PROBLEMA DE MINIMIZACIÓN DEL ANCHO DE BANDA EN ÁRBOLES DE PROFUNDIDAD 2

Daniel González Alonso

ENERO DE 2016

ESCUELA DE INGENIERÍA INFORMÁTICA DE VALLADOLID

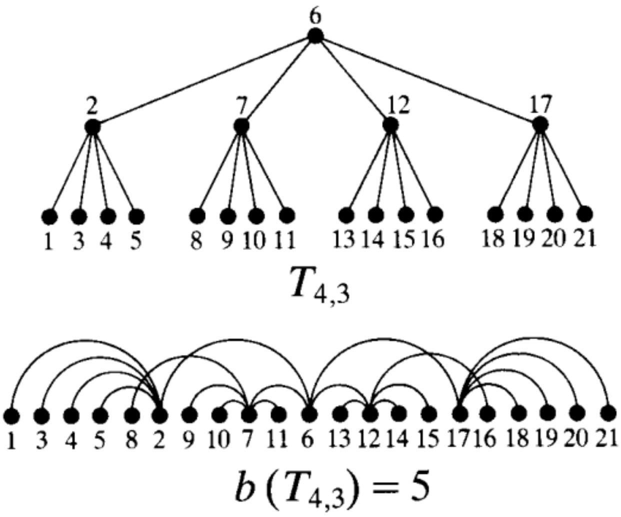
Este trabajo se basa en una publicación del [2] departamento de matemáticas de la Universidad de Luisiana.

Índice:

Descripción del problema

El problema de la minimización del ancho de banda (Graph Bandwidth Minimization, ***GBM***) de un grafo $G = (V, E)$, representado como $b(G)$, consiste en etiquetar los n vértices de V con distintos enteros mediante una función uno a uno $f: V \to \{1,2, \dots, |V|\}$ de forma que para todo $\{u, v\} \in E,\ |f(u) - f(v)|$ sea minimizado.

Este problema se puede visualizar también como encontrar un orden de los vértices del grafo G sobre el eje X que minimice la longitud de la mayor arista del grafo.



Ejemplo de disposición de un grafo sobre el eje X y del problema de GBM.

También se suele definir como la siguiente pregunta: Dado un grafo $G = (V, E)$ y un entero $k \leq |V|$, ¿Existe algún orden lineal de V cuyo ancho de banda sea k o inferior, es decir, una función uno a uno tal que para todo $\{u, v\} \in E,\ |f(u) - f(v)| \leq k$?

Este problema se ha demostrado ser NP-Completo para arboles de grado menor o igual que 3 así como otros grafos como Caterpillar de longitud 2, nuestra demostración se va a realizar para **árboles de profundidad 2**.

Las principales aplicaciones de este problema son el diseño de sistemas VLSI y la reducción del ancho de banda en matrices simétricas dispersas, para ello existen implementaciones bastante eficientes como el algoritmo Cuthill–McKee.

Definiciones

Problema NP: Conjunto de problemas de decisión cuya solución se puede comprobar en tiempo polinómico. Es el conjunto de problemas que pueden ser resueltos en tiempo polinómico por una máquina de Turing no determinista.

Problema NP-Completo: subconjunto de problemas NP tal que todo problema de NP se puede reducir a un NP-Completo, se suele decir que son los problemas más difíciles de NP.

> Si algún problema de este conjunto se puede resolver en tiempo polinómico, entonces se podrá para todos los NP, y ocurrirá que P = NP.
>
> Si algún problema NP no es resoluble en tiempo polinómico, entonces ningún problema de este conjunto podrá serlo.

Problema de la partición: Problema NP-Completo que consiste en, dado un multiconjunto de números enteros S, determinar si este puede particionarse en dos multiconjuntos S1 y S2, tal que la suma de los elementos de S1 sea igual a la de S2.

Notación

Dado un grafo $G = (V, E)$

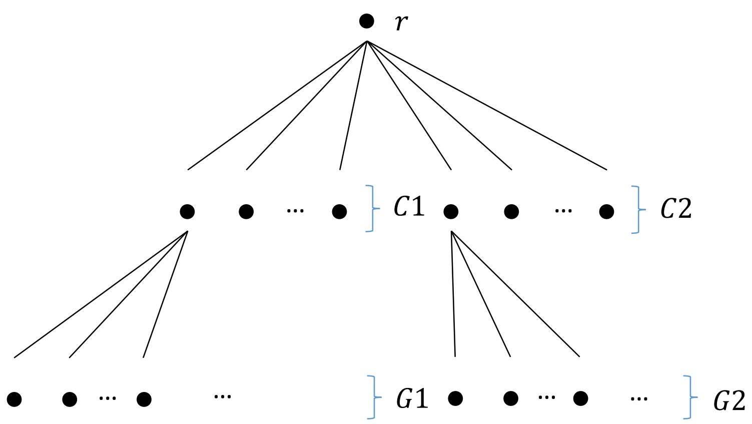
- $d(x)$ representa el grado del vértice $x \in V$
- $\Delta(V)$ Representa el máximo grado de los vértices V
- $|V|$ Representa el número de vértices de V

Demostraciones

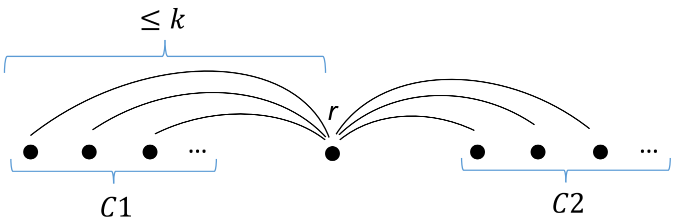
Tenemos un grafo $G = (V, E)$ en forma de árbol de dos niveles con raíz r, y $k \in \mathbb{N}$.

Teorema 1: $b(G) \leq k$ si y solo si se cumple:

1. $\Delta(V) \leq 2k$
2. $|V| \leq 4k + 1$
3. Podemos dividir G en dos particiones $C1$ y $C2$ tal que:
   3.1. $|C1| \leq k$ y $|C2| \leq k$
   3.2. $|C1| + |C2| + |G1| \leq 3k$ y $|C1| + |C2| + |G2| \leq 3k$ donde G1 y G2 son los hijos de C1 y C2 respectivamente.



Demostración: Suponemos que $b(G) \leq k$ gracias a la función $f: V \rightarrow \{1,2, \dots, |V|\}$, y que en $C1$ se encuentran los vértices hijo de r cuya distribución en f es inferior a r (f les sitúa a la izquierda de r en el eje x), y $C2$, los hijos con una distribución superior a r. Se puede observar que $|C1| \leq k$ y $|C2| \leq k$ puesto que $|f(c) - f(r)| \leq k$ para cualquier hijo de r, viéndolo sobre el eje-x, la separación máxima de los hijos de r ha de ser inferior o igual a k:



Con lo cual, el punto 3.1 queda demostrado.

Sucede algo similar entre los hijos $G1$ de $C1$, con lo cual, la separación máxima en la distribución entre los vértices de $G1$ y r ha de ser $2k$, y por tanto, el número máximo de vértices de $G1$ con una distribución inferior a r ha de ser como mucho $2k - |C1|$, por otro lado, el número de vértices de $G1$ con una distribución superior a r ha de ser como mucho $k - |C2|$ (Tenga en cuenta que si $|C2| \leq |C1|$, puede que parte de *G1* tuviese que estar a la derecha de r junto a $C2$), y entonces, $|G1| \leq 2k - |C1| + k - |C2| \quad \rightarrow \quad |G1| + |C1| + |C2| \leq 3k$, y como con $G2$ sucede lo mismo, los puntos 3.2 y 3 quedan demostrados.

Ahora gracias a los datos anteriores podemos concluir que $\left(\frac{|V|-1}{4}\right) \leq p(G) \leq b(G) \leq k$, con lo cual $|V| \leq 4k + 1$ demostrando así el punto 2.

Finalmente, como la separación máxima de los vértices de cada partición ha de ser inferior a k y tenemos dos particiones, $\left(\frac{\Delta(V)}{2}\right) \leq p(G) \leq b(G) \leq k \quad \rightarrow \quad \Delta(V) \leq 2k$, demostrando así el punto 1 y el Teorema 1.

Ahora, vamos a hacer una **transformación del problema de la partición a GBM**, para ello suponemos que disponemos un árbol el cual satisface el punto 3 y su partición respecto a r: $(C1, C2)$, y queremos obtener una función $f: V \rightarrow \{1,2, \dots, |V|\}$ para la cual $b(G) \leq k$. Para ello, supongamos que $C1 = \{x_{11}, x_{12}, \dots, x_{1n1}\}$ con $d(x_{11}) \geq d(x_{12}) \geq \dots \geq d(x_{1n1})$, $C2 = \{x_{21}, x_{22}, \dots, x_{\,2n2}\}$ con $d(x_{21}) \geq d(x_{22}) \geq \dots \geq d(x_{2n2})$, y además $|C1| + |G1| \geq |C2| + |G2|$, el siguiente algoritmo producirá f de forma que $b(G) \leq k$:

Algoritmo:

Entrada: G, árbol de altura como mucho 2 que satisface las condiciones (1)-(3)

Salida: $f(v)$ que ordena los vértices de G desde $\{1, \ldots, 4k+1\}$

Etiquetar con el menor entero sin asignar empezando por 1:

(1) Como mucho k "nietos" de r, empezando desde x_{11} hasta x_{1n1}.

(2) x_{11}hasta x_{1i}, donde x_{1i} es el último vértice cuyo hijo tiene un número ya asignado.

(3) El resto de hijos de x_{1i}.

(4) Hasta etiquetar $2k$: la mitad de los hijos de x_{1j} (redondeando hacia arriba), x_{1j}, y después el resto de hijos de x_{1j} desde $j = i + 1$ hasta $n1$.

(5) r.

(6) Continuar con (4), sin el límite superior de $2k$, hasta completar el resto de x_{1j} y sus hijos.

Etiquetar con el mayor entero sin asignar empezando por $f(r) + 2k$:

(7) Como mucho k "nietos" de r, empezando desde x_{21} hasta x_{2n2}.

Etiquetar con el mayor entero sin asignar empezando por $f(r) + k$:

(8) x_{21}hasta x_{2l}, donde x_{2l} es el último vértice cuyo hijo tiene un número ya asignado.

(9) El resto de hijos de x_{2l}.

(10) la mitad de los hijos de x_{2j} (redondeando hacia arriba), x_{2j}, y después el resto de hijos de x_{2j} desde $j = l + 1$ hasta $n2$.

Demostración**:** Cada $x_{11}, x_{12}, \ldots, x_{1i-1}$ tiene asignado un entero superior al de sus hijos. Como de estos "nietos" de r tenemos como mucho k, y los x_{1j} están ordenados por su grado, $|f(x_{1j}) - f(y)| \leq k$ para todo $j < i, y$ siendo los hijos de x_{1j}.

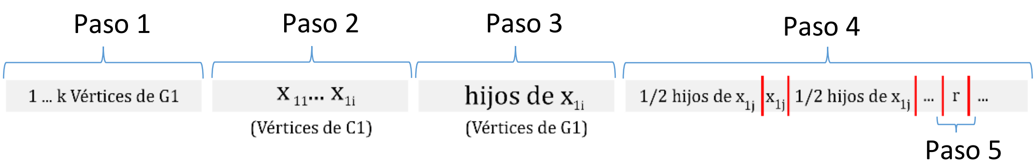


Ilustración para comprender mejor la siguiente demostración

Ahora probaremos que para cada hijo y de x_{1i}, $|f(x_{1i}) - f(y)| \leq k$. Debido a las etapas (1)-(3), los hijos de x_{1i} pueden tener asignado un entero tanto superior a $f(x_{1i})$ como inferior. Como hemos dicho anteriormente, cada hijo y de x_{1i} con una etiqueta $f(y) < f(x_{1i})$ tiene $|f(x_{1i}) - f(y)| \leq k$. Ahora asumiremos por contradicción que x_{1i} tiene más de $k-1$ hijos con una etiqueta superior a $f(x_{1i})$, entonces el grado de cualquier x_{1j} con $j < i$ seria de al menos $k+1$ y x_{1j} tendría al menos k hijos. Como cada uno de estos hijos y de x_{1j} tiene asignado un valor $f(y) < f\left(x_{1j}\right)$, y como hay como mucho k de estos hijos debido a (1), i ha de valer 1. Pero entonces, x_{11} tendría k hijos con valor inferior a $f(x_{11})$ y otros k hijos con valor superior a $f(x_{11})$, los cuales junto con el hecho de que x_{11} es un hijo de r implica que $d(x_{11}) \geq 2k+1$, pero esto no es posible puesto que $\Delta(G) \leq 2k$. Entonces, como mucho puede haber $k-1$ hijos de x_{1i} con un valor asignado superior a $f(x_{1i})$ y por tanto $|f(x_{1i}) - f(y)| \leq k$ para cualquier hijo y de x_{1i}.

Todos los x_{1j} con $j > i$ se encuentran centrados entre sus hijos, por lo tanto, el único problema posible se encuentra con algunos de sus vértices, llamémosles $x_{1j'}$, con $j' > i$, los cuales pueden tener un valor asignado tanto superior como inferior a $f(r)$. En $x_{1j'}$ debe haber como mucho $k-1$ hijos con un valor asignado inferior a $f(r)$ y a su vez, como mucho $k-1$ hijos con un valor asignado superior, ya que si no fuese así, $x_{1j'}$ tendría al menos $2k-1$ hijos y entonces $d(x_{1j'}) = 2k$ debido a la condición (1). Además, al ordenar los vértices por el grado, tendríamos $d(x_{1j}) = 2k$ para todos los $j < j'$, y como existe x_{1j} tal que $j' > i$, entonces $j' > 1$, y entonces $|C1| + |C2| + |G1| \geq 4k$, lo cual sería una contradicción. Entonces $x_{1j'}$ tiene por lo menos $k-1$ hijos con un valor inferior a $f(r)$ y como mucho $k-1$ hijos con valor superior a $f(r)$, dándonos $|f(x_{1j}) - f(y)| \leq k$ para cualquier $j > i$ y cualquier y hijo de x_{1j}.

También x_{2j} y sus hijos tienen que tener su propio valor, pero debido a la similitud entre los puntos (1)-(4) y (7)-(10) del algoritmo no lo vamos a mostrar, ya que a los x_{2j} se les da valor de la misma forma, aunque sea más sencillo ya que x_{2j} no tiene hijos con un valor inferior a $f(r)$.

Finalmente, si $|G1| \geq k$, y $|C1| + |G1| \geq 2k$, entonces $f(r) = 2k+1$, y $k+1 \leq f(x_{1j}) \leq 3k+1$ para todo j. Si $|G1| \geq k$, y $|C1| + |G1| < 2k$, entonces $k+1 \leq f(x_{1j}) \leq f(r) \leq 2k+1$ para todo j. Finalmente si $|G1| < k$, entonces $f(r) - k \leq f(x_{1j}) \leq f(r)$ para todo j. En todos los casos, $|f(r) - f(x_{1j})| \leq k$ para cualquier j. Ahora, como $f(x_{2j})$ está entre $f(r)+1$ y $f(r)+k$, tenemos que $|f(r) - f(x)| \leq k$ para cualquier x hijo de r.

Este algoritmo es capaz de funcionar en $O(|V|)$ dándole las particiones que cumplen los puntos (1)-(3) y los hijos ya ordenados por su grado, con lo cual, podemos concluir que **Partición $\leq_p$ GBM en árboles de profundidad 2**, y por lo tanto, es NP-Completo.

Referencias

[1] COMPUTERS AND INTRACTABILITY. A Guide to the Theory of NP-Completeness. Michael R. Garey / David S. Johnson

[2] BANDWIDTH OF TREES OF HEIGHT AT MOST TWO - LSU Vigre Combinatorics Crew

[3] Partition Problem – Wikipedia

[4] Graph bandwidth – Wikipedia

[5] NP complexity – Wikipedia

[6] Diapositivas de Algoritmos y Computación – M.Barrio Solorzano