4.- Weighted vertex cover with penalties

Consider a connected undirected graph G = (V,E), having an edge cost $c_e \geq 0$, for each e $\in$ E, and a weight $p_i \geq 0$, for each i $\in$ V . The values c_e, for e = (u, v), represent the cost of not covering edge e. The value p_u is an estimation of the cost of including vertex u in a set.

We want to solve a variant of the vertex cover problem, the partial vertex cover problem in which we look for a set S of vertices that minimizes the total weight of its vertices, plus the total weight of the edges that S does not cover.

Provide a 3-approximation for this problem.

Solution:

Primero, formulamos el problema como un problema de programación entera y más tarde aplicaremos una relajación a programación lineal para dar una 3-aproximación al problema.

Para ello introducimos una variable X_i para cada vértice v_i, tal que $X_i = 1$ si y solo si $v_i \in S$. En otro caso X_i=0.

De la misma manera incluimos la variable Y_{ij} para cada arista e$\in$E, de manera que $Y_e = 1$ si y solo si la arista está cubierta. En otro caso Y_{ij}=0.

La formulación en IP del problema sería la siguiente:

$$\min \sum_{i=1}^{n} p_i X_i + \sum_{(i,j)\in E,\, i,j\notin S} c_{(i,j)} (1 - Y_{ij})$$
$$s.t.\ X_i + X_j \geq Y_{i,j}\ \forall(i,j) \in E$$
$$X_i \in \{0, 1\}\ \forall i \in V$$
$$Y_{i,j} \in \{0, 1\}\ \forall(i,j) \in E$$

Cabe destacar que la primera condición es la que nos garantiza las propiedades de un vertex cover, es decir, si añadimos un vértice a S, entonces cubrimos las aristas incidentes en ese vértice.

Ahora, aplicamos la relajación a LP, cambiando la dos últimas condiciones por las siguientes:

$$X_i \in [0, 1]\ \forall i \in V$$
$$Y_{i,j} \in [0, 1] \forall(i,j) \in E$$

Lo siguiente que debemos hacer es encontrar un valor α a partir del cual redondear cuando encontremos una solución al problema de LP. Sea X_i* y $Y_{i,j}$* los valores de la solución óptima del problema en LP y X_i' y $Y_{i,j}$' los resultados redondeados.

De este manera podemos decir que si $Y_{i,j}^* >= \alpha$, donde $\alpha \in [0, 1]$, entonces $Y_{i,j}' =1$ y si no $Y_{i,j}' = 0$. Para la variable X debemos ajustar un poco más ya que para cumplir la primera condición se tiene que cumplir que $X_i' = 1$ cuando $X_i^* >= \alpha/2$. Es fácil ver que $\alpha/2$ es el máximo valor que podemos tomar para asegurar que siempre se cumpla que $X_i' + X_j' >= Y_{i,j}'$.

Con esto podemos acotar los valores aproximados en relación a los valores óptimos de la siguiente manera:

$$1 - Y'_{i,j} \leq (1-\alpha)^{-1} \cdot \left(1 - Y^*_{i,j}\right) \quad \forall (i,j) \in E$$

$$X'_i \leq \frac{2}{\alpha} \cdot X^*_i \quad \forall i \in V$$

Resumiendo, si calculamos los valores de los nuevos sumatorios cogiendo estos valores para cada arista de E y cada vértice de V, obtenemos lo siguiente:

$$\sum_{i \in S} P_i \cdot X'_i + \sum_{(i,j) \in E,\, i,\, j \notin S} c(i,j) \cdot \left(1 - Y'_{i,j}\right) \leq \left(\frac{2}{\alpha} \sum_{i \in S} P_i \cdot X^*_i\right) + (1-\alpha)^{-1} \cdot \sum_{(i,j) \in E,\, i,\, j \notin S} c(i,j) \cdot \left(1 - Y^*_{i,j}\right)$$

Podemos apreciar que esto es optimizado cuando $\frac{1}{\alpha} = (1 - \alpha)^{-1}$. Por lo tanto, si resolvemos esta igualdad, llegamos a la conclusión de que $\alpha = 2/3$. Si sustituimos este valor en la la desigualdad de sumatorios de arriba, obtenemos lo siguiente:

$$3 \cdot \sum_{i \in S} P_i \cdot X^*_i + 3 \cdot \sum_{i,j \in E,\, i,j \notin S} c(i,j) \cdot \left(1 - Y^*_{i,j}\right) \leq 3 \cdot LP_{OPT} \leq 3 \cdot IP_{OPT}$$

Entonces, podemos observar que si redondeamos los valores de las variables X* i Y*, obtendremos como mucho 3 veces el óptimo. Por lo tanto, concluimos que tenemos una 3-aproximación, y como lo hemos resuelto para la mejor α, entonces tenemos la mejor aproximación posible.