

Mourad Baïou

Bureau:

CUST

Campus des Cézeaux -BP- 206

63174 Aubière Cedex

Tèl : 04 73 40 52 92

Laboratoire d'Econométrie

Ecole Polytechnique

1 rue Descartes, 75005 Paris.

Né le 17-07-1969 à Alger

Formation universitaire

- 1991-1992 : DEA de recherche opérationnelle à l'IMAG, Institut de Mathématiques Appliquées de Grenoble, Université Joseph Fourier et Institut National Polytechnique de Grenoble.
- 1992-1996 : Thèse de Doctorat en Informatique, Université Rennes 1.
Titre : Le problème du sous-graphe Steiner 2-arêtes connexe : Approche polyédrale.
Soutenue le 17 décembre 1996 avec mention très honorable.

Activités professionnelles

- 1995-1997 : Attaché Temporaire d'Enseignement et de Recherche, Département d'informatique de l'université de Brest.
- 1997-1998 : Attaché de Recherche à l'Ecole Polytechnique, Laboratoire d'Econométrie, Paris.
- 1998-199 : Attaché de Recherche à Universidad de Chile, Departamento de Ingenieria Matematica, Santiago, Chili.
- Depuis le 1^{er} septembre 1999 : Maître de conférences en recherche opérationnelle au CUST, Centre Universitaire des Sciences et Techniques, Université Blaise Pascal, Clermont-Ferrand.

Enseignements

Mes enseignements se déroulent principalement au département de Génie Mathématique et Modélisation (GMM) de l'école CUST de l'université Clermont II.

- Troisième année GMM: “optimisation combinatoire”
Approche Polyédrale pour les problèmes d’optimisation combinatoire: Description partielles de l’enveloppe convexe des solutions, contraintes valides et leur séparation, algorithmes des plan-coupanants. Applications: le problème du couplage et celui du voyageur de commerce.
- Deuxième année GMM: “optimisation combinatoire”
la complexité des problèmes d’optimisation combinatoire; flots maximum et coupes de capacité minimum; flots de coût minimum: la méthode du simplexe pour les réseaux; programmation dynamique; programmation linéaire en nombres entiers: coupes de Gomory, méthode duale fractionnaire, méthodes par séparations et évaluations, relaxation Lagrangienne; algorithmes d’approximation: algorithmes gloutants, algorithmes de recherche locale, algorithmes d’approximation pour le problème du voyageur de commerce Euclidien et le problème du sac-à-dos.
- Première année GMM: “programmation linéaire et théorie des graphes”
méthode du simplexe; dualité et le théorème des écarts complémentaires, la méthode révisée du simplexe, la méthode primale duale; Introduction à la théorie des graphes, connexité et forte connexité; les problèmes du plus court chemin; les arbres de poids minimum; l’ordonnancement simple: représentation par un graphe potentiels tâches et par un graphe PERT.
- Cours initiation à l’optimisation combinatoire et approche polyédrale donné à l’Instituto de Matematica y Ciencias Afines, Universidad Nacional de Ingenieria, Lima, Pérou (décembre 2002): Polyhedral techniques to solve combinatorial optimization problems: a mini-course.

Synthèse des activités de recherche

Recherche fondamentale

- **Allocations stables**

Le problème des allocations stables a été introduit par Baïou et Balinski. Il peut être interprété comme celui du problème de transport classique mais *ordinal*. En plus clair, dans un problème de transport classique (cardinal) chaque agent $i \in I$ est muni d’une quantité $s(i)$ d’un bien à offrir; et chaque agent $j \in J$ demande une quantité $d(j)$ du même bien. Si un agent i affecte une quantité x_{ij} à l’agent j , alors un coût égal à $c_{ij}x_{ij}$ est généré. Il s’agit de trouver une affectation qui minimise la somme des coûts générés. Dans un problème de transport ordinal il n’est plus question de coûts mais d’ordres. Chaque agent $i \in I$ a un ordre de préférences sur les agents de J et vice-versa. $x = (x_{ij})$ est une *allocation stable*, si aucune paire d’agents (i, j) , $i \in I$ et $j \in J$, ne peut améliorer sa situation. La première question qui se pose est : *existe-t-il toujours une allocation stable?* Un algorithme fortement polynomial établit l’existence d’allocations stables. Aussi, nous étudions leurs structures : il est montré que l’ensemble des allocations stables est un treillis distributif en général; mais dans le cas “nondégénéré” il forme un ordre linéaire total. Dans le cas générique, quand un problème est “fortement nondégénéré”, il existe une unique allocation stable.

- **Typage approximatif des données semi-structurées**

Les données *semi-structurées*, par exemple XML, se caractérisent par une structure qui peut être absente, irrégulière, implicite ou partielle. Dans beaucoup de domaines d’applications, tels que les bibliothèques digitales, le commerce électronique et les bases de données du génôme, les données résident sous forme semi-structurée et ne peuvent pas être contraintes par un schéma rigide comme dans les systèmes classiques de base de données.

Les données semi-structurées sont représentées par un graphe orienté $G = (V, A, L)$ où V est l’ensemble des sommets, A est l’ensembles des arcs et L est une fonction qui associe un label

à chaque arc. Les sommets du graphe représentent des objets. Les relations entre les objets sont représentées par des arcs. Ces relations permettent de définir le type parfait d'un objet. Le problème est de limiter le nombre de types à une constante préfixée p , tout en minimisant la *distance* par rapport au type parfait. Cette distance est la distance de Manhattan.

Ce problème a été déjà traité dans littérature. Il est ramené à un problème d'optimisation combinatoire, *le problème de k -clustering*. Seules des heuristiques ont été appliquées pour sa résolution.

Dans notre approche le problème est ramené à une instance du problème k -median. Dans une étude préliminaire, nous montrons que pour plusieurs instances de données semi-structurées générées aléatoirement la relaxation linéaire du problème du k -median suffit pour obtenir des solutions optimales.

- **Problème de localisation multipériode**

Parmi les problèmes qui se posent lors de la conception d'un réseau de télécommunication est celui de la localisation des sites d'installations des concentrateurs pour servir un ensemble de sites demandeurs. Soit $G = (I, J, E)$ un graphe biparti complet. I est l'ensemble de sites offreurs, susceptibles de recevoir des concentrateurs, et J est l'ensemble des sites demandeurs. Chaque site offreur peut servir n'importe quel site demandeur. Chaque concentrateur a une capacité fixe K et le coût d'installation d'un concentrateur dans un site $i \in I$ est égal à f_i . Chaque site demandeur j lui est associé une demande d_j . Aussi il faut ajouter un coût unitaire c_{ij} de transport d'une unité du site i au site j . Le problème est de *localiser* les sites où installer les concentrateurs et comment servir les sites demandeurs afin de minimiser le coût total.

Le terme *mutipériode* provient du fait que les demandes fluctuent d'une période à l'autre, et donc le modèle doit répondre à ces changements.

- **Sous-graphes 2-arête connexes : Topologie des réseaux de télécommunication**

Soit un réseau formé d'un ensemble de noeuds et de connexions, appelées aussi arêtes, où chacune est munie d'un coût d'installation. On dit qu'un réseau est 2-arête connexe s'il existe au moins deux chemins arêtes-disjointes entre chaque paire de noeuds. Etant donné un sous-ensemble de noeuds appelés *terminaux*, *le problème d sous-réseau 2-arête connexe Steiner* est de trouver un sous-réseau 2-arête connexe contenant tout les terminaux tel que la somme des coûts de ses connexions soit minimum. *Le problème de l'arbre Steiner* est de trouver un arbre contenant les noeuds terminaux tel que la somme des coûts sur les arêtes de l'arbre soit minimum. Ces deux problèmes se posent dans la conception des réseaux de télécommunication.

Une vaste littérature traite le problème de l'arbre Steiner mais considère seulement des variables sur les arêtes. D'autres formulations dites *étendues* ajoutent aussi des variables sur les sommets. Ces formulations sont beaucoup plus simple et s'avèrent plus efficaces. L'étude du problème dans une dimension plus grande que celle des variables originales se simplifie.

Dans notre travail, nous avons considéré une formulation étendue pour le problème du sous-graphe 2-arête connexe. Nous avons introduit de nouvelles inégalités valides et des procédures de séparation. Cette formulation étendue permet de résoudre des problèmes plus généraux quand les noeuds sont aussi munis d'un coût ce qui est le cas en pratique. Une question se pose : Est-ce-que la formulation étendue est efficace seulement pour les arbres et autres structures sous-jacentes?

Recherche appliquée

- **Etude de faisabilité pour la transformation pluie-débit**

Le SIAAP (Syndicat Interdépartemental pour l'Assainissement de l'Agglomération Parisienne), dont la mission consiste à épurer les eaux usées de l'agglomération parisienne (8.5 millions de franciliens répartis sur 1900 km²), envisage de se doter d'un système d'aide à la décision pour optimiser la stratégie de contrôle des effluents arrivant dans le réseau d'assainissement .

Dans une première étape, le SIAAP a souhaité vérifier une idée originale.

Serait-il possible de prédire les débits arrivant dans le réseau SIAAP en temps réel sur la base de deux types d'information : les images radars délivrées par Météo France enregistrant toutes les 5 minutes les pluies, et les débits connus du passé?

Si cette idée était réalisable — c'est à dire s'il est possible de trouver une transformation mathématique pluie-débit suffisamment précise sur un bassin versant — le SIAAP pourrait alors généraliser l'approche à l'ensemble du réseau puis envisager de construire sur ces bases un système pour optimiser la gestion en temps réel des effluents.

Notre étude a permis de donner une réponse positive : il existe bien une transformation mathématique simple permettant de prédire les débits à attendre une demi-heure plus tard.

Les données ont été fournies dans 365 fichiers (extension .dta). Chaque fichier contient les pluies d'un jour enregistrées toutes les 5 minutes et un seul fichier contenant les débits de l'année 1999 enregistrés toutes les demi-heures. Ces données ont été récupérées sous le Système de Gestion de Bases de Données (SGBD) MS ACCESS. Ce logiciel a été choisi pour l'implémentation de notre modèle compte tenu de son prix (inférieur à 1000FF), sa simplicité (apprentissage facile) et des ressources requises pour son installation (12 Mo de RAM et 40 Mo d'espace disque).

Cette étude a donné lieu à un rapport confidentiel de 14 pages fourni au SIAAP :

Mourad Baïou, Michel Balinski et Farouk Toumani, "Etude de Faisabilité pour la Transformation Pluie-Débit".

- **Découpage électoral**

Le découpage actuel des départements Français en un nombre de circonscriptions égal au nombre de députés de chaque département, a été effectué en 1986 sous la base du recensement de 1982. Depuis, il y eut deux recensements en 1990 et 1999. Mais le découpage reste inchangé au point où la 6^{ème} circonscription du Var compte 2,4 fois le nombre d'habitants de la 1^{ère} circonscription, ce qui représente une inégalité de 144%. Et ce n'est qu'un exemple!

Prenons le département du Var. Il contient 43 cantons et 7 députés. Du point de vue mathématique un découpage du Var n'est rien d'autre qu'une partition des nombres 1,2,...,43 en 7 sous-ensembles disjoints. En gros, un découpage est équitable si le nombre d'habitants de chaque circonscription est plus égal autant que possible, l'allure des circonscriptions est raisonnable (n'est pas biscornue) et que chaque député doit pouvoir se déplacer de n'importe quelle localité de sa circonscription à n'importe quelle autre, sans jamais en sortir.

Le problème de découpage électoral est un *problème d'optimisation combinatoire*. Dans notre modèle nous avons redécoupé un échantillon formé de 12 départements. Ses résultats préliminaires sont à apprécier dans la table ci-après.

Cette étude a été publiée dans *Pour la science* du mois d'avril 2002, et diffusée avec illustrations et animations dans l'émission Archimède (ARTE télévision) du 9 avril, d'après un script rédigé par moi même et Michel Balinski.

Département	Nombre de députés	Inégalité		Inégalité calculée
		1982	1999	
Hauts-de-Seine	13	44,17%	50,97%	18,41%
Val-d'Oise	9	32,87%	78,87%	17,70%
Alpes-Maritimes	9	26,63%	66,41%	18,94%
Var	7	31,66%	143,63%	8,76%
Puy-de-Dôme	6	24,53%	31,66%	1,72%
Vaucluse	4	17,85%	40,96%	5,71%
Tarn	4	25,11%	45,71%	1,69%
Mayenne	3	26,99%	28,82%	0,97%
Corrèze	3	18,47%	37,08%	0,22%
Meuse	2	28,20%	28,94%	0,17%
Hautes-Alpes	2	25,66%	30,64%	0,23%
Lozère	2	9,82%	13,85%	0,07%

Les 3ème et 4ème colonnes représentent l'écart entre la population la plus peuplée et la moins peuplée de chaque département, selon le découpage actuel. Dans la 3ème colonne les calculs sont réalisés sur la base du recensement de 1982 et dans la 4ème colonne sur celui de 1999. La dernière colonne donne l'écart calculé d'après le découpage fourni par notre modèle, utilisant les populations du recensement de 1999.

Articles scientifiques de recherche publiés

- M. Baïou et A. R. Mahjoub, “Steiner 2-edge connected subgraph polytopes on series-parallel graphs,” *SIAM Journal on Discrete Mathematics* 10 (1997) 505-514.
- M. Baïou, F. Barahona et A. R. Mahjoub, “Separation of partition inequalities,” *Mathematics of Operations Research* 25 (2000) 243-254.
- M. Baïou et M. Balinski, “Many-to-many matching : Stable Polyandrous Polygamy (or Polygamous Polyandry),” *Discrete Applied Mathematics* 101 (2000) 1-12.
- M. Baïou et M. Balinski, “The stable admissions polytope,” *Mathematical Programming* 87 (2000) 427-39.
- M. Baïou, “On the dominant of Steiner 2-edge connected subgraph polytope,” *Discrete Applied Mathematics* 112 (2001) 3-10.
- M. Baïou et M. Balinski, “The stable allocation (or ordinal transportation) problem,” *Mathematics of Operations Research* 27 (2002) 485-503.
- M. Baïou et M. Balinski, “Erratum: The stable allocation (or ordinal transportation) problem,” *Mathematics of Operations Research* 27 (2002) 662-680.
- M. Baïou et A. R. Mahjoub, “The Steiner traveling salesman polytope and related polyhedra,” *SIAM Journal on Optimization* 13 (2002) 498-507.
- M. Baïou et M. Balinski, “Student Admissions and Faculty Recruitment,” à paraître dans *Theoretical Computer Science*.
- M. Baïou et M. Balinski, “Admissions and Recruitment,” à paraître dans *The American Mathematical Monthly*.

Articles soumis ou en préparation

- M. Baïou et M. Balinski, “Stable Allocation Mechanisms”.
- M. Baïou et J. R. Correa, “The node-weighted 2-edge connected subgraph problem: linear relaxation, facets and separation”.
- M. Baïou et Farouk Toumani, “Semistructured Data and Optimization”.
- M. Baïou et Hervé Kerivi, “The Multiperiod Capacitated Concentrator Location Problem”.

Autres articles et travaux

- M. Baïou et M. Balinski, “Flots et Couplages,” mars 2000, à paraître dans *Graphes et Applications*, édité par Jean Claude Fournier et Michel Las Vagnas, Hermès Sciences.
- M. Baïou et M. Balinski, “Découpage électoral,” *Pour la science* 294 (avril 2002) 60-64.
- M. Baïou et M. Balinski, “Charcuterie Fine,” *Emission Archimède, ARTE Télévision* (9 avril 2002).