Bounds for Edge-Cover by Random Walks

Athina Panotopoulou *

February 24, 2013

Abstract

Let $G = (m, n)$ be an undirected graph with m edges and n vertices. For a random walk on G it is known that the time to cover all its edges is bounded by $O(mn)$ [2]. In a later work the bound $O(m^2)$ [3] is proved, which holds even for graphs with weighted edges. Here, we briefly discuss these results along with their proofs.

1 Introduction

The work [2] proves a bound $O(mn)$ for random walks on general graphs, and a bound $O(n^2)$ for regular graphs.

The work [3] proves that the bound $O(m^2)$ holds even for graphs with weighted edges; then m is the total weight of all the edges in the graph, instead of their number. Specifically, the authors prove that the expected time to cover each edge in just one direction is at most $2m^2$, while the expected time to cover each edge in both directions is at most $3m^2$. In order to prove both results, they bound the expected edge cover time of random walks that start and conclude at the same vertex.

2 Preliminaries

A random walk on a connected, undirected graph without weights is a walk that starts at a vertex x of the graph and at every time step chooses one of the neighbours of the current vertex with equal probability and the time to traverse every edge is one. On the other hand, on graphs with weights $l(e)$ on edges the time to traverse an edge of length l is l^2 and the

*Dartmouth College, Hanover NH. Final paper for the course Random Walks.

transition probability from a vertex x to its neighbour y is $\frac{1/l(x,y)}{\Sigma_{z \circ x} 1/l(x,y)}$, where $z \circ x$ denotes the vertices z that are neighbours of x.

Let $\mathbb{E}_x(T_u)$, $\mathbb{E}_x(T_{(u,v)})$ to be the expected time starting from x to visit vertex u, and traverse edge (u, v), respectively. Also, $\mathbb{E}_x C_e$ is the expected time starting from vertex x to visit (i.e., cover) all edges in both directions, $\mathbb{E}_x C_a$ visit all edges in at least one direction, $\mathbb{E}_x C_e$ visit all edges in both directions. Let $\mathbb{E}_x CR_e$ be the expected time starting from vertex x to cover all the edges in both directions, and then return to the vertex x, $\mathbb{E}_x CR_a$ the same as before but cover only one direction. Moreover, let $c_v = max_{u \in V}\{\mathbb{E}_u C_v\}$ be the vertex cover time. Finally, $cr_e = max_{u \in V}\{\mathbb{E}_u CR_e\}$ be the edge cover and return time, $cr_a = max_{u \in V}\{\mathbb{E}_u CR_a\}$ be the arc cover and return time.

In the following, we will interpret any undirected graph with its equivalent directed representation; i.e., we replace each edge of the undirected graph with two bi-directed edges, (and the same holds even for loops, or multiple edges). Additionally, we will consider $m := |E|$ when the edges are unweighted (or, simply, have weight one), and $m := \Sigma_e \in E(G) l(e)$ when each edge e is associated with a weight $l(e)$.

3 Results and Proofs

3.1 On the results of paper [2]

We consider random walks on a simple un-weighted graph. Aleliunas et al. showed that the vertex cover time is $O(mn)$ [5].

Lemma 3.1. *For any edge (v, w) it is $E_v T_{(v,w)} < m$.*

Proof. We have: $\mathbb{E}_v T_{(v,w)} < \mathbb{E}_w T_v + \mathbb{E}_v T_{(v,w)} \Rightarrow \mathbb{E}_v T_{(v,w)} < \mathbb{E}_w T_{(v,w)} \Rightarrow \mathbb{E}_v T_{(v,w)} < \mathbb{E}_{(v,w)} T_{(v,w)}$. Now, if we take the equivalent random walk on the edges, we would have that the stationary probability for every edge is $1/m$. Consequently, if we take a walk the expected time starting from an edge to return to that edge would be m. Thus, $\mathbb{E}_v T_{(v,w)} < m$. □

Now, we have the following theorem which is based on this observation: if we are at a specific vertex it is easy to cover all its adjacent edges; thus, if we cover all the vertices of a graph, it is also easy to cover all its edges.

Theorem 3.2. *For any positive k it is $cr_a = O(k(c_v + m^{1+\frac{1}{k}}))$.*

Proof. We are taking a random walk starting at x, with which we want to cover all the edges. We want to bound the length of this walk. We divide the walk into T walk pieces each with duration S_i, $\forall i = 1, \dots, T$. The division of the initial walk is as follows: at every piece we start a random walk at x, we cover all vertices, we walk for another $m^{1+\frac{1}{k}}$ steps, and we continue until we reach x again. When we arrive at x a new piece of walk starts. Observe that the S_i's are i.i.d., because every walk piece starts from the same vertex, so it does not depend to the previous piece-walks and comes from the same distribution. We need to compute now how many such pieces we need until we are able to cover all directed edges, namely $T = \min\{t|$ all edges are traversed after $S_1, S_2, ..., S_t$ steps$\}$: $\mathbb{E}_x C_e \leq \mathbb{E}_x[\Sigma_{i=1}^{T} S_i] \Rightarrow \mathbb{E}_x C_e \leq \mathbb{E}_x S_i \mathbb{E}_x T$ since S_i's are i.i.d., and using Wald's identity [4]. Thus, we only need to bound $\mathbb{E}_x S_i$ and $\mathbb{E}_x T$. We know from the definition of S_i that: $\mathbb{E}_x S_i \leq c_v + m^{1+\frac{1}{k}} + \max_u\{\mathbb{E}_u T_x\} \leq 2c_v + m^{1+\frac{1}{k}}$. Next, to bound $\mathbb{E}_x T$ we need the following: Let $B_{i,u,v}$ be the event in which the i-th walk fails to traverse the directed edge (u, v). We have the following Lemmas:

Lemma 3.3. $Pr(B_{i,u,v}) < m^{-\frac{1}{k}}$.

Proof. We define p_j as the probability of going from x to (u, v) in j steps. We have that, $\mathbb{E}_u T_{(u,v)} = \Sigma_{j=0}^{\infty} p_j j = \Sigma_{j=0}^{S_i} p_j j + \Sigma_{j=S_i+1}^{\infty} p_j j \Rightarrow \mathbb{E}_u T_{(u,v)} > \Sigma_{j=S_i+1}^{\infty} p_j j \Rightarrow m > \Sigma_{j>S_i}^{\infty} p_j j$, using Lemma 3.1. We have that $S_i > m^{1+1/k}$ by definition so: $m > m^{1+1/k} \Sigma_{j>S_i}^{\infty} p_j \Rightarrow m^{-\frac{1}{k}} > P(B_{i,u,v})$, and the proof is complete. □

Lemma 3.4. $Pr(T > 2kj) < 1/m^j$.

Proof. It is $Pr(T > 2k) = Pr(\exists(u, w), \forall i \in \{1, \dots, 2k\} : B_{i,u,w})$, i.e., $Pr(T > 2k) < \Sigma_{j=1}^{m} Pr(\forall i \in \{1, \dots, 2k\} : B_{i,e_j})$, or $Pr(T > 2k) < \Sigma_{j=1}^{m} Pr(B_{1,e_j}, \dots, B_{2k,e_j}) = \Sigma_{j=1}^{m} Pr(B_{1,e_j})^{2k}$ because S_i's are i.i.d. Using the previous lemma 3.3: $Pr(T > 2k) < m(m^{-\frac{1}{k}})^{2k} = 1/m$. Finally, $Pr(T > 2kj) = Pr(T = T_1 + T_2 + \cdots + T_j > 2kj) = Pr(T_1 > 2k, \dots, T_j > 2k) < 1/m^j$. □

Now, we can prove that $\mathbb{E}_x T < 2k/(1 - 1/m)$ since $\mathbb{E}_x T = \Sigma_{j=0}^{\infty} P(T > j)$ equals $\Sigma_{j=0}^{2k-1} P(T > j) + \Sigma_{j=2k}^{4k-1} P(T > j) + \dots$, which is less than $2kP(T > 0) + 2kP(T > 2k) + \dots$, and using Lemma 3.4 this is less than $2k\Sigma_{j=0}^{\infty} 1/m^j$.

So far, we established the bounds $\mathbb{E}_x S_i \leq 2c_v + m^{1+\frac{1}{k}}$, $\mathbb{E}_x T < 2k/(1 - 1/m)$, and as a result, for any positive k it is $cr_e = O(k(c_v + m^{1+1/k}))$. □

We note that, using results for the vertex cover problem as obtained by [5], as well as, Theorem 3.2, we have that for $k = 2$ it is $O(mn)$. On the other hand, for regular graphs where $m = n^{2-\delta}$, if we take $k > (2-\delta)/\delta$, it is only $O(n^2)$.

3.2 On the results of paper [3]

Now, we have a random walk on a multi-graph with weights l on the edges (a multi-graph is a graph that may have multiple edges and loops). Note that we can interpret our scheme as an electrical network where the resistance of every edge is equal to the length of that edge. We define the commute time $T^{x\leftrightarrow y} = \mathbb{E}_x(T_y) + \mathbb{E}_y(T_x)$ as the time it takes for a random walk to travel from x to y and back to x.

Lemma 3.5. *The $\mathbb{E}T^{x\leftrightarrow y} = 2mR^{xy}$ holds both for weighted and un-weighted graphs.*

We note that the above was proved for unit edge lengths in [6]. Using Theorem 2.2 from [6] and substituting the time you need to traverse an edge e $(l(e)^2)$ we have that the same result holds even for graphs with weighted edges.

The basic idea of the entire proof is the use of four different expressions that represent four different types of commute tour using as characteristic of each type the sub-networks the tour traverse.

Suppose that G is the union of two sub-networks A, B such that $A \cap B = \{x, y\}$, and $A \cup B$ is the set of all vertices: We start from vertex x on G and we define four different types of $T^{x\leftrightarrow y}$ related to closed walks start at x, pass through y, and return to x. Specifically:

1. An $\overrightarrow{A}-$commute is a closed walk that passes through A as the walk goes from x to y.
2. An $\overleftarrow{A}-$commute is a closed walk that passes through A as the walk returns from y.
3. An $A-$commute is a closed walk that passes through A.
4. An $\overleftrightarrow{A}-$commute is a closed walk that passes through A both ways.

We compute the expected times for each of the above commute times: $\mathbb{E}T^{xy}_{A}$, $\mathbb{E}T^{xy}_{\overrightarrow{A}}$, $\mathbb{E}T^{xy}_{\overleftarrow{A}}$, $\mathbb{E}T^{xy}_{\overleftrightarrow{A}}$.

Proof. We start the walk from x and stop after T steps until we have the first commute we are interested in from the above list, and we want to find $\mathbb{E}T$. To this end, let Y be the number of commute walks from x to y until we stop; X_i be the duration of every commute walk, $i = \{1, 2, \dots, Y\} \Rightarrow T = \Sigma_{1 \le i \le Y} X_i$. Now, $\mathbb{E}T^{x \leftrightarrow y} = 2mR^{xy}$ is the expectation of every X_i, from Lemma 3.5, so

$$\mathbb{E}T = 2mR^{xy}\mathbb{E}Y, \tag{1}$$

because X_i's are i.i.d., from Wald's identity [4]. Therefore, to find $\mathbb{E}T$ we only need to compute $\mathbb{E}Y$ for each commute type:

1. $\mathbb{E}T^{xy}_{\overrightarrow{A}} = 2mR_A$: Specifically, let p_A be the probability that starting from x the first time you go to y is through A. Disconnect the network at y and now we have two points for y, y_A and y_B. We know that starting from x the probability that a random walk hits y_A before y_B is $1/V_x$ [7]. Consequently, using Kirchoff's Law and the laws for computing R_{eff} from [7] we compute that $p_A = R^{xy}/R^A$. Now, we can think of Y as the number of Bernoulli trials until the first success with probability of success p_A; i.e., Y follows a geometric distribution. Therefore $\mathbb{E}Y = \frac{1}{p_A}$. Using this result, and Eq. 1 we have the result $\mathbb{E}T^{xy}_{\overrightarrow{A}} = 2mR_A$.

2. $\mathbb{E}T^{xy}_{\overleftarrow{A}} = 2mR_A$: As above.

3. $\mathbb{E}T^{xy}_{A} = 2mR_A\frac{R_A+R_B}{2R_A+R_B}$: Here, first denote as E_1 the event that starting from x the first time you go to y is through A, and as E_2 the event that starting from y the first time you go to x is through A. Then, observe that we need the success probability of the event $E_1 \cup E_2$, i.e., $Pr(E_1 \cup E_2) = Pr(E_1) + Pr(E_2) - P(E_1 \cap E_2)$, and since the trips from x to y and from y back to x are independent, we have that $Pr(E_1 \cup E_2) = 2p_A - p_A^2$. Thus by a similar argument as the above and using again the laws for computing R_{eff} [7] and Eq. 1 we have the result.

4. $\mathbb{E}T^{xy}_{\overleftrightarrow{A}} = 2mR_A\frac{3R_A+R_B}{2R_A+R_B}$: Let Y_i be the number of tries until the first $A-$commute is achieved. Then, if at this A-commute both directions through A are always done with probability one, then $\mathbb{E}Y = \mathbb{E}Y_i$. However, there is a probability q that we will have only the one direction through A at this $A-$commute, and as a result we will need to do an additional Y_{ii} tries until we go though A in the other direction also, (i.e., $\mathbb{E}Y_{ii}$ is as in the cases of $\overrightarrow{A}$, $\overleftarrow{A}$). In other words, q is the probability that the first A-commute is not also an $\overleftrightarrow{A}$-commute, i.e., $q = Pr(\neg\overleftrightarrow{A}|A) = \frac{Pr(\neg\overleftrightarrow{A} \cap A)}{Pr(A)}$. Now, the numerator is $Pr((E_1 \cap \neg E_2) \cup (\neg E_1 \cap E_2)) = 2p_A(1 - p_A)$, while the denominator is $Pr(A) = Pr(E_1 \cup E_2) = 2p_A - p_A^2$. Thus, $q = 2(1 - p_A)/(2 - p_A) = 2R_A/(2R_A + R_B)$.

Thus, since $\mathbb{E}Y = \mathbb{E}Y_i + q\mathbb{E}Y_{ii}$ using our previous computations and Eq. 1 we have the result. $\square$

Theorem 3.6. *Let $N = (G, l)$ be an undirected network, where $m = \Sigma_{e \in E(G)} l(e)$. Then,*

$$cr_e \leq 2m^2, \; cr_a \leq 3m^2$$

- It is $cr_a \leq 2m^2$, were we have to cover just a predefined direction for each edge.

Proof. Let σ be a closed walk in G starting at r and traversing each edge once in each direction. Such a walk exists: Take a spanning tree of G and do Depth First Search and expand the walk so that it includes the remaining edges. We can divide this closed walk in epochs so that it will be easier to prove our bound. For every undirected edge we have an epoch, that is defined through two different time steps that each one concern each direction of that edge. We enumerate the time steps $\tau(a)$ with the order the σ traverses the arcs from 1 to $2|E(G)|$. For every edge e an epoch is defined using the two time steps of the two arcs a and a' that come from e; the beginning is related to the first time an edge e is visited, let this be through arc a, the end is then related to the time step of a'. If the edge is directed from x to y and we traverse it through the same direction the time step of arc (x, y) is the time we are at y using arc (x, y), differently if we traverse it through the other direction the time step of arc (y, x) is the time we are at x after the previous time step. Let $\tau(a) < \tau(a')$ for two arcs that come from the edge e, then the epoch of e can be defined as the addition of the following two pieces $[\tau(a-1), \tau(a)] + [\tau(a'-1), \tau(a')]$, this will be a commute time between the two vertices of that edge. We consider that the sub-network A consist of e and $B = G - A$. We mention that the expected time of the epoch is the expected time of an $\overleftarrow{A}$ or an $\overrightarrow{A}$ commute that depends of the direction of the predefined direction of e. The expected value will be $2ml(e)$ from Lemma 3.5. The total time of the walk σ will be no more than all the commute times: $cr_a \leq \Sigma_{e \in E(G)} 2ml(e) = 2m^2$.

$\square$

- $cr_e \leq 3m^2$, were we have to cover both directions for each edge.

Proof. Similarly with the above, we sum instead the expected time of the $\overleftrightarrow{A}$ $-$commute time for every edge. Consequently, we take the quantity $cr_e \leq \Sigma_{e \in E(G)} 2ml(e) \frac{3l(3)+R_B}{2l(e)+R_B} \leq 3m^2$. $\square$

Finally, we note that the above bounds are tight: The first is tight for a path of length m, and the second is tight for a single vertex consisting of a loop of length m.

References

[1] R. Ahlswede and D. E. Daykin, An inequality for the weights of two families of sets, their unions and intersections, *Z. Wahrsch. Verw. Gebiete* **43**, 183–185.

[2] D. Zuckerman, On the time to traverse all edges of a graph, *http://www.sciencedirect.com/science/article/pii/002001909190091U, 1991.*

[3] A. Georgakopoulos and P. Winkler, New Bounds for Edge-Cover by Random Walk, *http://arxiv.org/abs/1109.6619, 2011.*

[4] R. Durrett, Probability: Theory and Examples, *Wadworth and Brooks/Cole Advanced Books and Software, Monterey, CA, 1991*, 156.

[5] R. Aleliunas, R. M. Karp, R.J¿ Lipton, L.Lovasz, C. Rackoff, Random walks, universal traversal sequences and the complexity of maze problems, *Proc. 20th Ann. IEEE Symp. on Foundations of Computer Science(1979), 1979*, 218–233.

[6] A.K. Chandra, P.Raghavan, W. L. Ruzzo, R. Smolensky and P. Tiwari, The electrical resistance of a graph captures its commute and cover times, *Proc. 21st ACM Symp. Theory of Computing, 1989,*576–586.

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