COMP 7720 - Online Algorithms

Online Bin Packing

Shahin Kamali

Lecture 15 - Oct. 25, 2018

University of Manitoba
Review & Plan
Today’s objectives

- Competitive ratio of Next Fit and Worst Fit
- Lower bound for competitive ratio of any algorithm
- Lower bound for competitive ratio of Best Fit and First Fit
(our beautiful) Bin Packing problem
Bin Packing Problem

- The input is a **multi-set** of items of various sizes in range \((0,1]\).
- The goal is to pack these items into a minimum number of bins of uniform capacity.
- E.g., \(S = \{0.1, 0.2, 0.2, 0.3, 0.3, 0.4, 0.4, 0.5, 0.5, 0.5, 0.5, 0.6, 0.8, 0.8, 0.9\}\)
Offline Bin Packing

- The problem is NP-Hard
  - reduction from the partition problem
  - **Partition**: decide whether a multiset $S$ of positive integers can be partitioned into two subsets $S_1$ and $S_2$ s.t.
    - sum of the numbers in $S_1 = \text{sum of the numbers in } S_2 = X$
    - $S = \{3, 1, 3, 2, 3, 2, 3, 3, 4, 1\} \rightarrow S_1 = \{3, 2, 3, 3\} \quad S_2 = \{1, 3, 2, 4, 1\}$
  - The answer to partition is yes, if the items/integers can be placed in 2 bins of size $X$

- It is NP-hard to see whether a set can be packed in two or three bins!
Online Bin Packing

- The input is a **sequence** of items of various sizes which are revealed in a sequential, online manner.
- Item sizes are in range \((0, 1]\), and the goal is to pack these items into a minimum number of bins of uniform capacity.
- An online algorithm places items into bins, one by one, with no knowledge of future items.
- Decisions of an online algorithm are irrevocable.
Next Fit Algorithm

- Next Fit: Maintain one open bin at any given time.

- Place an incoming item in the open bin if there is space there; otherwise close the bin and open a new bin.

< 0.9 0.3 0.8 0.5 0.1 0.1 0.3 0.2 0.4 0.2 0.4 0.5 0.5 0.8 0.6 0.4 0.5 >
First Fit Algorithm

- First Fit: place an incoming item in the first bin which has enough space for the item.
- Open a new bin if such bin does not exist.
Best Fit Algorithm

Place an incoming item in the bin with the highest level (used space) which has enough space for the item.

< 0.9  0.3  0.8  0.5  0.1  0.1  0.3  0.2  0.4  0.2  0.4  0.5  0.5  0.8  0.6  0.4  0.5 ... >
Harmonic Algorithm

- Harmonic Algorithm classes: \((\frac{1}{2}, 1], (\frac{1}{3}, \frac{1}{2}], \ldots, (\frac{1}{K}, \frac{1}{K-1}], (0, \frac{1}{K}].\)
- Place members of each class separately from others.

Harmonic \(K = 4\)

\(< 0.9 \ 0.3 \ 0.8 \ 0.5 \ 0.1 \ 0.1 \ 0.3 \ 0.2 \ 0.4 \ 0.2 \ 0.4 \ 0.5 \ 0.5 \ 0.8 \ 0.6 \ 0.4 \ 0.5 \ldots >\)
Analysis Measures

- Competitive ratio of $A$ is the maximum value of $A(\sigma)/\text{OPT}(\sigma)$ among all sequences $\sigma$.
  - We are interested in the \textit{asymptomatic} competitive ratio where $\text{OPT}(\sigma)$ is arbitrary large.
- For a sequence of items with total size $S$, the cost of $\text{OPT}$ is at least $S$
  - It is equal to $S$ if opt can pack items in a way that all bins are completely full (which is not always possible)
  - For example, consider $\langle 0.51, 0.51, \ldots, 0.51 \rangle$. The cost of opt is $n$, while $S$ is $0.51n$. 
Competitive Analysis of Next Fit

- Competitive ratio of NextFit is at most 2.
  - In the final packing, total size of items in each two consecutive bins is larger than 1.
  - The cost of NextFit for serving $\sigma$ is smaller than $2S(\sigma)$ where $S(\sigma)$ is the total size of items in $\sigma$.
    - Assume $\text{cost}(\text{NextFit}) = k$
    - Each two consecutive bins have total size $> 1 \rightarrow$ total-size $S(\sigma)$ of items in $\sigma$ is more than $k/2 \rightarrow S > k/2 \rightarrow k < 2S$
  - $\text{OPT}(\sigma) \geq S(\sigma)$: Even when $\text{OPT}$ packs items tightly (with no wasted space), $S(\sigma)$ bins are required.

- Competitive ratio of NextFit is at most 2.
Competitive Analysis of Next Fit

- Competitive ratio of NextFit is at least 2.
  - Consider sequence $\sigma = \langle 0.5, \epsilon, 0.5, \epsilon, \ldots \rangle$.
  - The cost of NextFit for serving $\sigma$ is roughly $n/2$ ($n$ is the length of $\sigma$).
  - The cost of $\text{OPT}$ is roughly $n/4$.

\[ \begin{array}{cccccc}
\epsilon & \epsilon & \epsilon & \epsilon & \epsilon & \epsilon \\
5 & 5 & 5 & 5 & 5 & 5 \\
\end{array} \]

\[ \begin{array}{cccc}
5 & 5 & 5 & \phi \\
5 & 5 & 5 & \phi \\
\end{array} \]

**Theorem**

*Competitive ratio of NextFit is exactly 2.*
Competitive ratio of First Fit and Best Fit are both 1.7
- We see the proof later

Any **Almost Any Fit** strategy has a competitive ratio of 1.7
- Any Fit algorithm: algorithm which avoid opening new bin when one of the currently open bins have enough space
- Almost Any Fit algorithm: an AnyFit algorithm which avoid WorstFit strategy (i.e., avoid placing item in the least full bin)

Competitive ratio of Harmonic is \(1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{42} + \ldots \approx 1.691\)
An Easy Lower Bound ...

- Consider the following input
  \[\sigma = \langle 1/2 - \epsilon, 1/2 - \epsilon, \ldots, 1/2 - \epsilon, 1/2 + \epsilon, 1/2 + \epsilon, \ldots, 1/2 + \epsilon \rangle\]
  - \(m\) items
  - \(m\) items

- Consider the sub-sequence formed by the first \(m\) items
  - Cost of \(OPT\) for the sub-sequence is \(m/2\)
  - Cost of Alg is \(\alpha m\) for some \(\alpha\) so that \(1/2 \leq \alpha \leq 1\).
  - Competitive ratio will be at least \(\frac{\alpha m}{m/2} = 2\alpha\).

- Consider the whole sequence \(\sigma\)
  - Cost of \(OPT\) is \(m\)
  - Alg has opened \(\alpha m\) bins for the first \(m\) items, out of which \(m - \alpha m\) bins have two items (of size \(1/2 - \epsilon\)).
  - So, \(\alpha m - (m - \alpha m) = 2\alpha m - m\) bins have one item
  - Alg has to open \(m - (2\alpha m - m) = 2m - 2\alpha m\) new bins for second half.
  - The total cost will be \(\alpha m + 2m - 2\alpha m = 2m - \alpha m\)
  - The competitive ratio will be at least \(\frac{2m - \alpha m}{m} = 2 - \alpha\).
An Easy Lower Bound

Consider the following input
\[ \sigma = \langle 1/2 - \epsilon, 1/2 - \epsilon, \ldots, 1/2 - \epsilon, 1/2 + \epsilon, 1/2 + \epsilon, \ldots, 1/2 + \epsilon \rangle \]

To summarize, any sequence that opens \( \alpha m \) bins for the first half has competitive ratio at least \( \max\{2\alpha, 2 - \alpha\} \)

The value of \( \max\{2\alpha, 2 - \alpha\} \) is minimized for \( \alpha = 2/3 \), and the competitive ratio will be at least 4/3.

Theorem

No online bin packing algorithm (deterministic or randomized) can have a competitive ratio better than 4/3.
An Easy Lower Bound

- **Input:**
  \[ \sigma = \langle \frac{1}{2} - \epsilon, \frac{1}{2} - \epsilon, \ldots \frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon, \frac{1}{2} + \epsilon, \ldots, \frac{1}{2} + \epsilon \rangle \]
  \[ m \text{ items} \quad m \text{ items} \]

  - No algorithm can be within a ratio less than \( \frac{4}{3} \) of \( \text{OPT} \).
  - The worst-case sequence that we formed is **fixed** and not **adversarial**

- What if the online algorithm knows the whole input?
  - The lower bound still holds even if the algorithms knows the whole input
Shortcomings of Online Algorithms

- Online algorithms have two shortcomings against $\text{OPT}$

I) **Online constraint**: online algorithms do not know the future requests/items
   - We do not know the future items in bin packing, future points in clustering, etc.
   - Often randomization helps to cope with this!

II) **sequential constraint** online algorithms have to build their solution sequentially
   - They cannot change their previous decisions, e.g., an item placed in a bin, two points placed in the same cluster, etc.
   - Even if the algorithms knows the whole input, adversary just needs to decide where to end the input
   - Randomization does not help!
   - This is the main problem of online bin packing algorithms!
A better lower bound

- $\sigma_1 = \langle 1/2 - \epsilon, \ldots 1/2 - \epsilon, 1/2 + \epsilon, \ldots, 1/2 + \epsilon \rangle \rightarrow$
  
  competitive ratio $> 4/3$

- $\sigma_2 = \langle 1/6 - \epsilon, \ldots 1/6 - \epsilon, 1/3 - \epsilon, \ldots 1/3 - \epsilon, 1/2 + 2\epsilon, \ldots, 1/2 + 2\epsilon \rangle \rightarrow$
  
  competitive ratio $> 3/2$

- More complicated sequences results better lower bounds

Theorem

No online bin packing algorithm can have a competitive ratio better than 1.54037.

In the next class, we study the **weighting technique** to prove that Best Fit and First Fit have competitive ratio 1.7.