Lecture 12: Hitting Time and Cover Time

Lecturer: Shayan Oveis Gharan *May 4th*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications.*

12.1 Properties of Effective Resistance

We continue our discussion of effective by giving an overview of basic properties of effective resistances.

Lemma 12.1 (Metric Property)**.** *For any triple of vertices s, t, u,*

$$\text{Reff}(s,t) + \text{Reff}(t,u) \geq \text{Reff}(s,u).$$

Proof. By (**??**),

$$\begin{aligned}
\text{Reff}(s,u) &= b_{s,u}^\intercal L_G^\dagger b_{s,u} \\
&= (b_{s,t} + b_{t,u})^\intercal L_G^\dagger (b_{s,t} + b_{t,u}) \\
&= b_{s,t}^\intercal L_G^\dagger b_{s,t} + b_{t,u}^\intercal L_G^\dagger b_{t,u} + 2 b_{s,t}^\intercal L_G^\dagger b_{t,u} \\
&= \text{Reff}(s,t) + \text{Reff}(t,u) + 2 b_{s,t}^\intercal L_G^\dagger b_{t,u}.
\end{aligned}$$

So, we just need to show that the last term in the RHS is non-positive. The last term is equal to $p(t) - p(u)$ when we send one unit of flow from s to t. But, this means that t has the lowest potential in the network, so $p(u) \geq p(t)$ as required. □

Lemma 12.2 (Rayleigh Monotonicity Property)**.** *For a weight function $w : E \to \mathbb{R}_+$ let $\text{Reff}_w(.,.)$ be the effective resistance function when the conductance of each edge $e \in E$ is $w(e)$. For any w, w' such that $w \leq w'$ and any $s, t \in V$,*

$$\text{Reff}_w(s,t) \geq \text{Reff}_{w'}(s,t).$$

Proof. Let x be the one unit electrical flow from s to t with respect to w. Since $w \leq w'$,

$$\text{Reff}_w(s,t) = \sum_{e \in E} \frac{x(e)^2}{w(e)} \geq \sum_{e \in E} \frac{x(e)^2}{w'(e)}.$$

Since x is a feasible flow that sends one unit of flow from s to t, by Thompson's law, the RHS is at least the energy of the electrical flow that sends one unit from s to t w.r.t. w'. □

Lemma 12.3 (Convexity)**.** *The effective resistance is convex w.r.t. the conductances and is concave w.r.t. resistances. In particular, for any s, t,*

$$\frac{1}{2}(\text{Reff}_{1/w_1}(s,t) + \text{Reff}_{1/w_2}(s,t)) \leq \text{Reff}_{2/(w_1+w_2)}(s,t). \quad (12.1)$$

$$\frac{1}{2}(\text{Reff}_{w_1}(s,t) + \text{Reff}_{w_2}(s,t)) \geq \text{Reff}_{(w_1+w_2)/2}(s,t), \quad (12.2)$$

12.1.1 Bounding the Effective Resistance

?? gives a simple method to upper bound the effective resistance of a pair of vertices: All we need to do is to find a flow from s to t of small ℓ_2^2 energy. Conversely, the following lemma gives a lower-bound on the effective resistance of a pair of vertices.

Lemma 12.4 (Nash Williams Inequality). *Let $S_1, S_2, \dots, S_k \subseteq V$ such that for all $1 \leq i \leq k$, $s \in S_i, t \notin S_i$. If for all $1 \leq i < j \leq k$, $E(S_i, \overline{S_i}) \cap E(S_j, \overline{S_j}) = \emptyset$, then*

$$\text{Reff}(s,t) \geq \sum_{i=1}^{k} \frac{1}{\sum_{e \in E(S_i, \overline{S_i})} w(e)}.$$

Proof. Suppose x sends one unit of flow from s to t. We lower bound $\mathcal{E}(x)$ with the expression in the RHS. Since the cuts corresponding to $S_1, \dots, S_k$ are disjoint, we can write

$$\mathcal{E}(x) \geq \sum_{i=1}^{k} \sum_{e \in E(S_i, \overline{S_i})} \frac{x^2(e)}{w(e)}.$$

Therefore, it is enough to show that for each $1 \leq i \leq k$,

$$\sum_{e \in E(S_i, \overline{S_i})} \frac{x(e)^2}{w(e)} \geq \frac{1}{\sum_{e \in E(S_i, \overline{S_i})} w(e)}. \tag{12.3}$$

Since $(S_i, \overline{S_i})$ separates s, t, $\sum_{e \in E(S_i, \overline{S_i})} |x(e)| \geq 1$. Therefore, by Cauchy-Schwarz inequality,

$$\begin{aligned} 1 &\leq \left(\sum_{e \in E(S_i, \overline{S_i})} \frac{|x(e)|}{\sqrt{w(e)}} \cdot \sqrt{w(e)} \right)^2 \\ &\leq \left(\sum_{e \in E(S_i, \overline{S_i})} \frac{x(e)^2}{w(e)} \right) \cdot \left(\sum_{e \in E(S_i, \overline{S_i})} w(e) \right). \end{aligned}$$

This proves (12.3) and completes the proof of the lemma. □

As a simple application of the above lemma we can show that in a $\sqrt{n} \times \sqrt{n}$ grid there is a pair of vertices s, t such that $\text{Reff}(s,t) \geq \Omega(\log(n))$.

Next, we discuss methods for upper bounding the effective resistance between a pair of vertices s, t. Note that to upper bound the effective resistance it is enough to construct a feasible flow that sends one unit of flow from s to t, then the energy of the flow will give an upper bound on $\text{Reff}(s,t)$.

For example, suppose there are k edge disjoint paths each of length at most ℓ from s to t. Then we can construct x by sending $1/k$ amount of flow on each path and

$$\mathcal{E}(x) \leq \sum_{i=1}^{k} \sum_{e \in P_i} x(e)^2 = \sum_{i=1}^{k} \sum_{e \in P_i} \frac{1}{k^2} \leq \frac{k\ell}{k^2}.$$

Unfortunately, we may not be able to find many edge disjoint paths between s, t even though $\text{Reff}(s,t)$ is small. For example, in a k-dimensional hypercube there are at most k edge disjoint paths between each pair of vertices because the degree of each vertex is k. But because the length of each path between $s = 00\dots0$ and $t = 11\dots1$ is at least k, the best upper bound that we can get is $O(1)$.

12.2 Hitting Time

Consider a simple random walk on a (weighted) graph G, that is at a given vertex u it goes to random neighbor v with probability proportional to $w_{u,v}$, i.e.,

$$\mathbb{P}\left[X_1 = v | X_0 = u\right] = \frac{w_{u,v}}{d_w(u)},$$

where $d_w(u) = \sum_{v' \sim u} w_{u,v'}$ is the weighted degree of u. In this section we study the hitting time: for two vertices $u, v \in V$, we define the hitting time $H_{u,v}$ from u to v as the expected number of steps for the random walk to hit v (for the first time) when started at u. Formally, define the random variable

$$T = \min\{t \geq 0 : X_t = v\}.$$

Then $H_{u,v} = \mathbb{E}\left[T | X_0 = u\right]$. Note that his quantity is not necessarily symmetric, i.e., $H_{u,v} \neq H_{v,u}$ in general.

Lemma 12.5. *Let b be the vector which injects $d_w(u)$ unit at any vertex u and extracts $d_w(V)$ at v, and let p be the corresponding potential vector. For any vertex $u \neq v$,*

$$H_{u,v} = p(u) - p(v).$$

Proof. Fix a vertex v, and let $h : V \to \mathbb{R}_{\geq} 0$ be the hitting time vector. Then, for any $u \neq v$, we can write the following system of linear equations for the Hitting time:

$$h(u) = 1 + \sum_{u'} \frac{w_{u,u'}}{d_w(u)} h(u')$$

Equivalently,

$$d_w(u) = \sum_{u'} w_{u,u'}(h(u) - h(u')) = L_G h(u)$$

Adding up the above equalities for $u \neq v$ we obtain,

$$d_w(V) - d_w(v) = \sum_{u'} w_{u',v}(-h(v) + h(u')),$$

i.e., the constraint for v is redundant.

Now, if we consider the demand vector $b = d_w - \mathbf{1}_v d_w(V)$, i.e., we inject $d_w(u)$ at any vertex $u \in V$ and we extra all the flow at v and solve for p then p exactly satisfies all of the above equations:

$$L_G p = b$$

As alluded to before these equations have a unique solution (when the graph is connected) up to a shift, i.e., if h is a solution then so is $h + \alpha\mathbf{1}$ for any $\alpha \in \mathbb{R}$. So, we add one extra constraint that $h(v) = 0$. Equivalently, $p(u) - p(v) = h(v) - h(u) = h(v)$ as desired. $\square$

12.3 Commute Time

Given a pair of vertices u, v, the commute time $C_{u,v}$ is defined as:

$$C_{u,v} := H_{u,v} + H_{v,u}.$$

Theorem 12.6 (Commute Time)**.** *For any weighted graph $G = (V, E)$ and any pair of vertices u, v,*

$$C_{u,v} = Reff(u, v)d_w(V).$$

Proof. We build on the proof of Theorem 12.6. We consider 3 different demand vectors:

- Let b^A be the demand vector that injects $d_w(x)$ at any vertex x and extracts $d_w(V)$ at v.
- Let b^B be the demand vector that injects $d_w(V)$ at u and extracts $d_w(x)$ at any vertex x.
- Let $b^C = b^A + b^B$. In particular, b^C injects $d_w(V)$ at u and extracts it at v.

By, Theorem 12.6,

$$H_{u,v} = p_{b^A}(u) - p_{b^A}(v).$$

By symmetry,

$$H_{v,u} = -(p_{b^B}(v) - p_{b^B}(u))$$

So,

$$C_{u,v} = p_{b^A}(u) + p_{b^B}(u) - (p_{b^A}(v) + p_{b^B}(v)) \qquad = p_{b^C}(u) - p_{b^C}(v).$$

Most importantly, in the second equality we used that p can be obtained by solving a linear system involving the same matrix $L_G^{\dagger}$.

But by definition of effective resistance

$$p_{b^C}(u) - p_{b^C}(v) = d_w(V)(p_{b_{u,v}}(u) - p_{b_{u,v}}(v)) = d_w(V)\text{Reff}(u, v).$$

□

As an application, we can use **??** to give another proof of Lemma 12.1. Say for a triple of vertices s, t, u we want to show

$$\text{Reff}(s, t) + \text{Reff}(t, u) \geq \text{Reff}(s, u)$$

By **??** this is equivalent to

$$C(s, t) + C(t, u) \geq C(s, u).$$

Now, by the definition of expected Hitting time this is equivalent to

$$H(s, t) + H(t, s) + H(t, u) + H(u, t) \geq H(s, u) + H(u, s).$$

Rearranging the terms we need to show

$$(H(s, t) + H(t, u)) + (H(u, t) + H(t, s)) \geq H(s, u) + H(u, s).$$

But, for an triple of vertices s, t, u we have

$$H(s, t) + H(t, u) \geq H(s, u).$$

The expected number of steps to go from s to u is smaller than the expected number of steps to go from s to u while visiting t along the way.

12.4 Cover Time

The cover time of $G = (V, E, w)$ starting from u is the quantity $cov_u(G)$ which is the expected number of steps needed to visit **every** vertex of G started at u. Again, we can define this formally: Let

$$T = \min\{t \geq 0 : \{X_0, X_1, \ldots, X_t\} = V\}.$$

Then,

$$cov_u(G) = \mathbb{E}\left[T | X_0 = u\right].$$

Finally, we define the cover time of G as $cov(G) = \max_{u \in V} cov_u(G)$.

We can now use Theorem 12.6 to give a universal upper bound on the cover time of any graph.

Theorem 12.7 (Cover Time)**.** *For any unweighted connected graph $G = (V, E, w)$, we have $cov(G) \leq 2|E|(|V| - 1)$.*

Proof. Fix a spanning tree T of G. Then we have

$$cov(G) \leq \sum_{\{x,y\} \in T} C_{x,y}.$$

The right-hand side can be interpreted as a very particular way of covering the graph G: First, given the tree T, we can turn it into a TSP tour by visiting every edge twice, we start at some node x_0 and "walk" around the edges of the T in order $x_0, x_1, x_2, \ldots, x_{2(n-1)}$. In such a case the expected time from x_0 to get to x_1, is H_{x_0,x_1}, and similarly to cover G it takes at most,

$$\sum_{i=0}^{2(n-1)} H_{x_i,x_{i+1}} = \sum_{\{x,y\} \in T} C_{x,y}.$$

Note that this is just one particular way to visit every node of G, so it gives an upper bound on the cover time. Finally, we note that if $\{x, y\}$ is an edge of the graph, then by Theorem 12.6 we have $Cxy = 2|E|\text{Reff}(x, y) \leq 2|E|$. Here we use the fact that for every edge $\{x, y\}$ of G, the effective resistance is at most the resistance of the edge connecting x to y. A candidate flow is just send one unit of flow directly from x to y on the edge. □

12.5 Examples of Cover Time

First observe that if G is a d regular graph by a direct application of Theorem 12.7,

$$cov(G) \leq dn(n - 1).$$

The path Consider first G to be the path on vertices $\{0, 1, ..., n\}$. Then

$$H_{0,n} + H_{n,0} = C_{0,n} = 2n\text{Reff}(0, n) = 2n^2.$$

Since $H_{0,n} = H_{n,0}$ by symmetry, we conclude that $H_{0,n} = n^2$. Note that Theorem 12.7 implies that $cov(G) \leq 2n^2$, and clearly $cov(G) \geq H_{0,n} = n^2$, so the upper bound is off by at most a factor of 2.

The lollipop. Consider next the "lollipop graph" which is a path of length $n/2$ from u to v with an $n/2$ clique attached to v. We have

$$H_{u,v} + H_{v,u} = C_{u,v} = \Theta(n^2)\text{Reff}(u,v) = \Theta(n^3).$$

On the other hand, we have already seen that $H_{u,v} = \Theta(n^2)$. We conclude that $H_{v,u} = \Theta(n^3)$, hence $cov(G) = \Omega(n^3)$. Again, the bound of Theorem 12.7 is $cov(G) \leq O(n^3)$, so it's tight up to a constant factor here as well.

The complete graph. Finally, consider the complete graph G on n nodes. In this case, Theorem 12.7 gives $cov(G) \leq O(n^3)$ which is way off from the actual value $cov(G) = \Theta(n \log n)$ (since this is just the coupon collector problem).

12.6 Mathews Bound

Theorem 12.8 (Mathews Bound)**.** *If $G = (V, E)$ is an unweighted connected graph and $R_{\max} := \max_{x,y \in V} \text{Reff}(x,y)$ is the maximum pairwise effective resistance in G, then*

$$|E|R_{\max} \leq cov(G) \leq O(\log n)|E|R_{\max}.$$

Proof. One direction is straightforward:

$$cov(G) \geq \max H_{u,v} \geq \frac{1}{2}\max_{u,v} C_{u,v} \geq \frac{1}{2}|E| \max \text{Reff}(u,v) = |E|R_{\max}.$$

For the other direction, we will examine a random walk of length $2c|E|R_{\max} \log n$ divided into $\log n$ epochs of length $2c|E|R_{\max}$. Note that for any vertex v and any epoch i, we have

$$\mathbb{P}\left[v \text{ unvisited in epoch } i\right] \leq \frac{1}{c}.$$

This is because no matter what vertex is the first of epoch i, we know that the hitting time to v is at most $\max_u H_{u,v} \leq max_u C_{u,v} = 2|E|R_{\max}$. Now Markov's inequality tells us that the probability it takes more than $2c|E|R_{\max}$ steps to hit v is at most $1/c$. Therefore, the probability that we don't visit v in any epoch is at most $c^{-\log n} \leq n^{-\log c}$, and by a union bound, the probability that there is some vertex left unvisited after all the epochs is at most $n^{1-\log c}$. We conclude that

$$cov(G) \leq 2c|E|R_{\max} \log n + n^{1-\log c} n^3$$

where we have used the weak upper bound on the cover time provided by Theorem 12.7. Choosing c to be a large enough constant makes the second term negligible, yielding

$$cov(G) \leq O(|E|R_{\max} \log n),$$

as desired. □