Steiner Tree Problem Optimization

Louis Cohen

February 11, 2018

Contents

1 Abstract

Dans ce rapport je presenterai différentes approches choisies pour approximer le probleme des arbres de Steiner. Basé sur le papier ref ref pour les choix des voisins.

2 Introduction

Le problème des arbres de Steiner est un problème de graphe NP Complet. Il consiste à, pour un graphe pondéré et un sous ensemble de noeuds (les terminaux), trouver le sous ensemble d'aretes de poids minimal connectant ces terminaux. Ce problème largement étudié trouve de nombreuses applications blablabla. Il n'est évidemmnet pas possible de résoudre completement

le problème dans le cas général. Cependant dans certains cas le calcul de l'optimum est possible, par exemple avec un nombre faible de terminaux ou encore pour un graphe avec faible treewidth. Enfin dans le cas général il reste comme approche l'approximation de l'optimum. Dans la majeure partie des cas en effet une approximation est largement appréciable. Dans un premier temps certaines approximations peuvent être deterministes. Nous nous baserons sur l'une d'entre elle **ref ref** . Nous randomiserons ensuite nos approches. Dans la section 1 je décrirais le problèmes et lui donnerais un cadre théorique. La section 2 sera dédiée à la description d'une approximation déterministe. La section 3 expliquera la notion de voisin de solution qui servira dans pour toutes les approximations suivantes. La section 4 se penchera sur des recherches locales tandis que la section 5 présentera des algorithmes biologiques (principalement un algorithme moléculaire). Enfin la section 6 comparera les différentes approximations.

3 Problèmes des arbres de Steiner

Soit $G = (V, E)$ un graphe (V son ensemble de noeuds et E son ensemble d'arêtes, muni d'une fonction de poids $w : E \to \mathbb{R}$. Soit de plus un sous ensemble de noeuds $T \subset V$, les terminaux. Le but est de trouver le sous ensemble d'arêtes de poids minimal connectant tous les terminaux. On dira qu'un sous ensemble est admissible si il connecte bien les terminaux. $\underset{S \subset E/S \quad admissible}{\arg\min} \sum_{e \in S} w(e)$.

On appellera graphe induit par une solution le graphe composé des arêtes de l'ensemble solution S et des noeuds appartement a cet ensemble d'arêtes.

On dira par la suite qu'une solution est optimale si elle est admissible et que la suppression de n'importe quel arête n'en fait plus une solution optimale. On peut remarquer qu'une solution est optimale si et seulement si le graphe induit par la solution est un arbre dont les feuilles ne sont que des terminaux.

4 Optimisation déterministes

La plus-part de nos approximations auront pour solution de base une 2-approximation déterministe (dans le code fonction first_solution() dans le fichier local_search()). Son principe est relativement simple. On va créer un nouveau graphe $G_{ter} = T, E_T$. Les noeuds sont donc les terminaux et les arêtes sont étiquetées par le poids du plus court chemin entre les deux terminaux. On va ensuite calculer un arbre couvrant minimal $Tree_{ter}$ de G_{ter}. Enfin va créer notre solution (2-approximation) du problème d'arbre de steiner $First\_solution$: pour chaque arête présente dans $Tree_{ter}$ on ajoute toutes les arêtes du plus court chemin correspondant dans G_{ter} a $First\_solution$. On remarque donc que $First\_solution$ est bien une solution admissible vu que tous les terminaux seront ajoutés. On peut prouver que cet algorithme donne une deux approximation (voir ...). Je ne ferai pas ici la démonstration mais me contenterai de présenter un cas ou la borne est atteinte : FIGURE A RAJOUTER... On voit que notre algorithme va choisir les arêtes $(1, 2)$ $(2, 3)$ et $(1, 3)$ sans passer par le noeud 4 alors que l'optimum (si $2a < b$ est de choisir les arêtes $(1, 4)$ $(2, 4)$ et $(3, 4)$. Quand b tend vers 2a le ratio entre les deux tends vers 2.

5 Notion de voisin

5.1 Presentation du précédent papier

Par la suite la pluspart de nos optimisation se baseront sur une notion de "voisin" des solutions. Il faut évidemment que ces "voisins" soit proche d'un point de vue de score face au problème. Cette proximité dans le cas de graphes comme les notre sera logiquement proche d'une métrique comparant les ensembles de sommets. Cependant comme expliqué dans REFERENCE, il n'est pas suffisant de regarder les ensembles de d'aretes directement. L'ajout ou la suppression d'arete ne suffit pas. On a donc implémenté un nouveau "modificateur" : l'ajout de chemin. Pour résumer nous avons quatre "modificateurs" d'une solution courante :

L'addition d'aretes : on ajoute un certain nombre d'arêtes "voisine" de notre graphe. C'est à dire d'arêtes dont au moins un des noeuds fait déjà partie de la solution. En effet sinon on risque de souvent rajouter des arêtes complétement déconnectées de notre graphe qui n'ont que très peu D'Intérêt.

La suppresion d'aretes : on supprime des arretes qui laisse notre solution admissible. Apres cette suppression on recalcule les composantes connexes et on ne garde que la composante connexe contenant tous les terminaux (ils sont tous dans la meme car la solution reste faisable)

Add a path : We randomly choose two nodes of our current solution and we add all the edges of a shortest path between this ndoes

Clear : clear make the solution minimal. While it's possible it delete an edge and calculate the connected componnents and only keep the one containing the terminals.

5.2 Neighbour choice

We randomly perform k of the first 3 modifications and then clear the results. The resulting solution is a solution at distance k of the first solution. In most of our test k is set to 7. We tried to not to randomized it but to make some precalculated serie of operation like :add nodes x3, add path x3, clean (problem was our deletion was really costly).

6 Local Searchs

On va vouloir parcourir l'ensemble des solutions a la recherche de minimum locaux. Tout le problème de ces parcours est d'arriver a trouver de "bons" minimums locaux, il faut donc en général continuer a trouver un bon trade off entre l'exploration et l'exploitation des points visités