Institut für Informatik
Dr. Melanie Schmidt
Clemens Rösner

UNIVERSITÄT BONN

Cluster Analysis
Winter 2018/19

Problem Set 6

Please hand in your solutions for this problem set via email (roesner@cs.uni-bonn.de) or personally at Room 2.060 until *Tuesday, 27th of November.*

Problem 1

Let us look at instances of the fair k-center problem. Show that the factor between the value of the optimal fair solution and the optimal unfair solution can be unbounded.

Problem 2

What changes when we instead of calling **farthest-first-traversal**$((P^0, d), k)$ in step 3 of the **fair-k-center** algorithm, we call **farthest-first-traversal**$((P, d), k)$ to compute C?

Problem 3

We would like to extend the fair k-center problem to more general settings. Specifically we would like to replace the restriction that the clusters are $(\frac{1}{t}, 1)$-balanced for some $t \in \mathbb{N}$. Instead we would prefer to require that the clusters are (ℓ, u)-balanced for arbitrary $\ell \leq u \in \mathbb{Q}$. Unfortunately we will see that this scenario seems to be more difficult.

- Assume that we know how the optimal fair clustering clusters the points in P^0. Show that we can then compute a 3-approximation.

One approach could now be to first compute an approximate unfair solution on P^0 and then try to add the points from P^1.

- Show specifically for the case $\ell = ratio(P, 0) = u$ that there are instances where it is impossible to make such a clustering obtained on P^0 fair by adding the points from P^1.

We want to keep the focus on the exact case where we have $\ell = ratio(P, 0) = u$. Assume that $ratio(P, 0) = \frac{a}{a+b}$ for some coprime integers $a \leq b$.

- Show that in every fair cluster the number of points from P^0 must be an integer multiple of a.

For the next task we assume that we know an approximation algorithm for the capacitated k-center problem. The capacitated k-center problem is in addition to P, d and k given a capacity cap and demands that each cluster contains at most cap points.

- Show how to compute a clustering on P^0, where the number of points in each cluster is an integer multiple of a and whose maximal radius is in $O(opt)$, where opt denotes the value of the optimal fair clustering.

- Use this clustering on P^0 to compute a fair clustering. What approximation factor do you obtain?