Graphes couvrants éparses

Master 2 Informatique - UFR S.A.T

Pr. Ousmane THIARE

ousmane.thiare@ugb.edu.sn
http://www.ousmanethiare.com/

12 septembre 2025

Chapitre 8 : Graphes couvrants éparses

Introduction au calcul de spanner dans le modèle LOCAL. Ce sont des sous-graphes couvrant les sommets d'un graphe possédant un nombre limité d'arêtes. De plus ils garantissent que les distances dans les sous-graphes ne sont pas trop éloignées des distances originales.

L'utilisation de sous-graphe couvrant, c'est-à-dire qui contient tous les sommets du graphe d'origine, est banale en informatique. Typiquement un arbre couvrant un graphe connexe G à n sommets garantit la connexité alors qu'il ne possède que n-1 arêtes. Cela peut être particulièrement intéressant en calcul distribué de minimiser de nombre de liens sans pour autant déconnecter G.

On a vu dans les chapitres précédents que le nombre de messages nécessaire à résoudre certaines tâches peut être réduit si un arbre préexistant était connu des processeurs. Au chapitre 5, on a vu que la connaissance d'un sous-graphe couvrant peu dense pouvait être utilisé pour construire des synchroniseurs.

De manière générale, on souhaite construire un graphe couvrant peu dense préservant les distances d'origines. Bien sûr il n'est pas possible de supprimer des arêtes sans modifier au moins une distance. Néanmoins on souhaite limiter ces modifications. La définition suivante s'applique aussi bien au cas des graphes valués que non valués.

Definition

Soient G un graphe et H un sous-graphe couvrant de G, c'est-à-dire avec V(G)=V(H).L'étirement de H est le plus petit réel s tel que $d_H(u, v) \leq s.d_G(u, v)$ pour tout sommets u, $v \in V(G)$.

Exemple : Dans la littérature on appelle aussi s-spanner tout sous-graphe couvrant d'étirement s (stretch en Anglais).

Le problème qu'on se fixe est donc de calculer de manière distribué, dans le modèle LOCAL, des sous-graphes couvrant avec peu d'arêtes et si possible un étirement faible. On se base essentiellement sur l'article [DGPV08].

Formellement, calculer un sous-graphe H signifie que chaque sommet u doit sélectionner un ensemble d'arêtes qui lui sont incidentes, disons H(u). Le graphe calculé est alors le graphe $\bigcup_{u \in V(G)} H(u)$ composé de l'union de tous ces choix.

Notons qu'il n'est pas formellement imposé que les deux sommets extrémités d'une arête s'accordent sur leur décision, c'est-à-dire $uv \in H(u)$ n'implique pas forcément $uv \in H(v)$. Cependant, avec une ronde supplémentaire, on peut toujours faire en sorte que tous les sommets se mettent d'accord (par exemple sur le fait de garder l'arête si l'un des deux l'a sélectionnée, ou le contraire, de ne pas la garder si l'un des deux ne l'a pas garder).

Introduction
Calcul d'un 3-spanner

On va voir un algorithme permettant de calculer un sous-graphe couvrant d'étirement 3 avec significativement moins de n^2 arêtes, en fait $2n^{1.5}$. Il suffit de seulement deux rondes. La décision de sélectionner ou non les arêtes est donc locale.
Avant de présenter l'algorithme, remarquons que le problème n'a rien d'évident. Décider localement de supprimer une arête est a priori très risqué, puisque par exemple un sommet n'a aucun moyen de deviner s'il est dans un arbre ou pas (ils ne connaissant pas, par exemple le nombre d'arêtes du graphe). Bien sûr pour avoir un l'étirement borné, aucun sommet d'un arbre ne peut supprimer d'arête. D'un autre coté, si aucun sommet ne supprime d'arête on pourrait avoir de l'ordre de n2 arêtes si le graphe G était très dense.

Jusqu'à récemment, aucun algorithme déterministe n'était connu pour ce problème. Ils existent des algorithmes en deux rondes, mais qui ne garantissent le nombre d'arêtes seule- ment en moyenne [BKMP05]. Le problème des algorithmes aléatoires de ce type (Monté Carlo) est qu'on est pas certain du résultat, contrairement aux algorithmes de type Las Vegas. Et contrairement au calcul séquentiel, en distribué il n'est pas toujours facile de vérifier la condition (comment vérifier que le nombre d'arêtes est correct sans une diffusion globale ?). En général, en distribué, on ne peut jamais recommencer !
Dans la suite, on rappelle que B(u,1) représente l'ensemble formé de u et des voisins de u, la boule centrée en u et de rayon 1. Évidemment, $|B(u,1)| = deg(u) + 1$.

Algorithme Span3 (code du sommet u)

1. NewPulse
2. Envoyer son identifiant et recevoir l'identifiant de tous ses voisins.
3. Choisir un ensemble R_u composé de son identifiant et d'un sous-ensemble arbitraire des identifiants reçus tel que $|R_u| = min\{|B(u,1)|, \lceil\sqrt{n}\rceil\}$.
4. NewPulse
5. Envoyer R_u et recevoir la sélection R_v de chaque voisin v.
6. Poser $H(u) := \{uv : v \in R_u, u \neq v\}$ et $W := B(u,1) - \{v \in B(u,1) : u \in Rv\}$.
7. Tant qu'il existe $w \in W$
   (a) $H(u) := \{uw\} \cup H(u)$.
   (b) $W := W - \{v \in W : R_v \cap R_w \neq \emptyset\}$.

Il y a une variante où u ne connaît pas le nombre $\lceil \sqrt{n} \rceil$, mais qui nécessite une ronde de plus. L'algorithme peut aussi être adapté pour fonctionner avec des poids arbitraire sur les arêtes. Il faut choisir R_u parmi ses plus proches voisins et le sommet w parmi les plus proches restant de W.

Théorème

Pour tout graphe G à n sommets, l'algorithme Span3 calcule en deux rondes un sous-graphe couvrant de G d'étirement ≤ 3 et avec moins de $2n^{\frac{3}{2}}$ arêtes.

Preuve : L'algorithme comprend deux rondes de communications.

Etirement. Soit uv une arête de G qui n'est pas dans le sous-graphe couvrant $H = \bigcup_u H(u)(Figure(a))$.
Clairement toutes les arêtes de la forme uw avec $w \neq u$, $u \in R_u \cup R_w$ et $w \in B(u, 1)$ existent dans H. Donc si cette arête a été supprimée c'est que v a été supprimé de W à l'étape 7b de l'algorithme (Figure (b)). Si v est supprimé de W à l'étape 7b, c'est que R_v et R_w s'intersectent (Figure (c)). Or dans H toutes les arêtes de v vers R_v et de w vers R_w existent. Donc il existe dans H un chemin de u à v de longueur trois.
Il suit qu'un plus court chemin de x à y dans H est de longueur au plus trois celle d'un plus court chemin P de G, chaque arête de P qui n'est pas dans H étant remplacée par un chemin de H d'au plus trois arêtes.

Introduction
Calcul d'un 3-spanner

Nombres d'arêtes Montrons que tout sommet w de l'ensemble W calculé à l'étape 6 est tel que $|B(w,1)| > \lceil\sqrt{n}\rceil$. En effet, si $|B(w,1)| > \lceil\sqrt{n}\rceil$, alors $|R_w| = min\{|B(w,1)|, \lceil\sqrt{n}\rceil\} = |B(w,1)|$. Donc R_w contiendrait tous les voisins de w, en particulier u. Or ce n'est pas le cas, donc $|B(w,1)| > \lceil\sqrt{n}\rceil$. Il suffit que $|R_w| = \lceil\sqrt{n}\rceil$

Les sommets sélectionnés dans la boucle « Tant que » ont, par construction, des ensembles deux à deux disjoints, c'est-à-dire $R_w \cap R'_w = \emptyset$ (Figure (e-d)) puisque ceux qui intersectent R_w sont supprimés à jamais de W. Il suit que le nombre de sommets sélectionnés lors de la boucle « Tant que » est au plus $|B(u,2)|/\lceil\sqrt{n}\rceil \leq n/\sqrt{n} = \sqrt{n}$ chaque ensemble R_w étant inclus dans B(u, 2) qui est de taille au plus n.

Nombres d'arêtes Le nombre d'arêtes sélectionnées par u est donc au plus $|R_u| - 1 + \sqrt{n} \leq \lceil\sqrt{n}\rceil - 1 + \sqrt{n} < 2\sqrt{n}$. Au total H n'a pas plus de $2n^{\frac{3}{2}}$ arêtes.

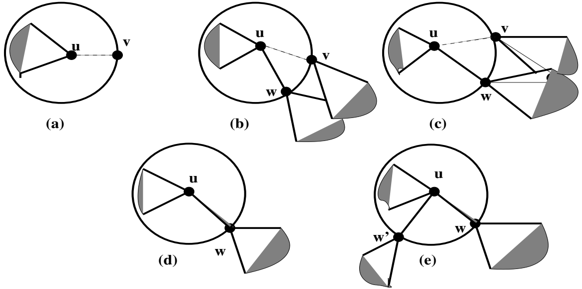

FIGURE: Illustration de la preuve du théorème

Exercice étendre l'algorithme à un algorithme qui ne nécessite pas la connaissance de n. Combien de rondes cela nécessite ? Solution, il faut choisir une sélection de $\lfloor |B(u, 3)| \rfloor$ voisins au lieu de $\lfloor \sqrt{n} \rfloor$. Pour cela il faut au moins 3 rondes.
Problème ouvert : Peut-on réaliser la tâche précédente avec seulement deux rondes ?