Cover Times of Random Walks

R. Teal Witter

April 3, 2019

Abstract

abstract

Contents

1 Introduction and Background

"white screen" problem numerous applications many mathematical tools

1.1 Motivation

universal travel sequences graph connectivity protocol testing

1.2 Definitions

We consider an undirected graph G without self loops or multi-edges. Let $V(G)$ be the set of vertices and let $E(G)$ be the set of edges on G. Then n is the number of vertices $|V(G)|$.

We say $i \sim j$ for i and j in the vertex set $V(G)$ if (i, j) in $E(G)$. Let $c(i, j)$ be the non-negative weight of the edge between vertices i and j. If the graph is unweighted then $c(i, j) = 1$ if $i \sim j$.

Define the weight of vertex i

$$c(i) = \sum_{j:i\sim j} c(i, j).$$

We consider a random walk (X_t) on G. Call X_t the vertex that the walk is on at time t. The random walker begins at vertex X_0. At each step, the walker at vertex i moves to neighboring vertex j with probability $\frac{c(i,j)}{c(i)}$.

Let T_j be the number of steps until the first visit to j. Formally,

$$T_j = \min \{t : X_t = j\}.$$

Let the hitting time $\mathrm{t_{hit}}$ be the maximum expected time between two vertices on G. Formally,

$$\mathrm{t_{hit}} = \max_{i,j\in V(G)} \mathrm{E}_i[T_j].$$

Note that $\mathrm{E}_i[T_j] = \mathrm{E}[T_j|X_0 = i]$. Let the cover time $\mathrm{t_{cov}}$ be the expected time until all vertices have been visited by the random walk (X_t). Formally,

$$\mathrm{t_{cov}} = \mathrm{E}[\max_{j\in V(G)} T_j].$$

1.3 Outline

2 Cover Times of Structured Graphs

2.1 Small Example

Always possible to calculate exactly but the number of equations grows exponentially Special graph [2] full cover time from vertex 1 In symmetric cases, we can calculate it In other cases, we have to rely on bounds

2.2 Complete Graph

We call the graph with n vertices and $\binom{n}{2}$ edges the complete graph.

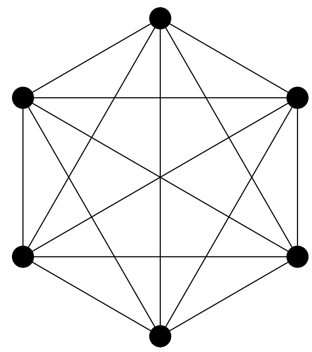

Figure 1: The complete graph with six vertices (and ten edges).

Our goal is to find the expected number of steps until we have visited all n vertices in the complete graph. Observe that the random walk can start at any vertex without loss of generality because the graph is symmetric. The strategy is to write the cover time t_{cov} in terms of the expected value of more simple random variables. Let X_i be the random variable that represents the number of steps to go from $i-1$ to i unique vertices (excluding the current vertex).

We write the cover time in terms of X_i and use linearity of expectation.

$$t_{cov} = E[X_1 + X_2 + X_3 + ... + X_{n+1}] = \sum_{i=1}^{n} E[X_i]$$

The random variable X_i takes value k when $k-1$ steps lead us to already visited vertices and the k^{th} step is to previously unvisited vertex. It follows that X_i is geometric with distribution

$$P(X_i = k) = (1-p_i)^{k-1} p_i,$$

where p_i is the probability of moving to a previously unvisited vertex. Then $E[X_i] = \frac{1}{p_i}$. We now find p_i: the probability that we go to a previously unvisited vertex is $\frac{n-i}{n-1}$. This is because there are $n-i$ unvisited vertices and a total of $n-1$ adjacent vertices. It follows that $E[X_i] = \frac{n-1}{n-i}$. Then

$$\begin{aligned} t_{cov} &= \sum_{i=1}^{n} \frac{n-1}{n-i} \\ &= (n-1)\left(\frac{1}{n-1} + \frac{1}{n-2} + ... + 1\right). \end{aligned}$$

A natural interpretation of the cover time on a complete graph is the so-called coupon collecting problem. There are r different coupons that are randomly packaged in cereal boxes. An avid fan buys a box of cereal every day. We want to know how many days until the fan has collected each coupon. The strategy is to think of every coupon as a vertex on the complete graph. We then move from vertex to vertex with uniform probability. The difference between the coupon collecting problem and the complete graph is that we now have self-edges. Before, we had to leave our current vertex at each step. Now, we can stay in the same vertex (provided our collector found the same coupon two days in a row). We apply our approach to the complete graph and substitute the r possible coupons we can find for the $n-1$ vertices we could move to. Then

$$t_{\text{cov}} = r\left(\frac{1}{r} + \frac{1}{r-1} + ... + 1\right).$$

Another more subtle application of the cover time on a complete graph is the cover time on an n-star.

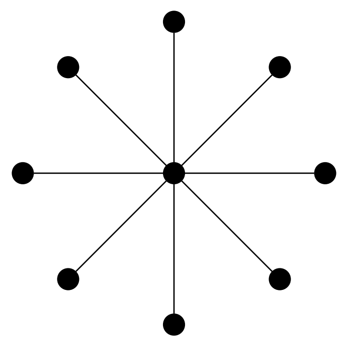

Figure 2: The nine-star.

An n-star is a graph with one central vertex and $n-1$ adjacent vertices each with a single edge to the center. A random walk begins at the central vertex and moves with uniform probability to one of the $n-1$ adjacent vertices. Whereas in the complete graph and coupon collecting problems we restarted in one step, a random walk on the n-star probabilistically restarts in two steps: one step to each leaf and one step back. We apply our approach to the complete graph and conclude that

$$t_{\text{cov}} = 2(n-1)\left(\frac{1}{n-1} + \frac{1}{n-2} + ... + 1\right).$$

2.3 Simple Cycle

A simple cycle is a connected graph where each vertex has an edge to exactly two other vertices. (Fig. 4 is an example.)

Our goal is to determine the cover time of a simple cycle. However, we begin by considering the cover time of a subset of the integers on the number line from 0 to N. We then use our analysis to build our way up to a simple cycle.

Consider a random walk on the finite linear graph. (For the finite linear graph in Fig. 3, note that $n = N+1$.) From $X_t = i$, we move with equal probability to $i+1$ and $i-1$ until we reach either 0 or N.

Call e_i the expected number of steps until we reach either end of our random walk from state i. Our strategy is to write and solve a recurrence relation that uses what we know about the random walk. If $i = 0, N$, we are already at the end point and $e_i = 0$.

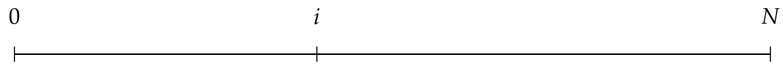

Figure 3: A finite linear graph with a vertex at each integer from 0 to N.

Otherwise, we take a step. There's a half probability our step is to the right and a half probability it is to the left.

$$\begin{aligned}
e_i &= 1 + \frac{1}{2}\left(e_{i+1}\right) + \frac{1}{2}\left(e_{i-1}\right) \\
2e_i &= 2 + e_{i+1} - e_{i-1} \\
(e_{i+1} - e_i) &= (e_i - e_{i-1}) - 2
\end{aligned}$$

While the final line may look unnecessarily complicated, notice that the structure of the equality lends itself to telescoping.

To solve the telescoping equations, notice that the walk is symmetric: we could flip the walk over the vertical axis without changing e_i. This makes sense because our intuition tells us that the only identifying feature of i is its respective distances to the endpoints. The logical conclusion is that $e_i = e_{N-i}$. In particular, we have already seen that $e_0 = e_N = 0$ and it naturally follows that $e_1 = e_{N_1}$. We use these observations to find e_1.

$$\begin{aligned}
(e_2 - e_1) &= (e_1 - e_0) - 2 \\
(e_3 - e_2) &= (e_2 - e_1) - 2 = (e_1 - e_0) - 4 \\
(e_4 - e_3) &= (e_3 - e_2) - 2 = (e_1 - e_0) - 6 \\
&\vdots \\
(e_N - e_{N-1}) &= (e_1 - e_0) - 2(N-1) \\
(0 - e_1) &= (e_1 - 0) - 2(N-1) \\
e_1 &= N - 1.
\end{aligned}$$

(Note that the vertical dots denote inductive reasoning.) We now know e_1. However, our goal is to find the expected number of steps from an arbitrary point i. We use e_1 to find the general solution.

$$\begin{aligned}
e_2 &= e_1 + (e_1 - e_0) - 2 = 2(N-2) \\
e_3 &= e_2 + (e_2 - e_1) - 2 = 2\left[2(N-2)\right] - (N-1) - 2 = 3(N-3) \\
&\vdots \\
e_i &= i(N-i) \qquad (1)
\end{aligned}$$

Theorem 1 follows from Eq. (1) and the observation that any linear finite graph can be shifted so that the left end point is at 0 and the right end point is at N.

Theorem 1. *The expected number of steps from a state to either end of a finite linear graph is the product of the distances from that state to each end point.*

We now return to cover times. Consider a random walk on a simple cycle as in Fig. 4. (Note that the starting vertex is arbitrary since the graph is symmetric.) Think of (X_t) as the walk on the number line (between $-n$ and n) where $X_0 = 0$. The cover time t_{cov} of the simple cycle is the expected number of steps until we visit n distinct integers.

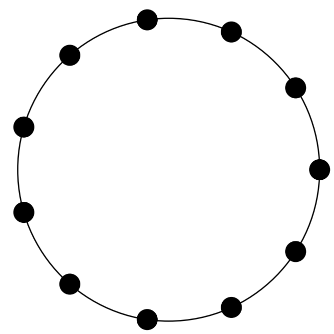

Figure 4: A cycle graph with $n = 11$ vertices.

Let $\{a, a+1, \ldots, b-1, b\}$ denote the set of vertices we have visited. Since the walk is on a line, we can write $[a, b]$ for the range of the walk where a is the smallest and b is the largest vertex we have visited. Notice that if $a \leq -n$ or $b \geq n$, we have covered all n vertices on the simple cycle and conclude the random walk.

In a slight abuse of notation, let T_k be the number of steps until we reach k distinct vertices. Formally, $T_k = \min t : b - a + 1 = k$. Then by telescoping,

$$t_{cov} = \mathrm{E}[T_n] = \mathrm{E}[(T_2 - T_1) + (T_3 - T_2) + \cdots + (T_k - T_{k-1}) + \cdots + (T_n - T_{n-1})]$$

At time T_k, we are either at integer a or b. Then $\mathrm{E}[T_{k+1} - T_k]$ is the expected time until we reach $a-1$ or $b+1$. Notice that we can shift our end points to 0 and $b+1-(a-1) = b-a+2 = k+1$. The expected number of steps to either $a-1$ or $b+1$ is $1(k+1-1) = k$ by Theorem 1.

By linearity of expectation,

$$\begin{aligned} t_{cov} &= (\mathrm{E}[T_2] - \mathrm{E}[T_1]) + \cdots + (\mathrm{E}[T_{k+1}] - \mathrm{E}[T_k]) + \cdots + (\mathrm{E}[T_n] - \mathrm{E}[T_{n-1}]) \\ &= 1 + 2 + \cdots + k - 1 + \cdots + n - 1 = \frac{1}{2}n(n-1). \end{aligned}$$

2.4 Finite Linear Graph

We have already seen a finite linear graph. (For the finite linear graph in Fig. 3, note that $n = N + 1$). Theorem 1 tells us the expected time from an arbitrary point to one of the end points but we do not yet know the cover time.

Observe that the cover time is the earliest time that both end points have been visited. We can use our earlier result to find the time to one of the end points from an arbitrary position but we still need to find the expected time from that end point to the other.

Consider $\mathrm{E}_i[T_N]$ the expected time to N from vertex i. We want to find $\mathrm{E}_0[T_N]$. By symmetry, $\mathrm{E}_0[T_N] = \mathrm{E}_N[T_0]$.

3 Bounds

3.1 Matthews Method

$\mathrm{t_{hit}} \leq \mathrm{t_{cov}}$

Theorem 2. *Let (X_t) be a random walk on a graph with n vertices. Then*

$$\mathrm{t_{cov}} \leq \mathrm{t_{hit}} \left(1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n-1}\right).$$

Proof. Without loss of generality, we may assume that the walk started at vertex n. Let σ be a uniform permutation of the unvisited $n-1$ vertices. The strategy is to look for the vertices in order σ. We call σ_i the i^{th} vertex we visit. Let T_k be the first time that vertices $\sigma_1, \sigma_2, ..., \sigma_k$ have all been visited. For $1 \leq s \leq n-1$,

$$\mathrm{E}_n[T_1 | \sigma_1 = s] = t_{n,s} \leq \mathrm{t_{hit}}$$

□

3.2 Spanning Tree Argument

4 Distributional Aspects

4.1 Complete Graph

5 Electrical Networks

5.1 Definitions

5.2 Commute Times

5.3 Balanced Trees

6 Simulation

[1] [2] [3] [4] [5] [6] [7]

References

[1] David Aldous and James Allen Fill. *Reversible Markov Chains and Random Walks on Graphs*. 2002.

[2] Gunnar Blom, Lars Holst, and Dennis Sandell. *Problems and Snapshots from the World of Probability*. Springer, 1994.

[3] Béla Bollobás. *Modern Graph Theory*. Springer, 1998.

[4] Peter G. Doyle and J. Laurie Snell. *Random Walks and Electric Networks*. Mathematical Assn of Amer, 1984.

[5] Rick Durrett. *Probability: Theory and Examples*. Cambridge University Press, 2011.

[6] Geoffrey Grimmett. *Probability on Graphs: Random Processes on Graphs and Lattices*. Cambridge University Press, 2010.

[7] David A. Levin and Yuval Peres. *Markov Chains and Mixing Times*. Cambridge University Press, 2017.