A Tight Lower Bound on the Cover Time for Random Walks on Graphs

Uriel Feige *

May 4, 2000

Abstract

We prove that the expected time for a random walk to cover all n vertices of a graph is at least $(1+o(1))n \ln n$.

1 Introduction

Let G be a simple connected undirected graph on n vertices and m edges. We consider 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 $E_v[G]$ denote the *cover time*, the expected number of steps that it takes a walk that starts at v to visit all vertices of G. It is a well known conjecture (see for example [1, 6, 14, 17]) that for connected graphs on n vertices, $\min_G \min_v E_v[G] \geq (1+o(1))n \ln n$, where $\ln n$ denotes the natural logarithm of n, and $o(1)$ denotes a (possibly negative) term that tends to 0 as n tends to ∞. We prove this conjecture.

This lower bound is best possible up to low order terms, as demonstrated by the complete graph on n vertices. The complete graph is not extremal with respect to $\min_v E_v[G]$, and faster cover times are possible by considering graphs in which a short path extends from a clique, and the random walk starts from the end point of the path. See [1, 3, 5] for more details.

1.1 Related work

To the best of the author's knowledge, the lower bound conjecture was first formulated by Linial. The conjecture has been proven for trees [14] (see [10, 5] for improved bounds). The

*Department of Applied Math and Computer Science, The Weizmann Institute, Rehovot, Israel. feige@wisdom.weizmann.ac.il. Research supported by a Koret Foundation fellowship.

conjecture has been proven, up to the leading constant factor, for graphs satisfying certain conditions on their eigenvalues [6, 17], and for regular graphs [3]. More results on special types of graphs can be found by tracing the references in [3].

On arbitrary graphs, Aldous [2, 3] proves an $\Omega(n \ln n)$ lower bound for walks that start at the stationary distribution (rather than any vertex).

In terms of upper bounds on connected graphs, Aleluinas *et al.* [4] prove that $\max_G \max_v E_v[G] \leq n(n-1)^2$. This upper bound was improved to $(1+o(1))4n^3/27$ [12], giving the best possible leading constant. As for min-max bounds, $\min_G \max_v E_v[G] = (1+o(1))n \ln n$ follows from the current paper, and $\max_G \min_v E_v[G] = (1+o(1))2n^3/27$ is shown in [13].

1.2 Useful technical background

For a vertex u, d_u denotes its degree (the number of vertices adjacent to it). For two vertices $u, v \in G$, the *hitting time* $H[u,v]$ is the expected number of steps it takes a walk that starts at u to reach v, and 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 *difference time* is defined as $D[u,v] = H[u,v] - H[v,u]$. A useful property of the difference time is its additivity (see [3, 9] for proof).

Lemma 1 *For any three vertices u, v, w, $D[u,w] = D[u,v] + D[v,w]$.*

A useful tool in the analysis of random walks is the notion of *electrical resistance* $R[u,v]$ between vertices, where a graph is regarded as an electrical network in which each edge has unit resistance (see [11]). Electrical resistance obeys the triangle inequality $R[u,v] + R[v,w] \geq R[u,w]$. Chandra *et al.* [7] show that

Lemma 2 *For any two vertices $C[u,v] = 2mR[u,v]$*

and Tetali [16] shows

Lemma 3 *For any two vertices*

$$H[u,v] = \frac{1}{2} \sum_w d_w (R[u,v] + R[v,w] - R[u,w])$$

.

We use the following lower bound on the resistance between two vertices.

Lemma 4 *If u is not adjacent to v, then $R[u,v] \geq 1/d_u + 1/d_v$. Otherwise, $R[u,v] \geq 1/(d_u+1) + 1/(d_v+1)$.*

Lemma 4 is proved by considering the new network in which all vertices (excluding u and v themselves) are shorted together, and applying the monotinicity property of electrical resistance. For a more detailed proof of the lemma, see [8].

An important technique for lower bounding the cover time was introduced by Matthews [15]. The following lemma is a refinement of Matthews' method, and is proved in [17].

Lemma 5 *Let S be a subset of vertices of G, and let t be such that for all $v \in S$, at most b of the vertices $u \in S$ satisfy $H[v,u] < t$. Then $E_v[G] \geq t(\ln(|S|/b) - 2)$.*

2 The lower bound

Theorem 6 *For any graph G on n vertices, and any starting vertex u*

$$E_u[G] \geq n \ln n + o(n \ln n)$$

Proof: Throughout this proof, we use the convention that ln denotes natural logarithms, and log denotes $\log_2$.

Consider the following two conditions:

1. There are two vertices, u and v, such that $H[u,v] \geq n \ln n$ and $H[v,u] \geq n \ln n$.

2. The assumptions of Lemma 5 hold with parameters $|S| > n/(\log n)^c$, $b < (\log n)^c$, and $t \geq n(1 - c/\log n)$, for some constant c independent of n.

If either one of the above conditions hold, then Theorem 6 trivially follows. For some graphs (e.g., the complete graph), Condition 1 does not hold. For some graphs (e.g., a star with $\sqrt{n}$ rays, each of length $\sqrt{n}$), Condition 2 does not hold. We will show that for any graph, one of the two conditions must hold. (In fact, a simple modification to our proof shows that if $E_u[G] \leq n(\ln n)^a$ for some constant $a \geq 1$, then Condition 2 must hold.)

We first identify a set S of the vertices of G that has favorable properties: the commute time between most pairs of vertices in this set is at least (roughly) $2n$. Furthermore, all vertices in S have approximately the same degree.

Lemma 7 *For some value of d, there exists a subset S of vertices, $|S| > n/2(\log n)^2$, such that for all $v \in S$:*

- $d \leq d_v \leq 2d - 1$,

- *at most* $2 \log n$ *of the vertices $u \in S$ satisfy* $C[v,u] < 2n(1 - 2/\log n)$.

Proof: Let $d_{ave} = 2m/n$. Then simple counting shows that there must be at least $n/(1+\log n)$ vertices of degree at most $d_{ave}(1+1/\log n)$. These are the vertices in a temporary set S_{temp}. We now use Lemmas 2 and 4. Consider two cases:

1. $d_{ave} > \log n$. Then for any two vertices $u, v \in S_{temp}$,

$$C[u,v] \geq n d_{ave} \frac{2}{d_{ave}(1+1/\log n)+1} > 2n(1-\frac{2}{\log n})$$

2. $d_{ave} \leq \log n$. Then each vertex has at most $(1+1/\log n)\log n$ neighbors in S_{temp}, and for any two vertices that are not adjacent,

$$C[u,v] \geq n d_{ave} \frac{2}{d_{ave}(1+1/\log n)} > 2n(1-\frac{1}{\log n})$$

To complete the proof, partition the set S_{temp} into at most $\log n$ sets, where the i^{th} set contains all vertices of degree between 2^{i-1} and $2^i - 1$. Let S be the largest of these sets. □

Order the vertices of S such that if $i < j$, then $D[v_i, v_j] \geq 0$. Such an ordering must exist, by Lemma 1. Now partition S into subsets S_i, where set S_i contains all vertices v_j such that $(i-1)n(\log n)^{-1} \leq D[v_1, v_j] < i \cdot n(\log n)^{-1}$. We now distinguish between two cases:

1. For some $i < j$, $H[v_j, v_i] \geq n \log n$. Then since $H[v_i, v_j] \geq H[v_j, v_i]$, Theorem 6 follows.

2. For some i, $|S_i| > n/16(\log n)^8$. By the additivity property of the difference time, for any two vertices in S_i, $|D[u,v]| < n/\log n$. Since the commute time between any two vertices in S_i (except for $2\log n$ neighbors of each vertex) is at least $2n(1-2/\log n)$, we get that the hitting time is at least $n(1-3/\log n)$. Now we can apply Lemma 5, proving Theorem 6.

It remains to show that one of the above two cases must hold. Assume otherwise. Then we shall show that $H[v_k, v_1] > n \log n$, for $k = |S|$, implying Case 1 above – a contradiction.

To compute $H[v_k, v_1]$ we use Tetali's formula (Lemma 3). Consider a single term T_i in this formula: $T_i = R[v_k, v_1] + R[v_1, v_i] - R[v_k, v_i]$. For the following lemma, recall that d specifies the range of degrees for vertices in S.

Lemma 8 *If cases 1 and 2 above do not hold, then for* $i \geq n/4(\log n)^2$,

$$T_i \geq (1+o(1))\frac{8(\log n)^3}{d}$$

Proof: Using Lemma 2 we obtain $T_i = \frac{1}{2m}(C[v_1, v_k] + C[v_1, v_i] - C[v_i, v_k])$. Using $C[u, v] = D[u, v] + 2H[v, u]$, and the assumption that Case 1 does not hold, we obtain:

$$T_i > \frac{1}{2m}(D[v_1, v_k] + D[v_1, v_i] - D[v_i, v_k] - 2n \log n) = \frac{1}{2m}(2D[v_1, v_i] - 2n \log n)$$

the latter equality following from additivity of the difference time.

Consider the sequence of vertices $v_{i_1}, ..., v_{i_\ell}$, where $\ell = 4(\log n)^3$, and for each $1 \le j \le \ell$, $i_j = 1 + (j-1)n/16(\log n)^5$ (we assume for simplicity that these numbers are integers). Since Case 2 does not hold, it follows that $D[v_{i_j}, v_{i_{j+1}}] \ge n(\log n)^2$. Since Case 1 does not hold, $D[v_{i_j}, v_{i_{j+1}}] \ge (\log n)H[v_{i_{j+1}}, v_{i_j}]$, and

$$C[v_{i_j}, v_{i_{j+1}}] \le (1 + \frac{2}{\log n})D[v_{i_j}, v_{i_{j+1}}]$$

implying

$$\frac{D[v_{i_j}, v_{i_{j+1}}]}{2m} \ge (1 + o(1))R[v_{i_j}, v_{i_{j+1}}]$$

By the degree bound for vertices in S, and by Lemma 4, it follows that for any two vertices in S, $R[u, v] \ge 1/d$. Hence for any vertex u beyond v_{i_ℓ}, we obtain:

$$T_u > \frac{1}{2m}(2D[v_1, u] - 2n \log n) > \frac{2 + o(1)}{2m} D[v_1, u]$$

$$= \frac{2 + o(1)}{2m}((\sum_{j=1}^{\ell-1} D[v_{i_j}, v_{i_{j+1}}]) + D[v_{i_\ell}, u])$$

$$\ge (2 + o(1)) \sum_{j=1}^{\ell-1} R[v_{i_j}, v_{i_{j+1}}]$$

$$\ge (2 + o(1))(\ell - 1)\frac{1}{d} \ge (1 + o(1))\frac{8(\log n)^3}{d}$$

□

Now use Tetali's formula, ignoring some of the terms T_i (these terms are nonnegative, by the triangle inequality for resistances).

$$H[v_k, v_1] = \frac{1}{2} \sum_{w \in G} d_w (R[u_k, v_1] + R[v_1, w] - R[v_1, w])$$

$$\geq \frac{1}{2} \sum_{i > n/4(\log n)^2}^{n/2(\log n)^2} d_{v_i} T_i$$

$$\geq \frac{1}{2} \frac{n}{4(\log n)^2} \cdot d \cdot (1 + o(1)) \frac{8(\log n)^3}{d} \geq (1 + o(1)) n \log n$$

This completes the proof of Theorem 6. $\square$

References

[1] D. J. Aldous. "An introduction to covering problems for random walks on graphs". *Journal of Theoretical Probability*, 2(1):87–89, January 1989.

[2] D. J. Aldous. "Lower bounds for covering times for reversible Markov chains and random walks on graphs". *Journal of Theoretical Probability*, 2(1):91–100, January 1989.

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

[4] 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.

[5] G. Brightwell and P. Winkler. "Extremal Cover Times for Random Walks on Trees". *Journal of Graph Theory, 14(5):547–554, 1990.*

[6] A. Broder and A. Karlin. "Bounds on the cover time". *Journal of Theoretical Probability*, 2(1):101–120, January 1989.

[7] 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.

[8] D. Coppersmith, U. Feige, and J. Shearer. "Random Walks on Regular and Irregular Graphs". *Technical report CS93-15*, the Weizmann Institute, Israel, 1993.

[9] D. Coppersmith, P. Tetali, and P. Winkler. "Collisions among random walks on a graph". *SIAM J. on Discrete Math.,* 6(3):363-374, August 1993.

[10] L. Devroye and A. Sbihi. "Inequalities for random walks on trees". In *A. Frieze and T. Luczak, editors, Random Graphs, Vol. 2, 35–45, Wiley, 1992.*

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

[12] U. Feige. "A Tight Upper Bound on the Cover Time for Random Walks on Graphs". To appear in *Random Structures & Algorithms.*

[13] U. Feige. "Collecting Coupons on Trees, and the Analysis of Random Walks". *Technical report CS93-20*, the Weizmann Institute, Israel, 1993.

[14] 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, January 1989.

[15] P.C. Matthews. "Covering Problems for Brownian Motion on Spheres. *Ann. Probab., 16:189-199, 1988.*

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

[17] D. Zuckerman. "A Technique for Lower Bounding the Cover Time". *SIAM J. Disc. Math., 5:81-87, 1992.*