

# AAP FR Agorantic 2016

## Titre

Optimisation et équilibre structurel dans les graphes signés : une application pour comprendre les marchés publics

## Equipe :

Rosa FIGUEIREDO (Porteuse) - Laboratoire Informatique d'Avignon (LIA)  
Pierre-Henri MORAND - Laboratoire Biens, Normes, Contrats (LBNC)  
Vincent LABATUT - Laboratoire Informatique d'Avignon (LIA)

## Table des matières

1	Description du Problème	2
2	Objectifs et attentes scientifiques	4
3	Innovation, Interdisciplinarité et Agorantic	4
4	Budget	5
5	Perspectives	5
	Références	6

## Résumé

Ce projet vise à étudier des problèmes de partitionnement définis sur des réseaux signés. Nous voulons déterminer et résoudre des variantes de ce problème pertinentes dans un certain domaine d'application. Nous proposons d'utiliser une méthode innovante combinant la *Recherche Opérationnelle* et les *Réseaux Complexes*, avec une application dans l'étude des marchés publics. Notre approche, de type ascendante, consiste à constituer un corpus de réseaux à partir de données réelles décrivant les collaborations et les concurrences dans un réseau d'entreprises, et à utiliser des outils issus de l'analyse de réseaux complexes pour les étudier. L'objectif de cette première phase est d'identifier les variantes du problème de partitionnement pertinentes pour nos données. Dans une deuxième phase, nous allons utiliser des outils d'optimisation pour proposer des méthodes de résolution de ces problèmes, et les appliquer à nos données.

# 1 Description du Problème

Ce projet vise à traiter des problèmes relatifs au partitionnement de graphes signés, ainsi qu'à identifier et à résoudre des variantes de ce problème qui soient pertinentes dans le cadre de l'étude de marchés publics. On propose de procéder en combinant deux approches complémentaires : d'une part, des méthodes d'optimisation conçues pour les graphes signés, et d'autre part, des outils développés pour l'analyse de réseaux complexes.

**Graphes signés.** Dans un *graphe signé*, chaque lien est annoté avec un signe : soit positif, soit négatif. Ce type de graphe a été défini par Heider en 1946 [31], dans le but de décrire des relations émotionnelles entre des personnes appartenant à un même groupe social. Depuis, les graphes signés ont été largement utilisés dans l'étude de réseaux sociaux [10, 21, 22, 33, 40], mais également dans d'autres domaines comme l'analyse de risque en gestion [30, 32] et l'étude de systèmes biologiques [19, 32], entre autres [13, 28]. L'élément commun à ces applications est le fait que chacune d'entre elles est définie dans un environnement caractérisé par des relations du type "conflit vs. collaboration", qui peuvent être modélisés naturellement via un graphe signé.

On dit qu'un graphe signé est *structurellement équilibré* s'il peut être partitionné en deux ou plus [20] sous-groupes mutuellement hostiles (i.e. reliés seulement par des liens négatifs) tout en exhibant une solidarité interne (i.e. contenant uniquement des liens positifs). Ce concept d'équilibre a été étendu de différentes façons dans la littérature. Mais quelle que soit la notion d'équilibre utilisée, un réseau réel (i.e. un graphe représentant un système du monde réel) est rarement *parfaitement équilibré* : on trouvera quelques liens positifs entre les groupes et quelques liens négatifs à l'intérieur. L'un des défis du domaine est de *quantifier* le niveau d'équilibre d'un tel réseau. Pour ce faire, on doit résoudre un problème d'optimisation sur le graphe : estimer la partition permettant d'obtenir l'équilibre maximal pour une mesure d'équilibre donnée. Chaque variante de la notion d'équilibre est susceptible de mener à un problème de partitionnement différent. Des formulations mathématiques ont été proposés avec succès pour résoudre ces problèmes, en particulier les formulations de programmation linéaire en nombre entiers [11, 16, 29, 34, 36]. Quelques travaux récents se sont intéressés aux problèmes relatifs à l'équilibre structurel [17, 15, T4, T3, T1], mais il y a encore beaucoup à faire, notamment en termes de revue des formulations mathématiques existantes et des méthodes de résolution proposées.

**Complex networks.** Un *réseau complexe* (ou *graphe de terrain*) est la représentation d'un système complexe du monde réel, sous forme de graphe. En tant que tel, ce type de graphe possède des propriétés topologiques non-triviales (par comparaison à des graphes complètement réguliers ou purement aléatoires), comme les propriétés *sans-échelle* (degré distribué en loi de puissance) [14] ou *petit-monde* (haute transitivité et faible distance moyenne) [39]. Par conséquent, des outils spécifiques ont été développés afin de les analyser et de mieux comprendre les systèmes qu'ils représentent [12, 37, 23].

Le domaine a été très actif pendant les deux dernières décennies, particulièrement sur le problème de la *détection de communautés* [26, 38]. Cette tâche consiste à partitionner l'ensemble des nœuds du réseau, avec pour contrainte d'obtenir des groupes de nœuds possédant de fortes connexions intra-groupe et de faibles connexions inter-groupes. Bien qu'il provienne d'un champ différent, ce problème peut être considéré comme une version plus générale de ceux présentés ci-dessus, relatifs à l'équilibre structurel. Cependant, dans le contexte de l'analyse de réseaux complexes, les efforts se sont concentrés presque exclusivement sur le traitement de réseaux non-signés, et seulement un petit nombre de travaux très récents se sont intéressés aux réseaux signés et à leur équilibre structurel.

L'un des intérêts de l'analyse des réseaux complexes pour ce projet est la méthodologie qu'elle propose, qui est principalement orientée données. En effet, de nombreux outils développés dans ce domaine ont été initialement conçus pour traiter un système bien spécifique. Mais en raison de la nature générique des graphes en tant qu'outil de modélisation, certains d'entre eux ont ensuite pu être utilisés pour étudier des systèmes complètement différents, tandis que certains autres ont simplement nécessité quelques ajustements. On peut par exemple citer le concept de *rôle communautaire*, initialement défini pour analyser des réseaux de protéines [27], et utilisé plus

tard pour caractériser certains utilisateurs particuliers dans le réseau Twitter [T2].

**Marchés publics.** Les marchés publics offrent un champ exploratoire extraordinairement riche [T5, T6, T8, T9]. Outre l'importance économique de ces modalités de transaction (on estime, en moyenne, à 15 pour-cent du PIB des principaux pays de l'OCDE qui relèveraient de la commande publique), l'impératif de transparence des données publiques (Open-Data) prend une signification toute particulière en ce domaine. Déjà initiée dans de nombreux pays, elle est même désormais, par exemple en France, une contrainte réglementaire. L'Ordonnance n°2015-899 du 23 juillet 2015 relative aux marchés publics et son article 56 précise ainsi que "les acheteurs rendent public le choix de l'offre retenue et rendent accessibles sous un format ouvert et librement réutilisable les données essentielles du marché public". Ce sont donc dès aujourd'hui une masse importante, et à court terme l'intégralité des données relatives aux marchés publics qui pourront faire l'objet d'analyses qui jusqu'à présent se heurtait à l'inexistence ou à la partialité de jeux de données. Quelques travaux pionniers ont montré l'intérêt des réseaux complexes dans l'analyse des marchés publics (voir par exemple [18]). En particulier, l'analyse des données ouvertes des marchés publics en Hongrie [25] permet de mettre en exergue des phénomènes de corruptions présents entre les initiateurs des appels d'offres et les entreprises sélectionnées. Pour autant, outre des *réseaux de corruption* (qui font apparaître des liens positifs entre certaines entreprises et les donneurs d'ordre), les données issues des marchés publics peuvent également faire apparaître des réseaux d'entente ou de cartel (qui constituent des liens positifs entre les entreprises partie prenante du cartel, au détriment de l'acheteur public), des *réseaux de concurrence*, lorsque sur certains marchés les mêmes entreprises se retrouvent en compétition répétée et enfin des réseaux de coopération lorsque, alloti, un marché public implique la sélection de plusieurs entreprises pour sa réalisation. Les graphes signés peuvent être utilisés pour combiner ces informations et les représenter sous forme de structure discrète.

**Approche.** Nous visons à attaquer le problème de l'équilibre structurel du point de vue des réseaux complexes, en adoptant une approche ascendante. Nous étudierons d'abord les réseaux signés modélisant des données décrivant l'activité de marchés publics, de manière à identifier les différentes formes que le problème de l'équilibre structurel peut prendre en pratique. Ceci nous permettra de développer une formation mathématique pour chaque variante, et de proposer des méthodes pour les résoudre.

Le point de départ de ce projet sera l'étude de la méthode de résolution de la version relaxée du problème de Correlation Clustering [T4]. Notre but est de répondre à des questions pratiques, telles que : Combien ce réseau est-il équilibré ? Est-ce qu'un réseau dynamique évolue vers un état d'équilibre ? Certains groupes de nœuds sont-ils plus équilibrés que d'autres ? Est-il possible d'identifier des groupes possédant un rôle spécifique (par exemple un groupe de médiation) dans de tels réseaux ?

L'autre aspect que nous voulons étudier est l'effet de la structure du réseau sur les méthode de résolution. Dans le cas de la détection de communautés, on sait que la présence de certaines propriétés topologiques affecte négativement certains algorithmes [35]. Nous sommes conscients du fait que nos objectifs soulèvent un certain nombre de problèmes méthodologiques, en particulier pour ce qui concerne la constitution du corpus, la génération de réseaux artificiels destinés à être utilisés comme banc d'essai, la mesure de la performance et le traitement de réseaux à grande échelle.

Ce projet peut être vu comme l'extension du travail déjà réalisé par certains membres de l'équipe dans le cadre d'un projet LIA. Dans [T7], nous avons commencé à étudier la valeur informative des liens négatifs pour des réseaux complexes définis dans un contexte d'activité législative. À cette fin, nous avons extrait et analysé une collection de graphes signés représentant des sessions de vote au Parlement Européen (EP). Nous avons d'abord traité des données collectées par le site web VoteWatch Europe pour la 7<sup>ème</sup> législature (2009-2014), en considérant les similarités entre les votes des différents députés européens (DE), dans le but de définir des liens pondérés signés. Nous avons ensuite appliqué à ces données d'une part une sélection de méthodes de détection de communautés (conçues pour traiter seulement les liens positifs), et d'autre part l'algorithme Parallel Iterative Local Search (défini pour identifier des partitions équilibrées dans les réseaux signés). Nos résultats ont montré que, contrairement aux conclusions d'une étude récente portant

sur d'autres données [24], les partitions détectées en ignorant les liens négatifs sont significativement différentes de celles obtenues en les considérant. La pertinence des liens négatifs dans le cadre du partitionnement de graphe est donc une question ouverte qui devrait être explorée de façon plus approfondie, et sur d'autres données expérimentales.

## 2 Objectifs et attentes scientifiques

Les objectifs affichés pour ce projet sont les suivants :

1. Définir une base de données des réseaux signés représentant la concurrence vs. collaboration dans les marchés publics ; analyser ces réseaux et identifier des problèmes d'optimisation liés à l'équilibre structurel.
2. Développer des modèles mathématiques et des méthodes de résolution approchée pour les problèmes d'optimisation identifiés ; une attention particulière sera donnée aux problèmes admettant la présence de groupes possédant un rôle spécifique dans le réseau.
3. Étudier l'effet de la structure du graphe sur les problèmes retenus et les méthodes de résolution proposées.

Du point de vue méthodologique, nous visons à élaborer et à mettre en œuvre une méthode efficace pour résoudre une version relaxée du problème de *Correlation Clustering*. Du point de vue applicatif, voulons extraire un corpus de réseaux signés représentant l'activité d'un ensemble de marchés publics. En se basant sur l'intégration de ces deux étapes, notre objectif est de proposer une réponse (et de tenter d'élargir) aux questions liées à la détection de la corruption, des ententes, de collaboration ou de marchés compétitifs dans le cadre des procédures des marchés publics.

## 3 Innovation, Interdisciplinarité et Agorantic

Outre ces objectifs scientifiques, nous visons à promouvoir la recherche interdisciplinaire entre le LIA et le LBNC. Ce projet réunit des chercheurs de deux laboratoires membres de la structure fédérative de recherche Agorantic :

- Laboratoire Informatique d'Avignon (LIA) : Rosa FIGUEIREDO est spécialiste dans le développement de formulations de programmation linéaire en nombre entiers, d'approches de résolution exactes, de formulations étendues et de formulations robustes pour des problèmes d'optimisation combinatoire. Au cours des dernières années, elle s'est intéressée aux problèmes d'optimisation étudiés dans ce projet. Vincent LABATUT est spécialisé dans l'analyse des réseaux complexes. Il a travaillé sur la détection de communauté (i.e. le partitionnement de graphes non-signés), et il est expérimenté pour ce qui touche à l'analyse de réseaux représentant des systèmes du monde réel.
- Laboratoire Biens, Normes et Contrats (LBNC) : Pierre-Henri MORAND consacre une partie importante de ses travaux à l'analyse des mécanismes d'enchères et des marchés publics. Depuis ses travaux doctoraux (analyse des mécanismes d'enchères combinatoires), il mobilise les outils de la théorie des jeux non-coopératifs en asymétrie d'information pour déterminer les stratégies optimales des entreprises dans les appels d'offres (*bidding strategies*) ou caractériser les propriétés souhaitables des procédures (*mechanism design*). Il s'intéresse en particulier à la question de la formation et au comportement de groupements d'entreprises dans les marchés publics, à l'alotissement des marchés publics, à la dématérialisation de la commande publique ainsi qu'au rôle de l'accès à l'information dans le comportement stratégique des entreprises et des acheteurs.

Thématiquement, ce projet intègre deux axes d'Agorantic :

- Axis 2 (Réseaux sociaux, structures, contenus et usages) : le but principal de ce projet est l'extraction de réseaux sociaux et la résolution de problèmes relatifs à leur analyse. De ce point de vue, notre projet est plus proche de cet axe pour ce qui est de la méthodologie.

- Axe 3 (Patrimoine, territoire et politiques publiques) : L'objectif à terme est également de questionner la stratégie des territoires (en l'espèce les différentes entités administratives recourant aux marchés publics, municipalités, départements, régions, états) sur la nature, l'exhaustivité, la disponibilité des jeux de données ouverts et notamment dans leur capacité à autoriser par les outils que nous souhaitons développer, la détection des phénomènes d'entente, de capture ou de concurrence effective.

## 4 Budget

**Stage de Master : 3324 Euros.** Ce projet implique une quantité importante de travail. Le plus grand besoin pour sa réalisation est donc la main d'oeuvre, et cela motive cette demande de financement. Plus précisément, nous avons l'intention d'encadrer un étudiant de Master 2 lors d'un stage de 6 mois. Le stagiaire serait en charge du développement et de l'implémentation des méthodes d'optimisation proposées dans le projet, ainsi que de l'extraction et de l'analyse des réseaux et partitions obtenus. Cet étudiant recevra 554 euros par mois pendant 6 mois. Un étudiant d'intérêt a déjà été effectivement identifié : Victor dos Santos Ponciano est un étudiant en deuxième année de master (modélisation mathématique et informatique) à l'Université Fédérale Rurale de Rio de Janeiro, au Brésil.

**Conférences : 1600 Euros.** Nous prévoyons de présenter nos résultats dans une conférence internationale (ASONAM 2016 –IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining), ainsi que dans une conférence nationale (JAM 2016 –Journées de Microéconomie Appliquée). Nous nous attendons à dépenser environ 1600 euros pour ces conférences.

**Montant total sollicité auprès de la FR Agorantic : 4924 euros.**

## 5 Perspectives

**Travaux futurs.** Le travail décrit dans ce projet doit être considéré comme une étape préliminaire, qui serait suivie par un projet de plus grande envergure. Si nous parvenons, comme c'est prévu, à concevoir des outils efficaces, nous comptons d'abord développer l'aspect applicatif, and collaborant avec d'autres chercheurs qui utilisent traditionnellement des réseaux signés pour modéliser les systèmes qu'ils étudient. Comme indiqué en section 1, c'est notamment le cas de nombreux réseaux sociaux utilisés notamment en sciences politiques.

Du point de vue théorique, nous pouvons explorer de nombreuses directions en généralisant différents aspects de notre travail. Premièrement, il est possible de traiter différents types de réseaux : les réseaux orientés (liens unidirectionnels), les réseaux multiplexes (plusieurs types de liens, avec des sémantiques différentes, cohabitant dans un même réseau). Deuxièmement, il est également possible de travailler sur la détection de couvertures de l'ensemble des nœuds, au lieu de partitions. Cela revient à relaxer la contrainte qui veut que les groupes détectés sont mutuellement exclusifs : un nœud pourrait alors apparaître dans plusieurs groupes simultanément.

**Doctorat et élargissement de l'équipe.** Nous voulons poursuivre ce travail en utilisant l'équipe de ce projet en tant que noyau central, et en intégrant des membres supplémentaires. Nous considérons le stagiaire de master comme une partie de ce noyau, et notre objectif est de continuer à travailler avec lui à travers une thèse de doctorat encadré en collaboration entre le LIA et le LBNC. Les membres supplémentaires pourraient provenir du LIA et du LBNC, mais aussi d'autres laboratoires associées à la structure Agorantic, en particulier pour les points liés aux applications susceptibles d'être représentées par des réseaux signés.

**Projet ANR.** Avec l’objectif de financer la thèse de doctorat mentionnée ci-dessus, ainsi que l’élargissement de l’équipe, nous visons à soumettre un projet ANR en 2016. Comme nous l’avons mentionné précédemment, ce projet est un prolongement des travaux financés par un projet courte durée (12 mois) développé au LIA dans l’année 2014-15. Ce projet va aboutir à une journée de séminaires prévue pour le 18 Mars 2016. Nous avons déjà la confirmation de la participation de 5 chercheurs venant de différents domaines et pays. La proposition de l’ANR émergera de ces deux années de collaborations et nous permettra de développer des collaborations pluridisciplinaires au niveau du LIA, du LBNC, de l’UAPV et d’autres universités européennes.

## Références

### — Publications de l’équipe du projet —

- [T1] L. Drummond, R. Figueiredo, Y. Frota, and M. Levorato. Efficient solution of the correlation clustering problem : an application to structural balance. *Proceedings of OTM 2013 Workshops in LNCS*, 8186 :674–683, 2013.
- [T2] Nicolas Dugué, Vincent Labatut, and Anthony Perez. Identifying the community roles of social capitalists in the twitter network. In *IEEE/ACM International Conference on Advances in Social Network Analysis and Mining (ASONAM)*, pages 371–374, Beijing, CN, 2014.
- [T3] R. Figueiredo and Y. Frota. An improved branch-and-cut code for the maximum balanced subgraph of a signed graph. *CoRR*, 2013.
- [T4] R. Figueiredo and G. Moura. Mixed integer programming formulations for clustering problems related to structural balance. *Social Networks*, 35 :639–651, 2013.
- [T5] François Maréchal and Pierre-Henri Morand. Free riding in combinatorial first-price sealed-bid auctions. *The BE Journal of Theoretical Economics*, 9(1), 2009.
- [T6] François Maréchal and Pierre-Henri Morand. The public release of information in first-price sealed-bid auctions. *Review of Economic Design*, 16(4) :323–330, 2012.
- [T7] Israel Mendonça, Rosa Figueiredo, Vincent Labatut, and Philippe Michelon. Relevance of negative links in graph partitioning : A case study using votes from the european parliament. *CoRR*, abs/1507.04215, 2015.
- [T8] Pierre-Henri Morand. Accès des pme aux marchés publics : allotissement ou groupement? *Économie publique/Public economics*, (10), 2003.
- [T9] Pierre-Henri Morand. Smes and public procurement policy. *Review of Economic Design*, 8(3) :301–318, 2003.

### — Autres publications —

- [10] P. Abell and M. Ludwig. Structural balance : a dynamic perspective. *Journal of Mathematical Sociology*, 33 :129–155, 2009.
- [11] G. Agarwal and D. Kempe. Modularity-maximizing network communities using mathematical programming. *The European Physical Journal B*, 66 :409–418, 2008.
- [12] Reka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1) :47–96, 2002.
- [13] N. Bansal, A. Blum, and S. Chawla. Correlation clustering. In *Proceedings of the 43rd annual IEEE symposium of foundations of computer science*, pages 238–250, Vancouver, Canada, 2002.
- [14] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439) :509, 1999.
- [15] M. Brusco, P. Doreian, A. Mrvar, and D. Steinly. Two algorithms for relaxed structural balance partitioning : linking theory, models and data to understand social network phenomena. *Sociological Methods & Research*, 40 :57–87, 2011.
- [16] M. Brusco and D. Steinley. Integer program for one- and two-model blockmodeling based on prespecified image matrices for structural and regular equivalence. *Journal of Mathematical Psychology*, 53 :577–585, 2009.
- [17] M. Brusco and D. Steinly. K-balance partitioning : an exact method with applications to generalized structural balance and other psychological contexts. *Psychological Methods*, 15 :145–157, 2010.
- [18] TS Chandrasekar and Y Narahari. Procurement network formation : A cooperative game approach. In *Planning Production and Inventories in the Extended Enterprise*, pages 185–206. Springer, 2011.
- [19] B. DasGupta, G. A. Encisob, E. Sontag, and Y. Zhanga. Algorithmic and complexity results for decompositions of biological networks into monotone subsystems. *BioSystems*, 90 :161–178, 2007.
- [20] J.A. Davis. Clustering and structural balance in graphs. *Human Relations*, 20 :181–187, 1967.
- [21] P. Doreian and A. Mrvar. A partitioning approach to structural balance. *Social Networks*, 18 :149–168, 1996.
- [22] P. Doreian and A. Mrvar. Partitioning signed social networks. *Social Networks*, 31 :1–11, 2009.
- [23] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks. *Advances in Physics*, 51(4) :1079–1187, 2002.
- [24] Pouya Esmailian, Seyed Ebrahim Abtahi, and Mahdi Jalili. Mesoscopic analysis of online social networks : The role of negative ties. *Physical Review E*, 90(4) :042817, 2014.
- [25] Mihály Fazekas and István János Tóth. From corruption to state capture : A new analytical framework with empirical applications from hungary. *Corruption Research Center Budapest Working Papers No. CRCB-WP/2014*, 1, 2014.

- [26] Santo Fortunato. Community detection in graphs. *Physics Reports*, 486(3-5) :75–174, Feb 2010.
- [27] Roger Guimerà and Luís A. Nunes Amaral. Cartography of complex networks : modules and universal roles. *Journal of Statistical Mechanics*, 02 :P02001, 2005.
- [28] N. Gülpinar, G. Gutin, G. Mitra, and A. Zverovitch. Extracting pure network submatrices in linear programs using signed graphs. *Discrete Applied Mathematics*, 137 :359–372, 2004.
- [29] P. Hansen and B. Jaumard. Cluster analysis and mathematical programming. *Mathematical Programming*, 79 :191–215, 1997.
- [30] F. Harary, M. Lim, and D. C. Wunsch. Signed graphs for portfolio analysis in risk management. *IMA Journal of Management Mathematics*, 13 :1–10, 2003.
- [31] F. Heider. Attitudes and cognitive organization. *Journal of Psychology*, 21 :107–112, 1946.
- [32] F. Huffner, N. Betzler, and R. Niedermeier. Separator-based data reduction for signed graph balancing. *Journal of Combinatorial Optimization*, 20 :335–360, 2010.
- [33] T. Inohara. On conditions for a meeting not to reach a deadlock. *Applied Mathematics and Computation*, 90(1) :1–9, 1998.
- [34] E. L. Johnson, A. Mehrotra, and G. L. Nemhauser. Min-cut clustering. *Mathematical Programming*, 62 :133–151, 1993.
- [35] A. Lancichinetti, S. Fortunato, and F. Radicchi. Benchmark graphs for testing community detection algorithms. *Physical Review E*, 78(4 Pt 2) :046110, Oct 2008.
- [36] A. Mehrotra and M. A. Trick. A column generation approach for graph coloring. *INFORMS Journal of Computing*, 8 :344–354, 1996.
- [37] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45 :167–256, 2003.
- [38] M.A. Porter, J.-P. Onnela, and P.J. Mucha. Communities in networks. *Notices of the AMS*, 56(9) :1082–1166, 2009.
- [39] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684) :440–442, 1998.
- [40] B. Yang, W.K. Cheung, and J. Liu. Community mining from signed social networks. *IEEE Transactions on Knowledge and Data Engineering*, 19(10) :1333–1348, 2007.

