EECS 495: Randomized Algorithms — **Lecture 7**
Using Chernoff

Reading: Text: Chapter 4; Balanced Allocations by Azar, Broder, Karlin, and Upfal

Balls and Bins

Problem: n balls, place obliviously into n bins

Algorithm: Random.

Claim: W/prob. $(1 - n^{-c})$, fullest bin has $(1 + o(1))\frac{\ln n}{\ln \ln n}$ balls.

Proof: Let E_{jk} be event that bin j has more than k balls

$$\begin{aligned}
\Pr[E_{jk}] &= \sum_{i=k}^{n} \binom{n}{i} \left(\frac{1}{n}\right)^i \left(1 - \frac{1}{n}\right)^{n-i} \\
&\leq \sum_{i=k}^{n} \left(\frac{ne}{i}\right)^i \left(\frac{1}{n}\right)^i \\
&= \sum_{i=k}^{n} \left(\frac{e}{i}\right)^i \\
&\leq \left(\frac{e}{k}\right)^k \sum_{i=0}^{\infty} \left(\frac{e}{k}\right)^i \\
&= \left(\frac{e}{k}\right)^k \left(\frac{1}{1 - e/k}\right) \\
&\leq 2\left(\frac{e}{k}\right)^k \\
&\leq O\left(\frac{1}{n^2}\right)
\end{aligned}$$

for $k = O(\log n / \log \log n)$. Result follows by union bound.

Oblivious Routing

Given:

- hypercube network
- permutation destinations

Output:

- set of oblivious routes with min # steps

Algorithm: (deterministic): Bit fixing (left to right)

Question: Randomized alg.?

[[*Recall load-balancing: m jobs, $n = m$ machines; to distribute load oblivioiusly, we randomly routed jobs to machines.*]]

Idea: Load-balance paths!

First try: random destination, bit-fixing

- $T(e_l)$ = # paths using e_l
- By symmetry, all $T(e_l)$ equal
- Expected path length $n/2$
- LOE, total expected path length $Nn/2$
- Nn edges in hypercube

So $E[T(e_l)] = 1/2$, so delay at most $n/2$.

Claim: Delay $\leq 6n$ with high prob.

Proof: Chernoff: Delay $X \leq \sum_l T(e_l)$, so

- $\Pr[X > (1+\delta)\mu] \leq \exp(-(1+\delta)\mu)$
- $\Pr[X > 6n] \leq \exp(-6n)$

WRONG!

Proof: Chernoff: Fix packet i, let H_{ij} indicate if routes for i and j share an edge. Independent and $\sum_j H_{ij} \leq \sum_i T(e_l)$.

- delay of fixed packet $i \geq 6n$ with prob. $\leq \exp(-6n) \leq 2^{-6n}$
- prob. any $N = 2^n$ packets gets delay more than 2^{-6n} at most 2^{-5n} (union bound)
- time to route any packet at most length plus delay, at most $7n$

$\rightarrow$ w/prob. $\geq 1 - 2^{-5n}$, every packet reaches random dest. in $7n$ or fewer steps.

But wanted to reach $d(i)$!

Idea: Route to random intermediate destination.

- doubles path length
- destroys bad perms.

Run backwards, same time bound, so packets fail to reach final dest. in at most $14n$ steps with prob. at most $2^{-5n+1} \leq 1/N$.

Claim: With prob. $\geq 1 - 1/N$, every packet reaches dest. in at most $14n$ steps.

Note: Didn't allow phase 2 to delay phase 1; must have packets wait at intermediate dest. for $7n$ steps.

Power of Two Choices

Algorithm: For each ball, pick two bins randomly, place in less-loaded bin.

Intuition:

- At most $n/4$ bins have 4 balls $\rightarrow$ prob. get bin with 5 balls at most $(1/4)(1/4) = 1/16$
- Expect $n/16$ bins have 5 balls $\rightarrow$ prob. get bin with 6 balls at most $1/16^2$
- Expect $n/2^{2^{k-3}}$ bins have k balls
- Whp, no bin has more than $\log\log n$ balls

Problem: Assume system behaves as expects to analyze next layer, must cope with conditioning.

Claim: Bound binomial: $\Pr[B(n,p) > 2np] < 2^{-np/3}$

Proof: $B(n,p) = \sum_{i=1}^{n} X_i$ where

$$X_i = \begin{cases} 1 & : w/prob.\ p \\ 0 & : otherwise \end{cases}$$

- Chernoff:

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

- $\mu = np, \delta = 1$, implies claim

Note: Let Y_i depend on $X_1, \ldots, X_{i-1}$. If $\Pr[Y_i | X_1, \ldots, X_{i-1}] < p$ for all i, then X *stochastically dominates* Y, so can use above bound.

Analysis:

- $v_i(t) = \#$ bins of height $\geq i$ at time t
- h_t = height of t'th ball
- bounds $\beta_4 = 1/4$, $\beta_i = 2\beta_{i-1}^2 = 2^{-2^i}$

Claim: Let Q_i be event that $v_i \le \beta_i n$. Then Q_i happens whp.

Proof: By induction.

- Let $Y_t = 1$ if $h_t \ge i+1$ and $v_i(t-1) \le \beta_i n$

  $\left[\left[\textit{$Y_t$ indicates t'th ball placed in a bad bin even though there were enough good bins.}\right]\right]$

- Then $\Pr[Y_i = 1] \le \beta_i^2$
- Then $\sum_t Y_t$ stochastically dominated by X with $p = \beta_i^2$
- By Chernoff,

  $$\Pr[\sum_t Y_t \ge 2n\beta_i^2 = \beta_{i+1}n] < 2^{-\beta_{i+1}n/6}$$

  which is $O(1/n^2)$ so long as $\beta_{i+1}n \ge c\log n$

When Q_i holds, then $\sum_t Y_t$ is # tall balls:

- # tall bins $\le$ # tall balls
- so

  $$\begin{aligned}\Pr[\neg Q_{i+1}|Q_i] &\le \Pr[\sum_t Y_t > \beta_{i+1}n|Q_i] \\ &\le \Pr[\sum_t Y_t > \beta_{i+1}n]/\Pr[Q_i] \\ &\le 1/(n^2\Pr[Q_i])\end{aligned}$$

Deal with conditioning:

$$\begin{aligned}\Pr[\neg Q_{i+1}] &= \Pr[\neg Q_{i+1}|Q_i]\Pr[Q_i] + \Pr[\neg Q_{i+1}|\neg Q_i]\Pr[\neg Q_i] \\ &\le \frac{1}{n^2} + \Pr[\neg Q_i] \\ &\le \frac{1}{n}\end{aligned}$$

by induction.

Deal with large i:

- ok until that i^* s.t. bound on # tall bins dips below $\log n$
- but then bins no longer grow because
  - prob. ball is tall for any $i \ge i^*$ at most $((\log n)/n)^2$
  - prob. two particular balls both tall at most $((\log n)/n)^4$ (events negatively correlated)
  - union bound, prob. two distinct tall balls $o(1)$

  and bins don't grown beyond $i^* + 1$ if only see one tall ball
- $\beta_{i^*}n = 2^{-2^{i^*}} \ge \log n \to i^* = \log\log n - \log\log\log n$
- so max load is $O(\log\log n)$

$\left[\left[\textit{Relationship to random graphs: bins are nodes, two choices are edges. Bound comes from low expected degree and “small” giant component.}\right]\right]$

Note: Extensions:

- Pick more bins: $O(\log_d \log n)$
- Pick more bins and consistent tie-breaking: $O(\log\log n/d)$