College of Computer & Information Science
Northeastern University
CS7880: Network Algorithms and Analysis

Spring 2016
Handout 2
14 February 2016

Sample Solution to Problem Set 1

Problem 1. Mixing time to cover time

We have shown that for any graph G, for any $\varepsilon > 0$, after $T = O\left(\frac{\ln(n/\varepsilon)}{1-|\lambda_2|}\right)$ steps of a random walk on G, we have for any vertex v:

$$\Pr[\text{the walk is at } v] \geq \frac{1-\varepsilon}{n}.$$

Using the above fact, show that the cover time for G is

$$O\left(\frac{n \ln n}{1-|\lambda_2|}\right).$$

Answer: In $T = O\left(\frac{\ln(2n)}{1-|\lambda_2|}\right)$, the probability that the walk hits v is at least $1/(2n)$. So if we run this T-step walk $4n \ln n$ times, the probability that v is not hit is at most

$$\left(1 - \frac{1}{2n}\right)^{4n \ln n} \leq \frac{1}{n^2}.$$

By a union bound, the probability that all the vertices are not covered in $4Tn \ln n$ steps is at most $1/n$. Thus, the expected cover time is at most

$$\frac{4Tn \ln n}{1 - 1/n} = O\left(\frac{n \ln^2 n}{1-|\lambda_2|}\right).$$

Problem 2. An extreme example for random walk in directed graphs

Show that there exists a strongly connected directed graph G with outdegree 2 such that the hitting time of a random walk on G is $\Omega(2^n)$.

Answer: Consider the graph formed by the edges $\{(i, i+1) : i < n\} \cup \{(i, 1) : i > 1\}$. Starting from vertex 1, to hit vertex n, the walk needs a sequence of $n-1$ consecutive selections of the edges of the form $(i, i+1)$; the probability of this event is $1/2^{n-1}$. So expected number of steps it takes for such an event to occur is $\Omega(2^n)$.

Problem 3. An extreme example for random walk in undirected graphs

Define a lollipop graph to be an undirected graph over a set V of n vertices consisting of a complete graph over a set A of $\lfloor n/2 \rfloor$ vertices, and a line graph over the set $V - A$ of $\lceil n/2 \rceil$ vertices, and an edge from an arbitrary vertex of A to one of the two endpoints of the line formed by the vertices in $V - A$.

Show that the cover time of a random walk on the lollipop graph is $\Theta(n^3)$.

Answer: Let the vertex in A that is adjacent to $V - A$ be labeled 0. Let its neighbor in $V - A$ be 1, and the vertices of the line graph be labeled $1, 2, \ldots, \lceil n/2 \rceil$. Let s be a start vertex for the walk, and let t be a target vertex. Let $H(x, y)$ denote the hitting time of walk from x to y.

We first show (a) $H(0, k)$ is $\Theta(n^3)$. We next show (b) $H(0, u)$ for any vertex u in the clique is $O(n)$. This implies that the expected time for visiting every vertex in the clique, starting from 0 is $O(n \log n)$; in fact, this is a high probability bound. Together, (a) and (b) show that the cover time starting from vertex 0 is $\Theta(n^3)$.

Let us calculate $H(0, k)$. This can be broken down as follows.

$$H(0, k) = H(0, 1) + H(1, 2) + \cdots + H(k - 1, k).$$

We calculate $H(i, i + 1)$ for $i \geq 1$ as follows.

$$H(i, i + 1) = \frac{1}{2} + \frac{1}{2} \left(1 + H(i - 1, i) + H(i, i + 1)\right),$$

which after rearranging yields

$$H(i, i + 1) = 2 + H(i - 1, i).$$

Let us calculate $H(0, 1)$. Note that the return time to 0 for a walk constrained to the clique of size $n - k$ is exactly $n - k$. The expected number of times the walk returns to 0 from within the clique of size $n - k$ before the walk takes the edge $(0, 1)$ is $n - k$ since the probability of taking the edge $(0, 1)$ is $1/(n - k)$. Therefore, $H(0, 1) = (n - k)^2$.

Substituting all the values to the summation for $H(0, k)$, we obtain

$$\begin{aligned} H(0, k) &= k(n - k)^2 + (2 + 4 + \cdots + 2(k - 1)) \\ &= k(n - k)^2 + k(k - 1). \end{aligned}$$

Setting $k = \lceil n/2 \rceil$ yields $H(0, k) = \Theta(n^3)$.

We are ready to prove (b). An easy upper bound of $O(n^2)$ on $H(0, u)$ for $u \in A$ is given by noting that the return time to 0 is $O(n)$ and the probability that the walk takes the edge from 0 to u is at least $2/n$. To obtain an upper bound of $O(n)$, we observe that the walk from 0 either goes to 1, or to u, or to a vertex in A not equal to u. In the last case, in expected $O(n)$ time, the walk visits one of u or 0, each with probability $1/2$. We thus have the recurrence

$$H(0, u) = \frac{1}{n - k} + \frac{1}{n - k} \left(1 + H(1, 0) + H(0, u)\right) + O(n) + \frac{n - k - 2}{2(n - k)} H(0, u),$$

which on simplification yields the upper bound

$$H(0, u) \left(1 - \frac{n - k - 1}{2(n - k)}\right) = O(n) + \frac{H(1, 0)}{n - k}.$$

By Problem 5(a) below, we have $H(1, 0) = \Theta((n - k)^2)$. We thus obtain $H(0, u) = O(n)$.

We have thus shown that the cover time starting from 0 is $\Theta(n^3)$. What about the cover time starting from any other vertex in the lollipop? Next, we show (c) $H(u, 0)$ is $O(n)$ for any vertex u in the clique. $H(u, 0)$ is $n - k - 1$ since from any vertex in the clique, the probability that the walk moves to 0 is $1/(n - k - 1)$. We thus obtain that the cover time from any vertex u in A is $\Theta(n^3)$.

Finally, we consider the cover time from any vertex v in the line. In Problem 5(a). we show that $H(v, 0)$ is $O(n^2)$. Putting this together with the claim that the cover time from 0 is $\Theta(n^3)$ completes the proof that the *cover time*, starting from any vertex in the lollipop is $\Theta(n^3)$.

Problem 4. We have seen in class that for an arbitrary d-regular connected undirected graph with n vertices, the second largest eigenvalue is at most $1 - 1/(dn^3)$.

Show that for any connected nonbipartite d-regular undirected graph, the smallest eigenvalue is at least $-1 + 1/(4dn^3)$. (Note that this implies that the spectral gap is at least $1/(4dn^3)$, leading to a polynomial bound on the mixing time for the random walk of any graph.)

The following statement on the smallest eigenvalue μ of any matrix M from the Courant-Fischer Theorem would be helpful.

$$\mu = \min_{x \neq \mathbf{0}} \frac{\|x^t M x\|}{\|x\|}.$$

Answer: Note that $x^T M x$ can also be written as follows.

$$\begin{aligned}
x^T M x &= \sum_i \Big(\sum_{(i,j) \in E} x_i x_j / d \Big) \\
&= \sum_{(i,j) \in E} 2 x_i x_j / d \\
&= \sum_{(i,j) \in E} ((x_i + x_j)^2 - x_i^2 - x_j^2)/d \\
&= -1 + \sum_{(i,j) \in E} (x_i + x_j)^2 / d.
\end{aligned}$$

Therefore, the smallest eigenvalue equals

$$\min_{x \perp \mathbf{1}, \|x\|=1} x^T M x = -1 + \min_{x \perp \mathbf{1}, \|x\|=1} \frac{1}{d} \sum_{(i,j) \in E} (x_i + x_j)^2.$$

Without loss of generality, assume $|x_1| \geq |x_n|$, so $x_1 \leq -1/\sqrt{n}$. If there is any vertex k such that $|x_k| \leq 1/(2\sqrt{n})$, then consider a shortest path from 1 to k: the sum of $(x_i + x_j)^2$ along the edges (i, j) of this path is at least $\sum_{(i,j) \in E} (|x_i| - |x_j|)^2$, which is at least $1/(4n^3)$. Otherwise, all vertices have $|x_k| \geq 1/(2\sqrt{n})$. Now, we consider two cases. First, there is an edge (i, j) with x_i and x_j having the same sign. In this case $(x_i + x_j)^2 \geq 1/n$, and we are done. Otherwise, all edges are from vertices with positive x_i to vertices with negative x_i; but this can only happen if the graph is bipartite.

Problem 5. Random walks that branch and meet

Consider the following variant of the random walk, which we call the *branch-meet walk*. Initially, an arbitrary start vertex s is *active*. Every subsequent step consists of two parts: (i) every active vertex *selects* two neighbors uniformly at random; (ii) a vertex is *active* if and only if it is selected by an active vertex in part (i). Note that an active vertex may become inactive, if it is not selected by another vertex.

(Such branch-meet walks could be considered as highly idealized models for the spread of epidemics: an active vertex corresponds to an infected node, which infects two neighbors at random; and an infected node recovers if it is not infected by any other node.)

(a) Show that the cover time of a *standard random walk* on a line of n vertices, starting from one of its endpoints, is $\Theta(n^2)$. (*Hint:* Find a formula for the time taken to traverse an edge in a certain direction.)

Answer: Let the n vertices of the line be labeled 0 to $n-1$, in order from left to right. We calculate $H(i, i+1)$, for $0 < i < n-1$, using the same analysis as in Problem 3, as

$$H(i, i+1) = 2 + H(i-1, i).$$

We also have $H(0,1) = 1$. So we obtain $H(i, i+1) = 2i + 1$. Thus, we have

$$H(0, n-1) = H(0,1) + H(1,2) + \cdots H(n-2, n-1) = n - 1 + 2\sum_{i=1}^{n-1} i = \Theta(n^2).$$

(b) Show that expected time taken for the branch-meet walk on a line of n vertices, starting from one of its endpoints, to activate every vertex at least once is $\Theta(n)$.

Answer: Let us follow the rightmost active vertex in the branch-meet walk. Let $A(i, i+1)$ denote the expected time it takes for the vertex $i+1$ to become the rightmost active vertex, starting from a state where the rightmost active vertex is i. We obtain the following recurrence for $i > 0$:

$$A(i, i+1) = \frac{3}{4} + \frac{1}{4}\left(1 + A(i-1, i) + A(i, i+1)\right).$$

Simplifying, we obtain

$$A(i, i+1) = \frac{4}{3} + \frac{1}{3} A(i, i-1).$$

We also have $A(0,1) = 1$.

Solving the recurrence yields $A(i, i+1) \leq 2$ for all i. (One can get a more accurate bound; here we are bounding the sum of a geometric series by its infinite sum.) We thus have

$$A(0, n-1) = A(0,1) + A(1,2) + \cdots + A(n-2, n-1) \leq 2n.$$

Clearly, $A(0, n)$ is at least n, so we have the desired $\Theta(n)$ bound.

(Tight bounds on such "branch-meet walks" and other variations of epidemic processes on *general graphs* are still unknown; this domain may offer potential ideas for a research project in the course.)