Some Math and Randomized Algorithms for the s-t Connectivity Problem[1]

Fei Li

January 30, 2012

[1] Some from Chapter 3 and Chapter 7 of "Probability and Computing: Randomized Algorithms and Probabilistic Analysis" by Michael Mitzenmacher and Eli Upfal and some from "An introduction to information theory and entropy" by Tom Carter

s-t Connectivity Problem

Given an undirected graph $G = (V,\ E)$ and two vertices s and t in G, determine if there is a **path connecting s and t**.

One Answer

Let $n = |V|$ and $m = |E|$. Standard BFS and DFS are deterministic algorithms **linear running time** $O(n + m)$.

Such algorithms **require** $\Omega(n)$ **space**.

A New Target

Develop a **randomized algorithm** that works with only $O(\log n)$ **bits of memory**.

$O(\log n)$ is even less than the number of bits required to write the path between s and t.

Math Background: Markov's Inequality

A Motivating Question

How do we bound the *tail distribution*, the probability that a random variable assumes values that are far from its expectation?

Theorem (Markov's Inequality)

Let X be a random variable that assumes only nonnegative values. Then, for all $a > 0$,

$$Pr[X \geq a] \leq \frac{E[X]}{a}.$$

Proof.

For $a > 0$, let

$$I = \begin{cases} 1, & \text{if } X \geq a, \\ 0, & \text{otherwise,} \end{cases}$$

and note that since $X \geq 0$,

$$(*)\ I \leq \frac{X}{a}.$$

Because I is a 0-1 random variable, $E[I] = Pr[I = 1] = Pr[X \geq a]$. Taking expectations in $(*)$ thus yields

$$Pr[X \geq a] = E[I] \leq E[\frac{X}{a}] = \frac{E[X]}{a}.$$

Math Background: Markov Chain and Random Walk

Definition (Markov Chain)

A discrete time stochastic process $X_0,\ X_1,\ X_2,\ \ldots$ is a Markov Chain if

$$\begin{aligned} & Pr[X_t = a_t \mid X_{t-1} = a_{t-1},\ X_{t-2} = a_{t-2},\ \ldots,\ X_0 = a_0] \\ = \quad & Pr[X_t = a_t \mid X_{t-1} = a_{t-1}] \\ = \quad & P_{a_{t-1}, a_t}. \end{aligned}$$

Let $G = (V,\ E)$ be a finite, undirected, and connected graph.

Definition (Random Walk)

A *random walk* on G is a Markov chain defined by the sequence of moves of a particle between vertices of G. In this process, the place of the particle at a given time step is the state of the system. If the particle is at vertex i and if i has $d(i)$ outgoing edges, then the probability that the particle follows edge $(i,\ j)$ and moves to a neighbor j is $\frac{1}{d(i)}$.

Math Background: Markov Chain and Random Walk

We might analyze a Markov chain to make predictions about system behaviors.

Let $\mathbf{P}$ is the one-step transition probability matrix of a Markov chain. If $\overline{p}(t)$ is the probability distribution of the state of the chain at time t, then

$$\overline{p}(t+1) = \overline{p}(t)\mathbf{P}.$$

Definition (Stationary Distribution $\overline{\pi}$)

A stationary distribution (also called an equilibrium distribution) of a Markov chain is a probability distribution $\overline{\pi}$ such that

$$\overline{\pi} = \overline{\pi}\mathbf{P}.$$

Note $\sum \pi_v = 1$.

Theorem

A random walk on G converges to a stationary distribution π, where

$$\pi_v = \frac{d(v)}{2|E|}.$$

Proof.

Math Background: Markov Chain and Random Walk

Theorem

A random walk on G converges to a stationary distribution π, where

$$\pi_v = \frac{d(v)}{2|E|}.$$

Proof.

Since $\sum_{v \in V} d(v) = 2|E|$, it follows that

$$1 = \sum_{v \in V} \pi_v = \sum_{v \in V} \frac{d(v)}{2|E|},$$

and π is a proper distribution over $v \in V$.

Let $\mathbf{P}$ be the transition probability matrix of the Markov chain. Let $N(v)$ represent the neighbors of v. The relation $\overline{\pi} = \overline{\pi}\mathbf{P}$ is equivalent to

$$\pi_v = \sum_{u \in N(v)} \left(\frac{d(u)}{2|E|} \frac{1}{d(u)} \right) = \frac{1}{2|E|} \sum_{u \in N(v)} 1 = \frac{d(v)}{2|E|},$$

and the theorem follows. $\square$

Math Background: Markov Chain and Random Walk

Theorem

A random walk on G converges to a stationary distribution π, where $\pi_v = \frac{d(v)}{2|E|}$.

Let $h_{v,u}$ denote the expected number of steps to reach u from v.

Corollary

For any vertex u in G, $h_{u,u} = \frac{2|E|}{d(u)}$.

Lemma

If $(u,\ v) \in E$, then $h_{v,u} < 2|E|$.

Proof.

Let $N(u)$ be the set of neighbors of vertex u in G. We compute $h_{u,u}$ in two different ways:

$$\frac{2|E|}{d(u)} = h_{u,u} = \frac{1}{d(u)} \sum_{w \in N(u)} (1 + h_{w,u}) .$$

Therefore,

$$2|E| = \sum_{w \in N(u)} (1 + h_{w,u}) ,$$

and we conclude that $h_{v,u} < 2|E|$. □

Math Background: Markov Chain and Random Walk

Definition (Cover Time)

The *cover time* of a graph $G = (V, E)$ is a maximum over all vertices $v \in V$ of the expected time to visit all of the nodes in the graph by a random walk starting from v.

Lemma

The cover time of $G = (V, E)$ *is bounded above by* $4|V| \cdot |E|$.

Proof.

Choose a spanning tree of G; that is, choose any subset of the edges that gives an acyclic graph connecting all of the vertices of G. There exists a cyclic (Eulerian) tour on this spanning tree in which every edge is traversed once in each direction: for example, such a tour can be found by considering the sequence of vertices passed through when doing a depth-first search. Let $v_0, v_1, \ldots, v_{2|V|-2} = v_0$ be the sequence of vertices in the tour, starting from vertex v_0. Clearly the expected time to go through the vertices in the tour is an upper bound on the cover time. Hence the cover time is bounded by above by

$$\sum_{i=0}^{2|V|-3} h_{v_i, v_{i+1}} < (2|V| - 2)(2|E|) < 4|V| \cdot |E|,$$

where the first inequality comes from $h_{v,u} < 2|E|$.

A Randomized s-t Connectivity Algorithm

1. Start a random walk from s.
2. If the walk reaches t within $4n^3$ steps, return that there is a path. Otherwise, return that there is no path.

Theorem

The s-t connectivity algorithm returns the correct answer with probability $\frac{1}{2}$ and it only errs by returning that there is no path from s to t when there is such a path.

Proof.

If there is no path then the algorithm returns the correct answer. If there is a path, the algorithm errs if it does not find the path within $4n^3$ steps of the walk. The expected time to reach t from s (if there is a path) is bounded by the cover time of their shared component, which is at most $4nm < 2n^3$. By Markov's inequality, the probability that a walk takes more than $4n^3$ steps to reach s from t is at most $\frac{1}{2}$. $\square$

A Randomized s-t Connectivity Algorithm

A Randomized s-t Connectivity Algorithm

1. Start a random walk from s.
2. If the walk reaches t within $4n^3$ steps, return that there is a path. Otherwise, return that there is no path.

The algorithm must keep track of its current position, which takes $O(\log n)$ bits, as well as the number of steps taken in the random walk, which also takes only $O(\log n)$ bits (since we count up to only $4n^3$).

Example: Boltzmann Economy

Boltzmann Economy Model

Suppose there is a fixed amount of money (M dollars), and a fixed number of agents (N) in the economy. Suppose that during each time step, each agent randomly selects another agent and transfers one dollar to the selected agent. An agent having no money does not go in debt. What will the long term (stable) distribution of money be?

Note

There is no growth, only a redistribution of money (by a random process). For the sake of argument, we can imagine that every agent starts with approximately the same amount of money, although in the long run, the starting distribution should not matter.

Example: Boltzmann Economy

For this example, we are interested in looking at the distribution of money in the economy, so we are looking at the probabilities $\{p_i\}$ that an agent has the amount of money i. We are hoping to develop a model for the collection $\{p_i\}$.
If we let n_i be the number of agents who have i dollars, we have two constraints:

$$\sum_i (n_i \cdot i) = M$$

$$\sum_i n_i = N.$$

Phrased differently (using $p_i = \frac{n_i}{N}$), this says

$$\sum_i (p_i \cdot i) = \frac{M}{N}$$

$$\sum_i p_i = 1.$$

A Maximum Entropy Principle

We define the *entropy* of the distribution P by

$$H(P) = \sum_{i=1}^{n} \left(p_i \log \frac{1}{p_i} \right).$$

Theorem (A Maximum Entropy Principle)

Suppose we have a system for which we can measure certain macroscopic characteristics. Suppose further that the system is made up of many microscopic elements, and that the system is free to vary among various states. Let us assume that with probability essentially equal to 1, the system will be observed in states with maximum entropy.

A Maximum Entropy Principle

What we have right now:

$$\sum_i (p_i \cdot i) = \frac{M}{N}$$

$$\sum_i p_i = 1.$$

We now apply Lagrange multiplier:

$$L = \sum_{i=1}^{n} \left(p_i \log \frac{1}{p_i} \right) - \lambda \left(\sum_i (p_i \cdot i) - \frac{M}{N} \right) - \mu \left(\sum_i p_i - 1 \right),$$

from which we get

$$\frac{\partial L}{\partial p_i} = -[1 + \ln(p_i)] - \lambda \cdot i - \mu = 0.$$

We can solve this for p_i:

$$\ln(p_i) = -\lambda \cdot i - (1 + \mu)$$

and so

$$p_i = e^{-(1+\mu)} e^{-\lambda \cdot i}.$$

A Maximum Entropy Principle

$$p_i = e^{-(1+\mu)} e^{-\lambda \cdot i}.$$

Putting in constraints, we have

$$\begin{aligned} 1 &= \sum_i p_i = \sum_i e^{-(1+\mu)} e^{-\lambda \cdot i} = e^{-(1+\mu)} \sum_{i=0}^{M} e^{-\lambda \cdot i} \\ \frac{M}{N} &= \sum_i (p_i \cdot i) = \sum_i e^{-(1+\mu)} e^{-\lambda \cdot i} \cdot i = e^{-(1+\mu)} \sum_i e^{-\lambda \cdot i} \cdot i \end{aligned}$$

We can approximate (for large M)

$$\begin{aligned} \sum_{i=0}^{M} e^{-\lambda \cdot i} &\approx \int_0^M e^{\lambda x} dx \approx \frac{1}{\lambda} \\ \sum_{i=0}^{M} e^{-\lambda \cdot i} \cdot i &\approx \int_0^M x e^{-\lambda x} dx \approx \frac{1}{\lambda^2}. \end{aligned}$$

From these we have (approximately) $e^{1+\mu} = \frac{1}{\lambda}$ and $e^{1+\mu} \frac{M}{N} = \frac{1}{\lambda^2}$.
We have $\lambda = \frac{N}{M} = e^{-(1+\mu)}$, and thus (letting $T = \frac{M}{N}$) we have

$$p_i = e^{-(1+\mu)} e^{-\lambda \cdot i} = \frac{1}{T} e^{-\frac{i}{T}} = \frac{1}{T} e^{-\frac{i}{T}}.$$