Ros Roger, Alexandre

May 19th, 2023

24. Clustering. In the CLUSTERING VERTEX DELETION problem, we are given a graph G and a positive integer k as input, and the objective is to check whether there exists an $S \subseteq V(G)$ with size at most k such that $G[V \setminus S]$ consists of a collection of cliques. A graph is a collection of cliques if, and only if, every connected component of the graph is a clique.

Given the parameterization $\kappa(G, k) = k$, show CLUSTERING VERTEX DELETION is in FPT.

Solution:

We are going to provide a polynomial-time reduction from CVD to HITTING. Let us remind ourselves of the HITTING problem:

Hitting. Given a set S with sets over a domain $U = \{1 \dots n\}$ as elements, and given a positive integer k as input, check whether there exists a subset $T \subseteq U$ such that $|T| \leq k$ and $\forall S' \in S$, $\exists x \in T$ s.t. $x \in S'$.

In this reduction, let S be the set of all non-cyclic paths of length 2 in G, where each path is represented by the set of the vertices in the path (set of size 3). Output S, $U = \{1 \dots |V(G)|\}$ and k.

- If there is a deletion subset T for CVD, then in HITTING the same subset will hit all sets. If there was a set that wasn't hit, then that means that there exists two vertices with a distance greater than one (not a deletion cluster).

- If there is a subset T that hits all sets of size 3, then the same subset is a deletion subset for CVD, since there exists no path of length 2 where none of the three vertices are hit (in the deletion subset).

The time-complexity of the reduction boils down to computing all non-cyclic paths of length two. This can be easily accomplished in polynomial time w.r.t. G by one BFS for every vertex. Therefore $\mathcal{O}(n(n+m))$. HITTING can be solved in $\mathcal{O}((2d^k - 1)cs)$ where $s = |U| + \text{sum}(S_i)$, $d = \max\{|A| \mid A \in S\}$ and c is a constant.

Using the same algorithm in our solution, we would be able to solve CVD in $\mathcal{O}(n(n + m) + 3^k\ p(|G|))$ where $p(|G|)$ is the number of possible paths of length 3, which must be polynomial w.r.t. the size of the graph (we gave an algorithm to generate them all).

Therefore CLUSTERING VERTEX DELETION is in FPT. ■