Application: Set Balancing

Given an $n \times n$ matrix $\mathcal{A}$ with entries in $\{0, 1\}$, let

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ \dots \\ b_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \dots \\ \dots \\ c_n \end{pmatrix}.$$

Find a vector $\bar{b}$ with entries in $\{-1, 1\}$ that minimizes

$$||\mathcal{A}\bar{b}||_\infty = \max_{i=1,\dots,n} |c_i|.$$

Theorem

For a random vector $\bar{b}$, *with entries chosen independently and with equal probability from the set* $\{-1, 1\}$,

$$Pr(||\mathcal{A}\bar{b}||_\infty \geq \sqrt{12 n \ln n}) \leq \frac{2}{n}.$$

- Consider the i-th row $\bar{a}_i = a_{i,1}, \ldots, a_{i,n}$. Let k be the number of 1's in that row.
- If $k \leq \sqrt{12n \ln n}$ clearly $|\bar{a}_i \cdot \bar{b}| \leq \sqrt{12n \ln n}$.
- If $k > \sqrt{12n \ln n}$, let

$$X_i = |\{j \mid a_{i,j} = 1 \text{ and } b_j = 1\}|$$

and

$$Y_i = |\{j \mid a_{i,j} = 1 \text{ and } b_j = -1\}|.$$

- Thus, X_i counts the number of $+1$'s in the sum $\sum_{j=1}^{n} a_{i,j} b_j$,
- Y_i counts the number of -1's
- $X_i + Y_i = k$.

if $|X_i - Y_i| \leq \sqrt{12n \ln n}$ then $|X_i - (k - X_i)| \leq \sqrt{12n \ln n}$ which implies

$$\frac{k}{2}\left(1 - \frac{\sqrt{12n \ln n}}{k}\right) \leq X_i \leq \frac{k}{2}\left(1 + \frac{\sqrt{12n \ln n}}{k}\right)$$

or equivalently

$$\frac{k}{2}\left(1 - \sqrt{\frac{12n \ln n}{k^2}}\right) \leq X_i \leq \frac{k}{2}\left(1 + \sqrt{\frac{12n \ln n}{k^2}}\right).$$

Using Chernoff bounds,

$$Pr\left(X_i \geq \frac{k}{2}\left(1+\sqrt{\frac{12n\ln n}{k^2}}\right)\right) \leq e^{-\left(\frac{k}{2}\right)\left(\frac{1}{3}\right)\left(\frac{12n\ln n}{k^2}\right)} \leq e^{-2\ln n} = n^{-2}$$

$$Pr\left(X_i \leq \frac{k}{2}\left(1-\sqrt{\frac{12n\ln n}{k^2}}\right)\right) \leq e^{-\left(\frac{k}{2}\right)\left(\frac{1}{2}\right)\left(\frac{12n\ln n}{k^2}\right)} \leq e^{-3\ln n} \leq n^{-2}$$

Hence, for a given row,

$$Pr(|X_i - Y_i| \geq \sqrt{12n\ln n}) \leq \frac{2}{n^2}$$

Since there are n rows, by union bound the probability that any row exceeds that bound is at most $\frac{2}{n}$.

Chernoff Bound for Sum of $\{-1, +1\}$ Random Variables

Theorem

Let $X_1, ..., X_n$ be independent random variables with

$$Pr(X_i = 1) = Pr(X_i = -1) = \frac{1}{2}.$$

Let $X = \sum_1^n X_i$. For any $a > 0$,

$$Pr(X \geq a) \leq e^{-\frac{a^2}{2n}}.$$

For any $t > 0$,

$$\mathbf{E}[e^{tX_i}] = \frac{1}{2}e^t + \frac{1}{2}e^{-t}.$$

$$e^t = 1 + t + \frac{t^2}{2!} + \cdots + \frac{t^i}{i!} + \ldots$$

and

$$e^{-t} = 1 - t + \frac{t^2}{2!} + \cdots + (-1)^i\frac{t^i}{i!} + \ldots$$

Thus,

$$\begin{aligned} \mathbf{E}[e^{tX_i}] &= \frac{1}{2}e^t + \frac{1}{2}e^{-t} = \sum_{i\geq 0}\frac{t^{2i}}{(2i)!} \\ &\leq \sum_{i\geq 0}\frac{(\frac{t^2}{2})^i}{i!} = e^{t^2/2} \end{aligned}$$

$$\mathbf{E}[e^{tX}] = \prod_{i=1}^{n} \mathbf{E}[e^{tX_i}] \leq e^{nt^2/2},$$

$$Pr(X \geq a) = Pr(e^{tX} > e^{ta}) \leq \frac{\mathbf{E}[e^{tX}]}{e^{ta}} \leq e^{t^2 n/2 - ta}.$$

Setting $t = a/n$ yields

$$Pr(X \geq a) \leq e^{-\frac{a^2}{2n}}.$$

By symmetry we also have

Corollary

Let $X_1, ..., X_n$ be independent random variables with

$$Pr(X_i = 1) = Pr(X_i = -1) = \frac{1}{2}.$$

Let $X = \sum_{i=1}^{n} X_i$. Then for any $a > 0$,

$$Pr(|X| > a) \leq 2e^{-\frac{a^2}{2n}}.$$

Application: Set Balancing Revisited

Theorem

For a random vector $\bar{b}$, with entries chosen independently and with equal probability from the set $\{-1, 1\}$,

$$Pr(||\mathcal{A}\bar{b}||_\infty \geq \sqrt{4n \ln n}) \leq \frac{2}{n} \tag{1}$$

- Consider the i-th row $\bar{a}_i = a_{i,1}, ..., a_{i,n}$.
- Let k be the number of 1's in that row.
- $Z_i = \sum_{j=1}^{k} a_{i,i_j} b_{i_j}$.
- If $k \leq \sqrt{4n \ln n}$ then clearly $Z_i \leq \sqrt{4n \ln n}$.

If $k > \sqrt{4n \log n}$, the k non-zero terms in the sum Z_i are independent random variables, each with probability $1/2$ of being either $+1$ or -1.
Using the Chernoff bound:

$$Pr\left\{|Z_i| > \sqrt{4n \log n}\right\} \leq 2e^{-4n \log n/(2k)} \leq \frac{2}{n^2},$$

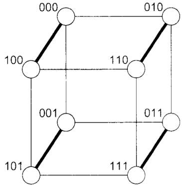
where we use the fact that $n \geq k$.
The result follows by union bound (n rows).

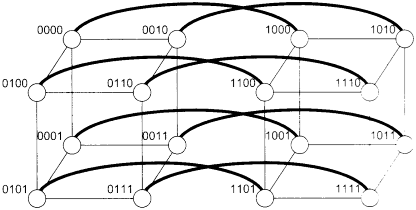
Packet Routing on Parallel Computer

Communication network:

- nodes - processors, switching nodes;
- edges - communication links.

The n-cube:
$N = 2^n$ nodes: $0, 1, 2, \ldots, 2^n - 1$.
Let $\bar{x} = (x_1, ..., x_n)$ be the number of node x in binary.
Nodes x and y are connected by an edge iff their binary representations differ in exactly one bit.
Bit-wise routing: correct bit i in the i-th transition - route has length $\leq n$.

The 3-cube: 

The 4-cube: 

A permutation communication request: each node is the source and destination of exactly one packet.
Up to one packet can cross an edge per step, each packet can cross up to one edge per step.
What is the time to route an arbitrary permutation on the n-cube?

Two phase routing algorithm:

1. Send packet to a randomly chosen destination.
2. Send packet from randomly chosen destination to real destination.

Path: Correct the bits, starting at x_1 to x_n.
Any greedy queuing method - if some packet can traverse an edge one does.

Theorem

The two phase routing algorithm routes an arbitrary permutation on the n-cube in $O(\log N) = O(n)$ parallel steps with high probability.

- We focus first on phase 1. We bound the routing time of a given packet M.
- Let $e_1, \ldots, e_m$ be the $m \leq n$ edges traversed by a given packet M is phase 1.
- Let $X(e)$ be the total number of packets that traverse edge e at that phase.
- Let $T(M)$ be the number of steps till M finished phase 1.

Lemma

$$T(M) \leq \sum_{i=1}^{m} X(e_i).$$

- We call any path $P = (e_1, e_2, \ldots, e_m)$ of $m \leq n$ edges that follows the bit fixing algorithm a *possible packet path*.
- We denote the corresponding nodes $v_0, v_1, \ldots, v_m$, with $e_i = (v_{i-1}, v_i)$.
- For any possible packet path P, let $T(P) = \sum_{i=1}^{m} X(e_i)$.

- If phase I takes more than T steps then for some possible packet path P,

$$T(P) \geq T$$

- There are at most $2^n \cdot 2^n = 2^{2n}$ possible packet paths.
- Assume that e_k connects $(a_1, ..., a_i, ..., a_n)$ to $(a_1, .., \bar{a}_i, ..., a_n)$.
- Only packets that started in address

$$(*, ..., *, a_i, ..., a_n)$$

can traverse edge e_k, and only if their destination addresses are

$$(a_1, ..., a_{i-1}, \bar{a}_i, *, ..., *)$$

.

- There are no more than 2^{i-1} possible packets, each has probability 2^{-i} to traverse e_i.

- There are no more than 2^{i-1} possible packets, each has probability 2^{-i} to traverse e_i.
- $$\mathbf{E}[X(e_i)] \leq 2^{i-1} \cdot 2^{-i} = \frac{1}{2}.$$
- $$\mathbf{E}[T(P)] \leq \sum_{i=1}^{m} \mathbf{E}[X(e_i)] \leq \frac{1}{2} \cdot m \leq n.$$
- **Problem:** The $X(e_i)$'s are not independent.

- A packet is *active* with respect to possible packet path P if it ever use an edge of P.
- For $k = 1, \ldots, N$, let $H_k = 1$ if the packet starting at node k is active, and $H_k = 0$ otherwise.
- The H_k are independent, since each H_k depends only on the choice of the intermediate destination of the packet starting at node k, and these choices are independent for all packets.
- Let $H = \sum_{k=1}^{N} H_k$ be the total number of active packets.
- $$\mathbf{E}[H] \leq \mathbf{E}[T(P)] \leq n$$
- Since H is the sum of independent $0-1$ random variables we can apply the Chernoff bound

$$\Pr(H \geq 6n) \leq \Pr(H \geq 6\mathbf{E}[H]) \leq 2^{-6n}.$$

For a given possible packet path P,

$$\begin{aligned}
\Pr(T(P) \geq 30n) &\leq \Pr(T(P) \geq 30n \mid H \geq 6n) \Pr(H \geq 6n) \\
&+ \Pr(T(P) \geq 30n \mid H < 6n) \Pr(H < 6n) \\
&\leq \Pr(H \geq 6n) + \Pr(T(P) \geq 30n \mid H < 6n) \\
&\leq 2^{-6n} + \Pr(T(P) \geq 30n \mid H < 6n).
\end{aligned}$$

Lemma

If a packet leaves a path (of another packet) it cannot return to that path in the same phase.

Proof.

Leaving a path at the i-th transition implies different i-th bit, this bit cannot be changed again in that phase. □

Lemma

The number of transitions that a packet takes on a given path is distributed $G\left(\frac{1}{2}\right)$.

Proof.

The packet has probability $1/2$ of leaving the path in each transition. □

The probability that the active packets cross edges of P more than $30n$ times is less than the probability that a fair coin flipped $36n$ times comes up heads less than $6n$ times.
Letting Z be the number of heads in $36n$ fair coin flips, we now apply the Chernoff bound:

$$
\begin{aligned}
& \Pr(T(P) \geq 30n \mid H \leq 6n) \leq \Pr(Z \leq 6n) \\
\leq \quad & e^{-18n(2/3)^2/2} = e^{-4n} \leq 2^{-3n-1}.
\end{aligned}
$$

$$
\begin{aligned}
\Pr(T(P) \geq 30n) \quad & \leq \quad \Pr(H \geq 6n) + \Pr(T(P) \geq 30n \mid H \leq 6n) \\
& \leq \quad 2^{-6n} + 2^{-3n-1} \leq 2^{-3n}
\end{aligned}
$$

As there are at most 2^{2n} possible packet paths in the hypercube, the probability that there is *any* possible packet path for which $T(P) \geq 30n$ is bounded by

$$2^{2n}2^{-3n} = 2^{-n} = O(N^{-1}).$$

- The proof of phase 2 is by symmetry:
- The proof of phase 1 argued about the number of packets crossing a given path, no "timing" considerations.
- The path from "one packet per node" to random locations is similar to random locations to "one packet per node" in reverse order.
- Thus, the distribution of the number of packets that crosses a path of a given packet is the same.

Oblivious Routing

Definition

A routing algorithm is **oblivious** if the path taken by one packet is independent of the source and destinations of any other packets in the system.

Theorem

Given an N-node network with maximum degree d the routing time of any deterministic oblivious routing scheme is

$$\Omega\left(\sqrt{\frac{N}{d^3}}\right).$$