Examen Final

2. **Node hubs**. The students of the MEI PTDMA course have to design an algorithm to provide a list of contacts pointing to a strong influential group in a social network. To pass the course it is enough to design an App that identifies a smallest *node hub* in the social network. A node hub is a subset of participants whose removal leaves a *fully connected* vertex subset. For this exercise you can assume that the social network is an undirected connected graph and that a vertex subset *X* is *fully connected* if and only if, for every pair of distinct vertices in *X*, either they are connected by an edge or they are both connected by an edge to a third vertex in *X*.

   (a). Design a constant factor approximation algorithm for the problem. Justify its correctness and efficiency.

   Suposem que tenim un graf $G = (V, E)$ en el que volem trobar el *node hub*. Així, el nostre problema es pot formalitzar com:

   $$\min_{\substack{S \subseteq V \\ S\ node\ hub \\ a\ G}} \{|S|\}$$

   A més, podem definir:

   $$S\ node\ hub \Leftrightarrow \forall x, y \in (V - S)(d_G(x, y) \leq 2)$$

   Ara, definim el *graf de distàncies 2* com:

   $$G_2 = (V, E_2) \qquad E_2 = \{(u, v) \in V | d_G(u, v) > 2\}$$

   En aquest nou graf, podem definir:

   $$S\ vertex\ cover \Leftrightarrow \forall (u, v) \in E_2 (u \in S \lor v \in S)$$

   Així, demostrarem que S és *node hub* a G si i només si S conforma un *vertex cover* a G_2. Formalment:

   $$S\ node\ hub\ a\ G \Leftrightarrow S\ vertex\ cover\ a\ G_2$$

   Primer demostrarem $S\ node\ hub\ a\ G \Rightarrow S\ vertex\ cover\ a\ G_2$.

   Si S és *node hub* a G, tindrem $\forall x, y \in (V - S)(d_G(x, y) \leq 2)$. Per tant, quan generem el graf G_2 tindrem que tots aquests vèrtex seran vèrtex aïllats. Així, totes les arestes que quedin seran entre vèrtex de S, de manera que S conformarà un *vertex cover*.

Ara demostrarem $S\ node\ hub\ a\ G \Leftarrow S\ vertex\ cover\ a\ G_2$.

Això ho demostrarem per reducció a l'absurd. Suposem que S no forma un *node hub* a G. Això implica que $\exists x, y \in (V - S)(d_G(x, y) > 2)$. Així, al graf G_2 tindrem que per a aquesta parella (x, y) hi ha una aresta, i com que $x, y \in (V - S)$, $x, y \notin S$. Així, com a mínim hi haurà una aresta a G_2 que no estarà recoberta per S, de manera que S no serà un vertex cover a G_2.

Així dons, queda demostrat que:

$$S\ node\ hub\ a\ G \Leftrightarrow S\ vertex\ cover\ a\ G_2$$

Per tant, podem expressar el problema com:

$$\min_{\substack{S \subseteq V \\ S\ vertex\ cover \\ a\ G_2}} \{|S|\}$$

Que és un problema conegut amb un algoritme de programació lineal entera que ens ofereix una 2-aproximació a la solució òptima.

(b). Prove that the decision version of the problem when parameterized by the size of the node hub belongs to FPT.

Si utilitzem la definició del problema de *node hub*, tenim:

$$NODEHUB = \{\langle G, k\rangle | \exists S \subseteq V, |S| \leq k \forall x, y \in (V - S)(d_{V-S}(x, y) \leq 2)\}$$

Si fixem el subconjunt $S \subseteq V$, el problema *NODEHUB* és clarament a P, ja que podem decidir si S és un *node hub* realitzant una cerca en profunditat acotada per a cada parella de vèrtex de $(V - S)$. Així, podem decidir *NODEHUB* generant els $2^{k+1} - 1$ subconjunts de mida $\leq k$ i decidir cada subconjunt en temps $O(n^2)$. Per tant, el temps d'execució d'aquest algoritme és en el cas pitjor $O(2^k n^2)$.

En global, el cost de l'algoritme és $O(2^k n^2)$, polinòmic en la mida de l'entrada i exponencial en el valor del paràmetre. Per aquest motiu, podem concloure que el problema de *NODEHUB* pertany a FPT, és a dir:

$$NODEHUB \in FPT$$