"A Roma invirtiendo aristas"

Hansel Blanco, Elena Rodríguez

1. Descripción del problema

Dado un grafo dirigido G y una función de costo de inversión definida sobre los arcos de G, determinar el menor costo para que sea accesible algún vértice desde el resto, donde el costo total está determinado por el del mayor arco que hay que invertir.

2. Transformando el problema

El problema en cuestión es equivalente a resolver el problema:

Dado un grafo dirigido G y una función de costo de inversión definida sobre los arcos de G, determinar el menor costo para que en el grafo G^T (donde cada arco tiene como costo de inversión el de su arco transpuesto en G) pueda desde un vértice accederse al resto, donde el costo total está determinado por el del mayor arco que hay que invertir.

2.1. Demostración

Para demostrar esto se demostrará que toda solución del problema original es una solución del nuevo problema y viceversa.

Sea s_1 una solución del problema inicial, es decir, el menor costo posible con el que puede lograrse invirtiendo arcos de G que sea accesible algún vértice de G desde el resto de vértices. Sea v un vértice de G que cumple esto y sean $c_1, c_2, ..., c_{|V|-1}$ los caminos desde todos los vértices u de G hacia v. Para todo camino c_i este puede tener arcos de G y arcos de G invertidos cuyo costo de inversión es menor o igual que s_1. Si el camino está compuesto solo por arcos de G : Se puede demostrar que existe un camino de u a v en G si y solo si existe un camino de v a u en G^T. Luego si c_i es un camino de u a v en G existe un camino ${c_i}^t$ de v a u en G^T.

Sea un par de vértices cualesquiera u, v y un camino de u a v en G nombrado c la secuencia de arcos $<< u, a_1 >, ..., < a_n, v >>$, en G^T existe un camino c' que es la secuencia $<< v, a_n >, ..., < a_1, u >>$ al todos los arcos del camino c' estar invertidos respecto al sentido en el grafo original G por definición de G^T, luego, existe un camino de v a u en G^T.

Si el camino c_i del vértice u al vértice v en G tiene arcos de G invertidos: Sea $e =< x, y >$ cualquiera de los arcos de G, tal que su transpuesto está en c_i, entonces el arco $< y, x >$ está en G^T y su costo de inversión es el mismo que el de $< x, y >$ en G, por tanto puede invertirse sin empeorar el costo. Invirtiendo todos los arcos en esta situación se obtiene un camino de v a u en G^T.

Además s_1 es una solución óptima para el nuevo problema, pues si no lo fuera, con un procedimiento similar al anterior se obtendría con un menor costo un vértice de G que fuera accesible desde todos los demás y esto no puede ocurrir porque s_1 es una solución óptima del problema original.

La demostración es análoga para el caso inverso.

3. Solución inicial

La solución que se propone consiste en, a partir de una ordenación ascendente de los costos de inversión de los arcos, explorar por cada costo c si existe algún vértice que llegue al resto invirtiendo potencialmente los arcos de costo menor o igual que c.

La modelación empleada consiste en construir un grafo G' donde para cada par de vértices u, v donde existe un arco e que va de u a v con costo de inversión w_e en G, se asignará en G' el arco en este mismo sentido con costo 0, además, se añadirá el arco $< v, u >$ con costo w_e.

Para cada costo c se trabajará con el grafo G_c que es el grafo G' quitándole las aristas de costo mayor que c. Esto garantiza no utilizar arcos de G' con costo mayor que c en los recorridos por el grafo.

3.1. Correctitud

Como se exploran todos los posibles costos, se garantiza encontrar la solución si existe. Además al hacerse en orden ascendente de los costos, una vez que se encuentre una solución factible, será la óptima, pues cualquier costo válido que se encuentre después será mayor que el primero factible encontrado.

A continuación se ofrecen detalles del algoritmo para determinar si un costo c es factible para solucionar el problema y su correctitud.

Para cualquier grafo G, donde V_r es el conjunto de vértices de G desde los que se puede llegar a todos los demás vértices de G y P es la primera componente fuertemente conexa (CFC) de un orden topológico cualquiera de las CFC de G puede demostrarse que $V_r \subseteq P$ y que si $V_r \neq \emptyset$, entonces $P \subseteq V_r$.

1. Lema 1: $V_r \subseteq P$
   Demostración:
   Si $V_r = \emptyset$, entonces $V_r = \emptyset \subseteq P$
   Si $V_r \neq \emptyset$:
   Sea $v \in V_r$, supongamos que $v \notin P$. Como hay en el grafo un camino desde v hasta todos los demás vértices, entonces hay un camino a cualquier vértice u de P. Sea e el último arco de dicho camino entre un vértice w que no está en P y un vértice de P, que se sabe que existe porque el camino parte de un vértice que no está en P y termina en un vértice de P. Entonces en el orden topológico hay un arco entre la CFC de w y P. Pero como P es la primera CFC del orden topológico esto no puede ocurrir. Se ha arribado a una contradicción y por tanto si $v \in V_r \implies v \in P$

   Queda demostrado que los vértices de P no pueden alcanzarse desde ningún vértice del grafo que no esté en P, por ser P la primera CFC en un orden topólogico de las CFC de G_c. Luego, los únicos vértices que pudieran alcanzar a todos los demás vértices del grafo son los que están en P, por tanto, $V_r \subseteq P$

2. Lema 2: Si $V_r \neq \emptyset$, entonces $P \subseteq V_r$
   Demostración:
   Como V_r no es vacío, existe un vértice $w \in V_r$ desde el cual se puede llegar a todos los demás vértices del grafo por definición de V_r. Por lo demostrado anteriormente, $w \in P$, luego existe un camino desde cualquier $v \in P$ a w por estar en la misma CFC. Luego, pasando por w, existe un camino desde v hacia cualquier vértice del grafo (y también un camino simple desde v hacia cualquier vértice del grafo), por tanto puede llegarse desde v a cualquier otro vértice del grafo y por tanto, $v \in V_r$ Como esto se cumple para cualquier vértive $v \in P$, entonces, si $V_r \neq \emptyset$, $P \subseteq V_r$

Para cada costo específico c se trabaja con un grafo G_c cuyos vértices son los vértices de G' y sus arcos son los arcos de G' cuyo costo de inversión sea menor o igual que c.

Esto puede hacerse porque como el costo total en el problema está determinado por el costo del arco invertido de mayor valor, entonces invertir cualquier arco de costo menor o igual que c no empeora el costo total. Luego, si se explora con un costo fijado c, cualquier arco con costo menor o igual que c se podría invertir.

Sea V_r el conjunto de vértices de G_c tal que $\forall v \in V_r$ existe un camino desde v hacia todos los demás vértices de G_c, se sabe que, $V_r \subseteq P$ ($Lema$ 1) y que si $V_r \neq \emptyset$, entonces $P \subseteq V_r$ ($Lema$ 2)donde P es la primera componente fuertemente conexa (CFC) de un orden topológico cualquiera de las CFC de G_c

Si se realiza un recorrido DFS sobre el grafo G_c, se obtendrá una partición de los vértices de G_c en t árboles, $a_1, a_2, ..., a_t$. Esto se garantiza por la correctitud del DFS. Se puede demostrar que:

Lema 3: El vértice v, que es la raíz de a_t pertenece a la primera CFC en algún orden topológico de las CFC de G_c

Demostración:

Sea T un orden topológico de G_c y sea T_v la CFC de G_c a la que pertenece el vértice v.

Si T_v es la primera CFC en el orden topológico T, entonces v pertenece a la primera CFC en un orden topológico de G_c.

Si esto no ocurre:

Sabemos que v no es alcanzable desde ningún vértice de G_c que no esté en el árbol de DFS del cual es raíz (a_t), pues todos los demás vértices están en árboles anteriores del DFS y si pudieran alcanzar a v, entonces por correctitud del DFS, el vértice v estaría en alguno de esos árboles anteriores. Por tanto v solo puede alcanzarse desde vértices que están en el árbol del cual es raíz. Luego, todos los vértices que pueden alcanzar a v están en su misma CFC. Además, ningún otro vértice de esa CFC puede alcanzarse por un vértice que no esté en ella pues de lo contrario, v también sería alcanzable desde otra CFC. Luego, en T, la CFC T_v tiene $indegree = 0$.

Luego puede construirse un orden topológico T' poniendo a T_v como primer vértice y manteniendo el resto igual a T. T' es un orden topológico de las CFC de G_c pues mantiene los mismos vértices y como el vértice que se cambió de posición tenía $indegree = 0$, todos los arcos siguen siendo "hacia adelante", por los que sigue siendo un orden topológico.

Luego, v pertenece a la primera CFC en un orden topológico de G_c

Luego, si desde v se puede llegar a todos los vértices del grafo, $v \in V_r$ y el problema puede solucionarse con costo c.

Además, si desde v no se puede llegar a todos los vértices del grafo, entonces, por el contrarrecíproco del *Lema* 2 el conjunto V_r del grafo G_c es vacío y por tanto el problema no puede solucionarse con costo c pues en G_c están todas los posibles arcos para este costo.

Por tanto, el algoritmo diseñado, realiza un recorrido DFS sobre G_c y luego desde el vértice raíz del último árbol, realiza un recorrido DFS para determinar si puede acceder a todos los vértices del grafo.

Puede notarse que el grafo G_c podría tener más arcos que el grafo resultante de invertir algunos arcos con costo menor o igual que c, pero solo serían casos en los que el arco inverso de un arco de dicho grafo tenga costo menor o igual que c. Sin embargo, como se trata de lograr alcanzar desde un vértice el resto, si se utiliza el arco $< u, v >$ en algún camino no se utilizará el $< v, u >$ por correctitud del algoritmo DFS. Luego, las modificaciones que habría que hacer en el grafo G para que cumpla lo que se quiere son las permitidas: invertir algunas aristas.

3.2. Complejidad temporal

El costo temporal de realizar un recorrido DFS sobre el grafo G' es $|V'| + |E'|$. Al realizar este recorrido dos veces, tiene costo temporal de $2(|V'|+|E'|)$, que pertenece a $O(|V'|+|E'|)$. Por la forma de construir el grafo del modelo G' se puede garantizar que $|E'| = 2 * |E|$ y que $|V'| = |V|$, por tanto, la complejidad temporal pertenece a $O(|V| + |E|)$.

El recorrido en profundidad se realiza por cada costo de inversión de arco del grafo original G, donde se tienen k valores, cardinalidad del conjunto de los costos de inversión, por tanto k es menor o igual que la cantidad de arcos de G: $|E|$ ($k \leq E$). Luego, la complejidad temporal pertenece a $O((|V| + |E|) * k)$.

También se realiza la ordenación de los pesos de los arcos de G, que pertenece a $O(k * \log k)$, que es menor que la complejidad temporal antes expuesta, por tanto está acotada.

4. Optimizando solución propuesta

Aprovechando las características del problema, y del algoritmo para buscar la solución implementado, se puede notar que si el problema no puede solucionarse con costo c entonces no puede solucionarse con costo menor que c y si puede solucionarse con costo c entonces solucionarlo con costo mayor no sería de interés pues se busca el menor costo factible. Es posible entonces, establecer sobre cada costo

c del listado de costos ordenados el predicado "¿Puede resolverse el problema con costo menor o igual que c?". Con esta pregunta, la estructura de la respuesta para todo el listado de costos será de la forma: $No, ..., No, Si, ..., Si$ donde en la posición i-ésima está la evaluación del predicado en el i-ésimo costo. La solución del problema será el costo de la posición del primer Si.

4.1. Correctitud

El predicado anterior es justamente lo que verifica el algoritmo descrito anteriormente, donde para cada costo c intenta buscar una solución invirtiendo arcos con costo menor o igual que c

Si el problema no puede solucionarse con costo menor o igual que c entonces no puede solucionarse con costo menor que c. Por tanto, cuando se determina que para el costo i-ésimo no es posible solucionar el problema, no tiene sentido analizarlo para costos anteriores en el listado ordenado de los costos.

Además si es posible resolver el problema con el costo i-ésimo, es posible resolverlo con cualquier costo mayor, por la definición del predicado.

Con esto se justifica que bajo el predicado propuesto la estructura de la respuesta para todo el listado de costos será de la forma: $No, ..., No, Si, ..., Si$ donde en la posición i-ésima está la evaluación del predicado en el i-ésimo costo.

Como se busca el menor costo posible, la solución será el menor costo que tenga elvaluación positiva del predicado, es decir, el costo de la posición del primer Si en el listado ordenado de costos.

Teniendo en cuenta lo anterior se utilizó un enfoque de búsqueda binaria sobre el arreglo de costos ordenados. Cada vez que se aplica el predicado sobre un costo c, si la respuesta es positiva, se descartan todos los costos mayores que c y si es negativa, se descartan todos los costos menores que c.

4.2. Complejidad temporal

En esta propuesta, se sustituye la búsqueda lineal sobre los costos por la búsqueda binaria sobre el arreglo de costos ordenados. Siendo $k \leq E$ la cantidad de costos de las aristas, es conocido que la búsqueda binaria sobre un arreglo ordenado de tamaño k tiene complejidad $\log k$. Además cada vez se realiza la doble exploración explicada en la propuesta anterior cuya compliejidad pertenece a $O(|V| + |E|)$. Esto tiene complejidad temporal $O((|V| + |E|) * \log k)$

Además como mismo en la propuesta anterior lo primero que se realiza es una ordenación de los costos de inversión de los arcos. El costo de ordenar los costos de inversión de los arcos no empeora la complejidad temporal puesto que sería $k * \log k$ y se sabe que $k \leq |E| \leq |V| + |E|$ y como $\log k \geq 0$, enconces $k * \log k \leq (|V| + |E|) * \log k$ por tanto, está acotado.