Balls and Bins with Structure

Brighten Godfrey
UC Berkeley
Soda 2008 • January 21, 2008
Nearby server selection

Servers in the unit square

Clients arrive, random locations

Probe some servers, connect to least loaded

Want a balanced allocation of clients to servers
It’s almost balls ‘n’ bins...

- \( n \) bins (servers), \( m \) balls (clients)
- Balls arrive sequentially: probe \( d \) random bins, placed in least loaded
- Classic results, when \( m=n \):
  - \( d=1 \): max load \( O(\log n / \log \log n) \)
  - \( d=2 \): max load \( O(\log \log n) \)
  - \( d=\log^c n \): max load \( O(1) \)
Want *structured* choices

- Standard balls-and-bins requires *uniform random* choices
- But probing *close* servers is better
In this paper

a balls and bins model
with arbitrary correlations
between a ball’s choices
Past work

- [Kenthapadi & Panigrahy, SODA’06]: Balanced allocations on graphs

\[ \text{Max load } O(\log \log n) \text{ when graph almost regular with degree } n^{\Theta(1 / \log \log n)} \]

- We allow stronger structure and primarily address \( d = \Theta(\log n) \) choices
Our model

- Given a distribution over sets of bins
- Each ball $i$ draws set $B_i$ from the distribution, put ball in random least loaded bin in $B_i$

Example: nearby server selection

- Pick random point $p$ in the plane
- $B_i =$ set of servers within some distance of $p$

What restrictions on the $B_i$s yield a good max load?
Main Theorem

If we have, for every ball $i$,

- enough choices
  
  \begin{align*}
  d := |B_i| &\geq \Omega(\log n) \\
  \forall \text{ bins } j, \Pr[j \in B_i] &= \Theta \left( \frac{d}{n} \right)
  \end{align*}

- “balance”

then

w.h.p. max load = 1 after placing $\Theta(n)$ balls

... $O(1)$ after placing $n$ balls

Power: arbitrary correlations among choices!
Ex. 1: arbitrary patterns

- Index the bins: $0, 1, \ldots, n-1$
- Adversary picks indexes $\{b_1, \ldots, b_d\}$
- Ball picks random offset $R$ and probes bins $\{b_1 + R, \ldots, b_d + R\}$ mod $n$

 enough choices

 Set $d = \Theta(\log n)$

 Due to random offset, $\Pr[\text{bin } j \in B_i] = \frac{d}{n}$

 $\Rightarrow$ max load $O(1)$ w.h.p.
Ex. 2: server selection

- $n$ servers at random locations in unit square
- Each client $i$ picks random point $p_i$ in the plane; $B_i = \text{set of servers within distance } r \text{ of } p_i$

<table>
<thead>
<tr>
<th>enough choices</th>
<th>Pick $r$ to cover area $(\log n)/n$. Chernoff shows w.h.p. about $\log n$ servers in any $B_i$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>balance</td>
<td>$p$ uniform random: servers have equal chance of falling within $r$ [\Rightarrow \text{max load } O(1) \text{ w.h.p.}]</td>
</tr>
</tbody>
</table>
Other cases in paper

• Application to load balance in peer-to-peer networks

• More general version of theorem
  • No need for same number of choices for each ball
  • No need for set of choices $B_i$ to come from same distribution for each ball
Remainder of the talk

1. Proof overview
2. Lower bound
3. Open problems
Intuition: regain independence

- Want to show each ball finds an empty bin

**Independent choices**
- Current allocation of balls is irrelevant
- \( \log(n) \) choices \( \Rightarrow \) find empty bin w.h.p.

**Correlated choices**
- Current allocation matters!
- Show current allocation almost uniform-random
Problem: allocation is not uniform-random

- Suppose one ball so far, sequential choices
  - Solution: show placement process is dominated by uniform process that places more balls

These bins have
- same chance of being in $B_i$
- greater chance of getting ball if in $B_i$ because they're picked along with filled bin
Proof structure

• Two processes:

<table>
<thead>
<tr>
<th>P1(i)</th>
<th>allocation after (i) balls with structured choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>P2(i)</td>
<td>allocation after (ki) balls put in uniform-random empty bins</td>
</tr>
</tbody>
</table>

• Show inductively \(P1(i)\) is dominated by \(P2(i)\):

\[ P1(i)_j \leq P2(i)_j \quad \forall \text{ bins } j \text{ w.h.p.} \]
Inductive step, ball $i+1$

- “Smoothness”: $\Pr[\text{bin } j \text{ gets ball}] = \Theta\left(\frac{1}{f n}\right)$ if $j$ empty, $\forall j$

- Show smoothness w.h.p., using balance and $O(\log n)$ size ($\# \text{ free bins in } B_i$ concentrates)

- Smoothness implies domination:
  - Set up bipartite graph, nodes = outcomes with structured and uniform choices, resp.
  - Show perfect fractional matching with vertex weights exists for suitable $k \Rightarrow$ domination preserved
Lower bound

• Main theorem: $\Omega(\log n)$ choices and balance are sufficient for $O(1)$ max load

• Are $\Omega(\log n)$ choices necessary? Yes, almost:

There exist balanced choices of bins ($B_i$) with $|B_i|=d$ for which max load is

$$\geq \frac{\ln n}{\ln \ln n} \cdot \frac{1}{d} \quad \text{w.h.p.}$$

At best linear decrease in max load: no power of two choices result!
Open problems

• Close gap between upper and lower bounds
• Conjecture: can improve number of placed balls from $\Theta(n)$ to $(1-\epsilon)n$ with max load 1
• Theorem requires placement in uniform random least-loaded bin among choices. Relax that requirement?
• Finding a job!