A tight upper bound on the cover time for random walks on graphs

Uriel Feige *

Abstract

We prove that the expected time for a random walk to visit all n vertices of a connected graph is at most $\frac{4}{27}n^3 + o(n^3)$.

1 Introduction

Let $G = G(V, E)$ be a simple connected undirected graph with n vertices and m edges. We consider simple random walks on G, where at each step the random walk moves to a vertex chosen at random with uniform probability from the neighbors of the current vertex. Let u and v denote two vertices. The *hitting time* $H[u, v]$ is the expected number of steps it takes a walk that starts at u to reach v. The *commute time* $C[u, v]$ is the expected number of steps that it takes a walk to go from u to v and back to u (that is, $C[u, v] = H[u, v] + H[v, u]$). The *cover time* $EC[v]$ is the expected number of steps it takes a random walk that starts at v to visit all vertices of the graph. For a graph $G(V, E)$ its *hitting time* $H[G]$ (*commute time* $C[G]$, *cover time* $EC[G]$, respectively) is defined as $H[G] = \max_{u,v \in V}[H[u, v]]$ $(C[G] = \max_{u,v \in V}[C[u, v]]$, $EC[G] = \max_{v \in V}[EC[v]]$, respectively). Aleliunas *et al.* [2] have shown that for any connected graph, $EC < n^3$. It has been conjectured that for any graph, $EC \leq \frac{4}{27}n^3 + o(n^3)$ (see [1]). This bound is best possible (up to low order terms): the hitting time for the *lollipop* graph (a path of length $n/3$ connected to a clique of size $2n/3$) is $\frac{4}{27}n^3 + o(n^3)$ (for a walk that starts at the clique and seeks to hit the endpoint of the path). Brightwell and Winkler [3] show that the lollipop is extremal with respect to hitting time, and Coppersmith *et al.* [5] show that the commute time of any graph is at most $\frac{4}{27}n^3 + o(n^3)$. (Coppersmith *et al.* also prove a similar bound on the *meeting time*. See definitions in [5].)

In this paper we prove the above conjecture regarding the cover time. In fact, we prove a somewhat stronger statement. The *cyclic cover time*, $ECC[G]$, is the expected number

*Department of Applied Math., The Weizmann Institute, Rehovot, Israel. feige@wisdom.weizmann.ac.il.
Supported by a Koret Foundation fellowship.

of steps it takes to visit all vertices of the graph in a prespecified cyclic order, for the best cyclic order (that minimizes this value). That is,

$$ECC[G] = H[v_{i_1}, v_{i_2}] + H[v_{i_2}, v_{i_3}] + ... + H[v_{i_{n-1}}, v_{i_n}] + H[v_{i_n}, v_{i_1}]$$

where $(i_1, i_2, ..., i_n)$ is a permutation that minimizes the above sum. Clearly, $ECC[G] > EC[G]$.

Theorem 1 *For any connected graph on n vertices, the cyclic cover time satisfies:*

$$ECC \leq \frac{4}{27}n^3 + O(n^{5/2})$$

2 Proof of main theorem

There is a well known analogy between random walks and electrical resistance [6, 4]. Interpret each edge of G as an electrical resistance of 1 ohm. The *effective resistance* between vertices v and u, denoted by $R[v, u]$, is the voltage that develops in v if a current of 1 amp is injected into v and vertex u is grounded. Chandra *et al.* [4] have shown that the expected commute time between any two vertices satisfies $C[u, v] = 2mR[u, v]$.

Let T be a spanning tree of G. Define $R[T]$, the effective resistance of T, as the sum of the effective resistances along its edges. Since T has $n - 1$ edges, and the effective resistance between two endpoints of an edge is at most 1, it follows that the effective resistance of a spanning tree is at most $n - 1$. Let R_{span} denote the effective resistance of the minimum resistance spanning tree of G. Then $ECC \leq 2mR_{span}$ [4]. This expression can be bounded from above by observing that $2m \leq n(n - 1)$ and $R_{span} \leq n - 1$, hence the bound $ECC \leq n(n - 1)^2$. We shall prove our stronger upper bound by showing that there is a tradeoff between the two quantities, m and R_{span}, and they cannot both attain their maximum value simultaneously.

We use the following lemma, which is due to Foster [7] (see also [8]).

Lemma 2 *The sum of effective resistances along the edges of a connected graph is $n - 1$.*

$$\sum_{(u,v)\in E} R[u, v] = n - 1$$

Now we are ready to prove Theorem 1.

Proof of Theorem 1: Consider T, the minimum resistance spanning tree of G. Edges in T are called *twigs*, and other edges are called *links*. If $R[T] \leq 4n/27$ then we are done. So assume that $R[T] = \alpha n$, for $\alpha > 4/27$. Let $\epsilon = 1/\sqrt{n}$. Since we may assume that $n > 50$, it follows that $\epsilon < \alpha$. Call an edge (u, v) *heavy* if $R[u, v] \geq \epsilon$, and otherwise *light*.

Lemma 3 *At least $(\alpha - \epsilon)n$ twigs are heavy. At most $(n-1)/\epsilon$ links are heavy.*

Proof: Assume that less than $(\alpha - \epsilon)n$ twigs are heavy. Then $R[T] \leq (\alpha - \epsilon)n + (n-1)\epsilon < \alpha n$, a contradiction. Assume that more than $(n-1)/\epsilon$ links are heavy. Then $\sum_{(u,v)\in E} R[u,v] > \epsilon \frac{n-1}{\epsilon} = n-1$, contradicting Lemma 2. $\square$

A *cut-set* is the set of edges that connect a specified set of vertices to its complement set. A *fundamental cut-set* is a cut set that contains exactly one twig (and arbitrarily many links). There are $n-1$ fundamental cut-sets. Denote the edges of T by $e_1, e_2, \ldots, e_{n-1}$. Each fundamental cut-set is denoted by the name of its twig. Select a root of T arbitrarily, and denote this vertex by r. Each fundamental cut-set e_i partitions V into two sets of vertices, and we denote the one that does not contain r by V_i.

Consider now only the fundamental cut-sets of heavy twigs. Without loss of generality, assume that $e_1, \ldots, e_k$, are the heavy twigs, for $k \geq (\alpha - \epsilon)n$, and let $V_1, \ldots, V_k$, be the sets of vertices associated with these cut-sets (they do not contain r). Define $U_0 = V \setminus \bigcup_{j=1}^{k} V_j$, and $U_i = V_i \setminus \bigcup_{V_j \subset V_i} V_j$.

Lemma 4 *The sets $\{U_0, U_1, \ldots, U_k\}$ are each nonempty, disjoint, and together contain all vertices of V.*

Proof: It can easily be verified that the following is an alternative characterization of the sets $\{U_0, U_1, \ldots, U_k\}$. For each vertex $v \in V$, consider the unique path in T connecting v to the root r. If this path does not use any heavy twig, then $v \in U_0$. Otherwise, let e_i be the first heavy twig that is used. Then $v \in U_i$.

It follows that each vertex is contained in exactly one of the sets, and that each set is nonempty (note that $r \in U_0$). $\square$

Lemma 5 *The number m of edges in G satisfies*

$$2m \leq (1-\alpha)^2 n^2 + O(n^{3/2})$$

Proof: The number of twigs is $n-1$. The number of heavy links is at most $(n-1)/\epsilon$ (by Lemma 3). It remains to bound the number of light links.

A fundamental cut set of a heavy twig cannot contain a light link, as otherwise this link could replace the twig in a new spanning tree T' for which $R[T'] < R[T]$, contradicting the fact that T is of minimum effective resistance. It follows that any light link of G must connect two vertices inside the same U_i, and the number of light links is at most $\sum_{i\geq 0}^{k} \binom{|U_i|}{2}$. Since $\sum_{i\geq 0}^{k} |U_i| = n$ and each U_i is nonempty (Lemma 4), the number of light links is at most $\binom{n-k}{2}$. By Lemma 3, $k \geq (\alpha - \epsilon)n$. Substituting $1/\sqrt{n}$ for ϵ, the proof follows by straightforward manipulations. $\square$

Using our assumption that $R_{span} = R[T] = \alpha n$, and the bound from Lemma 5, we obtain:

$$ECC \leq 2mR_{span} \leq \alpha(1-\alpha)^2 n^3 + O(n^{5/2}) \leq 4n^3/27 + o(n^3)$$

(using the fact that $\alpha(1-\alpha)^2$ is maximized when $\alpha = 1/3$). $\square$

References

[1] D. J. Aldous. "Reversible markov chains and random walks on graphs". *Draft of first six chapters of book*, January 26, 1993.

[2] R. Aleliunas, R. M. Karp, R. J. Lipton, L. Lovász, and C. Rackoff. "Random walks, universal traversal sequences, and the complexity of maze problems". In *20th Annual Symposium on Foundations of Computer Science*, pages 218–223, San Juan, Puerto Rico, October 1979.

[3] G. Brightwell and P. Winkler. "Maximum hitting times for random walks on graphs". *Random Structures and Algorithms*, 1:263-276, 1990.

[4] A. Chandra, P. Raghavan, W. Ruzzo, R. Smolensky, and P. Tiwari. "The electrical resistance of a graph captures its commute and cover times". In *Proceedings of the Twenty First Annual ACM Symposium on Theory of Computing*, pages 574-586, Seattle, WA, May 1989.

[5] D. Coppersmith, P. Tetali, and P. Winkler. "Collisions among random walks on a graph". *SIAM J. on Discrete Math.* (to appear).

[6] P. G. Doyle and J. L. Snell. "Random Walks and Electrical Networks". *The Mathematical Association of America*, 1984.

[7] R. Foster. "The average impedance of an electrical network". *Contributions to applied mechanics (Reissner anniversary volume)*, Ann Arbor, pp. 333-340, Edwards Brothers, Inc. 1949.

[8] P. Tetali. "Random walks and the effective resistance of networks". *Journal of Theoretical Probability*, 4:101-109, 1991.