The Online Median Problem

Guntash Arora

April 30, 2015

1 PROBLEM STATEMENT

Definition 1. *A λ - approximate metric d satisfies following relaxed triangle inequality*

For any sequence of points

$$x_0, x_1, x_2 \dots x_m$$

$$d(x_0, x_m) \le \lambda * \sum_{0 \le i < m} d(x_i, x_{i+1})$$

holds.

1.1 PROBLEM STATEMENT

Objective of online median problem is to output a total ordering on U. So,if we want to solve k-median problem, we pick first k elements of this ordering

- Cost function is same as that of k-median problem
- Input : A set of points U and λ - approximate metric d
- Output : A total ordering on U

Definition 2. *Competitive Ratio : It is maximum over all possible choices of input instances and k, of ratio of cost of center given by first k element of this ordering to that of optimal k centers.*

2 ALGORITHM

2.1 DEFINITIONS

- Let
$$\lambda, \alpha, \beta, \gamma$$
denote four real numbers satisfying following inequalities
$$\lambda \geq 1$$
$$\alpha > 1 + \lambda$$
$$\beta \geq \frac{\lambda(\alpha - 1)}{\alpha - \lambda - 1}$$
$$\gamma \geq (\frac{\alpha^2\beta + \alpha\beta}{\alpha - 1} + \alpha)\lambda$$
- Value of ball $A = (x, r)$ is
$$\sum_{y \in A} (r - d(x, y)) * w(y)$$
- Child of ball $A = (x, r)$ is any ball $(y, \frac{r}{\alpha})$ such that $d(x, y) \leq \beta r$. Note that $(x, \frac{r}{\alpha})$ is also a child of A.
- $isolated(x, \phi)$ is $(x, max_{y \in U} d(x, y))$
- $isolated(x, X)$,where X is non-empty set, is $(x, \frac{d(x,X)}{\gamma})$
- For any non-empty list σ, head(σ) and tail(σ) denote first and last element of list σ,respectively

2.2 ALGORITHM

Let $Z_0 = \phi$. For i = 0 to n-1, execute the following steps

1. list $\sigma_i = \{A\}$, where A is maximum value ball in $\{isolated(x, Z_i) | x \in U\}$
2. while $tail(\sigma_i)$ has more than one child, append its maximum value child at the end of σ_i.
3. $Z_{i+1} = Z_i \bigcup \{center(tail(\sigma_i))\}$.

3 COMPETITIVE RATIO ANALYSIS

We will try to prove the following theorem, which bounds the ratio of cost between chosen centers and any arbitrary set of centers by $2\lambda(\gamma + 1)$.

Theorem 1. *For any configuration X,* $cost(Z_{|X|}) \leq 2\lambda(\gamma + 1) cost(X)$

3.1 Some more defintions

- Let z_i is added in i^{th} iteration.
- Cost(X,Y) = $\sum_{y \in Y} d(y, X) * w(y)$
- Let cell(x,X) for any point $x \in X$ is $\{y | y \in U, d(y, x) = d(y, X)\}$
- For any configuration X, point x in X and a set Y, in(x,X,Y) is $cell(x, X) \bigcap isolated(x, Y)$ and out(x,X,Y) is $cell(x, X) \setminus in(x, X, Y)$.
- For any configuration X and a set Y, in(X,Y) is $\bigcup_{x \in X} in(x, X, Y)$ and out(X,Y) is $U \setminus in(X, Y)$.

3.2 Main Lemmas

Notice that we can rewrite $cost(Z_{|X|}, U)$ as $cost(Z_{|X|}, in(X, Z_{|X|})) + cost(Z_{|X|}, out(X, Z_{|X|}))$ and $cost(X, U)$ as $cost(X, in(X, Z_{|X|})) + cost(X, out(X, Z_{|X|}))$. Now Using lemma 2,4 and 5, we can obtain the theorem mentioned above.

Lemma 1. *For any configuration X,point $x \in X$, and point y in $out(x, X, Z_{|X|})$, $d(y, Z_{|X|}) \leq \lambda(\gamma + 1).d(y, X))$*

Lemma 2. $cost(Z_{|X|}, out(X, Z_{|X|})) \leq \lambda(\gamma + 1).cost(X, out(X, Z_{|X|}))$

Lemma 3. *For any configuration X and a point x in X,* $cost(Z_{|X|}, in(x, X, Z_{|X|})) \leq \lambda(\gamma+1)[cost(X, in(x, X, Z_{|X|})) + value(isolated(x, Z_{|X|}))]$

Lemma 4. $cost(Z_{|X|}, in(X, Z_{|X|})) \leq \lambda(\gamma+1)[cost(X, in(X, Z_{|X|})) + \sum_{x \in X} value(isolated(x, Z_{|X|}))]$

We can obtain lemma 1 and 3 by writing relaxed triangle inequality and using definitions of $in(X, Z_{|X|})$ and $out(X, Z_{|X|})$. Lemma 2 and 4 can be obtained by summing up the equations in lemma 1 and 3 over all $x \in U$.

Lemma 5. *For any configuration X,* $\sum_{x \in X} value(isolated(x, Z_{|X|})) \leq cost(X)$

.

3.3 Proof of lemma 5

3.3.1 Overview of proof

Definition 3. *A ball (x,r) is covered iff $d(x, X) < r$*

Lemma 6. *For any uncovered ball A,* $cost(X, A) \geq value(A)$

Now to prove lemma 5, we will try to map element of X to some uncovered ball in $\{\sigma_i, 0 \leq i < k\}$. Let π be this mapping. We will try to ensure that these uncovered ball satisfy following properties.

1. For any pair of distinct points x and y in X, $\pi(x) \bigcap \pi(y) = \phi$.

2. For any point x in X, $value(\pi(x)) \geq value(isolated(x, Z_k))$.

Notice that by property 1 and lemma 6, we have $cost(X) \geq \sum_{x \in X} value(\pi(x))$. By property 2, we have that $\sum_{x \in X} value(\pi(x)) \geq \sum_{x \in X} value(isolated(x, Z_{|X|}))$. This will give prove lemma 5.

In order to satisfy property 1, we first prune all the list such no two ball in two distinct list intersect with each other. By intersection of two balls (x,r) and (y,s), it is meant that $d(x, y) \leq r + s$. To do pruning, we make use of following lemmas,

Lemma 7. *Let A = (x,r) belong to σ_i. Then, d(x,Z_i) $\geq \gamma r$*

If A is $head(\sigma_i)$, then above lemma is true by defintion. For any arbitrary element in list, we can prove by induction on its position in the list.

Lemma 8. *Let A = (x,r) belong to σ_i and B = (y,s) belong to σ_j. If $i < j$ and $d(x, y) \leq r + s$,then following holds*

1. $radius(head(\sigma_i)) \leq \frac{r}{\alpha}$
2. $A \neq tail(\sigma_i)$
3. *the successor A in σ_i,call it C, satisfies* $value(C) \geq value(head(\sigma_j))$

Let τ_i be the list obtained after pruning. In a single pruning step, if some ball A in σ_i intersect with some ball B in σ_j, we set τ_i to suffix of σ_i starting at the succesor of A in $sigma_i$. Notice that since $A \neq tail(\sigma_i)$, succesor of any such A always exist. Then following holds,

Lemma 9. *Let A = (x,r) belong τ_i and B = (y,s) belong to τ_j.Then if i<j, d(x,y) > r+s*

Lemma 10. *Each sequence τ_i is non-empty*

Both lemmas follow from definition of the pruning.

Following lemmas establish relationship between value(head(τ_i)) and value(isolated(x,Z_k)).

Lemma 11. *Let x be a point and assume that $0 \leq i < j \leq n$.Then value(isolated(x,Z_i)) $\geq$ value(isolated(x,Z_j))*

This is trivially true because $Z_i \subset Z_j$.

Lemma 12. *Let x be a point and assume that $0 \leq i < k$.Then value(head(σ_i)) $\geq$ value(isolated(x,Z_k))*

If $x \in Z_i$, then RHs is zero, so we have nothing to prove. Else,
$value(head(\sigma_i)) \geq value(isolated(x, Z_i)) \geq value(isolated(x, Z_k))$, by using definition of head of σ_i and lemma 11.

Lemma 13. *Let x be a point and assume that $0 \leq i < k$.Then value(head(τ_i)) $\geq$ value(isolated(x,Z_k))*

We prove that the claim holds before and after each iteration of the pruning procedure. Initially, $\tau_i = \sigma_i$ and the claim holds by Lemma 12. If the claim holds before an iteration of the pruning procedure, then it holds after the iteration by part 3 of Lemma 8.

3.3.2 Mapping Construction

Let I denote set of all indices i in $\{k\}$ such that some ball in τ_i is covered.

Step 1:

1. Map each i in I to to a point x ∈ X belonging to last covered ball in τ_i
2. Map each i in $\{k\} \setminus I$ to any unmatched point in X.

Step 2:

1. Map each x that is matched to an index i in $\{k\} \setminus I$ to head(τ_i).
2. Map each x that is matched to an index i in I to successor of last covered ball in τ_i. If last covered ball is tail(τ_i), then map x to A = (x,0).

Let π be the final mapping. Now property 1 holds because in pruned lists, now two balls intersect. For property 2, it is each to see that for each x that is matched to an index in $\{k\} \setminus I$, property 2 is true using lemma 13. Otherwise, if the last covered ball in τ_i is the tail and $x \in tail(\tau_i)$, then tail will have another child. This implies that x is the center of tail and $x \in Z_{i+1}$, so RHS is 0. If not , then predecessor of $\pi(x)$, say (y,r), exists and contains x. Consider a ball B = $(x, \frac{r}{\alpha})$. Let C = (x,s) = $isolated(x, Z_k)$. Then, we claim that $\frac{r}{\alpha} \geq s$, which implies $C \subset B$ and $value(B) \geq value(C)$. First $d(x, z_i) \geq \gamma s$, by definition of C. Also ,

$$\begin{aligned} d(x, z_i) &\leq \lambda[d(x, y) + d(y, z_i)] \\ &\leq \lambda r + \beta\lambda(r + \frac{r}{\alpha} + ...) \\ &\leq (1 + \frac{\alpha\beta}{\alpha - 1})\lambda r \end{aligned} \tag{3.1}$$

Last quantity is less than $\frac{\gamma r}{\alpha}$ by definition of γ. This proves the fact that mapping that we created satisfies both properties and hence proof of lemma 5 is complete.