Random Walks

Instructor: Thomas Kesselheim

1 Random Walks

In today's lecture, we consider a random walk on an undirected graph. One way to think about a random walk is to imagine somebody walking through a city with no particular goal in mind. At each intersection, he or she chooses a random street to take next. We ask the questions: How long until this person reaches a particular point? How long until he/she has been to every intersection or every street of the city? While this sounds a bit like a toy problem, random walks are a very powerful tool to design randomized algorithms. Assume that you want to find a particular solution and you keep changing parts of it. This algorithm is super simple to implement. Under what circumstances can it be fast?

Definition 8.1. *A* random walk *on a connected undirected graph $G = (V, E)$ starting from v_0 is an infinite random sequence of vertices $X_1, X_2, \ldots \in V$ determined as follows. Vertex X_{t+1} is drawn uniformly among all neighbors of X_t. That is, for all $v, w \in V$, we have*

$$\mathbf{Pr}\left[X_{t+1} = w \mid X_t = v\right] = \begin{cases} \frac{1}{d(v)} & \text{if } \{v, w\} \in E \\ 0 & \text{otherwise} \end{cases} ,$$

where $d(v)$ denotes the degree of vertex v.

It is important to remark that the vertex can also go back via the same edge.

Definition 8.2. *The* cover time *of a graph $G = (V, E)$ is the maximum over all start vertices $v_0 \in V$ of the expected time to visit every vertex in V at least once.*

We will be interested in bounding the cover time of an arbitrary graph. To this end, it will be necessary to learn some foundations of Markov chains, which generalize random walks.

2 Markov Chains

A discrete-time stochastic process $X_0, X_1, X_2, \ldots$ is a Markov chain if

$$\mathbf{Pr}\left[X_t = a_t \mid X_{t-1} = a_{t-1}, X_{t-2} = a_{t-2}, \ldots, X_0 = a_0\right] = \mathbf{Pr}\left[X_t = a_t \mid X_{t-1} = a_{t-1}\right] .$$

In our case, the states are the vertices of the graph. As this set is finite, we speak of a *finite* Markov chain.

Our particular Markov chain is *time homogenous.* That is, for any states i and j, the probability that the Markov chain moves from i to j is the same in all steps. That is, for all t and t'

$$\mathbf{Pr}\left[X_t = j \mid X_{t-1} = i\right] = \mathbf{Pr}\left[X_{t'} = j \mid X_{t'-1} = i\right] .$$

Note that these conditional probabilities might not be well defined because, e.g. $\mathbf{Pr}\left[X_{t-1} = i\right] = 0$ depending on the initial state.

Therefore, we have a quadratic matrix P that encodes all relevant probabilities $P_{i,j} = \mathbf{Pr}\left[X_1 = j \mid X_0 = i\right]$. A probability distribution over states π can simply be represented as a vector such that $\sum_i \pi_i = 1$. Given such a probability distribution, the probability distribution after having performed one step in the Markov chain is given as πP. For this reason, P^k (i.e., multiplying P k times with itself) represents the probabilities of the respective state transitions in exactly k steps, i.e., $P^k_{i,j} = \mathbf{Pr}\left[X_k = j \mid X_0 = i\right]$.

Lemma 8.3. *State i is* accessible *from state j if for some integer $k \geq 0$, we have $P_{i,j}^k > 0$. A Markov chain is* irreducible *if for any two states i and j, i is accessible from j.*

The Markov chain corresponding to a random walk on a graph is irreducible if and only if the graph is connected.

Definition 8.4. *A state i has* period *k if any return to state i must occur in multiples of k time steps. Formally, the period of a state is defined as the greatest common divisor of all $t > 0$ such that $P_{i,i}^t > 0$. A state is* aperiodic *if it has period 1.*

Lemma 8.5. *The Markov chain corresponding to a random walk on a connected graph G is aperiodic if and only if G is not bipartite.*

Proof. First, we make the following observation. For every vertex $v \in V$ that has at least one outgoing edge, we have $P_{v,v}^2 > 0$. This is due to the fact that with positive probability the random walk leaves a v and gets back to it via the same edge in the following step.

Next, if G is bipartite, the period of all states is 2. Starting from any vertex in the bipartite graph, all odd steps lead to a vertex of the other side of the bipartition. Therefore $P_{v,v}^k = 0$ for all odd k.

Furthermore, if G is not bipartite, there has to be a cycle of odd length. It can be reached from any vertex v. Consider the tour that takes you from a fixed vertex v to this cycle, follow the cycle and then go back the same way you have come. This tour has length $k = 2\ell_1 + \ell_2$, where ℓ_1 is the length of the path to reach the cycle and ℓ_2 is its length. We have $P_{v,v}^k > 0$ and k is odd. As also $P_{v,v}^2 > 0$, the state has to be aperiodic. □

A state i is said to be positive recurrent if the expected time it takes to get from the state back to it is finite.

Definition 8.6. *For two states i and j, let the* hitting time *$h_{i,j}$ be the expected number of steps it takes to get from i to j, i.e.*

$$h_{i,j} = \sum_{t=1}^{\infty} t \cdot \mathbf{Pr}\left[X_t = j, \bigwedge_{t'=1}^{t-1} X_{t'} \neq j \;\middle|\; X_0 = i \right] .$$

State i is recurrent *if*

$$\sum_{t=1}^{\infty} \mathbf{Pr}\left[X_t = j, \bigwedge_{t'=1}^{t-1} X_{t'} \neq j \;\middle|\; X_0 = i \right] = 1 .$$

It is positive recurrent *if it is recurrent and $h_{i,i}$ is finite.*

Definition 8.7. *A distribution over states π in a finite Markov chain is called* stationary *if $\pi P = \pi$.*

Theorem 8.8. *Any irreducible, finite, and aperiodic Markov chain has the following properties.*

1. *All states are positive recurrent.*
2. *There is a unique stationary distribution π such that, for all i, we have $\pi_i > 0$.*
3. *For all i, $h_{i,i} = \frac{1}{\pi_i}$.*
4. *Let $N(i,t)$ be the number of times the Markov chain visits state i in t steps. Then, $\lim_{t\to\infty} \frac{1}{t} \mathbf{E}[N(i,t)] = \pi_i$.*

3 Stationary Distributions of Random Walks

By Theorem 8.8, we can deduce that a random walk converges in a non-bipartite graph. Furthermore, it allows us to state the hitting time of a vertex v when starting from this same vertex.

Theorem 8.9. *If G is not bipartite, the random walk converges to the stationary distribution π with $\pi_v = \frac{d(v)}{2|E|}$ and $h_{v,v} = \frac{1}{\pi_v}$.*

Proof. By Theorem 8.8, we know that the random walk converges to the unique stationary distribution. So, we only have to verify that π is indeed stationary.

Since $\sum_{v \in V} d(v) = 2|E|$ (we are counting each edge exactly twice), we have $\sum_{v \in V} \pi_v = 1$. That is, π is a feasible distribution.

The probability of entering v is given as

$$\mathbf{Pr}\left[X_t = v\right] = \sum_{u \in N(v)} \frac{1}{d(u)} \mathbf{Pr}\left[X_{t-1} = u\right] \quad .$$

If π is stationary, then $\mathbf{Pr}\left[X_t = v\right] = \pi_v$ and $\mathbf{Pr}\left[X_{t-1} = u\right] = \pi_u$. We have

$$\sum_{u \in N(v)} \frac{1}{d(u)} \mathbf{Pr}\left[X_{t-1} = u\right] = \sum_{u \in N(v)} = \frac{1}{2|E|} |N(v)| = \frac{d(v)}{2|E|} = \pi_v \quad .$$

So π is stationary. ☐

4 Bounding the Cover Time

Now, we have all the tools to bound the cover time for non-bipartite graphs. As a first step, we bound the hitting time of neighbors of the initial vertex. That is, we derive a bound on how long it takes in expectation for a random walk to get to a particular neighbor of the current vertex.

Lemma 8.10. *If G is not bipartite, for $\{v, u\} \in E$, we have $h_{v,u} < 2|E|$.*

Proof. By Theorems 8.8 and 8.9, we have

$$h_{u,u} = \frac{1}{\pi_u} = \frac{2|E|}{d(u)} \quad .$$

Another way to get $h_{u,u}$ is to observe how we can get from u back to u. In the first step, we leave u to one neighbor $w \in N(u)$, then we need to come back to u, which takes $h_{w,u}$ steps in expectation. Therefore, by linearity of expectation

$$h_{u,u} = \frac{1}{d(u)} \sum_{w \in N(u)} (1 + h_{w,u}) \quad .$$

In combination

$$2|E| = \sum_{w \in N(u)} (1 + h_{w,u}) = h_{v,u} + \sum_{\substack{w \in N(u) \\ w \neq v}} (1 + h_{w,u}) > h_{v,u} \quad .$$ ☐

Using this lemma, we can show our final bound.

Theorem 8.11. *If G is not bipartite, the cover time is at most $4 \cdot |V| \cdot |E|$.*

Proof. The argument is pretty similar to the one used in Christofides algorithm. Consider a spanning tree of G and double all edges $|V| - 1$ in this tree. In this multigraph each vertex has even degree, therefore there is an Eulerian tour (a tour that crosses every edge in the multigraph). Let this vertices on this tour be $v_0, v_1, \ldots, v_{2(|V|-1)} = v_0$. So, every vertex is among $v_0, \ldots, v_{2|V|-3}$. As $\{v_{i-1}, v_i\} \in E$, we can use Lemma 8.10 to upper-bound the cover time by

$$\sum_{i=1}^{2|V|-3} h_{v_{i-1},v_i} \leq (2|V| - 3)2|E| \leq 4 \cdot |V| \cdot |E| \ . \qquad \square$$

At first sight, this bound feels pretty weak. However, there are indeed graphs in which this bound is tight.

5 Further Reading

- Chapter 7 in Mitzenmacher/Upfal
- Chapter 6 in Motwani/Raghavan