least $1 - n/n^3 = 1 - 1/n^2$.

The probability that M has failed to reach a *specific node* v in $3n \log n$ steps is at most $1/n^2$. Thus, the probability that there exists a node v that M has failed to reach in $3n \log n$ steps is at most $1/n$ (using Boole's inequality).

□