

In the last lecture, we looked at random walks on line and used them to devise randomized algorithms for 2-SAT and 3-SAT. For 2-SAT we could design a randomized algorithm taking  $O(n^2)$  steps; for 3-SAT, we were able to reduce the number of steps from  $O(2^n)$  to  $O((\frac{4}{3})^n)$ . Today we will extend the concept of random walks to graphs.

## 1 Random Walks on Graphs

Consider a directed graph  $G = (V, E)$ ,  $|V| = n$ ,  $|E| = m$ . Each edge  $(u, v)$  of the graph has a weight  $M_{uv} > 0$ .  $M_{uv}$  denotes the probability to reach  $v$  from  $u$  in one step. A natural restriction, therefore is that for each vertex, the sum of the weights on outgoing edges is 1, i.e.,

$$\forall u \sum_{v \in N(u)} M_{uv} = 1 \quad (1.1)$$

where  $N(u)$  is the set of neighbor vertices of  $u$ . A random walk on graph, therefore implies starting at some vertex, and traversing the graph according to the probabilities  $M_{uv}$ . We are interested in the long term behavior of traversing like this on the graph, i.e., the probability distribution after a number of steps.

Let  $M$  denote the probability matrix. Then  $M$  is a Markov chain over graph  $G$ . Markov property implies that knowledge of previous states is irrelevant in predicting the probability of subsequent states. Therefore, the next state depends entirely on the current state.

(Aside: The above property is of memoryless Markov chains. Markov chains of order- $m$  have the property that the next state depends on  $m$  previous states, which also can be modeled as a classical Markov chain case by considering the set of  $m$  states as a single state.)

Mathematically,

$$\Pr[X_{t+1} = v | X_t = u, X_{t-1} = u_1, \dots, X_0 = u_t] = \Pr[X_{t+1} = v | X_t = u] = M_{uv} \quad (1.2)$$

The probability distribution over the graph  $G$  at any time gives for each vertex the probability of being at that vertex. Thus, the probability distribution is a 1-dimensional vector of size  $n$ . Let  $P^t$  denote the probability distribution after  $t$  steps.

$$P^t = (P_1, P_2, \dots, P_n) \quad (1.3)$$

where  $P_i$  denotes the probability of being at vertex  $i$ . Then,

$$P^{t+1} = P^t M \quad (1.4)$$

$$\Rightarrow P^t = P^0 M^t \quad (1.5)$$

where  $P^0$  is some initial distribution.

**Definition 1:** A distribution  $\Pi$  is *stationary* if  $\Pi M = \Pi$ .

which simply means that once we reach  $\Pi$ , the distribution will not change.

Some important issues that come up at this point are:

1. Does a Markov chain admit a stationary distribution?
2. If so, after how many steps is it reached?
3. Is there a unique stationary distribution?

In general, it may not be possible to reach a stationary distribution even if one exists for the given Markov chain, because of the initial distribution. For example, consider the Markov chain with two states as shown in Figure 1. It has a stationary distribution  $(\frac{1}{2}, \frac{1}{2})$  which is not obtainable if we start with an initial distribution of  $(1, 0)$ , or  $(0, 1)$ . The only way of attaining the stationary distribution of  $(\frac{1}{2}, \frac{1}{2})$  is when we start with this stationary distribution itself.

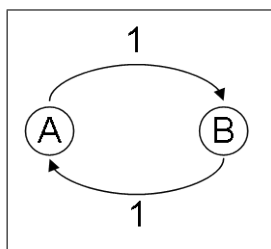


Figure 1

Here, we would like to investigate the conditions when a Markov chain admits a unique stationary distribution, and we can reach it starting from any initial distribution.

**Definition 2:** Let  $u, v$  be two vertices. We say that  $u$  and  $v$  *communicate* with each other if  $\exists n > 0, m > 0$  s.t.

$$M_{uv}^n > 0, M_{vu}^m > 0 \quad (1.6)$$

where  $M_{uv}^n$  denotes the probability that starting at  $u$ , we are at  $v$  after  $n$  steps.

**Definition 3:** A Markov chain over  $G$  is *irreducible* iff every state communicates with every other state.

**Observation 1:** A Markov chain over  $G$  is *irreducible* iff  $G$  is strongly connected.

It is easy to see that irreducibility is the minimal condition required for reachable unique stationary distribution. Otherwise, we may not be able to reach some vertices of the graph starting from some initial distribution.

Now, we need some condition to prohibit the case depicted in Figure 1.

**Definition 4:** A vertex  $v$  is *periodic* if there is a  $\Delta > 1$  s.t.  $\Pr[X_{t+s} = v | X_t = v] = 0$  unless  $s$  divides  $\Delta$ .

This means that starting at  $v$ , after every  $\Delta$  steps we may or may not come back at  $v$ , but we *cannot* come back at  $v$  when  $s$  is not a multiple of  $\Delta$ .

**Definition 5:** A Markov chain is *aperiodic* if all states are not periodic.

## 1.1 Fundamental Theorem of Markov Chains

Every finite, irreducible, aperiodic Markov chain has a unique stationary distribution  $\Pi$  that can be reached from any initial distribution. The stationary distribution is given by:

$$\Pi = (\Pi_1, \Pi_2, \dots, \Pi_n) \quad (1.7)$$

$$\Pi_i = \frac{1}{\text{Expected time to come back to vertex } i \text{ starting from } i} \quad (1.8)$$

It is important to note here that the phrase *can be reached* in the above theorem implies that we can reach such a stationary distribution in limit, i.e., we may not be able to reach such a distribution within finitely many steps.

In general, let  $u$  and  $v$  be two vertices in  $G$ . Let  $r_{uv}^t$  denote the probability of reaching  $v$  from  $u$  in *exactly*  $t$  steps s.t.  $v$  is not reached before  $t$  steps. Therefore, the expected time to go from  $u$  to  $v$  is:

$$h_{uv} = \sum_{t \geq 1} t \cdot r_{uv}^t \quad (1.9)$$

Therefore,  $h_{uu}$  is the expected time to return to  $u$

$$\Pi_i = \frac{1}{h_{ii}} \quad (1.10)$$

## 1.2 Reachability

We know that for undirected graphs, we have a non-deterministic log-space algorithm for reachability, which implies that we can solve reachability deterministically in  $\log^2$  space (by Savitch's theorem). Further, there is a randomized algorithm for reachability that works in log-space. Therefore, reachability is in randomized log-space. Recently, it has also been shown in 2005 that we can remove randomness from that algorithm, and hence, reachability in undirected graphs is in deterministic log-space. For directed graphs, the best we know so far is non-deterministic log-space, and hence, deterministic  $\log^2$  space.

Now we describe a randomized logspace algorithm for reachability in undirected graphs. Let  $G$  be a non-bipartite, undirected graph. Let the probability matrix be defined as follows:  $M_{uv} = \frac{1}{\deg(u)} \forall (u, v) \in E$ , and 0 otherwise. This means that from any vertex  $u$ , the probability to go to any of its neighbors is the same, and is  $\frac{1}{\deg(u)}$ .

**Claim:** A Markov chain over a non-bipartite, connected, undirected graph is aperiodic.

**Proof :** Since the graph is connected, it means that the Markov chain is irreducible (from Observation 1). Because the graph is not bipartite, it would have an odd-cycle. Finally, because the graph is undirected, there are paths starting and ending at same vertex of even lengths. Assume there is a node  $v$  with periodicity  $\Delta$ . Starting from  $v$  there is an even length path that comes back to  $v$ , and an odd length path that comes back to  $v$ . Using these two paths we can construct a path of length  $k\Delta + 1$  for some large enough  $k$ . Therefore, the Markov chain is not periodic.

**Claim:** Stationary distribution of the Markov chain over  $G$  is:

$$\Pi_v = \frac{\deg(v)}{2m} \quad (1.11)$$

where  $m$  is the number of edges in the graph. This further implies that the stationary distribution has at least  $\frac{1}{n^2}$  probability for every vertex.

**Proof:**  $\Pi$  is a stationary distribution if  $\Pi M = \Pi$

$\Pi = \left( \frac{\deg(1)}{2m} \frac{\deg(2)}{2m} \dots \frac{\deg(n)}{2m} \right)$   $M$  is the  $n \times n$  probability matrix.  $M_{uv} = \frac{1}{\deg(u)}$  if  $(u, v) \in E$ , i.e., if  $u$  and  $v$  are connected. This implies that each row  $i$  in the probability matrix  $M$  will have non-zero entries  $\frac{1}{\deg(i)}$ , and the number of non-zero entries will be equal to  $\deg(i)$ . Since it is an undirected graph, the zero/non-zero pattern will be symmetric, although the non-zero entries may be different because they are in different rows. Therefore, for vertex  $v$ , there will be exactly  $\deg(v)$  non-zero entries both in row  $v$  and column  $v$ . Multiplying the  $\Pi$  with  $M$ , we get

$$\begin{aligned} \Pi M &= \left( \frac{\deg(1)}{2m} \frac{\deg(2)}{2m} \dots \frac{\deg(n)}{2m} \right) \cdot \begin{pmatrix} \frac{1}{\deg(1)} & 0 & \frac{1}{\deg(1)} & \dots & 0 \\ 0 & 0 & \frac{1}{\deg(2)} & \dots & \frac{1}{\deg(2)} \\ \frac{1}{\deg(3)} & \frac{1}{\deg(3)} & 0 & \dots & \frac{1}{\deg(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \frac{1}{\deg(n)} & \frac{1}{\deg(n)} & \dots & \frac{1}{\deg(n)} \end{pmatrix} \\ &= \left( \left\{ \frac{1}{2m} + \dots + \frac{1}{2m} (\deg(1) \text{ times}) \right\} \dots \left\{ \frac{1}{2m} + \dots + \frac{1}{2m} (\deg(n) \text{ times}) \right\} \right) \\ &= \left( \frac{\deg(1)}{2m} \frac{\deg(2)}{2m} \dots \frac{\deg(n)}{2m} \right) = \Pi \end{aligned}$$

Therefore,

$$h_{uu} = \frac{1}{\Pi_u} = \frac{2m}{\deg(u)} \quad (1.12)$$

**Claim:** Say  $(u, v) \in E$ , then  $h_{vu} < 2m$ .

**Proof:** Consider an alternate representation of  $h_{uu}$ ,

$$h_{uu} = \sum_{v \in N(u)} \frac{1}{\deg(u)} [1 + h_{vu}] \quad (1.13)$$

i.e., take one step to go to a neighboring vertex of  $u$  with probability  $\frac{1}{\deg(u)}$ , and come back from  $v$  to  $u$  in  $h_{vu}$  steps. Also, we have just seen that

$$h_{uu} = \frac{2m}{\deg(u)}$$

$$\begin{aligned} \Rightarrow \sum_{v \in N(u)} \frac{1}{\deg(u)} [1 + h_{vu}] &= \frac{2m}{\deg(u)} \\ \Rightarrow \sum_{v \in N(u)} 1 + h_{vu} &= 2m \end{aligned}$$

Now,  $\deg(u)$  is at least 1, and thus from the above equation,  $h_{vu}$  is at most  $2m - 1$ . Therefore,

$$h_{vu} < 2m \quad (1.14)$$

**Definition 6:** Let  $v$  be a vertex. Let  $C_v$  denote the expected number of steps to visit every vertex of the graph starting at  $v$ .  $C_v$  is also called the *cover time* starting at  $v$ .

**Claim:**  $\forall v, C_v \leq 4nm$ .

**Proof:** Consider the spanning tree of the graph, starting at vertex  $v$ . Since it is a spanning tree, the number of edges in the graph is  $n$  (actually  $n - 1$ ). Therefore, the number of steps to traverse the tree can be at most  $2n$  in the worst case, considering that each edge of the spanning tree is traversed twice to start from and come back toward vertex  $v$ . Further, from the previous claim, we know that to traverse each edge (of the spanning tree), the expected number of steps to traverse (in the graph) is at most  $2m$  (since two adjacent vertices in the spanning tree are also adjacent in the graph). Therefore, an upper bound on the cover time of the graph starting at a vertex is given by  $2n \times 2m = 4nm$ , i.e.,

$$C_v \leq 4nm \quad (1.15)$$

It is important to note here that this is a very conservative upper bound on the cover time, as we have used the spanning tree to determine an ordering of the vertices in the graph. For specific graphs, however, we can do much better. For example, for complete graphs (all vertices connected to each other), the cover time is  $n \log n$ .

### 1.2.1 Randomized Algorithm for Reachability for Undirected Graphs

Starting at vertex  $s$ , is vertex  $t$  reachable?

ALGORITHM:

- 1 Do a random walk starting at  $s$  for  $8n^3$  steps.
- 2 if  $t$  is reached, ACCEPT; else REJECT.

It is easy to see that if  $s$  and  $t$  are not connected, the above algorithm will always reject. If, however,  $s$  and  $t$  are connected, the above algorithm accepts with probability  $\geq \frac{1}{2}$  (from Markov's inequality). Expected time to reach  $t = 4n^3$  ( $= 4nm$  with  $m = n^2$ ). And thus,

$$\begin{aligned} \Pr [t \text{ is not reached from } s \text{ in } 8n^3 \text{ steps}] &\leq \frac{4n^3}{8n^3} \\ &\leq \frac{1}{2} \end{aligned}$$

Therefore, the error probability is less than  $\frac{1}{2}$ .

### 1.2.2 Cover Time on Some Special Graphs

We have seen that for a general graph, the expected number of steps to visit vertex of the graph starting at a given vertex is  $4nm$ , where  $n$  is the number of vertices, and  $m$  is the number of edges.  $m$  can be at most of the order of  $n^2$ , and thus, the expected cover time on any graph is  $4n^3$ . We now look at some special instances of graphs and their cover times.

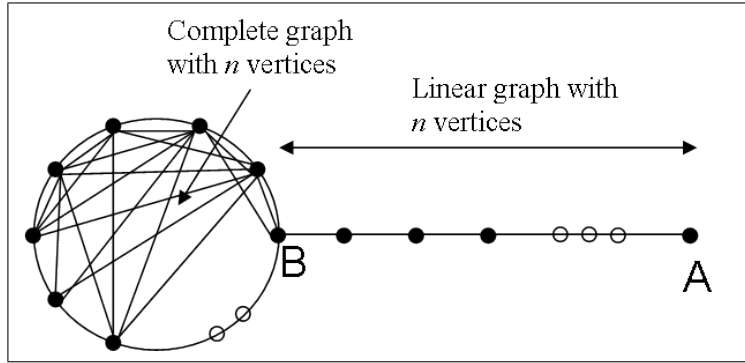


Figure 2

**Linear graphs:** In this case, the number of edges  $m = n$ , and hence, the cover time is  $4n^2$ . We have earlier seen that for random walks on a line, the expected number of steps to cover the line, and reach the other end of the line starting from one end in  $n^2$  steps, which was a tighter bound.

**Complete graph :** Consider a complete graph with  $n$  vertices, i.e., all the  $n$  vertices are connected to all  $n$  vertices. Then the problem of visiting every vertex starting at a given vertex reduces to the Coupon-Collector's problem, in which there are  $n$  coupons, and the expected number of selections (with replacement) to collect a complete set of all  $n$  coupons is  $n \log n$ . Hence, the cover time on complete graphs with self-loops is  $n \log n$ .

**Lollipop graph:** Consider a graph in Figure 2 with a complete subgraph of  $n$  vertices, attached to a linear subgraph of  $n$  vertices through a single vertex. The graph thus has a total of  $2n$  vertices, and  $2n - 2$  edges. This is one example of a graph where the cover time depends on the starting vertex. Consider the following two cases:

a) If we start at vertex  $A$ , then the cover time is  $O(n^2)$ , since it takes  $n^2$  steps to reach  $B$  from  $A$ , and then to traverse the remaining sub-graph (a complete graph within itself), it takes  $n \log n$  steps. Therefore, the expected number of steps to cover the graph is  $O(n^2)$ .

b) If we start at vertex  $B$ , then because of the  $n - 1$  edges from  $B$  into the complete sub-graph, with probability  $\frac{n-1}{n}$  we go somewhere into the complete graph and with only a small probability of  $\frac{1}{n}$  do we traverse toward  $A$  on the linear sub-graph. And then again, it has equal probabilities of moving toward the complete subgraph or toward  $A$ . Therefore, we can say that the expected number of times we will have to visit the complete subgraph and vertex  $B$  before reaching  $A$  is  $n$ . And this will happen for each of the  $n^2$  steps expected to reach from  $B$  to  $A$ . Therefore, the expected number of steps to cover the entire graph is  $n^3$ .