CSCI 2510 - Problem Set 3

Shay Mozes
shay@cs.brown.edu

Problem from class

Consider the approximation algorithm for the metric k-center problem suggested in class:

```
1: Let $e_1, e_2, \ldots, e_{\binom{n}{2}}$ be the enumeration of edges of $G$ in nondecreasing order of costs.
2: Let $G_i = (V, E_i)$ be the subgraph of $G$ with $E_i = \{e \in E : cost(e) \leq cost(e_i)\}$
3: Compute a maximal independent set $M_i$ in each graph $G_i$
4: Let j be the smallest index such that $|M_j| \leq k$
5: Return $M_j$
```

It turns out that my intuition in class was wrong. We will show that this is indeed a 2-approximation for the problem.

Lemma 1. *The value of a solution produced by the algorithm is at most e_j, where j is defined in line 4 of the algorithm*

Proof. We first note that j in line 4 is well defined since $G_{\binom{n}{2}} = G$ is the complete graph, so it has a maximal independent set of size 1. Hence j is well defined and the algorithm returns a set of at most k nodes in G. Since the algorithm returns a maximal independent set S in G_j, any node $v \in G$ is adjacent in G_j to some $s \in S$, so by the definition of G_j, $d(v, S) \leq d(v, s) \leq e_j$ as required. $\square$

Consider the optimal solution OPT, and let $e_{j^\star}$ be the edge whose cost is the value of OPT. This implies that $G_{j^\star}$ can be covered by at most k stars whose centers are the nodes in OPT. Let ℓ be the largest index such that $e_\ell \leq 2e_{j^\star}$.

Lemma 2. *G_ℓ can be covered by at most k cliques.*

Proof. Consider a cover of $G_{j^\star}$ by at most k stars and let u and v be two nodes incident to the same star whose center we denote by s. By the triangular inequality and by the definition of $G_{j^\star}$, the distance between u and v satisfies $d(u, v) \leq cost(u, s) + cost(s, v) \leq 2e_{j^\star}$. Hence, by the definition of ℓ we get that $d(u, v) \leq cost(e_\ell)$, so the edge (u, v) appears in G_ℓ. This implies that the nodes incident to the same star in $G_{j^\star}$ form a clique in G_ℓ and the lemma follows. $\square$

Since G_ℓ can be covered by at most k cliques, any independent set in G_ℓ has at most one vertex from each clique, so its size is at most k. This shows that the j in line 4 of the algorithm is at most ℓ, so by Lemma 1 the value of the solution produced by the algorithm is at most $e_\ell \leq 2e_{j^\star}$ which proves the approximation guarantee. We have already seen that this is the best possible, assuming $P \neq NP$.