CS270 Algorithms & Data Structures **Spring 2008**

Lecture 12 Path selection using linear programming. : 2.19.08

Lecturer: Satish Rao *Scribe: Rao, for now*

Disclaimer: *These are rough notes, with some exercises.*

12.1 A linear program for setting up communication.

Question: Given a set of pairs s_i, t_i that wish to communicate (say one unit of traffic) in a graph G what is the best way to route their communication paths?

Perhaps to minimize congestion.

Question: How to set this up?

A linear program perhaps. Lets give a variable, x_p, for each path $\mathcal{P}$. And indeed, say the s_i, s_t paths are $\mathcal{P}_{ij}$. Now we can do the following.

$$\min C$$

$$\forall e \sum_{e} \in p, p \in \mathcal{P} \leq C$$

$$\forall i \sum_{p \in \mathcal{P}_i} = 1$$

Question: Is there a problem here?

Well, the size of the program is large. That is, exponential. There are two ways around this, one being to use the ellipsoid algorithm, another is to rewrite the program to make it polynomial size.

Question: How?

Introduces flow variables $f_i(e)$ for each pair s_i, t_i such that one unit of flow is routed from each s_i and we have flow conservation, and the common capacity of an edge is obeyed. That is, we consider the following linear program.

$$\min C$$

$$\forall i \sum_{e=(s_i,u)} f_i(e) = 1$$

$$\forall v, \forall i \neq s_i \sum_{e=(u,v)} f_i(e) - \sum_{e=(v,u)} f_i(e) = 0$$

$$\forall e \sum_i f_i(e) \leq C$$

Question: How large is this program?

$O(km)$ unknowns, and around $O(kn)$ inequalities. Polynomial.

Question: Is there another problem?

Well, the flow is not integral so there is not a single path for each s_i and t_i.

Question: Can we at least convert the polynomial solution to a solution to the path version, so we have many paths.

Sure, we can use flow decomposition: for each i, find a set of paths that correspond to the flow. We do so, by finding a positive flow path from s_i to t_i (using say depth first search), finding the bottleneck edge, decrementing the flow variables $f_i(\cdot)$ by the bottleneck amount and continuing. Each path corresponds to zeroing one of km flow variables. Thus, we finish in time $O(km^2)$.

Question: How should we make it integral?

Question: Well, what is $\sum_{p \in \mathcal{P}_i} x_p$ for a particular i?

1. We have a bunch of numbers that sum to 1. And we want one number that is 1.

Question: What should we do?

View the set of a probability distribution. Choose one path for each i with probability x_p.

Question: What is the expected congestion on an edge?

We can express the congestion on an edge e as the sum of some random variables $X_{p_1}, \ldots, X_{p_l}$ where each X_p is 1 if the path p is selected and 0 otherwise, and $p_1, \ldots, p_l$ are paths that use edge e.

Question: OK, so what is the expected congestion on an edge?

Well, it is $E[\sum_k X_{p_k}] = \sum_k E[X_{p_k}]$. The expectation of a $0-1$ variable is, of course, the probability that it is 1. So, we get that the expected congestion is at most

$$\sum_{p_i \text{ that use } e} x_p.$$

But, this is upper bounded by the congestion of the flow.

That is our expected congestion is upper bounded by the optimal congestion!

Question: Cool! Are we done?

Well, there are a lot of edges. Perhaps some of them are congested.

Since we will only speak of one edge, let's simplify our notation to just be $X_1, \ldots, X_l$ be the variables corresponding to paths that use edge e.

Question: What is the chance that $\sum_i X_i$ is much higher than its expectation?

We can use a Chernoff bound. For example, we have the following.

Theorem 12.1 *Given $0-1$ valued independent random variables, $X_1, \dots, X_l$, and $X = \sum_i X_i$, with $E[X] = \mu$, then*

$$\Pr\big[X \geq (1+\delta)\mu\big] \leq \begin{cases} e^{-\delta^2\mu/3} & \text{if } \delta \leq 1, \\ e^{-\delta^2\mu/4} & \text{if } \delta \leq 2e-1, \\ 2^{-\delta\mu} & \text{if } \delta > 2e-1. \end{cases} \tag{12.1}$$

Question: Can we apply the lemma?

Well, we don't necessarily of independent variables, perhaps X_p and X_p' that use edge e both belong to the same source sink pair.

Question: Uh oh? How do we deal with this?

We can change the set of variables to be $X_1, \dots,$ to indicate whether any path from pair i that uses this edge is chosen. This again is $0-1$ and the variables are independent. The expectation remains the same.

Question: Now, what can we conclude from the Chernoff bound?

Well, that the probability that an edge has congestion more than $C + \sqrt{C \log N} + O(\log N)$ is at most $1/N^2$.

Question: Huh?

Well, we wish to bound the probability that the congestion is more than $(1+\delta)C = C + O(\sqrt{C \log N} + \log N)$. That is, $\delta = O(\sqrt{\log N/C} + \log N/C)$.

For $\delta > 2$ (say $C < \log N$), we get that the probability is less than $2^{-O(\log N)}$ which can be made to be less than $1/N^2$ by choosing the constant in the $O(\cdot)$ notation appropriately.

Say $C >> \log N$, then $\delta < 1$, and the probability is at most $e^{\delta^2\mu/3}$. That is, we get an something that is $\Theta(\log N)$ in the exponent. Again, by choosing the constant appropriately, we can get the probability to be bounded by $1/2N^2$.

Question: How do we know that every edge is good?

This is the union bound; the probability that any one event occurs is at most the sum of the probability of the events. That is there are at most N^2 edges thus, the probability that any edge is bad is at most $1/N^2$ times the number of edges, which is less than $1/2$.

Exercise 1: How could you modify the approach above so that one only produces paths of length at most say L to connect each pair?

12.2 Chernoff bound.

Question: Can we prove the Chernoff bound?

Perhaps, but we are computer scientist so let's first get some intuition from flipping n coins with equal heads probability, p.

Question: What is the probability that we get more than k heads?

$$\sum_{i>k} \binom{n}{i} p^i (1-p)^{n-i}$$

Question: That's intuitive? Sorry, let's continue.

Assuming $k > pn$, this is within a factor of n of

$$\binom{n}{k} p^k (1-p)^{n-k},$$

since later terms are decreasing (assuming $k > pn$.) And this is, in turn, less than

$$\binom{n}{k} p^k.$$

Using the fact that $\binom{n}{k} \leq (ne/k)^k$,

$$\left(\frac{ne}{k})\right)^k p^k.$$

Assuming $k > 2epn$, we obtain

$$(12)^k.$$

Thus, to defeat the factor of n above, and to make this polynomially small, $k > \Theta(\log N)$ for some large enough constant (say 3) to get this probability to be smaller $1/N^2$.

Question: Totalling up, what do we get?

$k \geq 2epn + 3 \log N$).

Chernoff gets a bit better. That is,

$$k \geq pn + O(\sqrt{pn \log N} + \log N).$$

12.3 Proving Chernoff.

We'll prove part of the following more general version.

Theorem 12.2 *For a random variable, $X = \sum_i X_i$, where X_i are $0-1$ random variables, and with mean $\mu = E[X]$, for $\delta > 0$*

$$Pr[X > (1+\delta)\mu] \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu. \tag{12.2}$$

and for $1 > \delta > 0$,

$$Pr[X < (1-\delta)\mu] \leq \left(\frac{e^-\delta}{(1-\delta)^{1-\delta}}\right)^\mu. \tag{12.3}$$

Question: Let's warmup with Markov's inequality.

For a positive random variable X, $Pr[X > c] \leq E[X]/c$.

Question: Proof?

Start with the definition of expectation.

$$\begin{aligned}
E[X] &= \sum_a Pr[X = a] \cdot a \\
&= \sum_{a \leq c} Pr[X = a] \cdot a + \sum_{a > c} Pr[X = a]a \\
&\geq \sum_{a > c} Pr[X = a]c
\end{aligned}$$

Last line is done. (We need positive value variable to drop first summation in line 2.)

Question: In words?

Can be more than 5 times better than average more than 1/5 of the time. For any constant 5.

Question: Now Chernoff's theorem.

Proof:

We only give the proof for bounding the probability of the "upper tail" or equation 12.2. Let t be an arbitrary positive constant.

$$\begin{aligned}
Pr[X > (1+\delta)\mu] &\leq Pr[e^{tX} > e^{(1+\delta)t\mu}] \\
&\leq \frac{E[e^{tX}]}{e^{(1+\delta)t\mu}} \\
&\leq \frac{\Pi_i E[e^{tX_i}]}{e^{(1+\delta)t\mu}} \\
&\leq \frac{\Pi_i (p_i e^t + (1-p_i)e^0)}{e^{(1+\delta)t\mu}} \\
&\leq \frac{\Pi_i (1 + p_i(e^t - 1))}{e^{(1+\delta)t\mu}} \\
&\leq \frac{\Pi_i e^{p_i(e^t-1)}}{e^{(1+\delta)t\mu}} \\
&\leq \frac{e^{\sum_i p_i(e^t-1)}}{e^{(1+\delta)t\mu}} \\
&\leq \frac{e^{(e^t-1)\sum_i p_i}}{e^{(1+\delta)t\mu}} \\
&\leq \frac{e^{(e^t-1)\mu}}{e^{(1+\delta)t\mu}} = \left(\frac{e^{(e^t-1)}}{e^{(1+\delta)t}}\right)^{\mu}
\end{aligned}$$

The third line follows from the fact that the X_i are independent. (That is, $E[XY] = E[X]E[Y]$ for independent random variables X and Y.) The sixth line follows from the fact that for positive x, $(1+x) < e^x$.

Choosing $t = \ln(1+\delta)$, we get equation 12.2.

$\blacksquare$

Exercise 2: Prove the first and the third case of theorem 12.1 using theorem 12.2.