Diffusion contrainte dans un groupe

Joanna Moulierac, Jean-Claude Konig, Miklos Molnar

IRISA/Université de Rennes I, Campus de Beaulieu, 35042 Rennes Cedex
LIRMM, 61 rue Ada, 34392 Montpellier Cedex 5
IRISA/INSA, Campus de Beaulieu, 35042 Rennes Cedex

Dans ce papier nous traitons de problèmes particuliers de communications de groupe dans les réseaux. Le problème de diffusion contrainte dans un groupe consiste à envoyer un message d'un émetteur à un ensemble de destinataires en imposant son passage par des nœuds spécifiques appelés les traducteurs. Dans un premier temps, nous prouvons que ce problème est NP-difficile et nous proposons une heuristique ρ-approchée pour le résoudre, où ρ est le facteur d'approximation de l'heuristique pour le problème de Steiner. Dans un deuxième temps, nous présentons une heuristique permettant de résoudre le problème de diffusion contrainte dans un groupe auquel on a rajouté une contrainte d'équilibrage.

Keywords: Problème de Steiner, Algorithme d'approximation, Multicast

1 Introduction

Les réseaux de communication ont connu un tel essor technologique ces dernières décennies que les applications possibles se sont diversifiées. La visioconférence et le télé-enseignement sont devenus aujourd'hui des services en forte croissance.

La solution naturelle pour gérer ces communications de groupe consiste à effectuer une inondation dans un sous-réseau du réseau physique, la duplication des messages étant gérée par des routeurs intermédiaires. Pour cela, on peut construire un arbre couvrant les membres du groupe tout en optimisant certains paramètres. Par exemple, l'arbre qui minimise le coût est un arbre de Steiner. Trouver un tel arbre étant un problème NP-difficile [Kar72], on cherche en pratique un arbre qui possède une solution dont le coût est le plus proche possible de celui de l'arbre de Steiner. Il existe des heuristiques ρ-approchées pour résoudre ce problème et le meilleur facteur d'approximation connu à présent est $\rho = 1,550$ [RZ00].

Nous étudions une variante du problème de Steiner : les communications entre un émetteur e et les membres du groupe S doivent passer par un ou plusieurs nœuds spécifiques capables de rendre un service particulier. Ces nœuds, appelés traducteurs par la suite, peuvent ainsi gérer l'authentification de messages ou un traitement spécifique des données envoyées par l'émetteur. Nous voulons gérer ces communications en minimisant le coût des communications sachant que la transmission des données avant et après la traduction implique des coûts différents. Dans le cas particulier où il n'y a qu'un seul traducteur t_1 dans le graphe, le problème est équivalent au problème de Steiner construit sur $\{t_1\} \cup S$. En effet, le problème revient à trouver une plus courte chaîne de e à t_1, puis à construire un arbre de coût minimum qui couvre t_1 et l'ensemble des destinataire. Notre problème est donc NP-difficile dans le cas général.

Dans ce papier, nous exposons des heuristiques avec garantie de performance permettant de résoudre le problème de diffusion contrainte dans un groupe ainsi qu'une variante équilibrée de ce problème qui impose un même nombre de destinataires reliés à chacun des traducteurs.

1.1 Travaux antérieurs

Le problème du k-fournisseur [LR00] est un problème voisin de celui traité dans ce papier, néanmoins les critères d'optimisation diffèrent. Le problème du k-fournisseur s'énonce de la façon suivante : un groupe de clients répartis sur des sites distincts veulent utiliser un certain service (une base de données par exemple) proposé par un ensemble de fournisseurs ayant chacun un coût d'utilisation. Le problème consiste à sélectionner un sous-ensemble de fournisseurs dont le coût total (somme des coûts d'utilisation de chaque fournisseur choisi) ne dépasse pas un budget k que les clients se sont fixés. Chaque client est affecté au fournisseur sélectionné qui est le plus proche, et on cherche à minimiser la plus grande distance d'un client à son fournisseur. Les auteurs se sont appuyés sur le travail réalisé par Hochbaum et Schmoys qui ont montré une 3-approximation de ce problème, sachant que ce facteur d'approximation est optimal à moins que P=NP [HS86].

Dans notre problème, nous n'utilisons pas la même métrique (nous cherchons à réduire le coût total) et nous nous autorisons dans un premier temps à utiliser autant de traducteurs que nécessaire.

La section 2 expose le problème abordé ainsi que les notations utilisées, la section 3 expose l'heuristique que nous proposons ainsi que son facteur d'approximation et la section 4 expose une variante équilibrée du problème de diffusion contrainte dans un groupe.

2 Préliminaires

2.1 Notations

Le réseau est modélisé par un graphe $G = (V,E)$ connexe, non orienté et valué. On note w la fonction qui associe une valeur à chaque arête du graphe. Soit $T \subseteq V$ l'ensemble des traducteurs du graphe et $S \subseteq V$ l'ensemble des destinataires des messages envoyés par un émetteur $e \in V$.

Nous notons par la suite ρ la garantie de performance d'une heuristique pour le problème de Steiner.

2.2 Définition du problème de diffusion contrainte dans un groupe

Le problème de diffusion contrainte dans un groupe consiste à trouver un sous-ensemble de traducteurs $T' \subseteq T$, un sous-graphe $A_0 = (V_0, E_0)$ de G couvrant $\{e\} \cup T'$ et un sous-graphe F de G contenant $|T'|$ graphes $A_i = (V_i, E_i)$. Chaque A_i de F contient au moins un traducteur de T' et F est tel que pour chaque destinataire $s \in S$, il existe un A_i de F auquel s appartient. Le but du problème est de minimiser la somme des valuations des arêtes utilisées par A_0 et F. Dans la solution optimale, A_0 est un arbre et F est une forêt dont chaque arbre contient exactement un traducteur.

Sur l'exemple de la Fig. 1, $T' = \{t_2, t_5\}$. L'arbre A_0 couvre $\{e\} \cup \{t_2, t_5\}$ et la forêt $F = (A_2, A_5)$ est telle que $A_2 = (V_2, E_2)$ où $V_2 = \{t_2, s_1, s_2, s_3, s_4\}$ et $A_5 = (V_5, E_5)$ où $V_5 = \{t_5, s_5, s_6, s_7\}$. E_2 et E_5 sont représentés en pointillés sur la figure.

Pour modéliser le fait que les données peuvent être éventuellement modifiées par les traducteurs (elles utiliseront plus ou moins de ressources après leur modification), nous attribuons un coefficient λ à la partie reliant l'émetteur aux traducteurs, l'arbre A_0, et un coefficient μ à celle connectant les traducteurs aux destinataires, la forêt F. Par exemple, si l'émetteur envoie une image compressée aux traducteurs et que les traducteurs décompressent cette image, on aura $\lambda \ll \mu$.

On doit donc minimiser la somme des deux parties auxquelles sont affectés les coefficients λ et μ c'est à dire : $\lambda w(|E_0|) + \mu \sum_{i=1}^{|T'|} w(|E_i|)$.

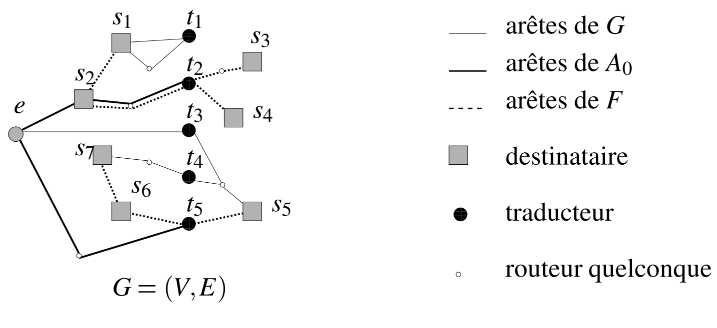

FIG. 1 – *Construction d'une solution dans le graphe $G = (V,E)$ de coût 14 pour notre problème avec $\lambda = 1$ et $\mu = 1$ (les arêtes sont toutes valuées à 1). Les traducteurs utilisés sont t_2 (relié à s_1, s_2, s_3, et s_4) et t_5 (relié à s_5, s_6 et s_7).*

3 Diffusion contrainte dans un groupe avec λ et μ quelconques

3.1 Heuristique proposée

L'heuristique pour résoudre le problème de diffusion contrainte dans un groupe est la suivante :

1. construction d'un nouveau graphe $G' = (V', E')$ qui contient $G_1 = (V_1, E_1)$ et $G_2 = (V_2, E_2)$ isomorphes à $G = (V, E)$. Chaque traducteur t_i^1 de V_1 est relié au traducteur correspondant t_i^2 de V_2 par une arête de coût nul comme indiqué sur la Fig. 2. Pour passer d'un sommet de G_1 à un sommet de G_2, on doit emprunter une de ces arêtes de coût nul. On a donc $V' = V_1 \cup V_2$ et $E' = E_1 \cup E_2 \cup \{t_i^1 t_i^2\}_{1 \leq i \leq |T|}$.
2. toutes les valuations des arêtes de G_1 sont multipliées par λ et toutes les valuations des arêtes de G_2 sont multipliées par μ.
3. construction d'un arbre couvrant $\{e\} \cup S_2$ dans G' où $S_2 = V_2 \cap S$ grâce à une heuristique ρ-approchée du problème de Steiner.

3.2 La solution

La solution pour notre problème est le graphe dans G construit à partir des arêtes de l'arbre couvrant $\{e\} \cup S_2$ dans G'. L'arbre de Steiner sur $\{e\} \cup S_2$ dans G' constitue une solution de même coût pour le problème de diffusion contrainte appliqué dans G, et réciproquement. On a donc équivalence des solutions pour ces deux problèmes[†]. Comme l'arbre de Steiner sur $\{e\} \cup S_2$ est construit grâce à une heuristique ρ-approchée, l'heuristique que nous proposons est elle aussi ρ-approchée.

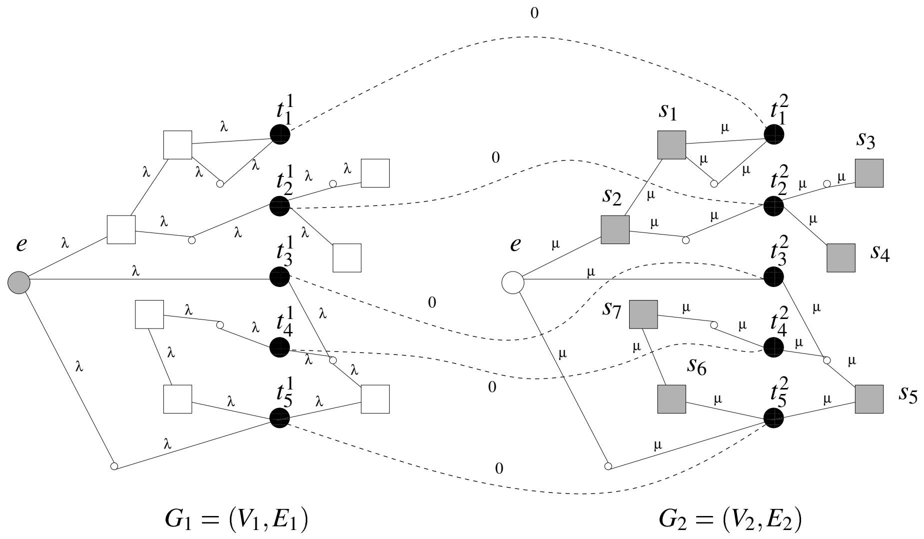

FIG. 2 – *Construction d'un arbre couvrant $\{e\} \cup S_2$ (les sommets grisés) dans G'*

4 Équilibrage des traducteurs

Nous considérons pour ce problème que chaque destinataire représente une charge de travail pour le traducteur auquel il est affecté. Dans le problème précédent, un traducteur peut être relié à beaucoup de destinataires alors que d'autres ont une charge de travail nulle et ne sont reliés à aucun destinataire : la forêt est déséquilibrée.

Nous nous intéressons au problème qui consiste à construire une forêt équilibrée dans laquelle les traducteurs sont tous utilisés et sont tous reliés au même nombre de destinataires : ils ont donc une charge de

[†] Cette réduction a été trouvée simultanément par Aubin Jarry (projet Mascotte, INRIA/Sophia-Antipolis).

travail équivalente (on suppose sans perte de généralité que $|S|$ est un multiple de $|T|$ et chaque traducteur est relié à $\frac{|S|}{|T|}$ destinataires).

4.1 Heuristique proposée

1. On calcule une forêt F dont $A_1, A_2, \dots A_k$ sont les arbres (non forcément équilibrée mais A_i contient t_i) avec l'algorithme ρ-approché présenté en section précédente en posant $\lambda = 0$ et $\mu = 1$. Soit ω_F son coût, c'est à dire la somme des valuations des arêtes de F.
2. On calcule un arbre A_0 couvrant $\{e\} \cup T$ avec l'algorithme ρ-approché du problème de Steiner. Soit ω_{A_0} son coût c'est à dire la somme des valuations des arêtes de A_0.
3. On construit une forêt équilibrée $F' = (A'_1, .., A'_k)$ à partir de F en reliant à des arbres déficitaires (reliés à moins de $\frac{|S|}{|T|}$ destinataires) des destinataires appartenant à des arbres majoritaires (reliés à plus de $\frac{|S|}{|T|}$ destinataires).

4.2 La solution

On note dans cette partie $|S| = m$ et $|T| = k$ et x_i le nombre de destinataires de l'arbre A_i. On pose $\forall i \in \{1, \dots, k\}$, $d_i = x_i - \frac{m}{k}$. d_i représente le nombre de destinataires manquants ($d_i < 0$) ou en trop ($d_i > 0$) dans l'arbre A_i. On a donc $\sum_{i=1}^{k} d_i = 0$.

Le but est maintenant d'équilibrer. Soit x_{ij} le nombre de destinataires que le traducteur i va transférer au traducteur j. Comme on veut l'équilibre on a $\sum_j x_{ij} - \sum_j x_{ji} = d_i$.

Lemme 1 *Il existe une solution ρ′-approchée, avec $\rho' = \rho\ max(k, 1 + \frac{\mu}{\lambda}(k-1))$, telle que au plus $(k-1)$ x_{ij} sont différents de 0.*

Preuve Si $k = 2$ alors $d_1 = -d_2$ (on suppose sans perte de généralité que $d_1 \geq 0$). On pose $x_{12} = d_1$.

Si $k > 2$, on prend i tel que $d_i > 0$ (si i n'existe pas alors la forêt est déjà équilibrée) et $d_j < 0$. Si $|d_i| > |d_j|$, on prend $x_{ij} = -d_j$, $d_i = d_i + d_j$, $d_j = 0$ et sinon on prend $x_{ij} = d_i$, $d_j = d_i + d_j$, $d_i = 0$. Dans tous les cas, on réduit le problème car l'un des d_i atteint la valeur 0.

On en déduit une solution à notre problème de la façon suivante :

Si d_i était initialement positif : on choisit $\frac{m}{k}$ destinataires de l'arbre A_i et on prend comme arbre A'_i final un sous-arbre de A_i permettant de les connecter et contenant t_i.

Si d_i était initialement strictement négatif alors l'ensemble $I = \{j | x_{ji} > 0\}$ est non vide. On choisit pour tout j dans I, x_{ji} destinataires non encore choisis et un sous-arbre A_{ji} de A_j permettant de les connecter et contenant t_j. On construit l'arbre A'_i associé au traducteur t_i en prenant A_0, A_{ji} pour tout $j \in I$ et A_i (éventuellement on peut supprimer des branches inutiles de A_0).

La solution finale contient donc au plus k copies de A_0 (dont une est utilisée pour connecter e aux traducteurs) et k copies des A_i. Elle a donc un coût de $\lambda\omega_{A_0} + \mu((k-1)\omega_{A_0} + k\omega_F))$. Comme nous utilisons une ρ-approximation pour construire A_0 et F, ce coût est inférieur à $\rho\lambda\omega^*_{A_0}(1 + (k-1)\frac{\mu}{\lambda}) + k\rho\mu\omega^*_F$.

L'algorithme est donc ρ′-approché avec $\rho' = \rho\ max(k, 1 + \frac{\mu}{\lambda}(k-1))$.

□

On remarque que la garantie de performance est $k\rho$ si $\mu \leq \lambda$. L'heuristique est donc efficace si λ est supérieur à μ. Dans le cas contraire (les traducteurs décompressent l'information transmise par exemple), une heuristique qui favorise la proximité des destinataires et des traducteurs est mieux adaptée. L'heuristique suivante peut être une réponse à ce but. L'idée est d'affecter d'abord les destinataires aux traducteurs puis de construire les arbres adaptés à cette partition.

4.3 Adaptation de l'heuristique

4.3.1 Équilibrage de deux traducteurs

Nous considérons dans un premier temps que le graphe ne possède que deux traducteurs. Les destinataires sont triés par rapport à leur distance à l'un des traducteurs, t_1 par exemple. Les $\frac{|S|}{2}$ premiers sont affectés à t_1

et les destinataires restants sont affectés à t_2. On construit ensuite A'_1 et A'_2 en calculant deux arbres couvrants sur les deux partitions trouvées. On remarque que l'on peut aussi répartir les destinataires autrement que par moitié.

4.3.2 Généralisation à k traducteurs

Soit $|T|$ quelconque. On partionne T en deux ensembles T_1 et T_2 les plus distants possibles, la distance entre deux ensembles étant la distance minimale entre deux sommets appartenant à des ensembles différents. Cette partition peut-être faite en temps polynomial [MKP03].

On fusionne les traducteurs de T_1 en un traducteur t_1 et ceux de T_2 en un traducteur t_2. On applique l'équilibrage sur 2 traducteurs comme vu précédemment en affectant un nombre de destinataires à t_1 (resp. à t_2) égal à $\frac{|T_1|}{|T|}|S|$ (resp. $\frac{|T_2|}{|T|}|S|$).

Une fois cette affection faite, on recommence récursivement sur T_1 et sur T_2 jusqu'à ce que tous les traducteurs aient été traités. On construit ensuite les arbres correspondants aux partitions.

L'heuristique présentée ci-dessus peut-être une alternative à l'heuristique précédente lorsque $\lambda \leq \mu$. Des simulations et une analyse poussée de l'heuristique nous permettront d'avoir des résultats sur sa garantie de performance.

5 Conclusion et perspectives

Le problème de diffusion contrainte dans un groupe que nous venons d'exposer dans ce papier nous permet d'entrevoir de multiples pistes de recherche. Nous avons exposé, en section 3, une heuristique ρ-approchée pour le problème général où ρ est le facteur d'approximation de l'heuristique utilisée pour construire un arbre de Steiner.

Nous avons travaillé sur plusieurs variantes du problème de diffusion contrainte dans un groupe et nous avons proposé pour chacun de ces problèmes des heuristiques avec garantie de performance pour les résoudre. Notamment, nous nous sommes intéressés au problème consistant à ne pouvoir choisir qu'un nombre limité de traducteurs en considérant que chaque traducteur induit un coût d'utilisation et nous avons montré l'existence d'une heuristique avec garantie de performance. Le problème de diffusion contrainte avec comme critère d'optimisation la distance des destinataires aux traducteurs est un problème auquel nous nous sommes aussi intéressés.

Pour le problème de diffusion contrainte dans un groupe comprenant la contrainte de l'équilibrage, nous exposons une heuristique ayant un facteur d'approximation qui dépend du nombre de traducteurs. Il reste à savoir si la garantie de performance peut être indépendante de k : ce problème est il approximable? Des simulations et une analyse de l'heuristique présentée en section 4.3 sont d'autres pistes de recherche à envisager.

Références

[HS86] D. S. Hochbaum and D. B. Shmoys. A unified approach to approximation algorithms for bottleneck problems. *J. ACM*, 33:533–550, 1986.

[Kar72] R. Karp. Reducibility among combinatorial problems. *Complexity of Computer Computations*, pages 85–103, 1972.

[LR00] C. Laforest and B. Rouits. Le problème du k-fournisseur: technique d'approximation et aspect incrémental. *INRIA algotel 2000*, 2000.

[MKP03] J. Moulierac, J.C. König, and J. Palaysi. Diffusion contrainte dans un groupe. Mémoire de DEA Université de Montpellier II, 2003.

[RZ00] G. Robins and A. Zelikovsky. Improved Steiner tree approximation in graphs. In *Proc. of SODA 2000*, 2000.