CS 710: Complexity Theory Date: Feb. 6th, 2024

Lecture 5: Random Walk for USTCONN

Instructor: Jin-Yi Cai Scribe: Hao Lin

1 Introduction

We have talked about the **GAP** (directed Graph Accessibility Problem) which can be solved by NDTM using $\log n$ space, i.e. GAP $\in$ NL. It is somehow trivial that the undirected version, which is named as **USTCONN**, should be no more difficult than GAP (as undirected graph can be turn into a directed graph).

This raises a question: can we solve this problem more aggressively, in deterministic log space, i.e. USTCONN $\in$ L? By saying "aggressively", it is based on the assumption that L $\subsetneq$ NL whilst if one can prove L $=$ NL, then GAP and USTCONN are of the same difficulty to solve in terms of space complexity.

Or, let us take one step back: can we solve it using deterministic log space, but with randomized methods to approximate it? We will show that the answer is yes with one-sided error. The intermediate results are charming as well.

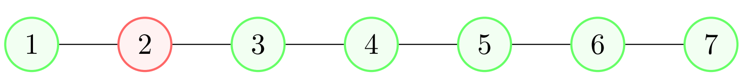
2 Random Walk

Definition 1. We focus on the symmetric random walk on an undirected graph $G = (V, E)$, to be a stochastic path $\{Y_i\}_{i \in \mathbb{N}^+}$, where $Y_i = j$ is a random variable representing node j is visited at i-th step, with the transition probability

$$Pr[Y_{i+1}|Y_i] = \begin{cases} \frac{1}{\deg(Y_i)} & (Y_i, Y_{i+1}) \in E \\ 0 & \text{otherwise} \end{cases}$$

Let us begin with a one-dimensional finite chain. We start at node 1, and the random walk stops at node n.

Example: For this specific example, $n = 5$, and suppose we are now at node 2, so it has $p = 0.5$ probability to reach node 1 and $1 - p = 0.5$ to reach 3 as well.



$\boxtimes$

2.1 Hitting Time for the One-dimensional Finite Chain

Definition 2. T_i is the number of steps spent at node i

$$T_i = \sum_{j=0}^{\infty} X_{ij}, X_{ij} = \begin{cases} 1 & Y_j = i \\ 0 & \text{otherwise} \end{cases}$$

For simplicity, we omit the constraints on the length should be no smaller than the subscription j, and similar for the rest in this scribe.

Proposition 1. The expectation t_i, for $2 \le i \le n-2$,

$$\begin{aligned} t_i = \mathbb{E}[T_i] = \sum_{j=0}^{\infty} X_{ij} &= \sum_{j \ge 0} Pr[Y_j = i] \\ &= \sum_{j \ge 1} Pr[Y_{j-1} = i-1] \cdot Pr[Y_j = i | Y_{j-1} = i-1] \\ &+ Pr[Y_{j-1} = i+1] \cdot Pr[Y_j = i | Y_{j-1} = i+1] \end{aligned}$$

For $3 \le i \le n-2$,

$$\begin{aligned} t_i =& \frac{1}{2}\Big[\sum_{j' \ge 0} Pr[Y_j' = i-1] \\ &+ \sum_{j'' \ge 0} Pr[Y_j'' = i+1] \Big] \\ =& \frac{1}{2}(t_{i-1} + t_{i+1}) \end{aligned}$$

Therefore, we have $t_i = \begin{cases} t_1 = n-1 \\ t_i = 2(n-i) & 2 \le i \le n-1 \\ t_n = 1 \end{cases}$

Claim 1. The sum of expectation of the number of steps spent at each node is the expected hitting time from node 1 to node n with the random walk.

We will show this after we defining the *hitting time* as follows and show its equivalence.

Definition 3. $Y_{i,j}$ is a random variable that equals to the number of steps for a random walk path $\mathcal{C}$, from node i to j.

$h_{i,j} = \mathbb{E}[Y_{i,j}]$, which is called the hitting time for a random walk from node i to j.

Theorem 1. For a one-dimensional finite chain, the hitting time from node 1 to node n is $h_{1,n} = (n-1)^2$.

Proof. First, by definition, we have $Y_{i-1,i+1} = Y_{i-1,i} + Y_{i,i+1}, \forall 1 < i < n$, then we can deduce that $Y_{1,n} = Y_{1,2} + Y_{2,3} + \cdots + Y_{n-1,n}$.

By linearity of expectation, we have $h_{1,n} = h_{1,2} + h2, 3 + \cdots + hn - 1, n$.

By definition, we have $h_{i,i+1} = \frac{1}{2} \cdot 1 + \frac{1}{2}(1 + h_{i-1,i+1}), \forall 1 < i < n$, since when we are on node i, we have two choices with equal probability: moving right to $i+1$, or moving left to $i-1$.

We also have $h_{i-1,i+1} = h_{i-1,i} + h_{i,i+1}$. Combining these two equations, we have $h_{i,i+1} = hi - 1, i + 2$. Since $h_{1,2} = 1$, we have $h_{1,n} = (n-1)^2$. □

Here is the proof of the claim.

Proof. The above claim is easily proved by adding t_i up and gives to $(n-1)^2$. □

We can see that the expected time traveling between two nodes, can be calculated via a direct expansion of definition, or by the summation of the visiting number of each node. These results are highly dependent on the assumption that each step's choice Y_i in the random walk is identically independent distributed (i.i.d).

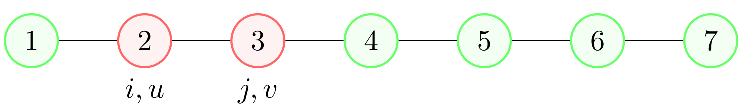
2.2 Random Walk on *ij-commute* Path

Definition 4. A path $\mathcal{C} = Y_i$ is called *ij-commute* if and only if for $i \neq j$, $Y_1 = i$ and there exists an index $k < n$ such that $Y_k = j$ and $Y_k' \neq j, \forall k' < k$, and beyond this, $Y_n = i$ with $Y_k'' \neq i, \forall k'' < n \wedge k'' > k'$.

Definition 5. For a random walk, C_{ijuv} is a random variable that counts the number of existence of directed edge $\overrightarrow{uv}$ in an *ij-commute* path $\mathcal{C}$.

$\theta_{ijuv} = \mathbb{E}[C_{ijuv}]$

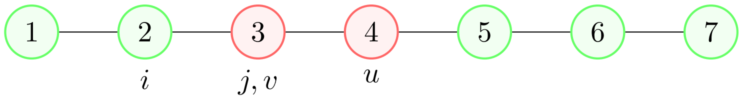
Corollary 1. θ_{ijij} in the one-dimensional finite chain is 1 if $(i,j) \in E$.



Proof. Since edge $(i,j) \in E$, there must exist $1 < k < n$ such that $Y_k = j$ in this *ij-commmute* path $\mathcal{C}$, and $Y_{k-1} = i$. We know that $\theta_{ijij} \geq 1$.

Now, let us prove it by contradiction. Suppose $\theta_{ijij} > 1$. This means that there exists these indices $1 < k < k' < k'+1 < n$, such that $Y_{k-1} = i$, $Y_k = j$, $Y_{k'} = i$, $Y_{k'+1} = j$. However, by definition of *ij-commute*, the path should stop at k', but is less than n, which is a contradiction. Therefore, $\theta ijij = 1$. □

Corollary 2. θ_{ijuv} in the one-dimensional finite chain is 1 if $u = v + 1$, $v = j = i + 1$.



Proof. By definition, $\theta_{ijuv} = Pr[\{v, u\} \text{ is crossed}] \cdot \mathbb{E}[C_{ijuv} | \{v, u\} \text{ is crossed}]$.

Since j must be reached, the former probability goes to $\frac{1}{2}$. The latter expectation is exactly the same regardless of how many times $\{v, u\}$ has been crossed (at least once). Each time $\{v, u\}$ is crossed, it has to pass the other direction $\{u, v\}$ once, so if we denoted this expectation as x, we have a 'fixed point' equation $x = 1 + \frac{1}{2}x$, so $x = 2$. Therefore, $\theta_{ijuv} = 1$.

$\square$

This gives an interesting guess that for a one-dimensional finite chain, will $\theta_{ijuv} = 1, \forall i, j, u, v$? Or even further, does this hold for all graphs? The answers are given in the next lecture scribe and we will see these results gives birth to a randomized log space algorithm for the USTCONN problem.