Lecture 7: Hitting Time and Cover Time of Random Walks

Lecturer: Thomas Sauerwald & He Sun

In the previous two lectures, we focused on the distribution of the visited vertices of the random walk. The goal of this lecture is to analyze the expected time for a random walk to visit a certain vertex, or, to visit all vertices of the underlying graph. We will show several combinatorial and spectral bounds on these parameters.

Let us formally define hitting and cover times:

- $\mathbf{H}(u,v) := \mathbf{E}\left[\min\{t \in \mathbb{N} \setminus \{0\} : X_t = v\} \mid X_0 = u\right]$,
- $\mathbf{C}(u) = \mathbf{E}\left[\min\{t \in \mathbb{N} \setminus \{0\} : \cup_{s=0}^{t} X_s = V\} \mid X_0 = u\right]$, $\mathbf{C} = \max_{u \in V} \mathbf{C}(u)$.

Recall that $\mathbf{H}(u,u) = 1/\pi_u = 2|E|/\deg(u)$. It can be easily verified that if $\mathbf{H}(u,v)$ were redefined as

$$\mathbf{H}(u,v) = \mathbf{E}\left[\min\{t \in \mathbb{N} \cup \{0\} : X_t = v\} \mid X_0 = u\right],$$

then $\mathbf{H}(u,v) + \mathbf{H}(v,u)$ forms a metric on $V \times V$.

Lemma 7.1. *There are regular graphs for which* $\mathbf{H}(u,v) \neq \mathbf{H}(v,u)$.

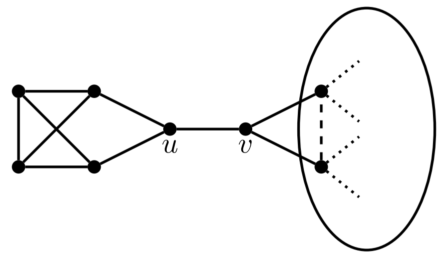

Figure 1: Illustration of the graph with an edge $\{u,v\} \in E$ such that $\mathbf{H}(u,v) = \mathcal{O}(1)$, but $\mathbf{H}(v,u) = \Omega(n)$. The ellipse on the right hand side represents a 3-regular graph with $\Omega(n)$ vertices, where the dashed edge has been removed.

Proof. Consider the graph in Figure 1 where an edge $\{u,v\} \in E$ connects a left graph with 5 vertices to a right graph with $n-5$ vertices. Then, $\mathbf{H}(u,v) = \mathcal{O}(1)$, since from every vertex in the left graph is a path of constant length to v. On the other hand,

$$n = \mathbf{H}(u,u) = \frac{1}{d} \cdot \sum_{w \in N(u)} (\mathbf{H}(w,u) + 1) = \frac{1}{3} \cdot (2 \cdot (\mathcal{O}(1) + 1) + \mathbf{H}(v,u) + 1),$$

from which it follows that $\mathbf{H}(v,u) = \Omega(n)$. □

Theorem 7.2. *The cover time of any graph G is upper bounded by*

$$2 \cdot (|V| - 1) \cdot 2|E| \leqslant 2n \cdot (n-1)^2.$$

Proof. We first observe that for any edge $\{u, v\} \in E$ that $\mathbf{H}(u, v) < 2|E|$, since $\mathbf{H}(u, u) = \frac{2|E|}{\deg(u)}$ implies

$$\frac{2|E|}{\deg(u)} = \mathbf{H}(u, u) = \frac{1}{\deg(u)} \sum_{w \in N(u)} (\mathbf{H}(w, u) + 1).$$

Now let T be any spanning tree of G. Consider a traversal of $T = (v_0, v_1, \ldots, v_{2|V|-2} = v_0)$ such that every vertex is visited at least once and every edge is traversed at most twice. Then the cover time is bounded by the expected time needed for this traversal of T:

$$\mathbf{C} \leqslant \sum_{i=0}^{2|V|-3} \mathbf{H}_{v_i, v_{i+1}} \leqslant (2|V| - 2) \cdot 2|E|.$$

□

An example for a graph with a cover time of $\Omega(n^3)$ is the so-called Lollipop graph which consists of a clique of size $(2/3)n$ and a path of length $n/3$ attached to it. In fact, it can be shown that for this graph the cover time is $(4/27 + o(1))n^3$ and moreover, the cover time is also always upper bounded by this value (see [Fei95]).

The next simple, but very important theorem connects the maximum hitting time to the cover time. It implies that the maximum hitting time approximates the cover time up to logarithmic factors.

Theorem 7.3. *For any graph G,*

$$\mathbf{C} = 6 \max_{u,v} \mathbf{H}_{u,v} \cdot \log n + 2n.$$

Proof. We divide the random walk of length $6 \max_{u,v} \mathbf{H}(u, v) \log n$ into $3 \ln n$ epochs each of length $2 \max_{u,v} \mathbf{H}(u, v)$. By Markov's inequality, the probability that a certain vertex is not visited within one epoch is at most $1/2$. Hence the probability that a vertex is not visited after $3 \ln n$ epochs is $2^{-3 \log n} = n^{-3}$. Taking the union bound, it follows that all vertices are visited with probability at least $1 - n^{-2}$. If this does not happen, then we use our upper bound of $4|V| \cdot |E| \leqslant 2n^3$ from Theorem 7. Thus the expected total time to cover all vertices is upper bounded by

$$6 \max_{u,v} \mathbf{H}(u, v) \cdot \log n + n^{-2} \cdot 2n^3 = 6 \max_{u,v} \mathbf{H}(u, v) \cdot \log n + 2n.$$

□

For deriving an upper bound on the cover time of regular graphs, we require the following lemma.

Lemma 7.4. *Let u, v be two vertices in a graph. Then any shortest path $P = (v_0 = u, v_1, \ldots, v_l = v$ between u and v satisfies*

$$\sum_{i=0}^{l-1} \deg(v_i) \leqslant 3n.$$

Proof. Let P be any shortest path between u and v. Every vertex w not lying on the shortest path can only be adjacent to at most three (consecutive) vertices on P. Hence,

$$\sum_{i=0}^{l-1} \deg(v_i) \leqslant l \cdot 2 + (n - l) \cdot 3 \leqslant 3n.$$

□

Theorem 7.5. *For any regular graph $G = (V, E)$,*

$$\max_{u,v} \mathbf{H}(u, v) \leqslant 9n^2.$$

Consequently, $\mathbf{C} = \mathcal{O}(n^2 \log n)$.

We point out that in [KLNS89] the stronger and tight bound $\mathbf{C} = \mathcal{O}(n^2)$ was shown.

Proof. Fix two arbitrary vertices u, v and consider a shortest path $P = (v_0 = u, v_1, \ldots, v_l = v)$ from u and v. Consider two consecutive vertices x and y on P. Let $\tilde{\mathbf{H}}(x, x)$ be the expected time to return to x conditioned on the event that the random walk does not move to y in the first step. Then,

$$n = \mathbf{H}(x, x) \geqslant \frac{\deg(x) - 1}{\deg(x)} \cdot \left(\tilde{\mathbf{H}}(x, x) + 1\right),$$

and hence,

$$\tilde{\mathbf{H}}(x, x) \leqslant 2n.$$

Using this and the fact that the random walk at x moves to y with probability $1/(\deg(x))$, we have

$$\mathbf{H}(x, y) \leqslant \left(1 + \tilde{\mathbf{H}}(x, x)\right) \cdot \deg(x) \leqslant 3n \cdot \deg(x).$$

Thus by the triangle inequality,

$$\mathbf{H}(u, v) \leqslant \sum_{i=0}^{l-1} \mathbf{H}(v_i, v_{i+1}) \leqslant \sum_{i=0}^{l-1} 2n \cdot \deg(v_i) \leqslant 3n \cdot 3n = 9n^2.$$

□

Let us now derive a spectral representation of the hitting time. To this end, we use the symmetric matrix

$$\mathbf{N} = \mathbf{D}^{-1/2} \mathbf{M} \mathbf{D}^{1/2}.$$

Let $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ be the eigenvalues of $\mathbf{N}$ and let $v_1, \ldots, v_n$ be the eigenvalues. Observe that if w_i is the i-th eigenvector of $\mathbf{M}$ corresponding to λ_i, then $v_i := \mathbf{D}^{-1/2} w_i$ is also an eigenvector of $\mathbf{N}$ with corresponding eigenvalue λ_i. Note that the entries of w_1 are proportional to the degree, hence the entries of v_1 are proportional to the square root of the degree. As usual, we may assume that $v_1, v_2, \ldots, v_n$ is an orthonormal set of eigenvectors of $\mathbf{N}$.

Theorem 7.6 (cf. [Lov93])**.** *For any two vertices $s, t \in V$,*

$$\mathbf{H}(s, t) = 2|E| \cdot \sum_{k=2}^{n} \frac{1}{1 - \lambda_k} \cdot \left(\frac{v_{k,t}^2}{\deg(t)} - \frac{v_{k,s} v_{k,t}}{\sqrt{\deg(s)\deg(t)}} \right)$$

Using this representation, we can derive a very famous formula, the so-called "Random Target Lemma".

Theorem 7.7 (Random Target Lemma)**.** *For any vertex $s \in V$,*

$$\sum_{t \in V} \pi(t) \cdot \mathbf{H}(s, t) = \sum_{k=2}^{n} \frac{1}{1 - \lambda_k}.$$

The point behind this theorem is that the right hand side is independent of the starting vertex s.

Proof. Using Theorem 7.6 and summing over all nodes $t \in V$ gives

$$\begin{aligned}
\sum_{t \in V} \pi(t)\mathbf{H}(s,t) &= \sum_{t\in V} \frac{\deg(t)}{2|E|} \cdot 2|E| \cdot \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \cdot \left(\frac{v_{k,t}^2}{\deg(t)} - \frac{v_{k,s}v_{k,t}}{\sqrt{\deg(s)\deg(t)}} \right) \\
&= \sum_{t\in V}\sum_{k=2}^{n} \frac{1}{1-\lambda_k} \cdot \left(v_{k,t}^2 - \frac{v_{k,s}v_{k,t}\sqrt{\deg(t)}}{\sqrt{\deg(s)}} \right) \\
&= \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\sum_{t\in V} v_{k,t}^2 - v_{k,s}\frac{1}{\sqrt{\deg(s)}} \cdot \sum_{t\in V} v_{k,t}\sqrt{\deg(t)} \right) \\
&= \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(1 - v_{k,s}\frac{1}{\sqrt{\deg(s)}} \cdot 0 \right) = \sum_{k=2}^{n} \frac{1}{1-\lambda_k},
\end{aligned}$$

where in the last line we used the fact that the eigenvectors $v_1, v_2, \ldots, v_n$ are orthonormal. □

Application to the Hypercube

Recall that the n eigenvalues of the $\log n$-dimensional hypercubes are given as follows:

$$\text{eigenvalue } 1 - \tfrac{2i}{\log n} \text{ with multiplicity } \tbinom{\log n}{i}, \qquad 0 \leqslant i \leqslant \log n.$$

Using the eigenvalues of the hypercube, let us first apply the random target lemma:

$$\begin{aligned}
\frac{1}{n}\sum_{t\in V} \mathbf{H}(s,t) &= \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \\
&= \sum_{k=1}^{\log n} \binom{\log n}{k} \frac{\log n}{2k} \\
&\leqslant 2 \sum_{k=1}^{\frac{1}{2}\log n} \binom{\log n}{k} \frac{\log n}{2k} \\
&\leqslant 2 \cdot \left(\log n \cdot \sum_{k=1}^{\frac{1}{4}\log n} \binom{\log n}{k} + 2 \sum_{k=\frac{1}{4}\log n+1}^{\frac{1}{2}\log n} \binom{\log n}{k} \right) \\
&\leqslant 2 \cdot \left(\log n \cdot (4e)^{\frac{1}{4}\log n} \binom{\log n}{k} + 2 \sum_{k=0}^{\log n} \binom{\log n}{k} \right) \\
&\leqslant c \cdot n,
\end{aligned}$$

where c is some constant independent of n.

Define now for any vertex $u \in V$,

$$\mathcal{S}_u := \{ v \in V \colon \mathbf{H}(u,v) \leqslant 4cn \}.$$

Note that for any vertex $u \in V$

$$|\mathcal{S}_u| \geqslant \frac{3}{4}n.$$

Now for $s \in V$ and $t \in V$, we have

$$\mathcal{S}_s \cap \mathcal{S}_t \neq \emptyset,$$

which implies that there is a vertex $w \in V$ so that

$$\mathbf{H}(s,w) \leqslant 4cn \qquad \text{and} \qquad \mathbf{H}(t,w) \leqslant 4cn.$$

Since the hypercube is vertex-transitive, $\mathbf{H}(t, w) = \mathbf{H}(w, t)$ which implies

$$\mathbf{H}(s, t) \leqslant \mathbf{H}(s, w) + \mathbf{H}(w, t) \leqslant 4cn + 4cn = 8cn.$$

Hence, $\max_{s,t} \mathbf{H}(s, t) = \mathcal{O}(n)$ and therefore $\mathbf{C} = \mathcal{O}(n \log n)$.

References

[Fei95] U. Feige. A Tight Upper Bound for the Cover Time of Random Walks on Graphs. *Random Structures and Algorithms*, 6(1):51–54, 1995.

[KLNS89] J.D. Kahn, N. Linial, N. Nisan, and M.E. Saks. On the Cover Time of Random Walks on Graphs. *Journal of Theoretical Probability*, 2(1):121–128, 1989.

[Lov93] L. Lovász. Random walks on graphs: A survey. *Combinatorics*, 2:1–46, 1993.