CS-E400204: Approximation Algorithms **Spring (Fifth Period) 2022**

Lecture 8 (May 18): Parametric Pruning for Metric $k-$Center

Lecturer: Kamyar Khodamoradi *Scribe: Sampo Niemelä*

8.1 Introduction

Definition 1 (metric) *Metric on set X is a function $d : X \times X \to \mathbb{R}_{\geq 0}$ for which all of the following holds for all $x, y, z \in X$:*

1. $d(x, x) = 0$
2. $d(x, y) = d(y, x)$
3. $d(x, z) \leq d(d, y) + d(y, z)$

Example 1 (Metric $k-$center Problem)

- *Input:*
  - *a complete graph $G = (V, E)$*
  - *a metric cost function $c : E \to \mathbb{Q}_{\geq 0}$*
  - *a parameter $k \in \mathbb{N}_{>0}$*
- *Output:*
  - *A set $S^* \subset V$ with $|S^*| \leq k$ and that minimizes $\max_{v \in V} c(v, S^*)$*

In the above definition of the problem, $c(v, S) \triangleq \min_{u \in S} c(v, u)$.

8.2 Metric $k-$center

Target of this section is to produce a 2-approximation algorithm for the METRIC $k-$CENTER problem.

Start the analysis of the problem by noting that since $OPT = c(v, u)$ for some distinct $v, u \in V$, we have that $OPT \in \{c(u, v) | (u, v) \in E\}$. Because G is a complete graph, i.e. $|E| = \binom{|V|}{2}$, this then shows that OPT is one of $|E| = \binom{|V|}{2}$ values.

Label the edges of G by $1, 2, ..., |E| = m$ and order them such that

$$c(e_1) \leq c(e_2) \leq .... \leq c(e_m)$$

and let j be such that $c(e_j) = OPT$. For any $i \in \{1, 2, ..., m\}$, let $G_i = \{V, E_j\}$ be a graph with $E_j = \{e \in E : c(e) \leq c(e_j)\}$.

Definition 2 (dominating set) *For graph $G = (V, E)$, set $D \subset V$ is a dominating set iff for all $v \in V \setminus D$, there is a $u \in D$ such that $(v, u) \in E$.*

Note that for G_j, with $c(e_j) = OPT$, any $k-$element dominating set D is a solution to the METRIC $k-$CENTER -problem in G. This turns the problem into the Dominating Set problem:

Example 2 (Dominating Set Problem)

- *Input:*
  - *a graph $G = (V, E)$*
  - *a parameter $k \in \mathbb{N}_{>0}$*
- *Output:*
  - *A dominating set $D \subset V$ with $|D| \leq k$.*

Unfortunately, this problem is NP-hard. Assuming P≠NP, it is also known to be hard to approximate with ratio better than $\log(|V|)$. Therefore we want to "relax" our definition of dominating set. This is done through the concept of a square graph.

Definition 3 (square graph) *Square of graph $G = (V, E)$ is the graph $G^2 = (V, E')$, where $E' = \{(u, v) \in V \times V : d(u, v) \leq 2\}$.*

Definition 4 (independent set) *For graph $G = (V, E)$, a set $I \subset V$ is independent iff for all distinct $u, v \in I$, $(u, v) \notin E$.*

A simple observation about independent sets is that any maximal independent set is also a dominating set. This is because if for independent set I there are no vertices that can be added to I, then every vertex outside the independent set must be adjacent to some vertex in I.

Maximal independent set can also be found greedily by the following algorithm:

Algorithm 1 Greedy for maximal independent set

Input: Graph $G = (V, E)$
Output: A maximal independent set $I \subset V$

```
I ← ∅
while V ≠ ∅ do
    choose v ∈ V
    I ← I ∪ {v}
    remove all neighbors of v from V
end while
return I
```

Algorithm 2 A 2-approximation for METRIC $k-$CENTER PROBLEM via maximal matching

---

Input: Graph $G = (V, E)$, metric $c : V \to \mathbb{Q}_{\geq 0}$
Output: A set $S^* \subseteq V$

sort edges such that $c(e_1) \leq c(e_2) \leq ... \leq c(e_m)$
for $i = 1$ to m **do**
  create G_i^2
  $I_i \leftarrow$ maximal independent set of G_i^2
  if $|I_i| \leq k$ **then**
    return $S^* \leftarrow I_i$
  end if
end for

Claim 1 *For maximal independent set I_i of G_i^2, and minimum size dominating set D_i of G_i, $|I_i| \leq |D_i|$*

Proof. For the sake of contradiction, assume $|I_i| > |D_i|$. Let $D_i = \{v_1, v_2, ..., v_d\}$. for each $v_j \in D_i$, consider its closed neighborhood $N_{G_i}[v_j]$. D_i is a dominating set so vertex in V is in some $N_{G_i}[v_j]$. By the pigeonhole principle, because $|I_i| \geq |D_i|$, there are two vertices $u, v \in I_i$ that are in the same closed neighborhood $N_{G_i}[v_j]$. uv_jv is a path on G, so $d(u, v) \leq 2$. Then u, v are adjacent in G_i^2, which contradicts I_i being an independent set. ■

For rest of this section, let $i, j \in \mathbb{N}$ be such that I_i is the output of algorithm 2 and $c(e_j) = OPT$.

Claim 2 $c(e_i) \leq c(e_j) = OPT$

Proof. Proof by contradiction. Assume $c(e_i) > c(e_j)$, i.e. $j < i$. Then, $|I_j| > k$. The optimal solution S^*_{OPT} is a dominating set of G_j. By claim 1 this then implies that $|S^*_{OPT}| \geq |I_j| \geq k$, which contradicts the feasibility of S^*_{OPT}. ■

Combining these claims gives us the following theorem.

Theorem 1 *Algorithm 2 is 2-approximation for* METRIC $k-$CENTER.

Proof. I_i is a dominating set of G_i^2 and therefore for all $v \in V \setminus I_i$ and some $u \in I_i$, in G_i we have that $d(u, v) \leq 2$. If $d(u, v) = 1$, $c(u, v) \leq c(e_i) \leq OPT$. If $c(u, v) = 2$, then there is some vertex $w \in W$ such that uwv is a path in G_i, i.e. $c(u, w) \leq OPT$ and $c(w, v) \leq OPT$. c is a metric and therefore $c(u, v) \leq c(u, w) + c(w, v) \leq 2 \cdot OPT$. This shows that

$$\max_{v \in V} c(v, I_i) \leq 2 \cdot OPT.$$

■

This bound is tight as can be seen from the following example.

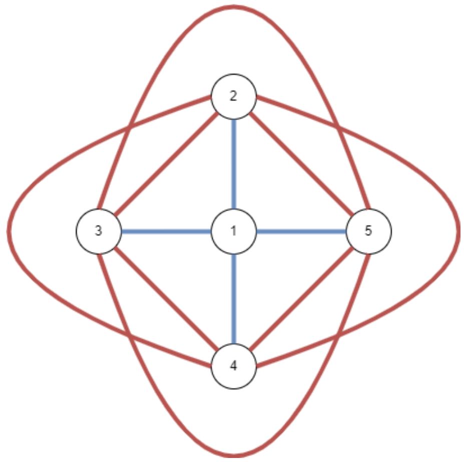

Figure 8.1: Blue edges have weight 1 and red edges weight 2

Optimal solution for the graph in figure 8.1 and any $k \geq 1$ is $S^* = \{v_1\}$ with $OPT = 1$. However, during the first iteration of the algorithm, $I_1 = \{v_5\}$ would be a possible output with and $d(v_2, v_5) = 2$, i.e. $ALG = 2$.

Theorem 2 *Assuming NP≠P, for any $\epsilon > 0$, there does not exist a $(2-\epsilon)-$approximation for the metric k-center problem.*

Proof. We use the fact that the dominating set problem does not have a $(2-\epsilon)-$ *approximation*. Assume that there is a algorithm that produces a $(2-\epsilon)-$approximation for the metric $k-$center problem.

Let $G = (V, E)$ be some graph and let $G' = (V, E \cup E')$ be a complete graph. Further let $c(u, v) = 1$ if $(u, v) \in E$ and $c(u, v) = 2$ otherwise. c satisfies the triangle equality, since $c(u, w) + c(w, v) \geq 2$ for all distinct $u, w, v \in V$. If G has a dominating set of size k, the algorithm will return a solution for G' with cost smaller than 2. Otherwise the algorithm will produce a solution with cost larger or equal to 2. Thus the algorithm can be used to deduce whether G has a dominating set of size k, which is known to be not possible in polynomial time. ■

8.3 Metric Weighted-Center

In this chapter we consider a generalisation of the metric k-center problem and produce a 3-approximation algorithm for the generalisation.

Example 3 (Metric weighted-center Problem)

- *Input:*
  - *a complete graph $G = (V, E)$*
  - *a metric cost function $c : E \to \mathbb{Q}_{\geq 0}$*
  - *a vertex weight function $w : V \to \mathbb{Q}_{\geq 0}$*
  - *a parameter $W \in \mathbb{Q}_{>0}$*

- *Output:*
  - *A set $S \subset V$ that satisfies $\sum_{v \in S} w(v) \leq W$ and that minimizes $\max_{v \in V} c(v, S)$*

Definition 5 (closed neighborhood) *In graph $G = (V, E)$, the closed neighborhood of $u \in V$ is $N_G[u] = N_G(u) \cup \{u\}$.*

Definition 6 *For graph $G_i = (V, E)$, let $s_i(u)$ denote the lightest vertex in $N_{G_i}[u]$.*

Algorithm 3 A 3-approximation for METRIC WEIGHTED-CENTER PROBLEM

Input: Graph $G = (V, E)$, metric $c : V \to \mathbb{Q}_{\geq 0}$, weight function $w : V \to \mathbb{Q}_{\geq 0}$, parameter $W \in \mathbb{Q}_{\geq 0}$.
Output: A set $S \subseteq V$

```
sort edges such that c(e_1) ≤ c(e_2) ≤ ... ≤ c(e_m)
for i = 1 to m do
  create G_i^2
  I_i ← maximal independent set in G_i^2
  S_j → {s_i(u) : u ∈ I_i}
  if w(S_i) ≤ W then
    return S ← S_j
  end if
end for
```

Theorem 3 *Algorithm 3 is 3-approximation.*

Proof. Let j be such that $c(e_j) = OPT$ and i be such that S_i is the output of the algorithm. Let D_j be the optimal solution to the problem. As we had in the previous section, for all $v \in V$, $c(v, I_i) \leq 2 \cdot OPT$. Also, $c(s_i(u), u) \leq c(e_i) \leq c(e_j)$. Therefore $c(v, s_i(u)) \leq 3 \cdot OPT$ and since this holds for all $v \in V$, $ALG \leq 3 \cdot OPT$. ■