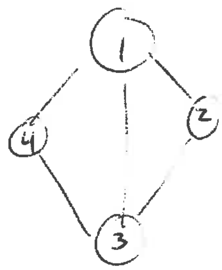
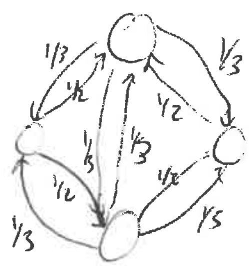
Lecture 23: Random Walks on Graphs II: Undirected Graphs

Recall: In a strongly connected Markov Chain with n states...

- $K[i,j] = \Pr[j \to i]$
- Invariant distribution π satisfies $K\pi = \pi$
- Thm: $E[\#\text{ steps to get back to } u \mid \text{starting at } u] = \frac{1}{\pi(u)}$.

Say G is a connected undirected graph, n vertices, m edges, d_i = degree of vtx i

e.g.: 

"standard random walk on G" is M.C.: 

[at each vertex, go to a random neighbor]

Adjacency matrix $A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$ (symmetric), $K = \begin{pmatrix} 0 & 1/2 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 0 \\ 1/3 & 1/2 & 0 & 1/3 \\ 1/3 & 0 & 1/3 & 0 \end{pmatrix}$

$\div$ by d_1, $\div d_2$, $\div d_3$, $\div d_4$

[not symmetric, unless it's a regular graph, all d_i's the same]

What is π? If no "stupid periodicity" [for undirected graphs $\equiv$ bipartite] [rand walk alternates between L & R sides], same as "limiting distribution".

[Higher degree $\equiv$ higher limiting probability

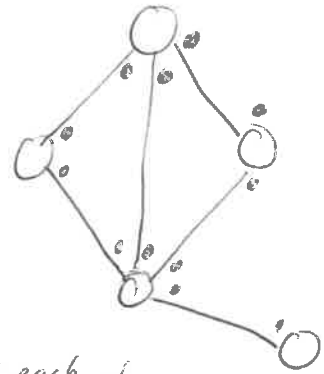
Could $\pi[u]$ just be proportional to degree?!

Yes, it's that easy!]

Theorem: $\pi = \begin{bmatrix} d_1/2m \\ d_2/2m \\ \vdots \\ d_n/2m \end{bmatrix}$ (check: $\sum_{i=1}^{n} d_i = 2m$)

why?

[[give each edge 2 tokens, one for each endpoint]]



$2m$ tokens $\leftarrow$ [[but each vtx has one token per degree]]

$\parallel$

Σd_i tokens

Proof: Need to show $K\pi = \pi$. [[Can just calculate, or ...]]

$\equiv$ if you start at vertex i with prob. $d_i/2m$, then walk to random nbr, then prob. you're at j is $d_j/2m$.

$\equiv$ { go sit on a random token, slide to token's twin }

equiv to going to a random token!

Cor: In standard rand walk on G, let $M_{vv} = E[\#\text{ steps to hit } v\text{, starting from } v] = \frac{2m}{v}$

[[Mean 1st Recurrence thm.]]

[[Kinda fun!!]] e.g.'s:

$m = 2, \quad 2m = 4$

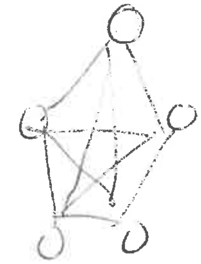
π:	$\frac{1}{4}$	$\frac{2}{4}$	$1/4$
M_{vv}	4	2	4

[[do you see why?]]

$n+1$ nodes, $m = n$

π:	$\frac{1}{2n}$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	...	$\frac{1}{n}$	$\frac{1}{2n}$
M_{vv}	$2n$	n	n	n	...	n	$2n$

Clique of n nodes:



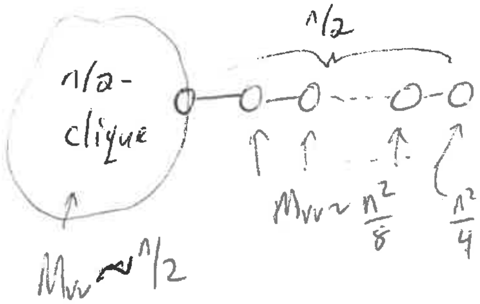
$$m = \binom{n}{2} = \frac{n(n-1)}{2}$$

$$d_v = n-1 \quad \forall v$$

$$\frac{d_v}{2m} = \frac{1}{n} \Longrightarrow \pi = \begin{bmatrix} 1/n \\ 1/n \\ \vdots \\ 1/n \end{bmatrix}, \quad M_{vv} = n \quad \forall n$$

[do you see why?]

Lollipop of n nodes:



$$m \approx \frac{(n/2)^2}{2} + \frac{n}{2} = \frac{n^2}{8} \pm O(n)$$

[Takes quadratic time, on average, to get out of the clique part]

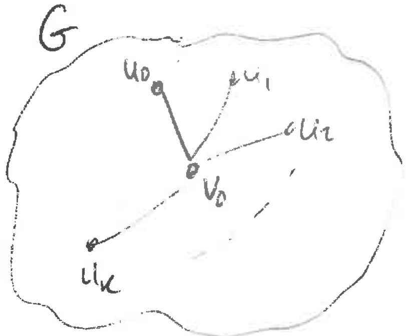
What about $M_{u,v} := E[\#\text{ steps to } \cancel{\text{get}} \text{ to } v, \text{ starting from } u]$?

[No easy formula. But we'll prove an upper-bound.]

Prop: Let (u_0, v_0) be an edge. [So v_0 is a nbr of u_0]
Then $M_{u_0,v_0} \le 2m-1 \le 2m$. [Seems weak, since you might go $u_0 \to v_0$ in one step! But it's actually tight.]

Proof:

G



Say v_0's neighbors are $u_0, u_1, \ldots, u_k$.

$$\frac{2m}{d_{v_0}} = E[\#\text{ steps } v_0 \to v_0]$$

$$= \sum_{i=0}^{k} \Pr[\text{first step } v_0 \to u_i] \cdot E[\#\text{ step } v_0 \to v_0 \mid \text{first step is } v_0 \to u_i]$$

$$= \sum_{i=0}^{k} \frac{1}{d_{v_0}} \left(1 + E[\#\text{ steps } u_i \to v_0]\right)$$

[drop all terms except $i=0$!]

$$\geq \frac{1}{d_{v_0}} \left(1 + E[\#\text{ steps } u_0 \to v_0]\right)$$

Now multip. both sides by d_{v_0}, subtract 1. $\square$

<u>Thm:</u> Let u, v be <u>any</u> two vertices.

Then $M_{uv} \le 2mn \le n^3$ $\quad \left(\because m \le \binom{n}{2} \le \frac{n^2}{2}\right)$

<u>Proof:</u> Pick a path $u, w_1, w_2, \dots, w_r, v$. At most n nodes

$$E[\#\text{ steps } u \to v] \le E[\#\text{ steps to go } u \to w_1 \to w_2 \to \cdots \to w_r \to v]$$
$$= E[\#\, u \to w_1] + E[\#\, w_1 \to w_2] + \cdots + E[\#\, w_r \to v]$$
$$\le 2m + 2m + \cdots + 2m$$
$$= 2mn. \qquad \square$$

[$\because$ $(u, w_1), (w_1, w_2), \dots$ all edges, we can use prev. prop.]

e.g. Path:

n nodes

$$E[\#\text{ step to hit } v, \text{ starting at } u] \le 2mn = 2(n-1)n = \Theta(n^2).$$

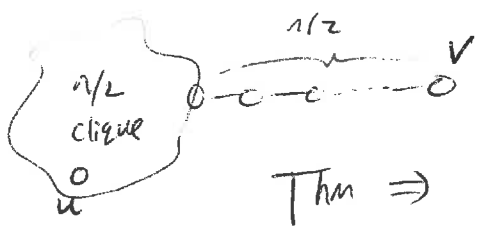
fact: Indeed, it's $\Omega(n^2)$. [! Kind of surprising takes quadratic time on average to walk the line.]

e.g. n-Clique

Thm $\Rightarrow$ $E[\#\text{ steps to get } u \to v] \le 2mn \le n^3$.

Actually: $\underline{\quad n-1 \quad}$. It's a "geometric" random var. with success prob. $\frac{1}{n-1}$.

e.g. lollipop:



Thm $\Rightarrow$ $M_{uv} \le n^3$

fact: It's indeed $\Omega(n^3)$.

[So cubic time to go anywhere $\to$ anywhere in a graph and that's tight!]

Application: [[Kind of dubious, but introduces some interesting ideas]]

Problem: Given undirected G, possibly disconnected, & $u, v \in V$:
~~...~~ Is there a path $u \rightsquigarrow v$? [[i.e., are they in same connected component]]

Solution: DFS/BFS: $O(m)$ time. $\square$

Difficulty: Make an algorithm Where: • input G is in read-only memory
• you may not alloc. any memory
(• local vars ok) [["low space/memory alg."]]

BFS/DFS require "marking" nodes → allocating 1 bit per node.

[[So with our restriction, you cannot keep track of where you've been, even!]]

[[As is so often the case, randomness to the rescue!]]

```
z := u
for t = 1....1000n^3
    z := random neighbor of z.
    if z=v return TRUE
return FALSE
```

2 local vars!

$O(n^3)$ time [[poly-time!]]

[[This alg is randomized, so it may fail, but...]]

Case 1: u & v are not connected.

Then Pr[Alg correctly says "FALSE"] = 100%

Case 2: u & v are connected.

Let T be # of steps it takes alg to go $u \rightsquigarrow v$ (if run indefinitely)

Thm says $E[T] \le n^3$.

$\therefore$ Markov's ineq $\Rightarrow \Pr[T > 1000n^3] \le .001 = .1\%$.

$\therefore \Pr[\text{Alg correctly says "TRUE"}] \ge 99.9\%$ ☺

Problem open for 25 years: Is there a poly(n)-time, $O(1)$ local vars, deterministic alg? Ans [Reingold'04]: Yes!