Outline

Set Balancing

Routing in Boolean Hypercube

Readings
Set Balancing

Let $A_1, \ldots, A_n$ be subsets of $[n]$ such that $|A_i| = n/2$. We want to partition $[n]$ into $B$ and $C$ such that

$$\max_i |A_i \cap B| - |A_i \cap C|$$

is minimized.

Hint: Use $\mathbb{P}[|X - \mathbb{E}[X]| < \delta \mu] \leq 2 \exp(-\mathbb{E}[X] \delta^2/4)$. 
Set Balancing Algorithm and Analysis

Algorithm: Consider a random partition!

Lemma
\[
\max_i \left| |A_i \cap B| - |A_i \cap C| \right| \leq 4\sqrt{n \ln n} \text{ with prob. at least } 1 - 2n^{-3}.
\]

Proof.

- Let \( X_j = 1 \) if \( j \)-th element of \( A_i \) is in \( C \) and 0 otherwise.
- Then \( X = \sum_j X_j = |A_i \cap C| \) and \( E[X] = n/4 \)
- \( \left| |A_i \cap B| - |A_i \cap C| \right| = \left| n/2 - 2|A_i \cap C| \right| = 2\left| E[X] - X \right| \)
- By an application of the Chernoff bound:
  \[
P \left[ 2\left| E[X] - X \right| \geq 4\sqrt{n \ln n} \right] = P \left[ \left| E[X] - X \right| \geq 8\sqrt{n^{-1} \ln n} \cdot E[X] \right] \leq 2e^{-(n/4)(64n^{-1} \ln n)/4} = 2n^{-4}
  \]
- Apply union bound over all \( i \).
Outline

Set Balancing

Routing in Boolean Hypercube

Readings
Routing in Boolean Hypercube

Boolean hypercube:
- $N = 2^n$ nodes where each is labeled by a length $\{0, 1\}^n$
- Edge between $x \in \{0, 1\}^n$ and $y \in \{0, 1\}^n$ iff $\Delta(x, y) = 1$.

Problem: Let $\pi$ be a permutation of $[N]$
- For $i \in [N]$, packet $v_i$ needs routed from $i$-th node to $\pi(i)$-th node.
- At each step, a packet may traverse an edge (or stay still).
- At most one packet can traverse an edge in the same step.

Want an “Oblivious Routing”:
- Pick route $\rho_i$ for $v_i$ that only depends only on $\pi(i)$.
- Pick rule for deciding who gets precedence when two packets want to use same edge.
The Algorithm

- For each $v_i$, pick random intermediate destination $\sigma(i)$
- **Phase 1**: Route each $v_i$ from $i$-th node to $\sigma(i)$-th node
- **Phase 2**: Route each $v_i$ from $\sigma(i)$-th node to $\pi(i)$-th node
- Use “first in, first out” queueing policy.

In each phase, route using “bit-fixing”:

- At each step, forward $v_i$ to neighboring node whose label agrees with longest prefix of label of $\sigma(i)$-th node
- E.g., to get from (1011) to (0110), the route would be $(1011) \rightarrow (0011) \rightarrow (0111) \rightarrow (0110)$

Once two paths diverge, they don’t merge again.
The Result

**Theorem**

*All packets get routed to final destination in $14n$ steps with probability at least $1 - 2/N$.***

For comparison (we won’t prove this):

**Theorem**

*For any deterministic oblivious routing algorithm, there is a permutation $\pi$ that requires $\Omega(\sqrt{2^m/n})$ time.*
Let $\rho_i$ be the path taken by $v_i$ in Phase 1.

Let $H_{i,j} = 1$ if $\rho_i$ and $\rho_j$ intersect and 0 otherwise ($H_{i,i} = 0$).

$H_{i,1}, H_{i,2}, \ldots, H_{i,N}$ are independent Poisson trials for each $i$.

**Thm:** Time for $v_i$ to get to intermediate destination is

$$\text{Length of Path} + \text{Delays} \leq n + \sum_j H_{i,j}$$

**Thm:** With probability at least $1 - 2^{-6n}$, $\sum_j H_{i,j} \leq 6n$

By union bound, with probability at least $1 - 2^{-5n} \geq 1 - 1/N$, all packets get to intermediate destination in at most $7n$ time.

Analysis of Phase 2 is identical.
Analysis Part 1: Expressing delay in terms of $H_{i,j}$ (1/2)

Theorem

Total delay incurred by $v_i$ is at most $\sum_j H_{i,j}$.

Proof.

- Let route of $v_i$ be $\rho_i = (e_1, e_2, \ldots, e_k)$.
- Let $S$ be the set of $\sum_j H_{i,j}$ packets (other than $v_i$) whose routes pass through an edge in $\rho_i$.
- Say $v \in S$ leaves $\rho_i$ at the time step at which it traverses an edge in $\rho_i$ for the last time.
- If a packet is ready to traverse edge $e_j$ at time $t$, we define its lag at time $t$ to be $t - j$. Final lag for $v_i$ is the delay.
- Claim: If lag of $v_i$ when it traverses $e_k$ is $\ell$ then for each $\ell' \leq \ell$ there exists a $v \in S$ that leaves $\rho_i$ with lag $\ell'$.
- Hence $|S| \geq (\text{final lag of } v_i) = (\text{total delay incurred by } v_i)$.

\[\square\]
Claim

If lag of \( v_i \) when it traverses \( e_k \) is \( \ell \) then for each \( \ell' < \ell \) there exists a \( v \in S \) that leaves \( \rho_i \) with lag \( \ell' \).

Proof.

- There exists a packet in \( S \) that has lag \( \ell' \) at some stage:
  1. Consider the step when \( v_i \) increases from lag \( \ell' \) to \( \ell' + 1 \) because it cannot use some edge \( e \)
  2. At this step, the packet that does use \( e \) has lag \( \ell' \).

- Let \( t' \) be the last time at which a packet in \( S \) has lag \( \ell' \).
  1. All packets waiting to traverse \( e_j \) for \( j = t' - \ell' \) have lag \( \ell' \).
  2. The packet that traverses \( e_j \) still has lag \( \ell' \) unless it leaves \( \rho_i \).
  3. So if it doesn’t leave \( \rho_i \), \( t' \) wasn’t the last time at which a packet in \( S \) has lag \( \ell' \).
Analysis Part 2: Bounding tail probability of delay (1/2)

Let $h_i = \sum_j H_{i,j}$. By Chernoff bound, $\mathbb{P}[h_i \geq 6\mathbb{E}[h_i]] \leq 2^{-6\mathbb{E}[h_i]}$.

Lemma

$\mathbb{E}[h_i] \leq n/2$ and so $\mathbb{P}[h_i \geq 6n] \leq 2^{-6n}$.

Proof.

- Let $T(e)$ be the number of routes that include edge $e$.
- $h_i \leq \sum_{e \in \rho_i} T(e)$ and, by linearity of expectation:

  \[
  \mathbb{E}[h_i] \leq \sum_{e \in \rho_i} \mathbb{E}[T(e)]
  \]

- Claim: $\mathbb{E}[T(e)] = 1/2$.
- Hence, $\mathbb{E}[h_i] \leq \sum_{e \in \rho_i} \mathbb{E}[T(e)] = |\rho_i|/2 \leq n/2$
Analysis Part 2: Bounding tail probability of delay (2/2)

Claim

\[ \mathbb{E} [T(e)] = \frac{1}{2}. \]

Proof.

- Expected number of bits to be fixed is \( n/2 \): \( \mathbb{E} [|\rho_j|] = n/2. \)
- By linearity of expectation: \( \mathbb{E} \left[ \sum_j |\rho_j| \right] = Nn/2. \)
- Each edge in a route contributes to a \( T(e) \) and vice versa:

\[
\sum_j |\rho_j| = \sum_e T(e)
\]

- Total number of edges is \( Nn \) and \( \mathbb{E} [T(e)] \) is the same for all \( e \).
Outline

Set Balancing

Routing in Boolean Hypercube

Readings
For next time, please make sure you’ve read:

- Chapter 4: Up to and including 4.2 (12 pages)