Cover Times of Random Walks

R. Teal Witter

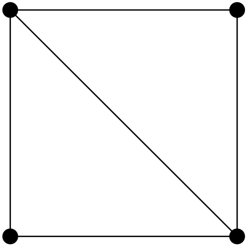
Middlebury College

May 9, 2019

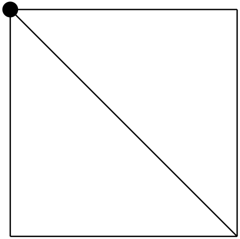
Overview

Overview

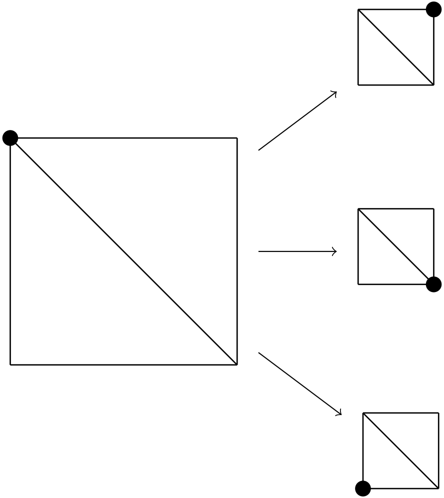
A Graph



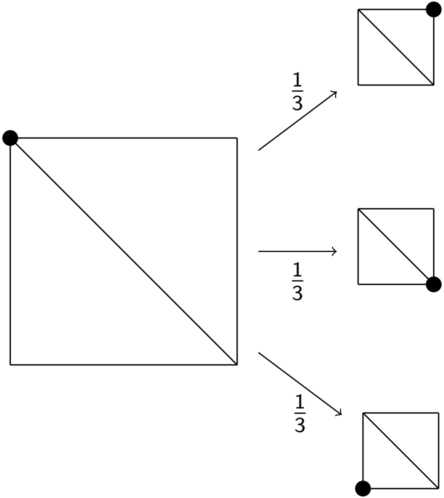
Random Walks



Random Walks



Random Walks



Cover Times

How many steps does a random walk take to visit every vertex?

Coupon Collector

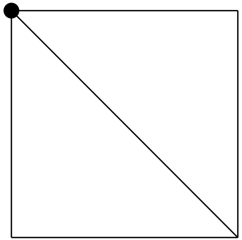
A collector receives one of n possible coupons in the mail everyday. How long until all n coupons have been collected?

Birthday Problem

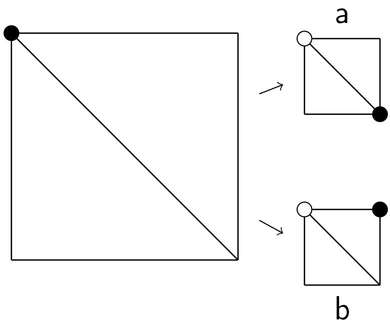
How many people do we need for each birthday (excluding February 29) to be represented?

Overview

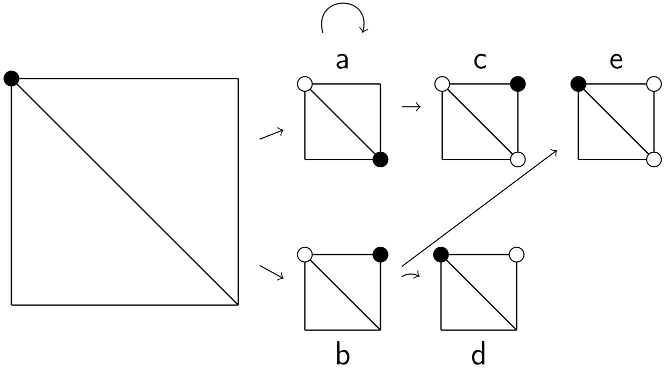
A Small Graph



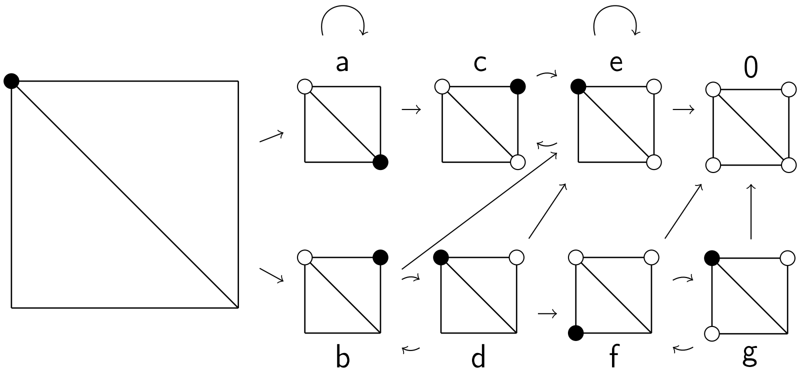
A Small Graph



A Small Graph



A Small Graph



A Small Graph

$$
\begin{aligned}
\mathrm{E}[\mathsf{t}_{\mathsf{cov}}] &= 1 + \frac{1}{3}a + \frac{2}{3}b \\
a &= 1 + \frac{1}{3}a + \frac{2}{3}c \\
b &= 1 + \frac{1}{2}b + \frac{1}{2}e \\
c &= 1 + 1e \\
d &= 1 + \frac{1}{3}a + \frac{1}{3}e + \frac{1}{3}f \\
e &= 1 + \frac{1}{3}c + \frac{1}{3}e + \frac{1}{3}0 \\
f &= 1 + \frac{1}{2}g + \frac{1}{2}0 \\
g &= 1 + \frac{2}{3}f + \frac{1}{3}0
\end{aligned}
$$

A Small Graph

$$
\begin{aligned}
\mathrm{E}[\mathsf{t}_{\mathsf{cov}}] &= 1 + \frac{1}{3}\mathrm{a} + \frac{2}{3}\mathrm{b} = \frac{43}{6} \\
\mathrm{a} &= 1 + \frac{1}{3}\mathrm{a} + \frac{2}{3}\mathrm{c} \\
\mathrm{b} &= 1 + \frac{1}{2}\mathrm{b} + \frac{1}{2}\mathrm{e} \\
\mathrm{c} &= 1 + 1\mathrm{e} \\
\mathrm{d} &= 1 + \frac{1}{3}\mathrm{a} + \frac{1}{3}\mathrm{e} + \frac{1}{3}\mathrm{f} \\
\mathrm{e} &= 1 + \frac{1}{3}\mathrm{c} + \frac{1}{3}\mathrm{e} + \frac{1}{3}0 \\
\mathrm{f} &= 1 + \frac{1}{2}\mathrm{g} + \frac{1}{2}0 \\
\mathrm{g} &= 1 + \frac{2}{3}\mathrm{f} + \frac{1}{3}0
\end{aligned}
$$

Overview

Complete Graph with Self-edges

Let X_i be the number of steps from $i-1$ to i distinct vertices.

Complete Graph with Self-edges

Let X_i be the number of steps from $i-1$ to i distinct vertices.

$$\mathrm{E}[\mathsf{t}_{\mathsf{cov}}] = \mathrm{E}[X_1 + X_2 + X_3 + \cdots + X_n]$$

Complete Graph with Self-edges

Let X_i be the number of steps from $i-1$ to i distinct vertices.

$$\mathrm{E}[t_{cov}] = \mathrm{E}[X_1 + X_2 + X_3 + \cdots + X_n]$$

$X_i \sim \mathrm{Geom}(\frac{n-i+1}{n})$

Complete Graph with Self-edges

Let X_i be the number of steps from $i-1$ to i distinct vertices.

$$\mathrm{E}[\mathsf{t}_{\mathsf{cov}}] = \mathrm{E}[X_1 + X_2 + X_3 + \cdots + X_n]$$

$X_i \sim \mathrm{Geom}(\frac{n-i+1}{n})$ so $\mathrm{E}[X_i] = \frac{n}{n-i+1}$

Complete Graph with Self-edges

Let X_i be the number of steps from $i-1$ to i distinct vertices.

$$\mathrm{E}[\mathsf{t}_{\mathsf{cov}}] = \mathrm{E}[X_1 + X_2 + X_3 + \cdots + X_n]$$

$X_i \sim \mathrm{Geom}(\frac{n-i+1}{n})$ so $\mathrm{E}[X_i] = \frac{n}{n-i+1}$ and

$$\mathrm{E}[\mathsf{t}_{\mathsf{cov}}] = n\left(\frac{1}{n} + \frac{1}{n-1} + \cdots + 1\right)$$

Coupon Collector

A collector receives one of n possible coupons in the mail everyday. How long until all coupons have been collected?

On average, we need about $n \log n$ days.

Birthday Problem

How many people do we need for each birthday (excluding February 29) to be represented?

On average, we need about 2365 people.

Overview

Matthews Method

Theorem

Consider a random walk on an arbitrary connected graph with n vertices. Then

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

Matthews Method Proof

Let σ be a uniform permutation of unvisited vertices from (worst-case) vertex n.

Matthews Method Proof

Let σ be a uniform permutation of unvisited vertices from (worst-case) vertex n.
Let T_k be the first time vertices $\sigma_1, \sigma_2, \ldots, \sigma_k$ have all been visited.

Matthews Method Proof

Let σ be a uniform permutation of unvisited vertices from (worst-case) vertex n.
Let T_k be the first time vertices $\sigma_1, \sigma_2, \ldots, \sigma_k$ have all been visited.

$$\begin{aligned} \mathrm{E}[\mathrm{t}_{\mathrm{cov}}] &= \mathrm{E}_n[T_{n-1}] \\ &\le \mathrm{E}_n[T_1 + (T_2 - T_1) + \cdots + (T_{n-1} - T_{n-2})] \end{aligned}$$

Matthews Method Proof

Let $L(k)$ be the final vertex to be visited among $\sigma_1, \sigma_2, \ldots, \sigma_k$ before the first time they are all visited.

Matthews Method Proof

Let $L(k)$ be the final vertex to be visited among $\sigma_1, \sigma_2, \ldots, \sigma_k$ before the first time they are all visited.

$$\mathrm{E}_n[T_k - T_{k-1}] = \sum_{i=1}^{k} \mathrm{E}_n[T_k - T_{k-1} | L(k) = \sigma_i] \, \mathrm{P}(L_k = \sigma_i)$$

Matthews Method Proof

Let $L(k)$ be the final vertex to be visited among $\sigma_1, \sigma_2, \ldots, \sigma_k$ before the first time they are all visited.

$$\begin{aligned} \mathrm{E}_n[T_k - T_{k-1}] &= \sum_{i=1}^{k} \mathrm{E}_n[T_k - T_{k-1} | L(k) = \sigma_i] \, \mathrm{P}(L_k = \sigma_i) \\ &= 0 + \mathrm{E}_n[T_k - T_{k-1} | L(k) = \sigma_k] \, \mathrm{P}(L_k = \sigma_k) \end{aligned}$$

Matthews Method Proof

Let $L(k)$ be the final vertex to be visited among $\sigma_1, \sigma_2, \ldots, \sigma_k$ before the first time they are all visited.

$$\begin{aligned}
\mathrm{E}_n[T_k - T_{k-1}] &= \sum_{i=1}^{k} \mathrm{E}_n[T_k - T_{k-1} | L(k) = \sigma_i] \, \mathrm{P}(L_k = \sigma_i) \\
&= 0 + \mathrm{E}_n[T_k - T_{k-1} | L(k) = \sigma_k] \, \mathrm{P}(L_k = \sigma_k) \\
&\leq t_{hit} \frac{1}{k}
\end{aligned}$$

Overview

Cover Time Distribution

After r steps of a random walk, let A_i be the event that vertex i is unvisited.

$$\mathrm{P}(A_i) = \mathrm{P}(\text{not visited in one step})^r = (1 - 1/n)^r$$

Cover Time Distribution

After r steps of a random walk, let A_i be the event that vertex i is unvisited.

$$\mathrm{P}(A_i) = \mathrm{P}(\text{not visited in one step})^r = (1 - 1/n)^r$$

Let N_n be the number of unvisited vertices.

$$\begin{aligned} \mathrm{E}[N_n] &= \sum_{i=1}^{n} \mathrm{P}(A_i) = n(1 - 1/n)^r \\ &= n[(1 - 1/n)^n]^{r/n} \sim ne^{-r/n} \end{aligned}$$

Cover Time Distribution

Introduce x and set $r = n \log n + nx$.

$$N_n \sim ne^{-r/n} = ne^{-\log n - x} = nn^{-1}e^{-x} \to e^{-x}$$

Theorem

If $ne^{-r/n} \to \lambda \in [0, \infty)$, then the number of unvisited vertices approaches a Poisson distribution with mean λ.

Cover Time Distribution

Introduce x and set $r = n \log n + nx$.

$$N_n \sim ne^{-r/n} = ne^{-\log n - x} = nn^{-1}e^{-x} \to e^{-x}$$

Theorem

If $ne^{-r/n} \to \lambda \in [0, \infty)$, then the number of unvisited vertices approaches a Poisson distribution with mean λ.

$$\mathrm{P}(N_n = 0) = \mathrm{P}(\mathrm{t_{cov}} \leq r) = \mathrm{P}(\mathrm{t_{cov}} - n \log n \leq nx) \to e^{-e^{-x}}$$

Birthday Problem

What is the probability all birthdays are covered in a town of 1825? How about a town of 2190?

Birthday Problem

What is the probability all birthdays are covered in a town of 1825? How about a town of 2190?

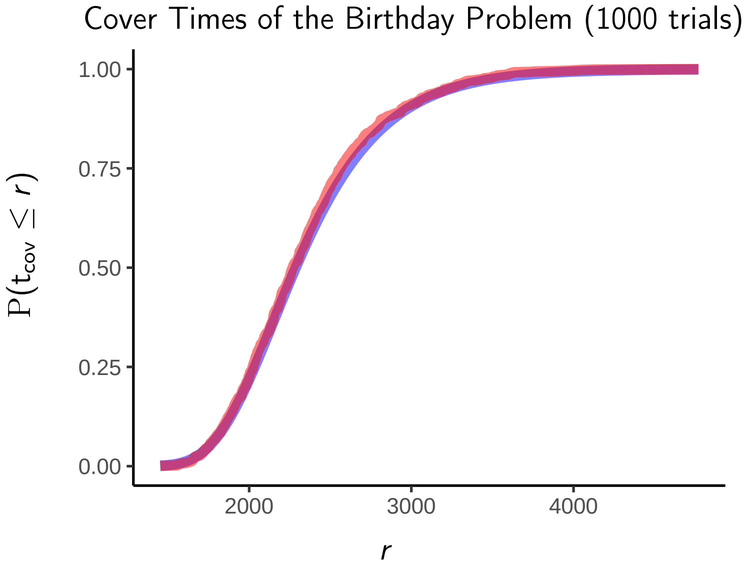
$$\mathrm{P}(\mathrm{t}_{\mathrm{cov}} \leq 1825) = \mathrm{P}\left(\frac{\mathrm{t}_{\mathrm{cov}} - 2153}{365} \leq \frac{-328}{365}\right)$$
$$\approx e^{-e^{-0.89863}} \approx 0.08576$$

Birthday Problem

What is the probability all birthdays are covered in a town of 1825? How about a town of 2190?

$$
\begin{aligned}
\mathrm{P}(t_{\mathrm{cov}} \leq 1825) &= \mathrm{P}\left(\frac{t_{\mathrm{cov}} - 2153}{365} \leq \frac{-328}{365}\right) \\
&\approx e^{-e^{-0.89863}} \approx 0.08576 \\
\mathrm{P}(t_{\mathrm{cov}} \leq 2190) &= \mathrm{P}\left(\frac{t_{\mathrm{cov}} - 2153}{365} \leq \frac{37}{365}\right) \\
&\approx e^{-e^{-0.10137}} \approx 0.40511
\end{aligned}
$$

Birthday Problem



Thank you!

github.com/rtealw/Cover-Times

- Aldous, D. and J. Fill. 2014. *Reversible Markov Chains and Random Walks on Graphs*.
- Blom, G., L. Holst and D. Sandell. *Problems and Snapshots from the World of Probability*, Springer Science & Business Media, 1994.
- Doyle, P. G. and J. L. Snell. 1984. *Random Walks and Electric Networks*, Mathematical Association of America.
- Levin, D. A. and Y. Peres. 2017. *Markov Chains and Mixing Times*, Vol. 107, American Mathematical Society.