HAL open science

Multigraphes, multiplexes, et réseaux couplés

César Ducruet

To cite this version:

César Ducruet. Multigraphes, multiplexes, et réseaux couplés. 2012. halshs-00746129

HAL Id: halshs-00746129

https://shs.hal.science/halshs-00746129

Preprint submitted on 27 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Multigraphes, multiplexes et réseaux couplés

César Ducruet, CNRS, UMR Géographie-cités 8504
cdu<at>parisgeo.cnrs.fr

Octobre 2012 - Version 1

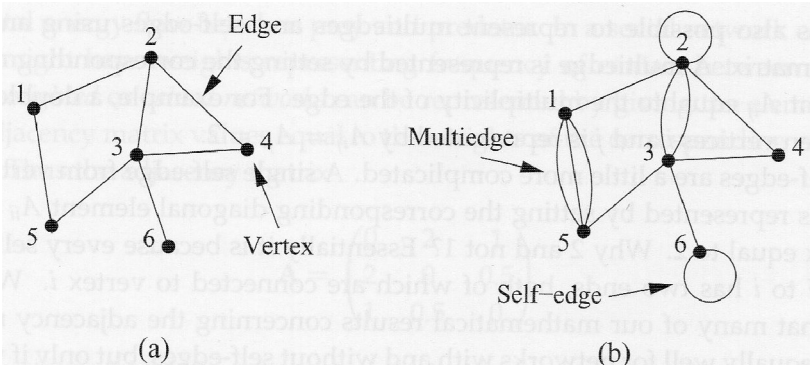
groupe f.m.r.

Introduction

La grande majorité des travaux sur les réseaux se caractérise par l'analyse de graphes simples (ou uniplexes), donc d'un seul type de liens entre les sommets. Cette synthèse s'attache donc plutôt à décrire les enjeux et méthodes propres à l'analyse d'un graphe aux relations multiples (Figure 1), plus communément appelé multigraphe (*multigraph* ou *multivariate directed graph*) ou réseau multiplexe voire multiplexé (*multiplex graph*). On appelle multigraphe orienté (*multidigraph*) un multigraphe ayant à la fois des liens orientés et de différente nature. Le pseudographe (*pseudograph*) est un multigraphe autorisant les boucles (*loops*, *self-loops* ou *self-edges*). La multiplexité, apanage des réseaux entretenant plusieurs types de relations entre ses composantes, serait apparue en sociologie suite aux travaux de Mitchell (1969 [20]) et de l'École de Manchester à travers ceux de Burt (1992 [8]) notamment sur les relations amicales et professionnelles entre individus dans le milieu de l'entreprise (voir sur ce point Eve, 2002 [13]). On parle aussi de relations unidimensionnelles ou bidimensionnelles en analyse de réseaux sociaux (Larson et Starr, 1993 [18] ; Degenne et Forsé, 1994 [10]).

L'absence de formalisation mathématique des propriétés spécifiques des multigraphes a longtemps motivé une simplification du réseau par remplacement des multiliens (*multiedges*) en liens simples (ou *single-stranded*) ayant plusieurs attributs afin de permettre l'application des méthodes conventionnelles (Wasserman et Faust, 1994 [27] ; Newman, 2010 [21]). D'autres travaux récents portent plutôt sur les effets globaux et locaux du couplage de réseaux différents afin de souligner les interdépendances, spécialisations et vulnérabilités en jeu entre plusieurs systèmes (Zhang *et al.*, 2005 [28]). Le multigraphe est alors analysé de façon implicite sous le terme de réseau couplé (*coupled network*) ou de « réseau de réseaux » (*network of networks*). Certains phénomènes comme l'accessibilité d'une ville dans un réseau de transport mul-

Figure 1 – Graphe simple (a) et multigraphe (b)



Source : Newman, 2010.

timodal (Ducruet *et al.*, 2011 [12]), les relations de confiance entre individus au sein d'une entreprise (Chollet, 2002 [9] ; Grossetti et Barthe, 2008 [15]), l'analyse d'un thésaurus (Boulet, 2011 [6]) ou encore l'apprentissage d'une langue (Gautier et Chevrot, 2012 [14]) ne peuvent ainsi être mis en lumière que sous l'angle combinatoire.

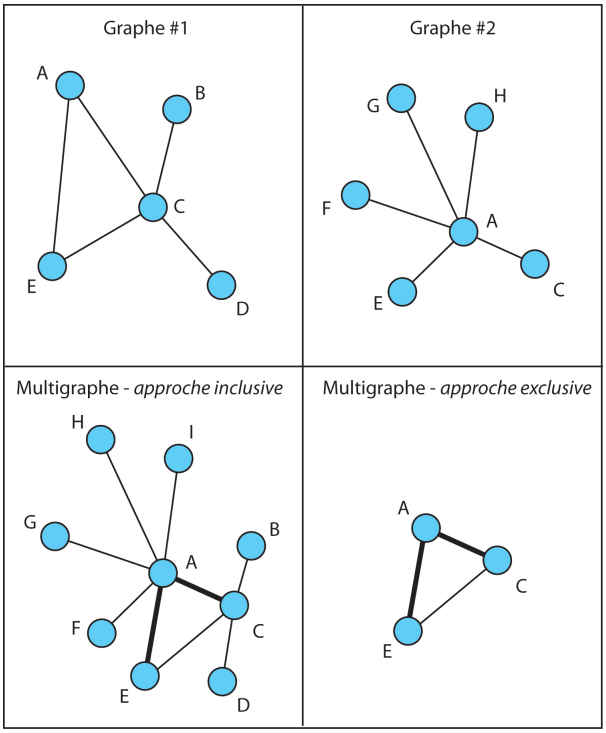
Sans prétendre à l'exhaustivité, cette synthèse présente quelques exemples d'analyse des multigraphes (le terme multiplexe étant équivalent nous n'emploierons que celui de multigraphe) afin de permettre un tour d'horizon des approches d'une discipline à une autre, l'ensemble des travaux examiné restant au final plutôt hétérogène.

1 Des graphes simples au multigraphe

Le passage au multigraphe n'est pas automatique. En fonction de l'objet de la recherche, on peut conserver l'intégralité des réseaux simples ou bien leurs seuls éléments communs (sommets et/ou liens) ; on parle d'approche inclusive (addition) ou exclusive (soustraction) (Figure 2). L'approche inclusive implique que tous les sommets soient concernés par la réunion des deux réseaux et que les sommets communs aux deux réseaux aient une fonction plurielle. L'approche exclusive supprime les sommets (et les liens) n'appartenant qu'à l'un des deux réseaux pour ne travailler que sur un échantillon réduit de sommets comparables.

Une réflexion sur la nature des sommets communs aux réseaux simples est parfois nécessaire et peut conduire à une agrégation, par exemple sur des bases géographiques (ex : gare routière et gare ferroviaire d'une même ville connectée aux deux réseaux). Deux réseaux de même type ne peuvent produire un multigraphe car leur réunion revient à augmenter la taille du graphe

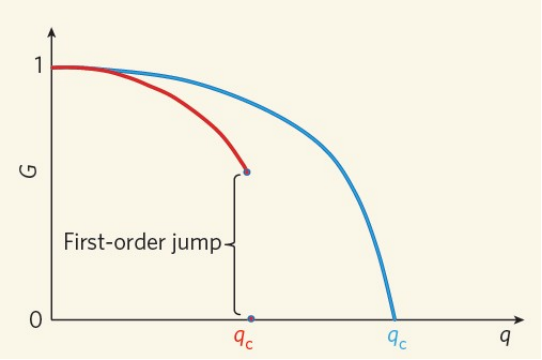
FIGURE 2 – Méthodes inclusive et exclusive



Seuls les sommets A, C et E sont communs aux deux graphes de départ, tout comme les liens AC et AE (chacun ayant deux arêtes). L'hétérogénéité des graphes d'origine peut nuire à l'intérêt de leur agrégation si très peu de sommets sont communs, rendant alors le multigraphe fortement vulnérable.

initial (ex : réseau Internet étatsunien combiné au réseau Internet canadien) et à modifier l'échelle d'analyse, l'étendue de l'aire d'étude. Quelle que soit la méthode choisie, on peut décider de compter le nombre d'occurrences, dans le cas d'un graphe non valué, ou bien d'additionner les quantités, dans celui d'un graphe valué, par lien et par sommet. Standardiser les valeurs initiales permet ainsi de calculer de nouveaux indicateurs de spécialisation (ex : part du flux ferroviaire dans le flux combiné rail-route entre la France et l'Allemagne). Enfin, une spécificité mesurable des multigraphes est leur degré (*degree centrality*) celui-ci correspondant au nombre d'arêtes et non au nombre de voisins adjacents.

FIGURE 3 – Impacts respectifs de la suppression des sommets dans un réseau individuel (bleu) et couplé (rouge)



G est le pourcentage de sommets interconnectés (non isolés) dans le réseau et Qc est la part des sommets dont la suppression mène à la dislocation du réseau. La dislocation est plus abrupte dans le cas des réseaux couplés (multigraphes) que dans celui des réseaux individuels (source : Vespignani, 2010).

2 L'analyse combinée aux niveaux global et local

2.1 Le niveau global

Les recherches sur les réseaux couplés ne mentionnent pas les multigraphes bien qu'il s'agisse en fait du même objet d'étude. Buldyrev *et al.* (2010 [7]) soulignent en premier lieu l'importance de la structure respective des réseaux simples en termes de taille et de propriétés générales. Ainsi, la fragilité du multigraphe s'accroît si les sommets couplés ont un degré comparable mais baisse si les sommets de fort degré dans le réseau A sont couplés à des sommets de faible degré dans le réseau B. Par simulation, Buldyrev *et al.* (*op. cit.*) montrent que la suppression itérative des sommets de fort degré dans les réseaux fortement couplés (dont la hiérarchie respective des degrés est équivalente) a un impact plus grand mesuré par la proportion de sommets appartenant au composant géant (sous-graphe incluant le plus grand nombre de sommets connectés ou *giant component*) qu'ils soient aléatoires ou sans échelle (Figure 3). La disparition de certains sommets peut ainsi avoir des effets variables sur l'un ou l'autre réseau dans leur globalité et sur les autres sommets selon leur situation respective dans l'un et l'autre réseau (sur la vulnérabilité des réseaux voir aussi Lhomme, 2012 [19]).

Les travaux empiriques évaluent de façon très diverse les effets et la robustesse du couplage. C'est par exemple le cas des réseaux électrique et Internet en Italie (Rosato *et al.*, 2008 [24]) ou bien des flux maritimes et aériens dans

le monde (Parshani *et al.*, 2010 [23]). Ces deux travaux concluent à l'existence de liens forts entre les systèmes étudiés sur la base de la proximité géographique de leurs sommets respectifs. Parshani *et al.* (*op. cit.*) montrent que les aéroports les mieux connectés (au réseau aérien) sont généralement situés à proximité des ports les mieux connectés (au réseau maritime) et concluent que les flux ne sont pas distribués aléatoirement. Un traitement en amont a été nécessaire : rendre les deux réseaux initiaux de taille identique par l'attribution (à partir des coordonnées géographiques) de chaque port à un aéroport à partir d'une distance minimale et par agrégation des ports et aéroports isolés (et donc reliés à un seul réseau) au nœud de même nature le plus proche. Les attributs (trafic) des sommets et des liens sont aussi sommés par la même occasion. Les auteurs proposent d'utiliser deux mesures globales spécifiques afin d'estimer le niveau de ressemblance entre les deux réseaux A et B :

- la corrélation inter degré-degré (*inter degree-degree correlation*) ou IDDC ;
- le coefficient d'inter-clustering global (*global inter clustering coefficient*) ou ICC.

La formule de l'IDDC est la suivante :

$$r^{AB} = \frac{1}{\sigma_q^2} \sum_{jk} jk(e_{jk} - p_j p_k)$$

où e_{jk} est la probabilité jointe qu'un lien de dépendance connecte un sommet de A au degré j et un sommet de B au degré k (assortativité / disassortativité) ; $p_j p_k$ est le produit des pentes respectives des droites de puissance associées à la distribution des degrés dans les réseaux A et B.

Les valeurs vont de -1 (les sommets de A ayant un fort degré connectent les sommets de B ayant un degré faible) à 1 (les sommets de degré élevé dans A et B sont connectés entre eux / identiques), la valeur 0 suggérant un couplage aléatoire des sommets. Cette mesure est identique à l'assortativité dans les réseaux simples (voir Ducruet, 2010 [11]) à ceci près qu'elle prend en compte la distribution des degrés (hiérarchie).

L'ICC se calcule ainsi :

$$c^{AB} = \frac{1}{M} \sum_j t_j$$

correspondant à la moyenne des coefficients d'inter-clustering locaux notés :

$$c_j^A = \frac{t_j}{k_j^A}$$

où t_j est le nombre de liens connectant les voisins du sommet A_j aux voisins du sommet B_j et k_j^A le degré du sommet A_j.

Le calcul est le même pour les sommets du réseau B. Ainsi la mesure globale au niveau des deux réseaux repose sur M, nombre total de liens de dépendance entre les deux réseaux.

L'obtention d'une corrélation inter degré-degré de 0.2 (contre 0 dans le cas de deux réseaux couplés aléatoirement) prouve que la proximité spatiale des réseaux simple est significative. L'application du modèle d'attachement préférentiel de Barabási-Albert (BA) montre aussi que le multigraphe invariant d'échelle est plus vulnérable aux attaques aléatoires qu'un multigraphe aléatoire.

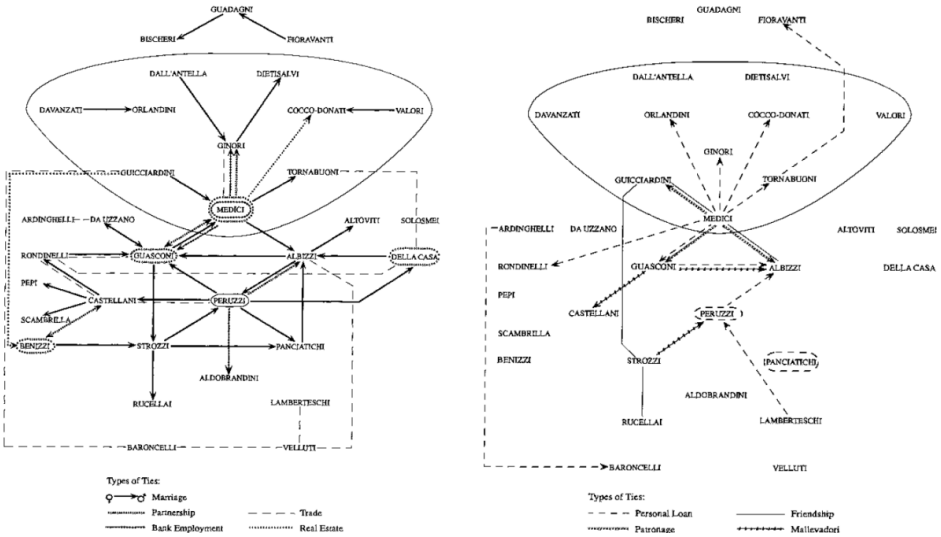
Bien que moins ciblée sur l'analyse de réseau proprement dite, l'application par Bogart (2009 [5]) d'un modèle *Vector Auto Regressive* (VAR) à plusieurs séries temporelles tirées d'archives britanniques (actes parlementaires) sur la création de routes, canaux et ports sur la période 1760-1830 permet de montrer les effets positifs de la croissance des routes sur la croissance des canaux.

2.2 Le niveau local

D'autres travaux insistent davantage sur les implications locales du couplage et utilisent d'autres méthodes d'analyse. Jin *et al.* (2010 [17]) par exemple proposent un indice de dominance pour les comtés chinois provenant de la somme de trois autres indices standardisés (densité du réseau de transports, proximité et accessibilité dans le réseau) qu'ils rapportent aux indicateurs classique de richesse et taille démographique, constatant alors sans surprise de fortes corrélations. L'originalité de leur indice tient dans le fait qu'il intègre plusieurs modes de transport.

L'analyse par Snyder et Kick (1979 [25]) des composantes du système-monde par *blockmodel* (pour une synthèse sur ces méthodes voir Beauguitte, 2011 [3]) à partir de quatre matrices de relations entre États (flux commerciaux, interventions militaires, relations diplomatiques et traités conjoints) s'en rapproche par le fait que leurs résultats proviennent de l'agrégation des quatre réseaux individuels en un seul, si bien que le rôle de chaque réseau dans l'émergence des sous-structure (et dans la position des États dans le réseau couplé) n'est pas analysé. Un travail du même genre a montré l'intérêt de comparer la structure des réseaux simples à celle du réseau combiné : l'émergence de l'Union Européenne est bien plus lisible et signifiante dans le réseau combiné des relations commerciales et des proximités diplomatiques entre États du monde que dans chaque réseau séparément (Beauguitte, 2010 [2]). L'étude des groupes familiaux basés sur le mariage par Alfani et Munno (2012 [1]) a aussi bien montré l'intérêt de combiner différents types de liens (parenté, témoins, beaux-parents) entre individus riches ou pauvres d'une ville moyenne italienne à l'époque moderne, même si leur analyse conclut

Figure 4 – Structure des multiliens interfamiliaux



Source : Padgett et Ansell, 1993.

à une faible surimposition des réseaux simples. Ce travail se situe dans la lignée de Padgett et Ansell (1993 [22]) étudiant neuf types de relations [1] chez 92 familles italiennes au début du 15e siècle (Figure 4). Ils appliquent notamment le *blockmodeling* aux multigraphes des relations de parenté et économiques puis à celui des relations économiques et amicales, permettant ainsi de mettre en lumière la structure en sous-groupes des liens et le rôle central des Médicis.

L'approche de Ducruet *et al.* (2011 [12]) veut vérifier l'adéquation entre la hiérarchie des sommets dans un réseau combiné et la spécialisation de ces sommets en fonction de leur participation inégale à l'un ou l'autre des réseaux, sur la base des flux aériens (vols passagers) et maritimes (escales de navires porte-conteneurs) mondiaux entre les villes, ceci à plusieurs niveaux d'agrégation des sommets (terminaux, agglomérations urbaines, corridors/mégalopoles). La méthode est inclusive. Si la part des sommets et liens communs aux deux réseaux augmente avec le niveau d'agrégation, c'est aussi le cas des corrélations linéaires (Pearson) entre les flux enregistrés au niveau des sommets et des liens ainsi qu'au niveau des réseaux dans leur totalité (méthode QAP : *Quadratic Assignment Procedure*). L'étude démontre aussi, certes sans surprise, l'importance du niveau géographique dans l'étude des

1. Relations de parenté (liens intermariages), économiques (liens commerciaux, propriétés conjointes, emplois bancaires, liens immobiliers), politiques (patronage et prêts personnels), et amicales (liens amicaux et de protection).

interdépendances entre réseaux différents mais imbriqués dans l'espace. En revanche, la corrélation entre niveau de flux et spécialisation n'est pas vérifiée partout dans le monde : si les *hubs* aériens et maritimes sont combinés en Asie par exemple (poids équilibré des flux aériens et maritimes dans le total), il sont découplés en Europe (ex : Rotterdam *hub* maritime et Londres *hub* aérien). Une typologie des villes est établie en fonction de la hausse ou de l'écart entre centralité dans un réseau individuel et centralité dans le réseau couplé. Ainsi certaines villes ou régions voient-elles leur centralité d'intermédiarité (*betweenness centrality*) augmenter ou diminuer suite au couplage par rapport à leur centralité dans le réseau simple. L'écart 1 (ratio entre centralité dans le réseau A et centralité combinée dans le réseau AB) peut être comparé à l'écart 2 (ratio entre centralité dans le réseau A et centralité combinée dans le réseau AB) afin de repérer les spécialisations. Enfin, l'analyse des flux majeurs (*single linkage analysis*) montre plutôt une diversité régionale des configurations obtenues.

Dans cette famille de travaux sur la spécialisation des sommets dans les multigraphes, une étude récente pour la DATAR (Berroir *et al.*, 2012 [4]) a proposé une analyse conjointe de sept réseaux reliant les aires urbaines françaises [2]. Le cumul des liens par paires d'aires urbaines a été considéré comme déterminant à partir de trois types d'échanges. Le partitionnement du réseau combiné produit un ensemble de sous-graphes plus ou moins polarisés selon que Paris soit inclue ou non dans l'analyse.

Conclusion

Malgré un manque évident de consensus sur les propriétés spécifiques des multigraphes en analyse de réseau, de nombreux travaux ont analysé différents réseaux simultanément et ce à travers une grande diversité de méthodes. La plupart du temps des mesures globales ou locales sont calculées sur le réseau couplé, fruit de la combinaison de deux ou plusieurs réseaux simples. Le rôle des réseaux simples dans la structure et la situation des sommets au niveau du multigraphe n'est pas toujours explicite ni même analysé pour autant, même si cette première étape permet l'analyse des spécialisations et interdépendances à l'œuvre. Il reste à vérifier en quoi ces spécialisations « réseau » diffèrent des spécialisations classiques mesurées à partir des stocks (ex : nombre d'emplois industriels dans le total de l'emploi *versus* connectivité des firmes industrielles dans le réseau des firmes tous secteurs). D'autres applications dans le domaine des transports cherchent plutôt à optimiser les déplacements de véhicules et de personnes à travers la prise en compte de

2. Mobilité domicile-travail, migrations résidentielles, flux à grande vitesse aériens et ferroviaires, liens établissement-siège dans les secteurs technologiquement innovants, partenariats scientifiques européens, liens établissement-siège dans tous les secteurs, et liens résidences principales - résidences secondaires.

modes différents (piéton, cycles, automobiles, transports publics etc.) lors des calculs d'accessibilité (voir Gueye, 2010 pour un état de l'art récent sur la question [16]). Il reste à affiner les implications pratiques de ces résultats eu égard au fait qu'en réalité, ports et aéroports ne sont pas directement interdépendants au moins d'un point de vue logistique ou circulatoire.

Références

[1] Alfani G., Munno C., 2012, "Godparenthood and social networks in an Italian rural community : Nonantola in the sixteenth and seventeenth centuries", in : Alfani C. and Gourdon V. (Eds), *Spiritual Kinship in Europe, 1500-1900*, Palgrave Macmillan, pp. 96–123.

[2] Beauguitte L., 2010, "Looking for European Union in the Word-System : a multi-graph approach", European Regional Science Association Conference, http ://www-sre.wu.ac.at/ersa/ersaconfs/ersa10/ERSA2010finalpaper698.pdf.

[3] Beauguitte L., 2011, « Blockmodeling et équivalences », Groupe fmr, http ://halshs.archives-ouvertes.fr/FMR/fr/.

[4] Berroir S., Cattan N., Guérois M., Paulus F., Vacchiani-Marcuzzo C., 2012, *Les systèmes urbains français*, Synthèse DATAR, Travaux en Ligne 10.

[5] Bogart D., 2009, "Inter-modal Network Externalities and Transport Development : Evidence from Roads, Canals, and Ports during the English Industrial Revolution", *Networks and Spatial Economics*, 9(3) : 309–338.

[6] Boulet R., 2011, « Introduction d'indices structuraux pour l'analyse de réseaux multiplexes. Application à l'analyse d'un thésaurus », Conférence MARAMI, Grenoble, 19–21 octobre.

[7] Buldyrev S.V., Parshani R., Paul G., Stanley H.E., Havlin S., 2010, "Catastrophic Cascade of Failures in Interdependent Networks", *Nature*, 464 : 1025–1028.

[8] Burt R., 1992, *Structural Holes : The Social Structure of Competition*, Cambridge (Mass.), Harvard University Press.

[9] Chollet B., 2002, L'analyse des réseaux sociaux : quelles implications pour le champ de l'entrepreneuriat ? Communication au 6e Congrès International Francophone sur la PME, HEC Montréal.

[10] Degenne A., Forsé M., 1994, *Les réseaux sociaux*, Paris, Armand Colin

[11] Ducruet C., 2010, « Les mesures globales d'un réseau », Groupe fmr, http ://halshs.archives-ouvertes.fr/FMR/fr/.

[12] Ducruet C., Ietri D., Rozenblat C., 2011, "Cities in worldwide air and sea flows : a multiple networks analysis", *Cybergeo*, http ://cybergeo.revues.org/23603.

[13] Eve M., 2002, « Deux traditions d'analyse des réseaux sociaux », *Réseaux*, 5(115) : 183-212.

[14] Gautier R., Chevrot J.P., 2012, « Influence des réseaux sociaux sur l'usage et l'acquisition du français langue seconde par des Américains lors de séjour d'étude en France : une étude exploratoire », Actes du 3e Congrès Mondial de Linguistique Française (CMLF), EDP Sciences.

[15] Grossetti M., Barthe J.F., 2008, « Dynamique des réseaux interpersonnels et des organisations dans les créations d'entreprises », *Revue Française de Sociologie*, 49(3) : 585–612.

[16] Gueye F., 2010, Algorithmes de recherche d'itinéraires en transport multimodal, LAAS-CNRS, communication, 14 décembre 2010.

[17] Jin F., Wang C., Li X., Wang J., 2010, "China's Regional Transport Dominance : Density, Proximity, and Accessibility", *Journal of Geographical Sciences*, 20(2) : 295–309.

[18] Larson A., Starr J.A., 1993, "A network model of organization formation", *Entrepreneurship Theory and Practice*, 17(2) : 5–15.

[19] Lhomme S., 2012, « L'analyse structurelle des réseaux techniques : modélisations, propriétés, vulnérabilités », Groupe fmr, http ://halshs.archives-ouvertes.fr/FMR/fr/.

[20] Mitchell C., 1969, "The Concept and Use of Social Networks", in Mitchell C. (ed.), *Social Networks in Urban Situations*, Manchester University Press.

[21] Newman M.E.J., 2010, *Networks : An introduction*, Oxford University Press.

[22] Padgett J.F., Ansell C.K., 1993, "Robust action and the rise of the Medici", *American Journal of Sociology*, 98(6) : 1259–1319.

[23] Parshani R., Rozenblat C., Ietri D., Ducruet C., Havlin S., 2010, "Intersimilarity between Coupled Networks", *Europhysics Letters*, 92(68002).

[24] Rosato V., Issacharoff L., Tiriticco F., Meloni S., De Porcellinis S., Setola R., 2008, "Modelling interdependent infrastructures using interacting dynamical models", *International Journal of Critical Infrastructures*, 4(1-2) : 63–79.

[25] Snyder D., Kick E.L., 1979, "Structural position in the world system and economic growth, 1955-1970 : A multiple-network analysis of transnational interactions", *American Journal of Sociology*, 84(5) : 1096–1126.

[26] Vespignani A., 2010, "Complex Networks : The Fragility of Interdependency", *Nature*, 464 : 984–985.

[27] Wasserman S., Faust K., 1994, *Social network analysis : Methods and applications*, Structural Analysis in the Social Sciences 8, Cambridge University Press.

[28] Zhang P.C., Peeta S., Friesz T., 2005, "Dynamic Game Theoretic Model of Multi-layer Infrastructure Networks", *Networks and Spatial Economics*, 5 : 147–178.