CSE525: Randomized Algorithms and Probabilistic Analysis	May 23, 2013
Lecture 7	
Lecturer: Anna Karlin	*Scribe: Svetoslav Kolev*

1 Random Walks on undirected graphs

We are given an undirected graph $G = (V, E)$. In this class we will consider random walks on this graph with transition probabilities of

$$P_{ij} = \frac{1}{d_i}$$

where d_i is the degree of vertex i. The Markov Chain will be periodic if and only if the graph is bipartite. In this setting the stationary distribution is given by

$$\pi_i = \frac{d_i}{2m}$$

where m is the total number of edges in the graph $|E|$.

Some key qntities for these random walks are:

$$\begin{aligned} \text{hitting time} \quad & h_{ij} = E(T_{ij}) \\ \text{commute time} \quad & c_{ij} = h_{ij} + h_{ji} \\ \text{cover time} \quad & C(G) = \text{time to visit all vertices} \end{aligned}$$

$$h_{ii} = \frac{1}{\pi_i} = \frac{2m}{d_i}$$

Lemma 1. *For all edges (i, j) we have $h_{ij} + h_{ji} < 2m$.*

Proof Consider a corresponding random walk on the directed graph that has $2m$ vertices, one for each edge and each orientation of that edge from the original graph. Being at vertex (i, j) means we are traveling from i to j. Hence the transition probabilities

$$Q_{(i,j)(j,k)} = p_{jk} = \frac{1}{d_j}$$

Claim 2. *Q is doubly stochastic (Both rows and columns sum up to 1)*

$$\sum_{i|(i,j)\in E} Q_{(i,j)(j,k)} = \sum_{i|(i,j)\in E} \frac{1}{d_j} = \frac{d_j}{d_j} = 1$$

Fact 3. *Stationary distribution of every doubly stochastic Markov Chain is uniform.*

Therefore we have that for Q, $\pi^Q_{(i,j)} = \frac{1}{2m} \quad \Rightarrow \quad h_{(i,j)(j,k)} = 2$. Since every random walk that is a cycle (also, starting from with vertices i and j and ending back in i) in the original graph corresponds to a random walk in Q from (i, j) to (i, j) with the same probability. The correspondence is that we must make the same choices at the corresponding places of the two graphs, therefore the probability will be the same. The only difference is that our first choice in Q is the move from (i, j) to (j, k) which corresponds to the second choice in the original walk, and so on until the last choice in Q $(w, i) \to (i, j)$ which corresponds to the first choice in the original walk. By choice we mean choosing among the same vertices as the next vertex to visit, therefore having the same probability.

So we have $h_{ij} + h_{ji} \leq h_{(i,j)(j,i)} = 2m$.

■

2 Cover time

$$C_u(G) \ : \ \text{ cover time starting at } u$$
$$C(G) = \max_u C_u(G)$$

Theorem 4. $E(C(G)) \leq 2m(n-1)$

Proof Construct any spanning tree T.

$$E(C(G)) \leq \sum_{(i,j)\in T} (h_{ij} + h_{ji}) \leq 2m(n-1)$$

. ■ WHYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY

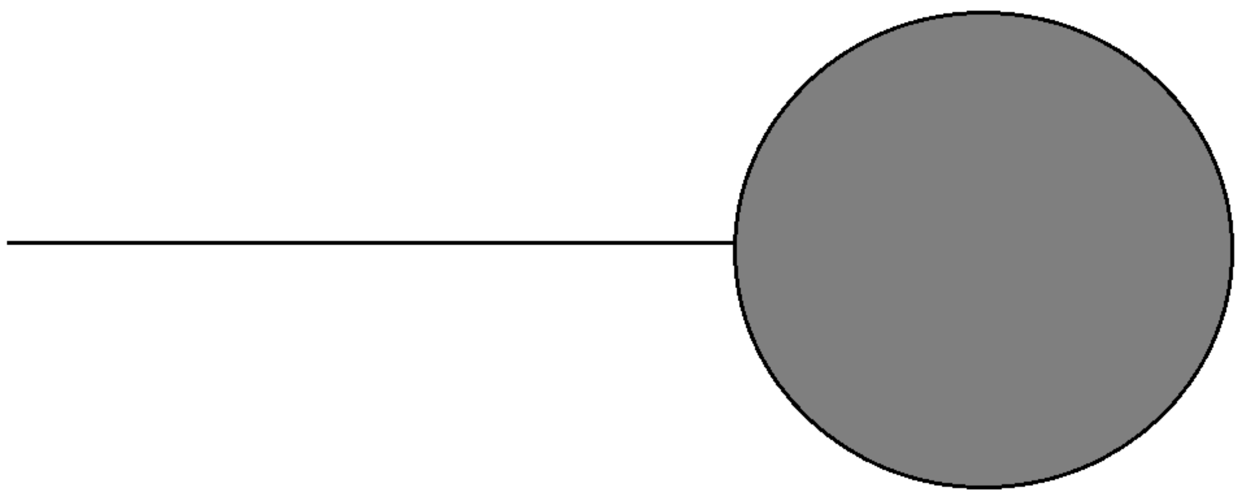

Figure 1: A lollipop graph. Half of the vertices are in the handle, the other half in the disc. The disc is a full graph.

For example the lollipop has covertime $C(G) = \Theta(n^3)$.

We already proved that the line graph has a covertime $C(G) = \Theta(n^2)$.

The complete graph has a covertime of $C(G) = \Theta(n \log n)$. Look up the coupon collector problem. It is exactly the same.

3 Application: s-t connectivity

We have the following problem: Given an undirected graph $G = (V, E)$ and $s, t \in V$, decide whether s and t are in the same connected component. We have DFS and BFS but both of them use $O(m)$ and $O(n)$ space for keeping track of which vertices are visited.

There is a very simple randomized algorithm that uses $\log V$ space. The input is on a separate read-only tape.

Simulate random walk of length $2n^3$ on G starting from s. We have then that:

$$\Pr(\text{doesn't reach } t \text{ when there is a path}) \leq \frac{1}{2}$$

.

4 Regular graphs

Here we look into random walks on regular graphs (each vertex has the same degree d.

$$P_{ij} = \frac{1}{d} \iff (i,j) \in E$$

.

$$\pi = (\frac{1}{n}, \cdots, \frac{1}{n})$$

5 Mixing time

How long does it take to converge to π? The mixing time of is related to the algebraic properties of P.

Suppose q^0 is the initial distribution over the vertices and q^t is the distribution after t steps. Also P^t is P raised to the power t. Then

$$q^t = q^0 P^t$$

.

How close is q^t to π.

Theorem 5 (Spectral Theorem)**.** *If $M \in R^{NxN}$ is symmetric, then:*

- *all N eigenvalues are real (solutions λ of $\det(A - \lambda I) = 0$.*
- *There exist orthonormal set of eigenvectors $v_1, \ldots, v_n$ corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$. Also $v_i \cdot v_j = 1$ when $i = j$ and 0 otherwise .*

Therefore we have that

$$M = \sum_{i=1}^{N} \lambda_i v_i v_i^T = \Phi^T \Lambda \Phi$$

where $\Phi^T = (v_1, v_2, \ldots, v_n)$ and we also assume that v_i are column vectors. $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$.

$$\Phi\Phi^T = \Phi^T\Phi = I$$

$$\Rightarrow M^t = \sum_i \lambda_i^t v_i v_i^T$$

We let superscript t to mean the corresponding eigenvalue of or the transition matrix itself at time t. This is equivalent to raising it to the power of t.

Theorem 6 (Penon-Frobenius Theorem)**.** *If $A > 0$ (positive definite) and $A^m \gg 0$ (elementwise greater) for all $m \leq M$.*

- *There exist $\vec{x} \gg 0$ such that $Ax = \lambda^* x$*
- *If $\lambda \neq \lambda^*$ is any other eigenvalue of A then $|\lambda| < \lambda^*$.*

If P is defined as follows (called lazy random walk) then it satisfies the theorem and $0 \leq$ eigenvalues ≤ 1.

$$P_{ij} = \frac{1}{2} i = j \tag{1}$$
$$P_{ij} = \frac{1}{2d}(i,j) \in E \tag{2}$$
$$P_{ij} = 0 \text{otherwise} \tag{3}$$

5.1 Mixing time

$$P^t = v_1 v_1^T + \sum_{i \geq 2} \lambda_i^t v_i v_i^T$$

where

$$v_1^T = (\frac{1}{\sqrt{n}}, \cdots, \frac{1}{\sqrt{n}})$$

$$v_1 v_1^T = \frac{1}{n} J \quad \text{(all 1's matrix)}$$

Let's write q^0 as $q^0 = \sum_i c_i v_i^T$ where $c_i = q^0 \cdot v_i$.

$$q^0 P^t = \sum_i c_i v_i^T \sum_j \lambda_j^t v_j v_j^T$$
$$= \sum_i c_i \lambda_i^t v_i^T$$

Using $c_i = q^0 \cdot v_i = \frac{1}{\sqrt{n}} \sum q_i^0 = \frac{1}{\sqrt{n}}$ and $\lambda_1 = 1$ we get

$$= \frac{1}{\sqrt{n}} v_1^T + \sum_{i \geq 2} c_i \lambda_i^t v_i^T$$
$$= (\frac{1}{n}, \cdots, \frac{1}{n}) + \sum_{i \geq 2} c_i \lambda_i^t v_i^T$$
$$= \pi + \sum_{i \geq 2} c_i \lambda_i^t v_i^T$$

Let's see how far we are from the stationary distribution at time t.

$$||q^0 P^t - \pi|| = ||\sum_{i=2}^n c_i \lambda_i^t v_i^T||$$
$$= \sqrt{\sum_{i=2}^n c_i^2 \lambda_i^{2t}}$$
$$\leq \lambda_2^t \sqrt{\sum_i c_i^2}$$
$$= \lambda_2^t ||q^0||$$
$$\leq \lambda_2^t$$

At

$$t = \Omega(\frac{1}{(a - \lambda_2)} \log n) \Rightarrow ABOVEWHY n^{-c}$$

It takes $O(\frac{1}{(1-\lambda_2)}\log n)$ steps to converge to the stationary distribution. When $(1-\lambda_2)$ is big, when $\lambda_2 \ll 1$, we have a fast convergence.

D-regular graphs with $\lambda_2 \ll 1$ are called expanders.

- Random d-regular graphs with even small d constant are expanders.

$$\lambda_2 = \Theta(\frac{1}{\sqrt{d}}) \;\Rightarrow\; \text{mixing time is } O(\log n)$$

  .

There exist explicit constructions with long mixing times.

- For hypercube: $n = 2^k$, $d = k = \log n$. $\lambda_2 = 1 - \frac{2}{\log n}$.
- Cycles: $\lambda_2 = 1 - \theta(\frac{1}{n^2})$.

Expanders find use in a wide variety of theoretical computer science:

- Comprexity Theory
- Design of robust computer networks
- Error correcting codes
- Pseudorandomness