

Diversité des approches de modélisation statistique en analyse de réseaux sociaux multiniveaux

Julien Brailly, Emmanuel Lazega

► **To cite this version:**

Julien Brailly, Emmanuel Lazega. Diversité des approches de modélisation statistique en analyse de réseaux sociaux multiniveaux. *Mathématiques et Sciences Humaines, Centre de Mathématique Sociale et de statistique, EPHE*, 2012, 50 (198), pp.5-32. <hal-01053557>

HAL Id: hal-01053557

<https://hal-sciencespo.archives-ouvertes.fr/hal-01053557>

Submitted on 31 Jul 2014

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.

DIVERSITÉ DES APPROCHES DE MODÉLISATION STATISTIQUE EN ANALYSE DE RÉSEAUX SOCIAUX MULTINIVEAUX¹

Julien BRAILLY, Emmanuel LAZEGA²

RÉSUMÉ – *Cet article présente différentes approches statistiques combinant l'analyse multiniveau et l'analyse de réseaux sociaux, allant des modèles p2 multiniveau pour l'analyse de données de réseaux égocentrés aux Exponential Random Graph Models (ERGM) pour des données multiniveaux néostructurales. Nous développons premièrement les modèles stochastiques basés sur une dépendance dyadique, intégrant l'approche hiérarchique par une mise en parallèle de différentes expériences. Après avoir présenté la modélisation de type ERGM, nous nous demandons explicitement comment introduire de la dépendance « multiniveau » dans ces derniers. Nous terminons alors sur la formalisation proposée par Robins et Wang [2010] qui semble la plus aboutie pour traiter la dépendance verticale entre ces deux réseaux de niveaux différents, mais partiellement emboîtés.*

MOTS CLÉS – Affiliation, Analyse Multiniveaux, Analyse Stochastique de réseaux, ERGM, Réseaux sociaux, Sous-structures multiniveaux

ABSTRACT – *Diversity of statistical modeling approaches in multilevel social network analysis. This article presents some different statistical approaches for multilevel network analysis including the multilevel-p2 models for egocentric network data analysis and the Exponential Random Graph Models (ERGMs) for neo-structural multilevel data. First, we study the stochastic models based on dependence assumptions between dyads, including a hierarchical approach using jointly different networks studies. After presenting the ERGMs, we investigate how multilevel dependence may be introduced in them. We conclude then by the Robins and Wang formalization [2010] which seems to be the most successful attempt to treat the dependencies between the two partially nested networks of different levels.*

KEYWORDS – Affiliation, ERGM, Multilevel Analysis, Multilevel Substructure, Social network, Stochastic Network Analysis

¹ Ce texte s'appuie notamment sur des contributions multiples apportées par les intervenants du séminaire Réseaux et Régulation (ORIO-IRISSO, Dauphine et CMH, ENS-Jourdan) de 2010-2011 et par la Conférence "Research on the Organizational Society: Advances in Multilevel and Dynamic Network Analysis", Dauphine, ORIO, le 16 juin 2011. Ce travail s'inscrit dans le cadre du projet *Multi-Level Social Networks* financé par l'Agence Nationale de Recherche à l'IRISSO-CNRS et à l'ORIO.

² Institut de Recherches Interdisciplinaires en Sciences Sociales (IRISSO) et Observatoire des Réseaux Intra et Inter-Organisationnels (ORIO), Université Paris-Dauphine, Place de Lattre de Tassigny 75775 PARIS cedex 16, picard_julien@hotmail.fr, emmanuel.lazega@dauphine.fr

1. INTRODUCTION

Certains processus en jeu dans un *système d'échange généralisé* [Lazega, Pattison, 1999] peuvent être appréhendés suivant deux axes complémentaires : une coupe *horizontale* cristallisant les interdépendances entre acteurs et les effets de structures engendrés ; ou via un angle *vertical* articulant des niveaux d'analyse conceptuellement différents mais pertinents quant à l'explication du phénomène observé. D'un côté, la séparation entre niveaux permet de se concentrer sur les mécanismes et processus se situant dans le passage d'un niveau à l'autre. De l'autre côté, l'*Analyse de Réseaux Sociaux* se centre sur les interdépendances entre acteurs au sein d'un même niveau. L'objet de cette contribution n'est pas d'étudier d'un point de vue théorique les mécanismes et logiques permettant de passer d'un niveau à l'autre et les processus en jeu dans chaque niveau, mais plutôt de recenser les modélisations permettant d'articuler ces deux dimensions résultant de la même complexité.

En suivant la première perspective, il est possible de différencier au sein d'un même système d'action plusieurs grandes catégories d'acteurs : l'individu, la famille et plus généralement le réseau constitué par les connaissances de l'individu, l'organisation, l'État... Ainsi, bien que la majorité des systèmes d'action soient constitués et animés par des individus, il n'en demeure pas moins que ces derniers sont insérés dans une *société organisationnelle* [Coleman, 1990 ; Perrow, 1991 ; Lazega, 2002], dans la mesure où ils agissent la majorité de leur temps dans des organisations (entreprise, association, famille,...) qui les dépassent et qui influencent en partie leur comportement³. Il peut alors être pertinent d'essayer de regarder cette même réalité (le système d'action) produite à plusieurs niveaux analytiquement différents par la prise en compte dans un même mouvement de ces différents niveaux. Pour le dire autrement, il s'agit alors d'adopter une position regardant à la fois l'individu mais aussi le groupe, supposant par la même d'alterner les deux positionnements méthodologiques polaires, à savoir *individualisme* et *holisme*. Ce constat nous amène à étudier plus en détail l'*approche statistique multiniveaux* qui, selon Courgeau [2002] :

... libère enfin la démographie [idem pour la sociologie] de l'opposition entre holisme et individualisme : elle introduit des niveaux d'agrégations multiples à considérer simultanément et permet une synthèse des approches précédentes, tout en les dépassant [p. 49].

Le dépassement de ces deux perspectives dans cette synthèse est notamment permis par la prise en compte d'une dépendance verticale entre les niveaux via les effets d'interaction. Les modèles statistiques multiniveaux [Bryk, Radenbush, 1992 ; Goldstein, 1995 ; Snijders, 1999 ; Courgeau, 2004] ont connu un fort développement depuis une quinzaine d'années notamment parce qu'ils constituent une tentative de synthèse originale entre *holisme* et *individualisme* méthodologique en expliquant un comportement individuel par des caractéristiques de la personne et des groupes auxquels elle est reliée. Les groupes sont assimilés à des contextes car ils constituent bien un contexte pour l'action individuelle, mais en parallèle les individus construisent les groupes. Il nous semble alors que ce type de modèle permet une contextualisation verticale d'un système d'action par la prise en compte de la dépendance entre les niveaux. Plusieurs auteurs ayant développés ces méthodes [Snijders, Kenny, 1999 ; Courgeau, 1999, 2004] insistent sur le fait que les caractéristiques du groupe sont aussi construits à l'échelle des individus, et sont le résultat de leur interaction au sein du

³ « À côté des individus (natural persons) apparaît une nouvelle forme d'acteurs, l'acteur organisationnel moderne (corporate actor) qui agit et sur lequel on agit, qui poursuit ses intérêts et contrôle des événements » [Coleman 1990].

système d'action. C'est en ce sens qu'une coupe horizontale du *méta-système* constitué par les acteurs des différents niveaux complète la précédente méthode en endogénéisant des processus sociaux expliquant des différences entre groupes.

Dans une perspective plus horizontale, la sociologie de l'action organisée de Crozier et Friedberg [1977] en appréhendant l'organisation comme un construit, développe l'existence de ce que l'on peut concevoir comme un niveau organisationnel, intermédiaire entre le niveau *micro*-social (l'acteur) et le *macro*-social. Partant de cela, l'analyse *des réseaux sociaux* permet de se centrer sur ce niveau *méso* par un positionnement à l'échelle d'un système d'acteurs pertinents [Lazega, 2007]. Le passage de l'individu au système d'interdépendances relationnelles dans lequel il s'insère permet de contextualiser l'action, en remplaçant l'acteur, par exemple, dans un *système d'échange généralisé* caractérisant le milieu social dans lequel il évolue. La sociologie de ces systèmes d'interdépendances sociales rend compte de la complexité *horizontale* du monde social en saisissant par la logique systémique les différents *feedbacks* cristallisés dans les interdépendances entre acteurs. En ce sens le contexte social n'est pas seulement un *facteur exogène* qui influe sur le comportement des acteurs, il est aussi engendré par des processus sociaux produits par ces mêmes acteurs [Pattison, Robins, 2004], ce qui nécessite de *l'endogénéiser* dans une perspective *top-down*.

L'existence d'une relation entre deux entités est à appréhender par une structure de dépendance dyadique et plus généralement par l'insertion de cette dyade dans un système social particulier, artificiellement représenté et synthétisé par un réseau social complet au sens de Wasserman et Faust [1994]. C'est la philosophie des modèles statistiques de réseaux qui endogénéisent l'effet de caractéristiques structurales du réseau (*e.g.* la fermeture d'une triade par cyclique) sur l'explication d'une variable de ce même réseau (la probabilité d'existence d'un lien entre deux acteurs) afin de révéler des processus sociaux théorisés par le chercheur (par exemple la solidarité dans le système ou *kula* de Malinowski). Afin d'analyser cela, une famille de modèles statistiques basés sur la fonction exponentielle (p_1, p_2, p^* ou *ERGM*) a été développé pour répondre à la spécificité des données de réseaux sociaux supposant une dépendance entre les différentes observations ou dyades, ce qui par conséquent ne permet pas l'utilisation en l'état de modèles *logit* traditionnels (ou l'indépendance des observations est postulée). L'un des enjeux de l'analyse de réseaux sociaux est le passage des processus et structures locaux à l'architecture du réseau global par la prise en compte d'une dépendance horizontale témoin de la complexité contenue dans le réseau.

Les questionnements relatifs à l'intégration de ces deux dimensions ont amené une grande hétérogénéité de modèles basés sur des appréhensions différentes de l'objet réseau (et donc mobilisant des définitions différentes de ce qu'est un niveau d'analyse) :

- *L'Analyse multiniveau de réseaux sociaux égocentrés* qui permet d'insérer des variables de contexte dans l'explication des comportements au voisinage direct ; ou *complets* qui permet d'insérer des variables de contexte et structurales dans la comparaison de différents réseaux indépendants de même niveau.
- *L'Analyse de réseau dite two-mode* qui étudie un système d'acteur comprenant deux niveaux (ou plus) par écrasement vertical de la structure.
- *L'Analyse stochastique de réseaux sociaux* à un niveau avec la prise en compte des attributs et de la structure d'un niveau et des attributs de l'autre niveau.

- *L'Analyse de réseaux sociaux multiniveaux* par une approche néostructurale [Lazega *et al.*, 2007, 2008] qui étudie un système composé de deux niveaux d'action (l'un individuel, l'autre organisationnel) par la reconstruction des interdépendances structurales à l'intérieur des niveaux et l'articulation verticale, dans l'analyse et le modèle, de ces deux niveaux.

L'enjeu est alors de savoir s'il est possible – et comment – de conjuguer ces deux types de modèles et de positionnement de façon à prendre en compte dans un même mouvement cette complexité en deux dimensions. De fait il s'agit d'introduire des hypothèses de dépendance entre ces deux dimensions dans un modèle d'analyse du réseau permettant d'endogénéiser les effets structuraux sur deux niveaux. D'un point de vue général, l'analyse de réseaux sociaux multiniveaux apparaît comme un ensemble de modèles hétérogènes quant à la nature des données prises en compte et la spécification des unités d'analyse pertinentes. Nous pensons que c'est parce que trois positionnements alternatifs sont possibles au sein d'un même système d'interdépendances (*ego* ou acteur individuel, voisinage ou *neighborhood* et réseau global) qu'il y a une telle diversité de méthodes. Nous commencerons par développer les approches d'analyse de réseaux directement inspirées par l'approche multiniveau traditionnelle (1 et 3 du Tableau 1), puis nous développerons les modèles ERGM traditionnels (4). Après avoir expliqué l'approche multiniveau néostructurale (6), nous expliquerons les limites des modèles bipartites (2) ou de l'approche de Boudourides (5) pour conclure avec les récentes et prometteuses avancées de Wang et Robins (6). Il s'agit en quelque sorte de passer d'une analyse multiniveau de réseaux sociaux à une analyse de réseaux sociaux et organisationnels multiniveaux.

	Attributs organisationnels	Réseau interorganisationnel (+ attributs organisationnels)
Attributs individuels	1. Analyse statistique multiniveau traditionnelle [Bryk, Radenbush, 1992 ; Goldstein, 1995 ; Snijders, 1999 ; Courgeau, 2004...] 2. Analyse de réseau bipartite [Breiger, 1974 ; Wang <i>et al.</i> , 2008]	<i>Idem</i> que réseau de niveau 1 et attributs de niveau 2
Réseau inter individuel (+ attributs individuels)	3. Analyse multiniveau de réseaux sociaux par mise en parallèle des réseaux égocentrés [Van Duijn <i>et al.</i> 1999], complets [Snijders, Baeverleedt 2003] 4. Analyse stochastique traditionnelle de réseaux sociaux: Modèle ERGM avec des attributs sur les nœuds de niveau 2 [Wasserman, Pattison, 1999 ; Robins <i>et al.</i> , 2007]	5. Analyse stochastique traditionnelle de réseaux sociaux: Modèle ERGM avec des attributs sur les dyades de niveau 2, dérivé du réseau de l'autre niveau [Boudourides, 2011] 6. Analyse de réseaux sociaux multiniveaux [Lazega <i>et al.</i> , 2008 ; Wang, Robins, 2011]

TABLEAU 1. Typologie des modèles d'analyse de réseaux sociaux multiniveaux selon les données utilisées

2. DE LA COMPLEMENTARITÉ : ANALYSE MULTINIVEAU ET ANALYSE DE RÉSEAUX SOCIAUX

Les modèles d'analyse multiniveaux en statistique tentent de réaliser une synthèse entre l'observation de l'individu et celle du contexte pour expliquer un phénomène social. Ce type d'analyse est justifié lorsque qu'un modèle estimé sur des données

individuelles diffère du modèle équivalent sur des données agrégées, ce qui laisse à penser que les variables explicatives du phénomène étudié se retrouvent à différents niveaux. Les modèles basés sur des variables agrégées reviennent à supposer que les différences entre les individus sont le résultat de différences entre groupes ou contextes. Cela correspond à l'*erreur écologique* mise en évidence par Robinson en 1950, étudiée par Lazarsfeld et Menzel [1961]. Autrement dit, l'*erreur écologique* consiste à expliquer les comportements individuels à partir d'une observation agrégée en posant l'hypothèse d'homogénéité de ces comportements au sein du groupe. De manière symétrique, l'*erreur atomiste* consiste à expliquer le comportement d'un individu uniquement en fonction de variables individuelles ce qui par définition sort alors l'individu du contexte, l'isole de la société. En effet, toutes choses égales par ailleurs, on peut considérer que deux individus d'un même groupe ont plus tendance à se ressembler que deux individus issus de groupes différents [Bressoux et al., 1997]. Cela peut conduire à expliquer les différences entre individus comme le résultat de différences entre groupes.

Afin d'étudier l'impact des groupes (G_j), mais à partir de modèles individualistes, il pourrait être possible d'introduire des variables muettes (γ_j) correspondant à l'affiliation des individus à un groupe (n : le nombre de groupe):

$$\gamma_{ij} = \beta_0 + \beta_1 x_{ij} + \gamma_1 G_1 + \dots + \gamma_n G_n + R_{ij} \quad (1)$$

Mais comme le souligne Bressoux et ses co-auteurs [1997] le problème de cette modélisation est de supposer une indépendance des observations, ... allant contre la justification du recours à ce type de modèle (dépendance entre les individus d'un même groupe). Les groupes sont alors considérés comme des effets fixes dans la mesure où ils n'ont pas de variance propre. Une autre solution serait de procéder par étapes en étudiant d'abord un modèle individualiste en agrégeant les résidus entre estimation et score individuel par groupe, de sorte à faire une régression sur ces résidus agrégés. Ce faisant, l'on suppose que le groupe a le même impact sur tous ses membres puisque l'intégration des niveaux est séquentielle et non simultanée [Bressoux et al., 1997]. Les modèles multiniveaux en posant l'effet du contexte comme aléatoire permettent d'introduire simultanément cette dimension stochastique aux différents niveaux: individus et contextes.

L'utilisation d'un modèle multiniveau permet de limiter le risque d'erreur écologique en faisant intervenir le groupe comme une caractéristique de la sous population qui influence par là même le comportement des individus qui la compose. En parallèle, l'erreur atomiste est évitée par la prise en compte des variables agrégées, à un niveau où la représentativité de l'échantillon est assurée. Ces dernières ne sont pas redondantes car elles vont permettre d'isoler les effets liés au groupe des effets plus strictement individuels, dans la mesure où comme le souligne Courgeau [2004] :

... les caractéristiques agrégées vont mesurer une construction différente de leur équivalent au niveau individuel [p. 10].

L'originalité de ces approches est de renverser la problématique des modèles traditionnels : le résidu n'est plus à combattre, il devient explicatif⁴. C'est dans l'introduction d'effets aléatoires, et donc de l'hypothèse de variabilité de ces effets en

⁴ L'un des buts est alors de comprendre la valeur des différents résidus qui correspondent aux niveaux agrégés aux effets des contextes. En effet, en analyse statistique multiniveau, l'introduction d'influence contextuelle (autres niveaux que le niveau individuel) permet de centrer l'analyse non sur les variables mesurées, mais sur la décomposition de la partie inexpliquée des phénomènes.

raison d'une distribution spécifique, propres à chaque niveau que réside la clé des modèles multiniveaux (la prise en compte d'effets fixes propres à chaque niveau peut se faire avec des modèles de régression plus simples). Dans une certaine mesure, on peut considérer qu'il s'agit d'une contextualisation verticale des phénomènes sociaux propres à une société organisationnelle. La forme structurelle du modèle, décrivant séparément chaque niveau permet de tenir compte de la dépendance des observations au sein de chaque groupe (corrélation intra-classe). Le lien entre niveau passe par l'introduction de terme d'erreur aléatoire expliqué en partie à un autre niveau. Ce lien entre les différentes équations permet d'écrire la forme réduite du modèle (une seule équation).

- *Niveau 1* : Variable à expliquer = Variables explicatives * coefficients + terme d'erreur. On va cependant poser que les coefficients de certaines variables ne sont pas fixes, mais sont des variables aléatoires.

$$Y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + R_{ij} \quad (2)$$

- *Niveau 2* : On explique chacun de ces coefficients par autant d'équations que de coefficients de niveau 1 comportant des effets aléatoires. Ces équations peuvent comporter une constante fixe et une ou plusieurs variables explicatives de niveau deux (caractéristiques des groupes, et ayant les mêmes valeurs pour tous les individus du même groupe). Enfin, des variables explicatives de niveau 2 peuvent être communes à plusieurs coefficients de niveau 1 que l'on veut expliquer.

$$\beta_{0j} = \gamma_{00} + \gamma_{01} z_j + U_{0j} \quad (3.1)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} z_j + U_{1j} \quad (3.2)$$

La forme réduite du modèle multiniveau est donc la suivante :

$$Y_{ij} = \gamma_{00} + \gamma_{10} x_{ij} + U_{1j} x_{ij} + \gamma_{01} z_j + \gamma_{11} z_j x_{ij} + U_{0j} + R_{ij} \quad (4)$$

L'hétérogénéité entre groupes est en fait déterminée par la variance du coefficient aléatoire de niveau 1 (la mesure du résidu au niveau groupe). La variabilité de la variable à expliquer peut alors être décomposée entre d'une part les différences entre groupes (variation inter-groupes) et d'autre part l'hétérogénéité au sein de chaque groupe (variation intra-groupe). Le fait de placer de l'aléa sur une caractéristique explicative revient à envisager un effet différent selon le groupe. Enfin, il est possible d'ajouter des effets d'interaction entre niveaux, consistant en un croisement de variables individuelles/contextuelles – ici via le coefficient $\gamma_{11} z_j$. Ce faisant, le modèle devient multiplicatif plutôt qu'additif en permettant directement une interaction entre les niveaux sur certaines variables. Le contexte n'influence pas uniquement l'acteur cible. Il influence aussi certaines de ses caractéristiques telle que son environnement relationnel direct (sa famille), et en parallèle il est raisonnable de penser que l'individu à conscience des différences de contextes et qu'il les inclut dans ses choix [Vallet, 2010]. L'introduction d'effets d'interaction entre les niveaux permet de limiter la sur- ou sous-influence du contexte sur l'acteur, comme cela peut être le cas dans un modèle linéaire. La complexité verticale du système est alors d'une certaine manière symbolisée par la rétroaction des niveaux endogénéisée par ces effets d'interaction.

Au-delà de cette complexité verticale, une complexité de nature plus horizontale peut être introduite. C'est d'ailleurs dans cet esprit que Courgeau conclut plusieurs de ses articles [Courgeau 2002, 2004] en insistant sur la nécessité de mieux prendre en

compte la structure sociale des groupes des différents niveaux étudiés dans la mesure où :

... la corrélation s'applique indifféremment à chaque couple d'individus de ce groupe. Il serait utile dans certains groupes, généralement d'effectif restreint, de tenir compte des interactions qui peuvent exister entre certains membres, par exemple les conjoints d'une famille [Courgeau, 2002, p. 73].

Il propose alors des modèles de chocs pour faire intervenir ces interactions, mais il est aussi possible d'introduire des données de réseaux sociaux, complets ou pas, de façon à différencier la valeur de la corrélation entre les membres du groupe suivant la position/statut type de l'acteur [Snijders, Kenny, 1999]. C'est un peu comme si l'approche multiniveau passait directement du *micro* au *macro*, sans passer par la case *méso*. Comme le souligne Bressoux *et al.* [1997], l'effet de contexte peut résulter de l'interaction sociale entre les individus au sein d'un même groupe, ce qui génère des processus sociaux spécifiques à ce groupe tels que le contrôle, la régulation, l'apprentissage ou encore la solidarité (pour reprendre les processus sociaux à l'échelle de la structure mis en évidence par Lazega [2001]). Toujours dans cet esprit, Christine Fauvelle-Aymar (2009) s'intéresse à l'impact du voisinage socio-géographique – entendu comme interaction sociale influençant les comportements sociaux et politiques – sur la participation électorale. Elle affirme qu'un modèle multiniveau permet de prendre en compte les effets de voisinage exogènes causés par l'environnement (au sens « écologique »). Encore une fois, en suivant son propos, il s'avère que l'insertion d'une logique néostructurale permettrait d'insérer les effets de voisinage endogènes propres à l'interaction sociale⁵. Analyse de réseaux sociaux et modèles multiniveaux apparaissent donc fortement complémentaires dès lors que l'on souhaite modéliser et estimer des effets propres à plusieurs niveaux d'analyse, mais interagissant.

3. L'ANALYSE MULTINIVEAU DE RÉSEAUX SOCIAUX : LA MISE EN PARALLÈLE DE DIFFÉRENTS RÉSEAUX

Partant de ce constat, les premiers modèles d'analyse multiniveau de réseaux sociaux ont été développés. Dans ces modèles, il s'agit de mettre en parallèle différents réseaux sociaux de même niveau afin de mieux calibrer les estimateurs et surtout d'introduire des dépendances spécifiques à chaque groupe. La complexité horizontale est donc rendue par une comparaison de plusieurs réseaux non reliés directement et la complexité verticale par l'introduction d'un effet propre à chaque réseau grâce à des variables de contextes.

Dans cette perspective et d'un point de vue *égocentré*, Van Duijn, Van Busschbach, et Snijders [1999] ont étudié la stabilité relationnelle dans des réseaux sociaux égocentrés en utilisant un modèle d'analyse multiniveau dans lequel *ego* constitue le niveau deux et *alter* le niveau un. Cette différenciation des unités d'analyse se justifie dans la mesure où deux liens émis par le même individu sont corrélés du fait même qu'ils partagent le même résidu provenant de cet *égo*. En parallèle, il est possible d'affecter à chaque *alter* un résidu spécifique ou dyade *ego-alter*. L'intérêt d'un modèle multiniveau *ego/alter* est alors de permettre de séparer deux sources de variabilité d'origine différente dans l'estimation du modèle qui est alors spécifié comme un multiniveau traditionnel puisqu'il ne comporte pas de

⁵ « Il y a les effets de voisinage endogènes qui sont médiatisés par le biais des interactions sociales et par l'interdépendance sociale existant entre les individus » [Fauvelle-Aymar, p. 3].

variables structurales. Miguel et Tranmer [2010] utilisent le même type de modélisation pour étudier les réseaux de soutien des immigrants. L'unité d'analyse est alors le lien entre un immigrant et un alter, immigrant récent ou plus établi (*spaniard*). L'une des questions de recherche est de savoir s'il existe des facteurs spécifiques pour expliquer un lien avec un immigré récent et d'autres facteurs pour un lien avec un *spaniard*. Pour ce faire ils vont notamment introduire l'homophilie au niveau 1 (*alter*) et comme effet d'interaction entre les niveaux. Ces deux contributions se limitent à l'étude de réseau égocentré et les variables de structures ne sont que très peu présentes.

En parallèle, d'autres modèles d'analyse multiniveaux sont développés pour des réseaux sociaux complets. Snijders et Kenny [1999] développent le *Social Relation Model* afin d'étudier la « chaleur émotionnelle » (variable continue) perçue par les différents membres d'une famille entendue comme un réseau complet. Cela nécessite alors d'avoir les relations entre ces différents membres, vus alors comme des « rôles » (père, mère, frère, sœur) plus ou moins présents dans chaque famille. Dans ces données, la famille, le groupe, l'individu ou encore les dyades constituent des unités d'analyse pertinentes pour la modélisation. Ils vont ainsi proposer par exemple un modèle où le niveau un est l'individu, le niveau deux de la dyade et le niveau trois la famille. L'effet de réciprocité est désormais représenté par un effet de la dyade, qui est l'effet aléatoire au niveau deux. Il s'agit ici d'une mise en parallèle de réseaux complets de petite taille.

La mise en parallèle de réseaux complets pose explicitement la question de savoir si parmi un ensemble de réseaux complets concernant les mêmes unités d'analyse, il existe des structures communes à l'ensemble de ces réseaux ou encore quelles variables exogènes (attributs des acteurs) expliquent les différences. Le modèle *p2 multiniveau* a été conçu pour répondre à ces types de questions [Zijlstra, van Duijn, Snijders, 2007]⁶. L'idée est de procéder en deux temps : dans un premier temps ils modélisent chaque réseau par un *p2* traditionnel, les différents résultats sont alors combinés pour estimer les paramètres. Le modèle *p2* [Lazega, Van Duijn, 1999] constitue une extension du modèle *p1* [Holland, Leinhardt, 1981] où les paramètres *sender/receiver* sont modélisés comme des effets aléatoires avec des covariances au niveau des acteurs. L'idée est alors de traiter la réciprocité et la différenciation comme des effets aléatoires et donc d'estimer les paramètres de la distribution sous-jacente. On pourrait ainsi voir le *p2* comme une sorte de multiniveau du *p1*. Les paramètres de régression fixes d'un *p2* standard deviennent des coefficients aléatoires en ajoutant des effets aléatoires à la régression des paramètres récepteur, émetteur, réciprocité et densité :

The multilevel p2 model can be regarded as a three-level random effects model where Level 1 is formed by the tie observations, cross-nested in the actors (Level 2), who are nested in the networks (Level 3) [Zijlstra, van Duijn, Snijders, 2007, p. 3].

Il s'agit d'estimer dans un premier temps les effets aléatoires à l'origine (au niveau du réseau) pour ensuite étudier la contribution de ces effets sur d'autres paramètres (ici genre/origine ethnique). L'intérêt est de pouvoir analyser simultanément plusieurs réseaux complets permettant une meilleure généralisation des résultats. Une question peut alors être qu'est-ce que les réseaux sociaux ont en commun ?

⁶ Lubbers [2003] a adopté la même logique pour comparer des réseaux sociaux dans des classes de manière multiniveau mais via une modélisation de type *p** ou ERGM.

Dans le même esprit de computation mais dans une perspective plus dynamique, Snijders et Baerveldt [2003] ont développé un modèle d'analyse multiniveau de réseaux d'amitié entre élèves de plusieurs classes dans la même école afin de comprendre l'influence respective de chaque niveau sur la criminalité. Dans cette conception les différents niveaux sont parfaitement emboîtés⁷ : des élèves dans des classes, dans des écoles, dans des académies,... et la frontière du système est évidente⁸. Les unités d'analyse sont reliées par un lien d'affiliation. Le niveau 1 est constitué d'élèves ; le niveau 2 de classes ; le niveau 3 de collèges... L'intégration des niveaux passe par une combinaison des réseaux de niveau 1, les relations intra-classes, pour les intégrer comme différences interclasses explicatives du niveau 2. Le niveau 2 de l'analyse permet de prendre en compte les différences entre classes. L'organisation est vue comme un contexte qui influence la création ou le maintien de la relation individuelle et il s'agit alors de combiner des données provenant de différentes expériences. Le niveau d'un élève est fonction de celui des autres élèves de la classe. Donc pour prendre en compte cet effet, il faut intégrer le niveau 2. L'idée générale est alors de décomposer cette part d'inexpliquée par l'introduction de variations aléatoires propres à chaque niveau, qui sont alors par définition supposées appartenir à des échantillons différents. Cela permet de prendre en compte la dépendance des observations provenant d'un même contexte. On fait comme si on supposait une homogénéité à l'intérieur d'un groupe et une hétérogénéité entre les groupes. D'une manière générale, cela permet d'éclairer sur le sens de la causalité (s'il y en a) entre acteur et contexte. Deux fonctions permettent d'expliquer l'évolution du choix du lien sortant d'un individu vers un autre individu. Il s'agit de maximiser les préférences relationnelles de l'acteur, l'élève, en fonction de variables structurales relatives à sa position dans la classe. C'est une somme pondérée de différents effets révélant l'asymétrie entre création et destruction du lien. À chaque pas, la structure relationnelle de l'acteur est censée résulter de la maximisation des fonctions objectif et gratification. Cette évolution est approchée par des attributs individuels et relationnels (*outdegree*). L'effet du niveau 2 est approché par la différence entre la moyenne globale par classe. C'est une analyse en deux temps (*Multiple parallel networks*). Les sous structures observées sont de nature horizontale dans la mesure où elles ne sont pas reliées par un lien, mais étudiées de manière comparative :

... *micro level is the study of the relational ties within each single network, and the macro level is the combination of these multiple network studies*
[Snijders, Baerveldt, 2003, p. 124].

Cette approche, basée sur la mise en parallèle de réseaux intra-organisationnels implique que les entités du deuxième niveau ne se rencontrent pas (absence de réseau interorganisationnel) ce qui n'est pas le cas de tous les systèmes d'action (cf. § 4).

⁷ Nous évoquons ici uniquement le cas des classifications hiérarchiques au sens strict. Il est cependant possible de construire des modèles multiniveaux dans lesquels les classifications (ou niveaux) sont croisées (par exemple travail et lieu de vie). Pour plus de détails, cf. [Courgeau, 2004 ; ou Snijders, Bosker 1999].

⁸ "The levels in the multilevel analysis are another name for the different types of unit of analysis. Each level of analysis will correspond to a population, so that multilevel studies will refer to several populations" [Snijders, 2004].

4. INTÉGRER LES DÉPENDANCES STRUCTURALES HORIZONTALES PAR LE VOISINAGE À L'ORDRE QUATRE

Dans les modèles précédents, la complexité horizontale est saisie uniquement à la dyade (le modèle précédent introduit les triades, mais dans une perspective dynamique). Les *Exponential Random Graph Models* – ERGM – introduisent les triades en considérant l'ensemble du réseau comme une structure complexe, plutôt que de le décomposer en un ensemble de dyades. Cela permet de contextualiser ces dernières en prenant en compte le voisinage à l'ordre supérieur, ce qui *in fine* permet de mieux saisir la complexité horizontale contenue dans le réseau. D'une manière générale, la philosophie du modèle est de comparer le réseau observé avec une distribution de graphe aléatoire finement calibrée afin de révéler l'impact des configurations locales (ou des structures de voisinage) dans l'émergence de la structure globale, pour un ensemble de nœuds donné. L'idée est d'endogénéiser l'influence des processus sociaux sur la création des liens en supposant que le réseau est généré par un processus stochastique particulier, ce qui revient à le considérer comme auto-organisé [Robins *et al.*, 2005 ; Robins *et al.*, 2007].

Pour comprendre la logique de ces modèles, il faut revenir aux travaux fondateurs des années 70. Holland et Leinhardt [1976] partent du postulat que l'on peut comprendre les mécanismes et processus en jeu dans un réseau en s'intéressant uniquement aux structures locales. Ils prouvent ainsi que les fréquences des différentes triades dans un réseau sont dépendantes entre elles. Or, comme le montre Frank et Strauss [1986] la prise en compte de la dépendance de structure conduit à différentes familles de modèles log-linéaires pour les graphes. Il s'agit en quelque sorte d'une réduction de l'ensemble des graphes possibles conditionnellement aux structures – ou configurations – spécifiées. Pour décrire cette dépendance de structure, ils utilisent comme Besag [1974] le théorème d'Hammersley-Clifford⁹ affirmant que les propriétés globales et au voisinage d'un graphe de Markov sont les mêmes. Un graphe est un graphe de Markov, si les liens possibles entre des paires de nœuds disjointes sont conditionnellement indépendantes du reste du graphe [Frank, Strauss, 1986]. C'est par la spécification des différentes dépendances structurales (*hypothèse de dépendance*) intégrées dans le modèle que l'ensemble des configurations aléatoires possibles va être calibrées¹⁰ car chaque paramètre restreint le sous-ensemble du nombre de liens possibles [Snijders 2002, Robins *et al.*, 2007]. On comprend dès lors que l'étape de spécification des dépendances dans le réseau est fondamentale¹¹. L'hypothèse de dépendance de Markov (ou *hypothèse de voisinage*) est alors que deux liens sont indépendants (conditionnellement au reste du réseau) s'ils n'ont aucun nœud en commun, ce qui explique que les structures prises en compte vont jusqu'à l'ordre trois (en terme de nœuds). Pour résumer :

The model then represents a distribution of random graphs which are assumed to be "built up" from the localized patterns represented by the configurations [Robins *et al.* 2007, p. 3].

⁹ "In fact, the Hammersley-Clifford theorem [Besag 1974] provides the important link between dependence structures and interactions" [Frank, Strauss, 1986, p. 833] dans la mesure où "In the Hammersley-Clifford terminology, the local and global Markovian properties are equivalent" [Besag, 1974, p. 198].

¹⁰ Un ERGM ne comprenant que l'effet densité, incarné par le paramètre edge correspond à un modèle de Bernoulli.

¹¹ "A critical step in model formulation is the specification of the relational topology, since the neighbourhood relation determines the form of the relational configurations parameterised in the model" [Pattison, Robins, 2004, p. 18].

Il s'agit alors d'estimer la probabilité d'observer un graphe, conditionnellement aux configurations spécifiées, grâce à un modèle représentant la distribution de probabilité des graphes possibles [Robins *et al.*, 2009]. Le modèle peut s'écrire sous la forme suivante [Wang, 2006] :

$$P(X = x) = \frac{1}{\kappa(\theta)} \exp \sum_p \theta_p z_p(x) \quad (5)$$

Avec $z_p(x)$ la statistique de réseau de la structure locale de type p (dyade, triade, ...) et θ_p le paramètre associé. $\kappa(\theta)$ est la constante de normalisation correspondant à l'ensemble des graphes possibles avec le même nombre de nœuds (*i.e.* $2^{n(n-1)}$). L'hypothèse de dépendance précédente permet de réduire cet ensemble¹². Afin de simplifier le modèle (trop de paramètres) il est posé l'hypothèse d'*homogénéité des configurations isomorphiques* : quels que soient les acteurs qu'elles concernent, l'impact des configurations est le même. L'avantage est que les statistiques deviennent uniquement des comptabilisations de l'occurrence des configurations dans le réseau.

C'est ensuite une procédure de *Monte-Carlo Markov Chain Maximum Likelihood Estimation* (MCMCMLE) qui permet de déterminer la distribution des paramètres du modèle en se basant sur l'ensemble des graphes possibles précédemment déterminé. Tant que les paramètres estimés par la simulation ne convergent et ne stabilisent pas, ils sont révisés [Snijders, 2002]. C'est en ce sens que la "Simulation is at the heart of Monte Carlo maximum likelihood estimation" [Robins *et al.*, 2007, p. 29]. Cette procédure fournit ainsi au final une distribution de probabilité pour l'ensemble des paramètres.

L'importance d'un paramètre dans la compréhension du réseau est à comparer à zéro qui serait une sorte d'effet neutre ou aléatoire. Ainsi, si le paramètre afférant à une configuration particulière est significativement positif cela signifie que l'on observe plus de fois cette configuration qu'au « hasard » (inversement). La configuration joue donc un rôle dans la structuration du système observé, et ce au-delà des dépendances hiérarchiques avec les autres configurations (l'intérêt des modèles de Markov pour les graphes est que les statistiques sont reliées entre-elles en raison d'une hiérarchie des configurations : les 2-star dans les triades,¹³).

Il est bien sûr possible d'intégrer des attributs (ou effets exogènes) dans ces modèles, qu'ils soient de niveau 1 ou de niveau 2. Cependant le propos de cette contribution est de recenser les différents moyens pour endogénéiser les effets structuraux du méta-réseau. Pour conclure sur la philosophie générale des modèles ERGM, l'introduction de variables structurales est totalement conditionnée par l'existence d'une dépendance avec les autres variables. Cette introduction de variables structurales, à l'ordre trois dans l'explication du lien permet de mieux saisir la complexité horizontale contenue dans un réseau social.

Cependant, on observe une dégénérescence¹⁴ des modèles ERGM, c'est-à-dire une concentration de la distribution des probabilités sur une petite sous-série du

¹² Pour un réseau symétrique binaire de 30 nœuds et sans boucle, le nombre de possibilités est de $8.87254302118661 e +130$.

¹³ C'est parce qu'il y a de la dépendance entre les différentes sous-structures que les paramètres afférents seront différents de zéro. Les valeurs observées constituent ainsi des effets marginaux (cf. Logit).

¹⁴ "A graph distribution is termed as near degenerate if it implies only a very few (possibly only one or two) distinct graphs with substantial non-zero probabilities" [Robins, 2010, WP, p. 11].

graphe, ce qui rend compliqué l'utilisation des modèles de Markov qui concentrent les triades impliquant des zones plus denses que le reste du réseau. Car comme le précise Snijders [2011], la distribution des graphes fondée sur un processus de Markov risque de dégénérer en une distribution bimodale (vide ou saturée) dès que le réseau observé dépasse 30 nœuds avec un degré moyen de 2 et un index de transitivité de 1. D'une manière générale, les paramètres estimés divergent lorsque le réseau présente une forte triangulation ou en présence d'acteurs avec une centralité importante (*outliers*), au regard de la distribution de degré dans le réseau [Robins *et al.*, 2009].

Ainsi, pour limiter la dégénérescence, Pattison et Robins [2002] proposent une distribution non linéaire des triades en donnant une plus petite probabilité aux cliques importantes. En effet, à la différence des modèles de Markov, le nombre de cliques non-complètes dans un graphe est souvent important. L'enjeu est alors d'étudier la transitivité dans des sous-structures de plus de 3 nœuds, ce qui permet de prendre en compte la structure plus globale des liens, donc de « contextualiser » la triade. Il est à noter que l'objet central de l'analyse est toujours l'étude de la triade, unité fondamentale des réseaux sociaux, mais Pattison et Robins rajoutent une partie du contexte, le voisinage. Une hypothèse importante est que les liens entre deux acteurs, quels que soient les niveaux, mais aussi et surtout quelle que soit leur longueur, servent à canaliser l'action sociale. L'idée est donc que c'est moins la structure dans son ensemble qui compte pour comprendre la position d'un acteur, mais plutôt son voisinage relationnel.

Le modèle de l'espace social qu'ils proposent est basé sur l'analyse stochastique des relations dans le voisinage. Pattison et Robins développent deux processus qui peuvent influencer la typologie des réseaux par l'étude des phénomènes de dépendances dans le réseau. Premièrement, le voisinage direct inscrit dans le réseau peut générer de nouveaux liens entre acteurs. C'est l'idée de proximité relationnelle comme support de la relation. Ils expriment cela par l'hypothèse de *dépendance conditionnelle partielle*. Deuxièmement, c'est l'appartenance à une sous structure qui peut créer de la proximité (étendue). Pour ce faire, ils étendent le concept de dépendance markovienne à celui d'*indépendance conditionnelle partielle* : deux liens sont indépendants, non seulement en fonction des nœuds qu'ils partagent, mais aussi en fonction de la structure des liens du reste du graphe [Snijders *et al.*, 2005 ; Robins, *et al.*, 2007 ; Robins, 2007]. Par conséquent, deux liens sont dépendants s'ils partagent un nœud (hypothèse de Markov traditionnelle) ou s'ils s'inscrivent dans une sous-structure quatre-pas (*four-cycle*). Donc, l'existence de deux liens est une pré-condition à la dépendance conditionnelle partielle du 4-pas. Cette sous structure peut être considérée comme un contexte social. L'idée est que l'absence de lien social entre deux personnes partageant un nombre important de relation n'est pas sans raison (de manière tendancielle). Il y a deux façons de voir la fermeture d'une sous-structure 4-pas¹⁵ :

- *Alternating k-triangles* : il s'agit alors d'étudier les sous-structures 4-pas constituées d'au moins deux triades qui partagent la même base, et de voir l'impact sur la fermeture, le dernier lien. Ce paramètre permet alors de voir dans quelle mesure les triades contribuent à générer une structure globale de type centre-périphérie car il permet de voir si les triades ont tendances à se regrouper dans des structures d'ordres supérieur.

¹⁵ Avec l'Alternating k-star, ces deux configurations permettent de limiter la dégénérescence du modèle ERGM. Elles sont mathématiquement équivalentes aux paramètres développés par Hunter et Handcock [2006] – Geometrically Weighted Dyadwise Shared Partner Distribution (GWDSPP) et Geometrically Weighted Edgewise Shared Partner Distribution (GWESP).

- *Alternating k-two-paths* : il s'agit d'étudier les sous-structures de 4 pas qui ne partagent aucun lien, et dont les points sont reliés par des chemins d'ordre 2. Cela peut constituer un indicateur de la pression de la structure à la fermeture transitive. Nous allons nous concentrer sur ce type de sous-structures, mais avec une approche multiniveau.

5. L'ANALYSE DE RÉSEAUX SOCIAUX MULTINIVEAUX : UNE PERSPECTIVE NÉO-STRUCTURALE

En se basant sur cette configuration d'ordre quatre, il est possible d'introduire la dépendance nécessaire à l'utilisation d'un ERGM sur des données de réseaux récoltées à plusieurs niveaux d'action différents mais partiellement imbriqués. Avant d'aller plus loin, il est donc important de préciser ce que nous entendons par niveau, et pourquoi cette différenciation est pertinente *a priori*. Dans la perspective de Lazega *et al.* [2007], nous posons comme présupposé l'existence de deux niveaux d'acteurs, l'individu et l'organisation, correspondant à deux niveaux d'action spécifiques. Comme nous l'avons précisé en introduction, pour nombre de systèmes d'action, les individus interagissent, mais sont aussi insérés dans (ou construisent) des organisations qui interagissent entre elles.

Contre la différenciation entre niveaux, on pourrait affirmer que l'organisation est toujours représentée et construite par des individus. Cependant, il est difficile de savoir qui représente l'organisation de manière générale, mais aussi et surtout de manière locale au sein d'organisations de taille conséquente. La complexification du processus productif, mais aussi du processus décisionnel implique l'intervention de plusieurs individus, dont l'analyse inter-organisationnelle révèle justement les mécanismes de coordination et les objectifs propres... justifiant ainsi la nécessité d'alterner les niveaux dans l'analyse. L'apport de la sociologie des réseaux sociaux intra-organisationnelle [Lazega, 2001] est justement de montrer l'existence de processus sociaux spécifiques à l'intérieur de l'organisation ou de tout milieu social organisé. D'une certaine manière, c'est considérer que les dilemmes de l'action collective et les interdépendances individuelles au sein d'une organisation font émerger des processus spécifiques orientés et orientant la rationalité organisationnelle. Ces interdépendances ont un impact sur la structure inter-organisationnelle : elles créent des ponts entre les organisations. Mais en parallèle, par la création de normes, par les processus de contrôle, de régulation, de solidarité, d'apprentissage [Lazega, 2001], l'organisation constitue un contexte pour l'action individuelle. Par conséquent, au-delà d'un niveau d'analyse, le niveau organisationnel constitue un niveau d'action et l'organisation peut être conçue comme un acteur collectif construit non-réductible à la simple concaténation des réseaux individuels de ses membres [Grossetti, Bès, 2003 ; Lazega *et al.*, 2007]. Il s'agit donc de dire que l'organisation agit comme un acteur, ce qui présuppose une théorie de l'action. Nous nous basons sur celle développée par Lazega [2009] :

Les entreprises gèrent leurs interdépendances de manière encore plus stratégique que les individus. Elles cherchent des contextes dans lesquels elles peuvent trouver ces ressources à moindre coût. Une fois dans ces contextes, elles tentent d'accumuler ces ressources afin de se mettre (ou de se maintenir) en position de définir les termes de leurs échanges. Tout comme les individus, elles cherchent des niches et du statut [p. 551].

Nous posons l'hypothèse de dépendance structurale horizontale propre à chaque niveau d'action dont le présupposé est que les acteurs de chaque niveau doivent

exister, agir dans un contexte social et pouvoir se rencontrer ce qui justifie par là même la nécessité d'une étude ethnographique à chaque niveau d'analyse afin de saisir les ressources structurant le système d'échange¹⁶. En parallèle, nous posons l'hypothèse de *dépendance structurale verticale* entre les niveaux. Le réseau (ou système d'interdépendances) de l'individu est en partie fonction du réseau (ou système d'interdépendances) de son organisation d'appartenance (en plus des attributs et de la structure observée au niveau interindividuel) et inversement.

Dans le même esprit que les effets d'interaction dans un multiniveau traditionnel, cette hypothèse symbolise le *feedback* interniveau. Nous pouvons supposer que les acteurs ont conscience du contexte engendré par le positionnement de l'unité d'affiliation dans son système propre (ou niveau ici). Il s'agit ici d'affirmer que les ressources circulant à un niveau sont pertinentes pour l'autre niveau et sont en partie nécessaires à l'action dans cet autre niveau puisqu'elles sont sources d'opportunités. La dimension stratégique du comportement se retrouve alors dans les deux dimensions : au sein des niveaux, mais aussi entre les niveaux. En effet, comme le soulignent Lazega *et al.* [2007] les acteurs essayent de pousser à une *reconfiguration* du réseau de(s) l'entité(s) d'affiliation. Parce que les acteurs contextualisent en partie leur action dans l'autre niveau, mais aussi parce qu'ils essayent de l'orienter que l'influence de l'autre niveau doit être endogénéisée.

Pour rendre compte de la complexité verticale du monde social contenu dans la cohabitation de plusieurs niveaux, se pose alors la question de l'articulation de ces niveaux : comment d'un point de vue théorique passer du réseau interindividuel au réseau interorganisationnel ?

Ronald Breiger [1974] développe l'approche duale permettant d'articuler les individus et les organisations en utilisant les liens d'affiliation. Ainsi, les groupes sont reliés entre eux s'ils partagent au moins un individu en commun et les individus sont reliés s'ils sont ensemble dans un groupe. Il est alors possible de construire deux réseaux (individus/groupes) dont les relations sont dérivées des liens d'affiliations, qui permettent de relier ces deux réseaux. L'espace social est représenté par des schémas de chevauchement des affiliations : les individus sont proches car ils partagent de nombreuses affiliations. Ces types de réseaux (bipartites ou *two-mode*) sont cependant pauvres en contenu relationnel dans la mesure où ils ne traduisent pas la circulation d'une ressource ou d'un engagement mais une co-affiliation qui implique au passage un lien non-dirigé.

Wang, Sharpe, Robins et Pattison [2008] partent des structures 4-pas pour appliquer le modèle ERGM aux réseaux *two-mode*. En effet, comme ils le soulignent, les triades ne se rencontrent pas dans une structure duale (non écrasée). Cependant, la densité et les configurations en étoile ne permettent pas de respecter les hypothèses de dépendance de Markov. C'est pour cette raison qu'ils introduisent alors les sous-structures d'ordre 4 pour capturer la force d'une coparticipation à deux événements dans la création d'un lien¹⁷. D'une certaine manière, le chemin d'ordre 3 et le cycle d'ordre 4 mettent en évidence la dépendance générée par la colocalisation ou coparticipation, ce qui constitue l'essence d'un graphe *two-mode*. Pour limiter les

¹⁶ « Les structures relationnelles à l'échelle inter-organisationnelle n'existent que parce qu'elles sont construites par des acteurs concurrents et interdépendants appartenant à des milieux sociaux observables ethnographiquement aussi bien dans leur relative spécificité que dans leurs points communs et recouvrement » [Lazega, 2009].

¹⁷ « Four-cycles are the simplest local closures representing the strength of ties when transferring to one-mode networks » [Wang, 2006, p. 13].

risques de dégénérescence, ils appliquent alors la même idée que les *alternating-two-paths*, *two-star* et *triangle* aux modèles bipartites. Ces nouvelles spécifications du modèle accroissent grandement la qualité d'ajustement. Cependant, la complexité verticale du système est prise en compte par l'écrasement de la structure.

Lazega *et al.* [2007, 2008] proposent une méthode permettant en partie cela. Ils en ont montré la faisabilité en reconstruisant et analysant des réseaux sociaux et organisationnels multiniveaux dans le monde de la recherche contre le cancer en France. Cette approche permet d'analyser de manière simultanée deux niveaux d'action synthétisés dans des réseaux complets partiellement emboîtés. Le premier niveau est constitué de l'élite des chercheurs français sur la recherche contre le cancer qui agissent guidés en partie par leurs propres buts dont la stratégie de poursuite dépend en partie des réseaux interindividuels de conseils (demandes d'avis de cinq sortes différentes). Les individus sont différenciés selon leurs scores d'*indegree centrality* (demi-degré intérieur, ou nombre de fois qu'un individu est cité), et classés en petits et gros « poissons ». Leur performance scientifique est appréhendée par les scores d'*Impact Factor* associés à leurs publications. Le deuxième niveau est constitué des laboratoires de recherche. Les réseaux interorganisationnels correspondent aux différentes relations structurant la performance dans le milieu de la recherche médicale (échange de matériels, recrutement de post-doctorants, etc...) Les organisations sont aussi différenciées selon leurs scores d'*indegree centrality* dans leur système d'interdépendances propres. Cette dernière permet d'appréhender le prestige des organisations, qu'il est alors possible de classer en « mares », petite ou grande.

Pour relier les deux niveaux Lazega *et al.* utilisent la méthode du *linked design* développée (dans un autre contexte, sans analyses de réseaux) par Parcel *et al.* [1991]. Ils procèdent de manière séquentielle en commençant par étudier dans les différents niveaux d'analyse de manière séparée (étude de réseaux classiques), puis ils articulent ces réseaux par les liens d'affiliation des individus aux organisations, plus généralement, d'un niveau i au niveau $i + 1$. On peut compter plusieurs liens d'affiliation pour un individu, s'il travaille dans plusieurs organisations. Dans les domaines où les unités d'analyse de niveaux 2 existent en propre et structurent le système, alors, cette méthode préserve la complexité verticale du monde social. Cet agencement de plusieurs réseaux acteurs différents, reliés entre eux par des liens d'affiliation peut permettre de comprendre les relations entre position, stratégie et performance :

Le linked design permettra d'évaluer l'importance de réseaux de relations inter-individuelles au sein d'une population d'individus lorsque cette population a les moyens de pousser à la reconfiguration du système inter-organisationnel qu'elle représente (en partie) [Lazega et al., 2007, p. 125].

Le graphique suivant constitue une visualisation possible de réseaux sociaux multiniveaux inspirée de la théorisation de Lazega *et al.* [2008, 2009] pour des données de réseaux sociaux multiniveaux récoltées sur un salon d'échange de contenus audiovisuels (enquête *MultiLevel Social Network* – ANR-IRISSO). Le lien entre les niveaux est réalisé par la méthode du *linked design*, basée sur les liens d'affiliation des individus aux organisations. Les liens interniveaux diagonaux sont engendrés par des organisations représentées par plusieurs individus. De plus, les nœuds et les liens sont coloriés selon leur niveau de centralité. S'ils ont une centralité de demi-degré intérieur supérieure à la moyenne de centralité du niveau considéré, ils sont considérés comme GROS (poisson ou mares). Cette visualisation illustre les dépendances en deux dimensions : à l'horizontale les dépendances au sein de chaque réseau ; à la verticale les dépendances entre réseaux.

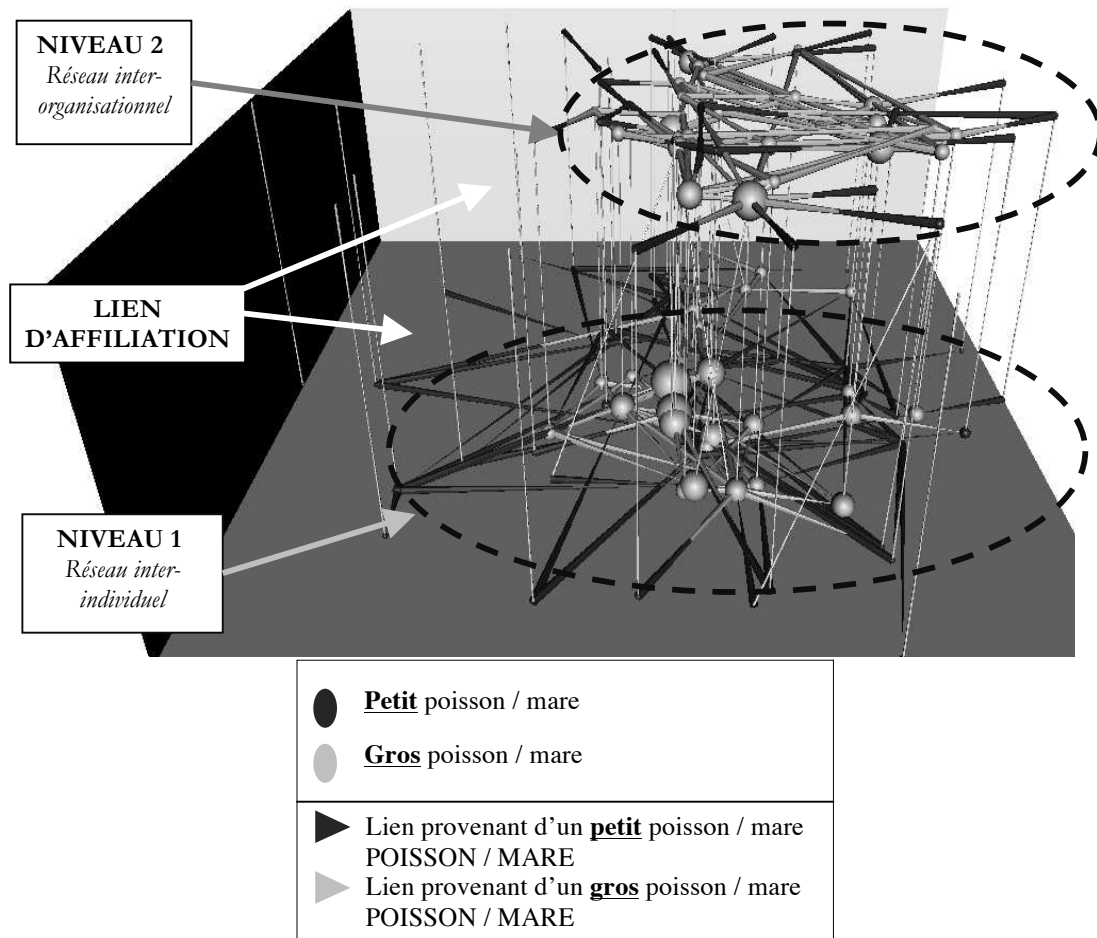


FIGURE 1. Réseaux sociaux multiniveaux dans un marché de programmes de télévision (enquête ANR-MLSN, 2010)

6. LA CONFIGURATION TÉTRADIQUE DANS L’INTRODUCTION D’UNE DÉPENDANCE VERTICALE

Afin d’agencer des réseaux sociaux de niveaux d’action différents, mais reliés selon la perspective du *linked design*, Lazega *et al.* [2010, 2011] et Brailly et Chatellet [2010] proposent d’introduire la *sous-structure tétradique*. Pour ce faire, ils s’inspirent de la structure 4-pas de Pattison et Robins permettant une contextualisation de la dyade dans le *métra-réseau*. L’idée est que le voisinage « vertical », entendu comme les relations entre les entités du niveau j , influence les relations entre les entités affiliées du niveau i . La structure multiniveau tétradique permet de prendre en compte l’existence de dépendance à l’ordre 4, ici au niveau supérieur, et c’est en ce sens que l’on peut l’ajouter comme ensemble de paramètre dans un ERGM.

En reprenant les hypothèses de dépendance du méta-système multiniveau dans une perspective néostructurale [Lazega, à paraître], il est possible de limiter l’ensemble canonique des sous-structures non-isomorphiques.

L’hypothèse de dépendance structurale interniveau implique que la configuration contienne au moins un acteur de chaque niveau et que l’action collective ait un sens à chacun des niveaux. Nous posons en outre que les liens interniveaux (L1

et L2) soient forcément présents dans la perspective du *linked design*, mais la pluri-appartenance et le cas de la multi-affiliation ne sont pas étudiés (L5 et L6). De plus on les considère comme non dirigés.

Les sous-structures d'ordre 2 comprennent un individu et une organisation, ce qui correspond au mécanisme d'articulation des deux niveaux, le lien d'affiliation (non orienté). Les sous-structures d'ordre 3 correspondent à des triades multiniveaux. S'il y a un individu et deux organisations ce sont des sous structures intra-organisationnelles révélant l'impact d'une pluri-affiliation sur le lien interorganisationnel. S'il y a deux individus et une organisation, il s'agit d'une dyade intra-organisationnelle, ce qui correspond à un écrasement de l'analyse. À partir de l'ordre quatre, l'hypothèse de dépendance structurale intra-niveau est respectée dès qu'il y a au minimum deux individus et deux organisations. C'est l'unité minimale pour étudier des processus multiniveaux. Cela permet de mettre en évidence le réseau « étendu » de l'individu [Lazega *et al.*, 2010], sans pour autant écraser la structure. Les liens intra niveaux peuvent être dirigés ; ils correspondent à des dyades traditionnelles (L3 et L4).

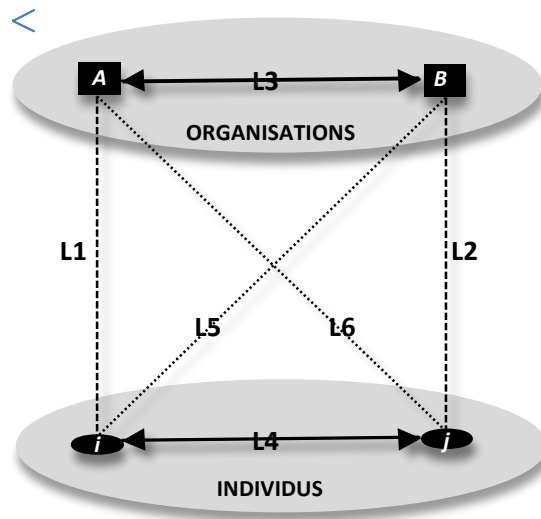


FIGURE 2. Sous-structure multiniveau d'ordre 4

En excluant la valuation des liens, on peut avoir 4 valeurs possibles pour un lien : absence de lien, lien entrant, lien sortant et lien réciproque. Nous ne considérons pas la multi-appartenance (*i.e.* les diagonales – L5/L6). De plus, nous posons que les individus sont toujours affiliés aux organisations ($L1 = L2 = 1$ *qlq* i, j, A, B) et que le lien d'affiliation est symétrique et vaut 1 dans les deux sens¹⁸. Par conséquent, il reste

¹⁸ Cette hypothèse est discutable dans la mesure où ce qu'envoie l'organisation à l'individu peut être considéré comme différent de ce qu'envoie l'individu à l'organisation. On peut en effet penser le lien d'affiliation comme un couple opportunité/contrainte. Ainsi, un directeur général peut relativement facilement mobiliser les ressources de son organisation et l'orienter. Au contraire, un employé de niveau hiérarchique de 3 ou 4 est plus fortement contrôlé par celle-ci. Il nous semble que derrière la question de la valuation du lien d'affiliation il y a deux dimensions : 1) La place de l'individu dans la structure formelle de l'organisation qui peut alors être approchée par le niveau hiérarchique, et 2) la place de l'individu dans la structure informelle pouvant être étudiée par l'analyse du réseau intraorganisationnel. Il est cependant important de noter qu'une valuation asymétrique du lien d'affiliation ajoutera de l'hétérogénéité dans le système hypothéquant la convergence des modèles statistiques. La meilleure solution est alors d'intégrer ces dimensions comme attribut sur le lien d'affiliation dans la modélisation.

deux liens qui peuvent prendre 4 valeurs : L3 et L4. Nous rajoutons que si le lien d'un niveau est absent ou réciproque, alors le sens de l'orientation du lien à l'autre niveau n'est pas importante, il suffit juste de savoir que le lien est orienté. L'ensemble des sous-structures multiniveaux d'ordre 4 non isomorphiques est donc :

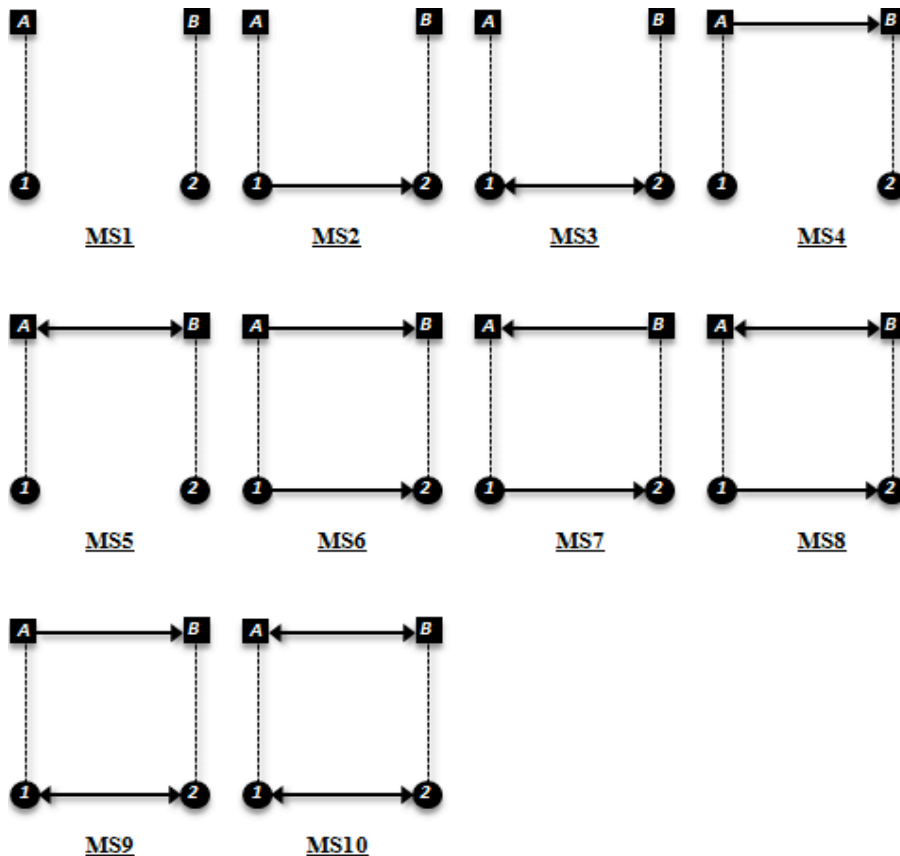


FIGURE 3. Sous-structure multiniveaux non isomorphiques d'ordre 4 avec mono-affiliation

Courgeau insiste sur l'importance des règles et d'une manière plus générale sur l'étude des processus de cohésion, d'évolution et de disparition des groupes restreints pour « pouvoir fonder plus solidement l'analyse multiniveau » [Courgeau, 2002, p. 72]. Les règles constituent alors une courroie de transmission entre les niveaux, elles influencent et sont influencées. Au-delà de la problématique encastrement/découplage, la capture de la réputation par l'un ou l'autre niveau constitue en soi un processus social par essence multiniveau. Dans son travail l'individu n'est pas que lui-même, il représente une organisation pour autrui ; par exemple, dans un marché, il endosse d'une certaine manière l'histoire et la réputation de l'entreprise (chez Disney, ils sont comme ça...), ce qui modifie à la fois ses possibilités et ses contraintes (toutes choses égales par ailleurs, représenter Disney n'ouvre pas les mêmes potentialités que représenter un petit producteur local). Ainsi, la performance de l'acteur est accélérée, ou ralentie par son organisation d'appartenance. L'analyse multiniveau peut permettre de mieux saisir comment se déroulent les processus de construction de la réputation pour chaque type d'acteur. C'est en ce sens que Lazega, Jourda, Mounier et Lazega écrivent :

In our view, this augmentation of individual networks is equivalent to a new approach to « borrowing » relational capital. Indeed we argue that this approach is useful to show that performance is collective even when it is measured at the individual level [Lazega et al., 2010].

L'apprentissage comme processus social peut être appréhendé de manière multiniveau. Comment le savoir de l'individu se diffuse dans une organisation ? Comment la connaissance d'un marché (au sens géographique) de tel membre d'une entreprise qui prospecte sur un nouveau territoire se diffuser à ses collègues ? Mais également comment le savoir de l'organisation, qui est d'une certaine manière socialisé, va-t-il être diffusé jusqu'à l'individu ? Le poids relatif de MS3, MS9 et MS10 dans un ERGM adapté peut nous apporter des éléments de compréhension sur la construction et la pérennité du lien réciproque interindividuel en fonction du lien interorganisationnel. Le poids relatif de MS5, MS8 et MS10 peut nous apporter des éléments de compréhension sur la construction et la pérennité du lien réciproque interorganisationnel en fonction du lien interindividuel. Le poids de ces différents paramètres peut alors témoigner de plusieurs types de processus sociaux pouvant être désormais appréhendés de manière multiniveaux.

7. L'INTRODUCTION DE LA DÉPENDANCE VERTICALE DANS LES MODÈLES STOCHASTIQUES

Afin de prendre en compte le réseau au niveau 2 dans l'explication du lien au niveau 1, Boudourides [2011] propose une typologie de la force des liens, en fonction d'un géodésique¹⁹ multiniveau. Le lien le plus fort est bien évidemment le lien direct, puis au fur et à mesure que le géodésique augmente, la force de ce lien diminue. Il propose d'intégrer le deuxième niveau via une structure 4-pas, quel que soit le lien d'en bas. Ainsi, le géodésique « transniveau » entre deux individus, passe par leurs organisations et si ces dernières sont en lien direct, cela génère une structure 4-pas. Il est alors possible d'estimer les modèles ERGM, en introduisant le géodésique 4-pas comme paramètre. Cependant, il s'agit plus d'un attribut sur la dyade interindividuelle qui s'apparente plus à un effet exogène. Par là même, la dépendance du niveau 1 au niveau 2 n'est pas prise en compte par une dépendance structurale. En l'état, cela ne permet pas réellement d'endogénéiser les effets de structures sur les deux niveaux. Dans le même esprit, nous pensons qu'il ne convient de considérer le réseau interorganisationnel comme une sorte de réseau interindividuel. L'avantage serait d'utiliser les modèles ERGM pour des réseaux multiplexes²⁰. Cependant, cela revient à écraser la structure verticale et donc à considérer que la différence entre un acteur et une entreprise est minime, voir négligeable ce qui *in fine* évacue les avantages de la différenciation entre niveaux. En parallèle, cela revient à nier la spécificité du lien d'affiliation, qui est plus qu'un lien mécanique, binaire et symétrique (simplification en première analyse).

Robins et Wang [2011] ont spécifié un modèle ERGM canonique pour un méta-réseau multiniveau qui permet d'endogénéiser dans un même mouvement les effets de structures au sein de chaque niveau, mais aussi ceux se déroulant sur deux niveaux (sous-structures tétradiques) :

¹⁹ Le géodésique entre deux nœuds correspond au nombre minimal de pas (d'arrêtes) nécessaire pour relier ces deux nœuds.

²⁰ Les réseaux sociaux multiplexes étudient la circulation de plusieurs types de ressources (donc de liens) entre les mêmes acteurs [Lazega, Pattison, 1999].

$$P(A = a, X = x, B = b) = \frac{1}{\kappa(\theta)} \exp \sum_p \left\{ \begin{array}{l} \theta_p z_p(a) + \theta_p z_p(x) + \theta_p z_p(b) + \\ \theta_p w_p(a, x) + \theta_p w_p(b, x) + \theta_p v_p(a, x, b) \end{array} \right\} \quad (4)$$

Dans cette formalisation, a concerne les structures de niveaux 1, b de niveau 2 et x le lien d'affiliation. Cette conceptualisation permet de rendre compte des effets de structure spécifique à chaque niveau ($z_p(a)$ et $z_p(b)$) et des effets d'interaction entre les niveaux, plus méso-sociaux ($w_p(a, x)$, $w_p(b, x)$ et $v_p(a, x, b)$). Dans une logique statique, ils se posent la question de savoir : “*For instance, with an exogenous meso-level X, we can ask: given the meso links, how do micro and macro structures relate?*” (p. 3). Il convient de noter que cette probabilité concerne au minimum trois liens potentiels (A, B, X). Ils posent en effet une condition importante : il y a toujours un lien entre les niveaux. Ils testent alors ce modèle en le simplifiant de façon à souligner uniquement les effets d'interaction entre niveaux. Les paramètres qu'ils utilisent sont alors le degré à chaque niveau, la multi-appartenance de niveau micro (structures multiniveau de rang 3 permettant de vérifier si les individus d'un même groupe collaborent ensemble), la multi-appartenance de niveau macro (structures multiniveaux de rang 3 permettant de vérifier si les groupes partageant un même individu collaborent ensemble) et la structure tétradique. Ils étudient une structure intra-organisationnelle de manière multiniveaux (individus et divisions) et montrent ainsi que les effets d'interaction entre les niveaux sont importants pour comprendre la structure informelle de l'organisation. Ils montrent ainsi que les liens entre divisions d'une structure organisationnelle sont dépendants des liens entre les divisions entendues comme entités de niveau supérieur qui ne peuvent être approché que via la structure tétradiques. Il est important de noter que le lien d'affiliation est traité ici comme un lien exogène, ce qui dans une perspective statique n'est pas problématique. Cette modélisation en cours de développement permet réellement de saisir et de rendre compte des *feedbacks* entre les deux niveaux d'actions.

Cette dernière modélisation empreinte plus à la conceptualisation duale qu'aux modèles hiérarchiques pour prendre en compte la dimension multiniveau. Ainsi, c'est parce qu'il y a de la dépendance entre les niveaux et donc entre les réseaux que cette modélisation est possible. A la différence de modèles hiérarchiques traditionnels, la dépendance verticale n'est pas rendue par la détermination des variabilités intra-classe et inter-classe.

L'articulation d'un ERGM et de la logique traditionnelle des modèles hiérarchiques implique de considérer le niveau deux comme un contexte de l'action. C'est la philosophie des modèles *Hierarchical-ERGM* (H-ERGM) de Schweinberger et Handcock [2009] basés sur les *blockmodels* stochastiques développés par Nowicki et Snijders [2001]. L'idée de cette modélisation est de poser que les estimateurs pour les différentes statistiques peuvent être différents selon les groupes pertinents dans le système social. Ces modèles permettent de réaliser une analyse multiniveaux de réseaux sociaux (*cf.* partie 2.) au travers d'une modélisation de type ERGM. Les différentes expériences ou niveaux proviennent alors du même jeu de données, qui est vu à différents niveaux simultanément (acteur, dyade, *blocks*, ensemble). Cela s'apparente donc plus à la mise en lumière de comportements différents selon les sous-parties du graphe, qu'il est alors plus souhaitable d'assimiler à des milieux, plutôt qu'à des acteurs. Nous retrouvons le principal critère différenciant les modèles articulant

analyse multiniveau et réseaux sociaux, à savoir si le niveau 2 est constitué d'acteur ou uniquement de contexte pour l'action.

8. CONCLUSION

Au final, bien que très hétérogènes quant aux unités d'analyse et à la méthode, les modèles d'analyse multiniveau de réseaux sociaux recensés dans cette contribution semblent tendre vers une prise en compte des différentes dépendances structurales, à la fois à l'intérieur des niveaux, mais aussi entre les niveaux. Il apparaît que le modèle de Robins et Wang est le plus adapté pour traiter des données de réseaux sociaux multiniveaux néostructuraux.

Cette conception permettant de considérer le niveau deux comme un système d'action, doit permettre à plus long terme les raffinements nécessaires à l'exploration plus en profondeur des mécanismes spécifiquement multiniveaux dans un *méta-système* d'action individus / organisations. D'un point de vue plus sociologique se pose la question du contenu des liens. Concernant les liens intra-niveaux, le contenu dépend des ressources révélant le mieux les interdépendances entre les acteurs du niveau dans la réalisation de leurs tâches et dans la poursuite de leurs objectifs. Reste alors en suspens la question du contenu du lien interniveau, qui comprend en plus de sa dimension formelle une dimension informelle dont la complexité et la spécificité des processus sociaux locaux sont justement mises en exergue par l'étude des systèmes d'interdépendances intraorganisationnels. L'approfondissement des processus multiniveaux que l'on souhaite révéler sera fortement conditionné par la compréhension de ce lien interniveau et de ce qui y circule, tel que les injonctions sur le niveau 1 venant du niveau 2, l'accès et la diffusion de l'information en interne, ...

Remerciements. Nous tenons à remercier les deux rapporteurs pour leurs remarques stimulantes. Nous tenons aussi à remercier Josiane Chatellet, Guillaume Favre, Dean Lusher, Garry Robins, Philippa Pattison, Peng Wang ainsi que l'ensemble des membres de l'Observatoire des Réseaux Intra- et Inter-organisationnel (ORIO) pour leurs critiques et suggestions.

BIBLIOGRAPHIE

- BOUDOURIDES M.A. (2011), "The strength of indirect relations in social networks" [soumis à publication].
- BRAILLY J., CHATELLET J. (2010), « Analyse de réseaux sociaux multiniveaux sur un salon de l'audiovisuel en Europe de l'Est. Quelques pistes pour une approche stochastique », *Cahier de Recherche de l'ORIO* 1, p. 64-81.
- BRAILLY J. (2010), "Comparing approaches to multilevel network analysis", *Sunbelt XXX Conference, International Network for Social Network Analysis*, Riva del Garda (Italie), July 2.
- BREIGER R. L. (1974), "The duality of persons and groups", *Social forces* 53, p. 181-190.
- BRESSOUX P., COUSTERE P., LEROY-AUDOUIN C. (1997), « Les modèles multiniveau dans l'analyse écologique : le cas de la recherche en éducation », *Revue française de sociologie* 38, p. 67-96.
- BRYK A. S., RAUDENBUSH S. W. (1992), *Hierarchical linear models*, Newbury Park (CA), Sage.
- COLEMAN J.S. (1990), *Foundations of social theory*, Cambridge (MA), Belknap Press.

- COURGEAU D. (1999), « De l'intérêt des analyses multi-niveaux pour l'explication en démographie », in D. Tabutin, C. Gourbin, G. Masuy-Stroobant et B. Schoumaker (éds), *Théories paradigmes et courants explicatifs en démographie*, Louvain-la-Neuve, Academia-Bruylant/L'Harmattan, p. 93-116.
- COURGEAU D. (2002), « Évolutions ou révolutions dans la pensée démographique ? », *Mathématiques et Sciences humaines* 160, p. 49-76.
- COURGEAU D. (éd.) (2003), *Methodology and epistemology of multilevel analysis*, Collection Methodos, Dordrecht, Kluwer Academic Publishers.
- COURGEAU D. (2004), *Du groupe à l'individu : synthèse multiniveau*, Paris, Éditions de l'INED.
- CROZIER M., FRIEDBERG E. (1977), *L'acteur et le système*, Paris, Le Seuil.
- FRANK O., STRAUSS, D. (1986), « Markov graphs », *Journal of the American Statistical Association* 81, p. 832-842.
- GOLDSTEIN H. (1995), *Multilevel Statistical Models*, London, Edward Arnold.
- GROSSETTI M., BES M.-P. (2003), « La dynamique des cercles et des réseaux. Encastrement et découplages », *Revue d'économie industrielle* 103, p. 43-58.
- HOLLAND P.W., LEINHARDT S. (1976), "Local Structure in Social Networks", *Sociological Methodology* 7, p. 1-45.
- HOLLAND P.W., LEINHARDT S. (1981), "An exponential family of probability distributions for directed graphs (with discussion)", *Journal of the American Statistical Association* 76, p. 33-65.
- HUNTER D. R., HANDCOCK M. S. (2006), "Inference in curved exponential family models for networks", *Journal of Computational and Graphical Statistics* 15, p. 565-583.
- LAZEGA E., VAN DUIN M.A.J. (1997), "Position in formal structure, personal characteristics and choices of advisors in a law firm: A logistic regression model for dyadic network data", *Social Networks* 19, p. 375-397.
- LAZEGA E., PATTISON P.E. (1999), "Multiplexity, generalized exchange and cooperation organisations: a case study", *Social Network* 21, p. 67-90.
- LAZEGA E. (2001), *The Collegial phenomenon*, Oxford, Oxford University Press.
- LAZEGA E. (2007), *Réseaux sociaux et structures relationnelles*, Presses Universitaires de France, coll. Que-sais-je ?
- LAZEGA E., JOURDA M.-T., MOUNIER L., STOFER, R. (2007), « Des poissons et des mares : l'analyse de réseaux multiniveaux », *Revue française de sociologie* 48, p. 93-131.
- LAZEGA E., JOURDA M.-T., MOUNIER L., STOFER, R. (2008), "Catching up with big fish in the big pond? Multi-level network analysis through linked design", *Social Networks* 30, p. 159-176.
- LAZEGA E. (2009), « Théorie de la coopération entre concurrents : Organisation, marché et analyse de réseaux », in P. Steiner et F. Vatin (éds), *Traité de sociologie économique*, Paris, Presse Universitaires de France, p. 547-585.
- LAZEGA E., JOURDA M.-T., MOUNIER L. (2009), "Borrowing social capital from one's organization: a case study", *Sunbelt XXIX Conference, International Network for Social Network Analysis*, San Diego (CA), March 15.
- LAZEGA E., JOURDA M.-T., MOUNIER L., LAZEGA D. (2010), "Network parachutes from tetradic substructures: 'Organizational expansion' and the measurement of the contribution of laboratories to their scientists relational capital and performance", *Cahiers de l'ORIO* 1, p. 116-133.
- LAZEGA E., JOURDA M.-T., MOUNIER L., LAZEGA D. (2011), "Network Lift from Dual Alters: Extended Opportunity Structures from a Multilevel and Structural Perspective" [soumis à publication].
- LAZEGA E., JOURDA M.-T., MOUNIER L., BRAILLY J., BRUNA M.-G., CHATELLET J., FAVRE G., PENALVA E., PINA-STRANGER A. (2011), "Dynamics of multilevel networks with two levels of agency", Communication à la Conférence *Research on the Organizational Society: Advances in Multilevel and Dynamic Network Analysis*, Université Paris-Dauphine, Observatoire des Réseaux Intra- et Inter-Organisationnels, le 16 juin 2011.
- LAZEGA E. (2012), « Sociologie néo-structurale », G. Bronner et R. Keucheyan (eds), *Théories sociales contemporaines*, Paris, Presses Universitaires de France.

- LUBBERS M.J. (2003), "Group composition and network structure in school classes: a multilevel application of the p^* model", *Social Networks* 25, p. 309-332.
- LUKEN V., TRANMER M. (2010), "Personal support networks of immigrants to Spain: A multilevel analysis", *Social Networks* 32, p. 253-262.
- LUSHER D., ROBINS G.P. (2012), "Statistical analysis of social networks: Exponential Random Graph Models", D. Lusher, J. Koskinen, G. Robins (ed.), *Exponential Random Graph Models for Social Networks: Theories, Methods and Applications*, New York, Cambridge University Press.
- NOWICKI K., SNIJDERS T.A.B. (2001), "Estimation and prediction for stochastic blockstructures", *Journal of the American Statistical Association* 96, p. 1077-1087.
- PARCEL T.L., KAUFMAN R.L., LEEANN J. (1991), "Going up the ladder: multiplicity sampling to create linked macro-to-micro organizational samples", P. Marsden (ed.), *Sociological methodology* 21, Oxford, Basil Blackwell, p. 43-79.
- PATTISON P.E., ROBINS G.L. (2002), "Neighbourhood-based models for social networks", *Sociological Methodology* 32, p. 301-337.
- PATTISON P.E., ROBINS G.L. (2004), "Building models for social space: Neighbourhood-based models for social networks and affiliation structures", *Mathematics and Social Sciences* 168, p. 11-29.
- PATTISON P. E., ROBINS G.L. (2006), *Probabilistic network theory*, [document de recherche].
- PERROW C. (1991), "A society of organizations", *Theory and society* 20, p. 725-762.
- ROBINS G.L., PATTISON P.E., WOOLCOCK J. (2005), "Social networks and small worlds", *American Journal of Sociology* 110, p. 894-936.
- ROBINS G.L., PATTISON P.E., KALISH Y., LUSHER D. (2007), "An introduction to exponential random graph (p^*) models for social networks", *Social Networks* 29, p. 173-191.
- ROBINS G.L., SNIJDERS T.A.B., WANG P., HANDCOCK M., PATTISON P.E. (2007), "Recent developments in exponential random graph (p^*) models for social networks", *Social Networks* 29, p. 192-215.
- ROBINS G.L., PATTISON P.E., WANG P. (2009), "Closure, connectivity and degree distributions: Exponential random graph (p^*) models for directed social networks", *Social Networks* 31, p. 105-117.
- ROBINS G.L., WANG P. (2001), *Multilevel networks in organizations*, Melbourne, University of Melbourne [document de recherche].
- ROBINSON W.S. (1950), "Ecological correlations and the behaviour of individuals", *American Sociological Review* 15, p. 351-357.
- SCHWEINBERGER M., HANDCOCK M. S. (2009), *Hierarchical exponential-family random graph Models*, Technical report, Pennsylvania State University [soumis à publication].
- SNIJDERS T.A.B., BOSKER R.J. (1999), *Multilevel analysis: An introduction to basic and advanced multilevel modeling*, London, Sage Publications.
- SNIJDERS T.A.B. (2002), "Markov Chain Monte Carlo estimation of exponential random graph models", *Journal of Social Structure* 3(21), p. 187-209.
- SNIJDERS T.A.B., BAERVELDT C. (2003), "A multilevel network study of the effects of delinquent behaviour on friendship evolution", *Journal of Mathematical Sociology* 27, p. 123-151.
- SNIJDERS T.A.B., PATTISON P.E., ROBINS G.L., HANDCOCK M. (2006), "New specifications for exponential random graph models", *Sociological Methodology* 36, p. 99-153.
- SNIJDERS T.A.B. (2001), "Statistical models for social networks", *Annual Review of Sociology* 37, p. 129-159.
- VALLET L.A. (2010), « Quelques aspects de la statistique en sociologie, 1950-2010 », *Mathématiques et Sciences humaines* 191, p. 65-80.
- VAN DUIJN M.A.J., VAN BUSSCHBACH J.T., SNIJDERS T.A.B. (1999), "Multilevel analysis of personal networks as dependent variables", *Social Networks* 21, p. 187-209.
- WANG P. (2006), *Exponential Random Graph (p^*) Models for Affiliation Networks*, Graduate Thesis, University of Melbourne.

WANG P., SHARPE K., ROBINS G. L., PATTISON P. E. (2009), "Exponential random graph (p^*) models for affiliation networks", *Social Network* 31, p. 12-25.

WASSERMAN S., FAUST K. (1994), *Social network analysis: methods and applications*, Cambridge, Cambridge University Press.

WASSERMAN S., PATTISON, P.E (1996), "Logit models and logistic regressions for social networks. I. An introduction to Markov graphs and p^* ", *Psychometrika* 61, p. 401-425.

WASSERMAN S., ROBINS G. L. (2005), "An introduction to random graphs, dependence graphs, and p^* ", P. Carrington, J. Scott, S. Wasserman (eds.), *Models and Methods in Social Network Analysis*, New York, Cambridge University Press, p. 148-161.

ZIJLSTRA B., VAN DUIJN M., SNIJDERS T.A.B. (2006), "The Multilevel p_2 Model A Random Effects Model for the Analysis of Multiple Social Networks", *Methodology* 2, p. 42-47.