

L'ANALYSE LONGITUDINALE DE RESEAUX SOCIAUX TOTAUX AVEC SIENA - METHODE, DISCUSSION ET APPLICATION

par

Ainhoa de Federico de la Rúa(1)

(CNRS-Clersé, IFRESI, 2 rue des Canonniers, 59000 Lille ;
federico@univ-lille1.fr)

Abstract - Longitudinal Analysis of Total Social Networks with SIENA – Method, Discussion and Applications: This article explains the functional basis of the SIENA statistical model for the analysis of longitudinal total social networks. In the first part, we explain the specificity of statistical models for analyzing network structure and give a brief history of the major models proposed in the past which produced dynamic models such as SIENA. We then present the basis on which SIENA is constructed and we discuss the theoretical implications. Model specification follows with a presentation of the principal effects which can be studied with this model. We end with an example of application on a network of international friendships. **Longitudinal analysis of social networks, SIENA.**

Résumé – L'objectif de cet article est d'expliquer les principes de fonctionnement du modèle statistique SIENA pour l'analyse longitudinale de réseaux sociaux totaux. Dans un premier temps, nous expliquons la spécificité des modèles statistiques pour analyser les structures des réseaux ainsi qu'une brève histoire des principaux modèles apparus pour y parvenir, débouchant sur les modèles dynamiques dont SIENA fait partie. Ensuite nous exposons les postulats sur lesquels est bâti le modèle SIENA et nous discutons leur réalisme théorique. Suit la spécification du modèle et les effets principaux qu'on peut étudier au moyen de ce modèle. Nous finissons avec un exemple d'application sur un réseau d'amitié international. **Analyse longitudinale des réseaux sociaux, SIENA.**

ANALYSER LA STRUCTURE D'UN RESEAU

Clyde Mitchel (1969) définit un réseau comme «un ensemble particulier d'interrelations (*linkages*) entre un ensemble limité de personnes, avec la propriété supplémentaire que les caractéristiques de ces interrelations, considérées comme une totalité, peuvent être utilisées pour interpréter le comportement social des personnes impliquées». Cette citation montre bien la perspective dominante des chercheurs qui analysent des réseaux sociaux : les caractéristiques d'un réseau sont mises en avant comme des variables expliquant le comportement social (conduites,

opinions, stratégies...) des personnes impliquées et ceci selon les *positions* qu'elles occupent dans leurs structures. Le réseau est dans ce cas-ci utilisé comme une variable contextuelle particulièrement sophistiquée et qui ne demande pas à être elle-même expliquée.

Une deuxième perspective, plus tardive dans l'analyse des réseaux sociaux, s'intéresse par contre aux relations elles-mêmes, elle se consacre à expliquer leurs émergences, leurs agencements, leurs évolutions. Dans cette perspective, les réseaux sociaux doivent eux-mêmes être expliqués à partir du comportement social (conduites, opinions, stratégies) des acteurs qui en font partie. Cet article privilégie cette deuxième perspective. Nous allons parler ici des outils méthodologiques, en particulier d'un modèle statistique (SIENA), qui ont été développés pour expliquer les structures des réseaux.

Analyser les facteurs intervenant dans l'établissement de relations à l'intérieur d'un ensemble d'acteurs (un réseau total) est une tâche complexe qui requière le développement de modèles statistiques spécifiques. La particularité et la difficulté des modèles aléatoires pour les réseaux sociaux sont que ces derniers doivent représenter *la dépendance statistique* entre observations (liens entre acteurs) et par conséquent, à l'opposé de la plupart des modèles aléatoires, comme les modèles de régression, ils ne peuvent pas être construits sur des postulats d'indépendance entre observations. Ceci rend la tâche bien plus complexe et spécifique. L'objectif de ces modèles est d'expliquer la probabilité de la structure observée d'un réseau total, c'est-à-dire, les agencements de liens parmi les acteurs. Cette structure est donc la variable dépendante. Elle doit être expliquée à partir d'un certain nombre d'effets et de variables.

Les différents modèles développés dans littérature des réseaux se sont améliorés considérablement avec le temps en permettant d'inclure plus d'effets différents et à des niveaux différents : niveau individuel (des propriétés des acteurs concernés), niveau dyadique (des propriétés des relations) et niveau triadique (des agencements typiques de relations entre trois acteurs) etc. De plus, étant donné que l'émergence et l'évolution d'un réseau est un processus dynamique, il est préférable d'utiliser un modèle qui tienne compte du temps. La tâche devient rapidement compliquée. Tel que le discute van de Bunt (1999, p27) la littérature à propos des réseaux sociaux ne fournit pas beaucoup de modèles formels qui puissent exprimer les aspects dynamiques du développement d'un réseau de relations. La plupart des modèles sont des modèles statiques et ce n'est que depuis peu que l'on dispose de modèles dynamiques.

Nous allons brièvement commenter des modèles statistiques existants pour les réseaux sociaux. Ce commentaire n'a pas vocation d'être exhaustif, il prétend simplement offrir un aperçu général des développements et des modèles disponibles dans la littérature des réseaux (pour plus de détails sur les modèles proposés antérieurement voir van de Bunt 1999, Snijders 1998, Leenders 1995, Zeggelink 1993) De fait nous allons commenter trois modèles pour les réseaux sociaux fondés sur des modèles loglinéaires qui ont été nommés : $p1$, $p2$ et p^* . Ensuite, nous allons

expliquer de façon plus détaillée le modèle dynamique SIENA. Il se différencie des précédents car il se fonde sur des processus de Markov, ce qui permet un traitement longitudinal des données. Pour ceux qui souhaitent plus de détails et spécifications mathématiques il vaut mieux se diriger vers la littérature de référence des auteurs des modèles(2), ici nous avons tenté de faire un exposé aussi « verbal » et non technique que possible.

BREVE HISTOIRE DES MODELES STATISTIQUES POUR LES RESEAUX

Un des premiers modèles statistiques pour les réseaux sociaux, développé par Holland et Leinhardt (1981) et Wasserman (1987), est connu sous le nom de modèle *p1* (1^{er} modèle de probabilité)(3). Les techniques d'estimation loglinéaires du modèle *p1* ont été développées par Wasserman et al. (Wasserman et Faust, 1994). Ce modèle incluait dans ses spécifications des effets individuels et dyadiques (caractéristiques des acteurs et de leurs relations deux à deux), mais pas des effets triadiques. Comme effets individuels, il permettait de tenir compte de la popularité (relations reçues ou *indegrees*) et de l'activité (relations émises ou *outdegrees*) d'un acteur, mais pas d'effets dus à des variables exogènes, comme le sexe, l'âge, le statut, ou d'autres caractéristiques individuelles qui peuvent cependant être cruciales pour le développement de la structure d'un réseau complet. Quant aux effets dyadiques, il ne permettait que de tenir compte de la réciprocité (probabilité que si *i* émet une relation vers *j*, *j* émet aussi une relation vers *i*), ce qui reste au niveau dyadique de la structure. Ce modèle peut être considéré comme un modèle de régression logistique bivariate (*bivariate logistic regression model*) pour des dyades. Ce modèle postulait l'indépendance entre dyades, ce qui empêche un traitement triadique et donc pleinement structural. Cependant ce modèle a été un point de départ important pour la modélisation statistique des réseaux sociaux.

Certaines des limitations du modèle *p1* ont été surmontées par le modèle *p2* développé par van Duijn (van Duijn et al. 2004, van Duijn 1995, Lazega et van Duijn 1997). Le modèle *p2* peut être considéré comme une version avec des effets aléatoires (*random effects*) ou comme une version multi-niveau du modèle *p1*. Les méthodes d'estimation pour le modèle *p2* sont assez compliquées et sont encore en cours d'amélioration par van Duijn et Zijlstra. L'apport principal du modèle *p2* est qu'il permet l'inclusion et le traitement dans le modèle, en tant que variables indépendantes, de variables individuelles exogènes (comme l'âge, le sexe, ou toute autre caractéristique des acteurs) et de variables dyadiques (comme des similarités, flux, fréquences, contenus ou toute autre caractéristique des relations). Il permet aussi d'étudier des effets d'interaction entre les variables explicatives et la réciprocité. Pour une analyse sociologique, le modèle *p2* est déjà bien plus intéressant que le modèle *p1*, car il permet d'incorporer des variables explicatives classiques de la sociologie (voir des exemples dans Lazega 2001). Cependant d'un point de vue structural, bien qu'il augmente les possibilités d'analyse au niveau dyadique par l'incorporation de variables relationnelles, et l'étude d'interactions

entre variables explicatives et réciprocité, le modèle p_2 n'incorpore toujours pas d'effets au niveau triadique.

En 1996 Wasserman et Pattison ont proposé le modèle p^* en généralisant la distribution de graphes de Markov proposée par Frank et Strauss (1986) et suivant la proposition déjà faite par Frank (1991). Ce modèle propose une formulation extrêmement générale capable de représenter n'importe quelle distribution de probabilités pour un graphe orienté. Ceci permet une grande souplesse et l'incorporation d'effets triadiques. Son avantage est aussi d'une certaine façon son inconvénient : cette formulation est tellement générale qu'il faut trouver les différentes formulations spécifiques pour « remplir » la formule et la rendre opérationnelle. Depuis son apparition en 1996 plusieurs chercheurs se sont consacrés à l'élaboration de ce modèle (Wasserman, Pattison, Robins et d'autres). D'ailleurs la procédure d'estimation des paramètres est encore provisoire : jusqu'à présent on utilisait une procédure appelée de *pseudo-likelihood* qui réduit l'estimation à une procédure de régression logistique. Cependant les propriétés statistiques de cette procédure d'estimation sont encore inconnues et donc il vaut mieux prendre les résultats avec réserve. Récemment Snijders (2002) a proposé une procédure plus fiable d'estimation fondée sur des chaînes Markov Monte Carlo, disponible aussi dans le logiciel StOCNET depuis juillet 2002.

Ceci dit, le modèle p^* a donc finalement rendu possibles la prise en compte, en plus d'effets dyadiques, de toutes sortes d'analyses triadiques par le dénombrement des types de triplets.

Une *triade* représente des relations réciproques (ou non) entre 3 acteurs (voir la figure 1). On appelle *triplet* la triade vue par un des acteurs qui la compose (ego). Chaque triade peut contenir de 0 à 6 fois un triplet donné. Un dénombrement de triplets représente simplement le nombre de fois qu'on retrouve un type de triplet particulier dans le réseau. Deux types particulièrement intéressants de triplet, pour des raisons théoriques, sont le *triplet intransitif* et le *triplet transitif* (voir la figure 2).

Comme le modèle p_2 , le modèle p^* permet l'incorporation de variables explicatives individuelles et l'étude des effets dyadiques exogènes.

LE PASSAGE AUX MODÈLES DYNAMIQUES

Jusqu'à présent, les modèles présentés sont des modèles statiques, cependant, étant donné que les réseaux sociaux sont des structures qui se développent et changent avec le temps, il parut nécessaire d'arriver à comprendre le développement de ces structures. Pour ceci il fallait développer des modèles avec des procédures permettant de tenir compte du passage du temps.

Une procédure alternative à celle des modèles loglinéaires proposés par Wasserman (1980) et Wasserman et Iacobucci (1988) et apte à étudier le développement des réseaux dans le temps, est l'usage de processus de Markov. L'usage de modèles de Markov présente de nombreux avantages et un nombre moindre d'inconvénients que les modèles loglinéaires que van de Bunt (1999, p 31) résume ainsi :

1. Le modèle loglinéaire permet de gérer des relations évaluées de façon ordinale. Le modèle de Markov, tel que les applications l'ont rendu disponible jusqu'à présent, ne peut considérer que des réseaux dont les relations sont caractérisées de manière dichotomique : on considère qu'une relation existe ou n'existe pas. Ceci dit, la méthode de Markov peut être étendue pour plusieurs valeurs. L'utilisation disponible actuellement du modèle de Markov contient ce petit inconvénient relativement aux modèles loglinéaires, mais il est compensé par des grandes avantages tel qu'on va le voir par la suite.

2. Le modèle loglinéaire produit trop de paramètres. Si g est la taille du groupe, le modèle pl estime $2g+1$ paramètres.

3. Le modèle loglinéaire ne permet que des attributs catégoriels ou nominaux pour les acteurs ; alors que le modèle de Markov permet l'usage d'attributs catégoriels et numériques. En plus l'analyse de Markov permet l'incorporation d'un nombre supérieur d'attributs.

4. Les modèles de Markov sont fondés sur des processus de temps continus, ce qui est plus réaliste et en même temps ceci permet des interprétations des paramètres moins complexes. Notamment ils permettent des interprétations directes des résultats sur l'évolution d'un réseau sur lequel on dispose de mesures avec une périodicité irrégulière (par exemple des mesures espacées d'une semaine, ensuite d'un mois, puis six mois et un an) sans que ceci constitue un problème.

5. Les effets structuraux, tels que la transitivité ou la densité, sont en général difficiles à représenter, d'ailleurs ils sont impossibles à représenter pour les modèles loglinéaires à l'exception de la réciprocité, mais on réussit à le faire avec les modèles de Markov. Ceci constitue la question clé pour l'application préférable des modèles de Markov à l'analyse des réseaux sociaux,

6. Les modèles de Markov permettent facilement la simulation par ordinateur. Tout cela rend, dans l'ensemble, les modèles de Markov plus appropriés pour l'analyse longitudinale de réseaux.

LE MODELE SIENA(4)

SIENA propose un modèle statistique d'analyse de données longitudinales pour des réseaux complets. Il est donc nécessaire de disposer, au moins, de deux

observations de l'état d'un réseau complet à deux moments consécutifs (au moins deux matrices, représentant la situation initiale et la situation finale du réseau).

La variable dépendante analysée est le changement de relations dans le réseau représenté par deux (ou plus) graphes orientés parmi un nombre fixe d'acteurs(5) dont les relations sont définies de façon dichotomique (présente ou absente). SIENA étudie le changement longitudinal des réseaux complets. SIENA étudie les facteurs qui influencent le passage de la première situation à la deuxième (et la troisième, et la quatrième...).

Le modèle repose sur la théorie de l'action orientée par un but (*purposive action*) et postule que les acteurs réalisent des choix (établissement, maintien, suppression des liens) pour maximiser leur utilité, ce qui détermine la structure ultérieure du réseau. L'état du réseau à un moment donné conditionne les états postérieurs de celui-ci ; ainsi le réseau est considéré comme une structure qui conditionne sa propre évolution (d'ailleurs ceci correspond très précisément à la notion de structure de Piaget 1968).

Le changement du réseau est modélisé en tant que résultant d'effets structuraux, d'effets exogènes, et d'effets aléatoires. Le modèle permet donc l'étude d'effets structuraux à des niveaux différents tels que, par exemple, la réciprocité à niveau dyadique, la transitivité ou la popularité à niveau triadique, la densité au niveau global, endogènes à la dynamique du réseau.

Il permet aussi l'inclusion de variables exogènes, donc pas définies en termes de la relation dont le changement est étudié, telles que les caractéristiques des acteurs (sexe, âge, taille de l'acteur) ; et certaines caractéristiques dyadiques (similarités entre acteurs) y compris des autres relations qu'ils entretiennent (fréquences d'interactions, multiplicité des liens, volume des transactions...) qui pourraient influencer l'évolution du réseau. Il est possible aussi d'étudier un certain nombre d'interactions entre les variables endogènes et exogènes.

Les modèles statistiques utilisés sont des chaînes de Markov de temps continu qui peuvent être utilisés en informatique dans des modèles de simulation. L'évolution du réseau est modélisée comme la conséquence des nouveaux choix des acteurs d'établir une relation ou de supprimer une relation existante sur la base de fonctions ayant des composantes fixes et aléatoires que les acteurs essayent de maximiser. Les paramètres du modèle doivent être estimés à partir de données empiriques. Pour estimer et tester ces modèles, SIENA propose des procédures statistiques fondées sur la méthode des moments. Les procédures statistiques sont mises en oeuvre en utilisant un algorithme d'approximation aléatoire fondé sur des simulations par ordinateur du processus d'évolution du réseau.

Un des intérêts de ce programme est qu'il permet de contrôler réciproquement les effets des différents facteurs qu'on inclut dans l'analyse, ce qui est particulièrement important pour des effets structuraux. Par exemple la transitivité est contrôlée par la

réciprocité, et toutes les deux par la densité du réseau, ou encore les effets des variables exogènes sont contrôlés par les effets structuraux.

SIENA travaille (pour le moment) seulement avec des relations définies de façon dichotomique, c'est-à-dire que chaque matrice de relations observées doit exprimer la présence ou l'absence d'une relation définie par le chercheur. Cependant il est possible d'inclure par ailleurs, en tant que variables dyadiques exogènes, différentes mesures d'intensité des relations et de les inclure dans le modèle. Par exemple on peut étudier la probabilité d'une relation de durer selon que la fréquence d'interaction (ou l'intensité subjective) ait été grande ou faible dans le passé. Il faut toutefois que les variables considérées soient exogènes par rapport à la relation étudiée.

Les relations sont des relations orientées (dans la matrice, la ligne d'un acteur représente les relations qu'il émet, et la colonne de ce même acteur, les relations qu'il reçoit). Les relations d'un acteur avec lui-même ne sont pas prises en compte.

Postulats du modèle

L'ensemble des socio-matrices (toutes les matrices $n \times n$ d'éléments 0-1) s'appelle X . X possède $2^{n(n-1)}$ éléments, relations entre acteurs, ce qui est un nombre énorme pour l'utiliser dans des calculs mathématiques.

Pour éviter les calculs impossibles le modèle assume un certain nombre de postulats plus ou moins réalistes, nécessaires pour la modélisation mathématique :

- 1) Les acteurs du réseau peuvent évaluer leur position dans le réseau et tenter d'acquérir une configuration relationnelle qu'ils considèrent positive.
- 2) Étant donné l'état actuel du réseau, le modèle postule que tous les acteurs agissent de façon indépendante, c'est-à-dire sans coordination. Par contre, l'évolution dynamique assure une dépendance dans les changements consécutifs effectués par les acteurs.
- 3) Le modèle postule que les acteurs ont une connaissance totale des agencements des liens actuels dans le réseau.
- 4) Chaque acteur choisit ses relations, il émet des choix sur des associés.
- 5) On suppose que les changements ont lieu un après l'autre et que tous les acteurs ont la même probabilité de réaliser un changement.
- 6) Des changements ont lieu de façon non observée entre les moments des observations du réseau.
- 7) Les états futurs du réseau sont une fonction seulement du présent, on considère que le passé pertinent est contenu dans le présent. SIENA considère que les acteurs n'ont pas de mémoire du passé ou de capacité d'apprentissage.
- 8) Les acteurs sont considérés « myopes » c'est-à-dire que leur choix est un essai de maximisation immédiate de leur utilité. On ne considère pas qu'ils aient des stratégies à long terme.

Tout cela représente l'idée que les acteurs poursuivent leurs propres objectifs sous les contraintes imposées par leur entourage, en même temps qu'ils constituent l'entourage changeant des autres. La fonction objective exprime les buts que les acteurs semblent vouloir réaliser à court terme.

Cette évaluation se définit comme la fonction de la position de l'acteur dans le réseau et dépend de paramètres estimés à partir des données empiriques. Le modèle inclut un élément aléatoire pour rendre compte de la déviation entre les précisions théoriques et les observations réelles. On peut interpréter que cet élément aléatoire inclut a) les variables qui jouent un rôle mais qui sont inconnues par le chercheur, donc pas présentes dans le modèle, et/ou b) les effets d'ignorance des acteurs sur la situation du réseau ou de mauvaise évaluation sur la meilleure action à entreprendre.

Les modèles peuvent être implémentés comme des modèles de simulation aléatoire. Il s'agit d'une procédure fréquentielle qui utilise la méthode des moments. L'algorithme d'approximation aléatoire est une dérivation de l'algorithme de Robbins-Monro (1951).

On postule que la magnitude des effets qui influencent le changement du réseau reste constante au fil du temps entre mesures. On ne considère pas, pendant des périodes brèves de temps dans des contextes précis, que des effets augmentent ou diminuent leur importance. Si on a des raisons de penser que différents effets ont une importance différant à des moments différents, il faut réaliser plusieurs mesures (comme le font van de Bunt 1999, Zeggelink et al. 1997, Zeggelink et al. 1999, van Duijn et al. 2003). Si le temps entre deux mesures est trop grand, il se peut que la situation initiale ne soit plus pertinente pour la situation finale. Ceci aurait la conséquence que SIENA serait moins habile pour analyser ces données.

Discussion des postulats du modèle

Le réalisme des postulats sur lequel SIENA est construit sont discutables d'un point de vue théorique, même si on admet que d'un point de vue technique il est difficile de procéder autrement. En général ces idées s'adaptent mieux à des relations de type volontaire ou affinitaires. On pourrait plaider que :

1) L'idée d'une évaluation par l'acteur et de l'essai d'optimisation de sa position dans un réseau implique une vision conséquentialiste et instrumentale des actions de l'acteur. Tel que le dirait Boudon (1984), ceci implique que l'acteur oriente son action en tenant compte des conséquences de ce qu'il fait, que parmi l'ensemble de conséquences qu'il perçoit, il ne fait attention qu'à celles qui l'intéressent. Une approche dans la théorie du choix rationnel classique affirmerait que l'acteur est capable de déterminer les coûts et les bénéfices de ses choix. Cependant il n'y a pas d'accord dans la littérature sociologique et anthropologique à propos des relations personnelles (amitié, parenté) sur ce point. Certains auteurs (Bloch et Buisson 1991,

Bidart 1991, Godbut et Charbonneau 1993) plaident que quand les acteurs centrent leur attention sur les équilibres entre coûts et bénéfices des relations, nous sommes face à des situations de crises. Ils affirment que l'état normal des relations personnelles est celui de la mutualité altruiste et non comptabilisée. D'autres, tels que van de Bunt 1999, diraient plutôt le contraire. L'approche du modèle Siena ne va pas si loin car elle considère que les acteurs sont myopes et pas vraiment capables de déterminer ces rapports de coûts et bénéfices à la façon de la théorie du choix rationnel classique, c'est un modèle d'action avec des buts à court terme.

Si nous admettons que les acteurs sont stratégiques et évaluent leurs positions pour essayer de les améliorer au moyen de calculs rationnels de coûts et bénéfices – même à court terme - alors ce postulat est mieux adapté à des types de relations plutôt volontaires sur lesquelles ne pèsent pas des normes sociales très restrictives (comme les relations d'amitié) et où les marges de manœuvre des acteurs sont relativement larges. Ceci-dit il est aussi possible d'inclure des variables pour tenir compte des effets des normes dans le modèle et les estimer parmi les autres effets.

Autrement, les auteurs du modèle n'affirment pas que l'évaluation en vue de l'optimisation du réseau que les acteurs réalisent soit forcément consciente ou intentionnelle. La seule chose qu'ils affirment est que le réseau peut être modélisé comme si chaque acteur cherchait à obtenir une configuration de ses relations évaluée au mieux. Cette façon de comprendre l'évolution des réseaux sociaux est en accord avec le principe théorique sociologique de l'individualisme méthodologique (Degenne et Forsé 1994) que Snijders appelle ici modèle aléatoire orienté par l'acteur (*stochastic actor-oriented model*). À partir des résultats de la modélisation, on peut retrouver les contraintes ou préférences fortes qui ont eu lieu quand elles ont été prises en considération par le chercheur et sont incluses dans l'analyse.

2) L'indépendance d'action des acteurs, le manque de coordination, s'adapte sans doute mieux aussi à des relations volontaires et de caractère privé. Et même si, dans la littérature anthropologique sur l'amitié, on trouve des phénomènes d'exclusion collectifs des groupes d'amis, par exemple dans les modèles d'amitié des « cuadrillas » (Cucó i Giner 1998) où l'acceptation des nouveaux membres par le groupe d'amis ou leur exclusion sont le fruit de décisions fortes et collectives. À nouveau, même si ces processus existent dans la réalité, pour le traitement mathématique on fait *comme si* chacun de ces choix avait lieu de façon indépendante et à des moments différents, qui peuvent être très proches dans le temps. Un tel traitement peut être considéré acceptable car quelque part il y a une série de décisions individuelles dans ce processus de groupe sur l'acceptation ou exclusion de quelqu'un (et de fait il est possible de trouver parfois des dissidents au sein de ces processus de groupe).

3) Le postulat de connaissance totale de la structure du réseau par les acteurs n'est pas réaliste du tout. Particulièrement pour des grands réseaux. Il est très difficile dans un grand réseau de connaître l'état précis des relations entre acteurs, ceux-ci peuvent au plus se représenter l'état du réseau, ce qui peut très bien ne pas correspondre à la réalité. Pour les relations de caractère privé, par exemple les

relations d'amitié, qui peuvent se détacher du contexte original de formation (Allan 1989), il n'est pas toujours possible de connaître l'existence d'un lien particulier entre des personnes au-delà de leur rapport simplement amical au sens de cordialité même quand leur interaction est observable. Certains types de liens sont plus ou moins visibles pour les autres membres du réseau(6). Aussi, tel que nous le dit Ferrand (1997, 44) les acteurs rencontrent des barrières cognitives et sémantiques pour se représenter les relations indirectes, y compris dans le langage de la parenté qui est un des plus riches. Une connaissance à une distance 2 (connaissance de ses relations et les relations de ses associés) serait peut-être plus réaliste.

Les postulats 4, 5, et 6 sont assez acceptables pour des relations de caractère volontaire.

4) Dans le cas des relations d'amitié, par exemple, on peut considérer qu'à l'intérieur du contexte dans lequel les acteurs se meuvent, ceux-ci peuvent choisir leurs amis parmi les relations possibles et disponibles.

5) Même si les changements n'ont pas lieu l'un après l'autre, de façon forcément dyachronique, mais aussi de façon synchronique, tel que pourraient plaider des théories sur l'ajustement réciproque, ce postulat n'est peut-être pas dérangeant. On peut faire comme si un intervalle existe entre le moment où un acteur entre en conflit avec un autre, et le moment où cet autre en fait autant, même si ce décalage est infinitésimale. On peut faire *comme si* ces intervalles étaient réguliers pour le modèle mathématique, même s'ils ne le sont pas dans la vie réelle. Rappelons que nous avons affaire à un instrument pour obtenir des indications qui nous aident à interpréter la réalité et non pas à une interprétation sur comment elles ont réellement lieu.

6) Le fait que des changements aient lieu entre les moments d'observation du réseau est parfaitement réaliste.

7) Le postulat que toute information pertinente du passé est contenue dans le présent peut être « corrigée » au moyen de variables additionnelles. Par exemple, si nous avons des raisons de penser que l'instabilité d'une relation dans le passé (son alternance entre présence et absence) peut expliquer les états futurs de celle-ci (p.e. que les relations instables aient plus de probabilité d'être éliminées que les stables), il est possible de créer des variables qui décrivent telle instabilité et les inclure dans le modèle.

8) Le postulat de l'intention non stratégique à long terme, mais de satisfaction immédiate s'adapte bien à certains types de relations, mais pas à toutes. Il s'adapte bien par exemple aux relations d'amitié, où les stratégies structurales à long terme même si elles pourraient exister (je veux devenir ami de X car il est ami de Y) seraient mal vues car elles vont s'opposer au modèle idéal de l'amitié. Pour des relations d'amitié, on peut se dire que le modèle est valable pour les acteurs « de bonne foi » ou qu'il fait comme si tous les acteurs étaient de bonne foi. Cependant

pour d'autres types de relations, les acteurs peuvent avoir des stratégies de distance 2 ou plus à moyen ou long terme.

Malgré toutes ces remarques et nuances, nous pouvons considérer que l'utilisation de ce modèle peut nous apporter des connaissances valables sur les processus d'émergence des liens dans un réseau. Il est possible de « minimiser » les problèmes que pourraient induire les postulats que nous venons de discuter grâce à d'une sélection théoriquement guidée des variables incluses dans le modèle.

LA SPECIFICATION DE SIENA

Le modèle utilisé par SIENA contient trois éléments :

- a) La fonction objective (*objective function*)
- b) La fonction de vitesse (*rate function*)
- c) La fonction de dotation (*endowment function*)

Dans cette section, nous parlons de la spécification simple du modèle, déterminée seulement par la fonction objective, avec une fonction de vitesse constante, et une fonction de gratification égale à 0. C'est-à-dire, en principe le modèle de base comprend seulement la fonction objective. Les deux autres fonctions seront discutées plus tard.

Composantes du modèle de base. La fonction objective

La fonction objective pour un acteur i indique le degré de satisfaction de l'acteur i inhérent à la configuration x du réseau. À un moment t l'acteur i a l'opportunité de changer une relation (commencer, maintenir ou dissoudre une relation). On postule que l'acteur i choisit l'action relative à l'acteur j qui maximise la valeur de la fonction objective. La fonction inclut un élément aléatoire qui indique les préférences de l'acteur qui ne sont pas représentées par les composantes qu'on est en train d'étudier.

Le modèle suppose que autant qu'il existe des différences entre les acteurs quant à leurs préférences, ces différences sont capturées (exprimées) par les variables exogènes. Par exemple, si certains acteurs ont une préférence plus grande de réciprocité que d'autres, il devrait exister une variable W observée de sorte que l'interaction entre W et l'effet de réciprocité peut représenter ces préférences différentielles.

On peut trouver curieux qu'on parle tellement en termes de préférences individuelles et que finalement le modèle extrait les paramètres du fonctionnement du réseau dans son ensemble sans faire de distinction entre les préférences des individus particuliers : ceci est sans doute dû à des complications mathématiques.

Pour notre interprétation, cependant, il vaut peut-être mieux parler de l'effet « moyen » qu'on retrouve dans le groupe pour chacune des variables considérées.

Principaux effets de la fonction objective

Ici on décrit quels sont les principaux effets que nous pouvons étudier avec SIENA. La fonction objective doit contenir les composantes substantifs du modèle, en incluant, par exemple les attributs de l'acteur et les propriétés structurales du graphe orienté qu'on veut étudier. La fonction est définie en tenant compte du fait que l'acteur contrôle seulement les choix qu'il réalise d'établir, maintenir, interrompre, des relations avec d'autres, et pas les choix faits par les autres.

On peut distinguer plusieurs types d'effets :

- 1) Des effets qui dépendent du réseau, ce qu'on appelle effets structuraux (réciprocité, transitivité...) ou variables endogènes.
- 2) Des effets exogènes qui peuvent être de deux types :
 - 2.1) Des effets de variables dépendantes de l'acteur, comme le sexe, la nationalité, l'âge, la productivité...
 - 2.2) Des variables dépendantes de paires d'acteurs ou variables dyadiques, par exemple la similarité sur un critère (avoir le même sexe ou voter au même parti) mais aussi des variables décrivant proprement des relations exogènes par rapport à la relation dont on étudie l'évolution (présence de certains échanges, fréquence de certaines interaction etc).

On peut considérer des variables constantes qui ne dépendent pas du temps, comme le sexe, mais on peut aussi considérer des variables changeantes comme des opinions ou le niveau de maîtrise d'une langue.

Effets principaux inclus dans la fonction objective :

- 1) Degré extérieur ou densité : définie par les relations émises par l'acteur i .
- 2) Réciprocité : définie par le nombre de relations réciproques (si i émet une relation vers j et j émet une relation vers i). La réciprocité exprime la préférence pour des liens réciproques : il est plus « intéressant » de choisir comme ami quelqu'un qui vous considère déjà comme ami que quelqu'un qui ne le fait pas.
- 3) Popularité : définie par l'addition des relations reçues par les acteurs j avec lesquels i est en rapport (degré intérieur du partenaire). L'effet de popularité reflète la préférence d'être lié à des acteurs populaires. Il peut refléter des effets d'aspiration de statut.
- 4) Activité : définie par l'addition des relations émises par les acteurs j avec lesquels i est en rapport (degré extérieur du partenaire). L'effet d'activité reflète la préférence d'être lié à des acteurs émetteurs de beaucoup de relations.

Effets de fermeture d'un réseau, on en distingue 3 :

5) Transitivité : définie par le nombre de triplets transitifs parmi les relations de i (i est en rapport avec j et h , et ceux-ci sont en rapport entre eux).

6) Relations indirectes : définie par le nombre d'acteurs avec lesquels i est en rapport indirect, avec une distance sociométrique de 2, c'est-à-dire avec un intermédiaire.

7) Équilibre (*balance*) : définie par la similarité entre les relations émises par l'acteur i comparés aux relations émises par l'acteur j avec lequel i est en rapport. La transitivité des triplets exprime qu'un acteur i sera plus attiré par un acteur j s'il existe plus de connexions indirectes $i \rightarrow h \rightarrow j$; un effet négatif des connexions indirectes exprime que i sera plus attiré par j s'il y a au moins un lien indirect entre les deux. L'effet de balance exprime que i préfère devenir ami avec ceux j qui font des choix d'amis semblables à lui-même. La liste peut augmenter, mais pour l'instant ceux-ci sont les effets inclus. Il est conseillé de toujours inclure dans le modèle la densité, car c'est un effet de contrôle pour tous les autres.

Pour toutes les variables exogènes dépendantes des acteurs, SIENA propose trois effets fondamentaux:

8) Popularité liée avec une variable. Elle est définie par la popularité relativement supérieure pour les acteurs ayant des valeurs élevés de la variable. Par exemple la popularité relative des hommes comparés à celle des femmes dans un réseau.

9) Activité liée avec une variable, définie par l'activité relativement supérieure pour les acteurs ayant des valeurs élevés de la variable. Par exemple l'activité des femmes

10) Similarité liée à une variable, définie par la somme des différences absolues entre la valeur de la variable pour i et tous les autres avec lesquels i est en rapport. Quand la similarité est positive ceci indique la préférence par les personnes de même caractéristique (sexe, nationalité, classe sociale...).

Effet pour les variables exogènes dyadiques (variables qui dépendent des paires d'acteurs) :

11) Les préférences liées à une variable dyadique définie par la probabilité d'un lien entre i et j étant augmentée si la variable dyadique est plus élevée pour ces acteurs.

12) Des interactions entre la réciprocité et les variables dyadiques.

Puis toute sorte d'interaction entre variables qu'on peut définir soi-même en créant des variables dichotomiques (*dummy variables*).

Les variables à inclure dans le modèle dépendent évidemment de l'objet d'étude et des orientations théoriques.

Le modèle étendu : la fonction de vitesse et la fonction de dotation

En plus de la fonction objective qui peut contenir les effets dont on vient de parler, SIENA comprend deux autres fonctions : la *fonction de vitesse* (*rate function*) et la *fonction de gratification* (*gratification function*). Ces deux fonctions sont ajoutées pour exprimer des effets qui ne peuvent pas être contenus dans la fonction objective.

Effets principaux inclus dans la fonction de vitesse :

La *fonction de vitesse* a comme but de rendre compte du fait que la vitesse de changements de la structure du réseau peut varier, et que ces changements de vitesse peuvent être liés à certaines caractéristiques des acteurs, des relations ou de la structure du réseau.

1) Effet de base. L'effet de base fait partie aussi du modèle de base et il rend compte de la vitesse à laquelle les changements ont lieu.

2) Effet de l'activité sur la vitesse. Il permet de tester si les acteurs qui sont très actifs, qui émettent plus de choix (*out degrees*), les changent plus (ou moins) vite que ceux qui émettent moins de choix.

3) Effet de la popularité sur la vitesse. Il rend compte de l'idée que les acteurs qui sont très populaires, qui reçoivent plus de choix (*in degrees*), pourraient les changer plus (ou moins) vite que ceux qui reçoivent moins de choix.

4) Effet de la réciprocité sur la vitesse. Il permet de tester si les acteurs aux relations réciproques changent plus (ou moins) vite que les acteurs ayant moins de relations non réciproques. Pour des relations d'amitié, par exemple, on pourrait s'attendre à ce que les acteurs ayant beaucoup de liens réciproques changent moins vite leurs liens.

5) Effet des attributs des acteurs sur la vitesse. Grâce à cet effet, on peut examiner comment les différentes caractéristiques des acteurs peuvent jouer sur la vitesse des changements des relations. Par exemple on pourrait examiner si les hommes changent plus vite leurs relations que les femmes.

Effets principaux inclus dans la fonction de dotation :

La *fonction de gratification* a comme but de rendre compte du fait qu'un même effet peut être plus ou moins important quand il s'agit de la création d'une relation, ou quand il s'agit de sa rupture. En effet, l'établissement et la rupture d'une relation

ne sont pas des actions tout à fait inverses. Théoriquement et empiriquement ceci est dû aux investissements en temps et autres ressources qu'on emploie dans les relations, ou aux coûts et sanctions, dyadiques ou triadiques, que peuvent impliquer les ruptures de relations. Dans la rupture d'une relation, les investissements non récupérables sont perdus, comme le temps et d'autres ressources. On prend aussi le risque de sanctions triadiques quand le lien est encadré.

On peut étudier grâce à cette fonction quels sont les facteurs qui rendent coûteuse la dissolution d'une relation.

1) Rupture des relations réciproques. Typiquement on peut considérer qu'il peut être plus coûteux de dissoudre une relation d'amitié réciproque qu'une relation asymétrique, c'est l'idée testée ici.

2) Accès par les relations indirectes. Il permet d'examiner si, dans certains modèles relationnels, il est plus facile de commencer une relation déjà encadrée qu'une relation isolée.

3) Rupture des relations selon une variable dyadique. Il montre comment des caractéristiques de la relation (p.e. flux, fréquences, intensités...) rendent sa rupture plus ou moins difficile.

4) Rupture des relations selon une dissimilitude. Il met à l'épreuve l'idée selon laquelle des individus différents sur une caractéristique (p.e. sexe, nationalité) termineraient leur relation plus ou moins facilement.

Quel que soient les composantes qu'on utilise dans le modèle, il doit être construit avec parcimonie, en incluant seulement 2 ou 3 effets additionnels à chaque fois qu'on fait tourner le modèle. A chaque tentative on élimine les effets non significatifs et on peut ajouter 2 ou 3 variables supplémentaires. Les effets étudiés se contrôlent mutuellement. Les résultats présentent un paramètre et son écart-type. Il faut lire ceux-ci comme des distributions T-test, c'est-à-dire que l'on considère un paramètre comme significatif quand sa valeur est au moins le double de son écart-type(7). Plus le rapport entre le paramètre et son écart type est grand, plus nous sommes sûrs de sa significativité, cependant il est difficile de parler de la magnitude des effets ou de les comparer entre eux, étant donné que celle-ci dépend de la distribution des valeurs sur chaque variable et celles-ci varient souvent d'une variable à une autre.

UN EXEMPLE D'APPLICATION. L'EVOLUTION D'UN RESEAU D'AMITIE INTERNATIONAL

Nous voulons illustrer les potentialités de SIENA au moyen d'un exemple d'application : l'évolution d'un réseau d'amitié d'étudiants Erasmus. Ces étudiants participent dans un programme d'échange universitaire de l'Union Européenne

appelé « Erasmus ». La présentation de cette question a déjà été effectuée ailleurs (voir de Federico 2003a, 2003b, 2004). Ici la présentation est réduite au nécessaire pour illustrer l'utilisation de SIENA.

Depuis les années 80 l'Union Européenne s'est efforcée pour développer le programme Erasmus avec l'objectif explicite de « renforcer l'esprit de la citoyenneté Européenne » et de « renforcer la compréhension parmi les peuples de l'Union Européenne ». Les efforts ont été conséquents, l'UE a investi des énormes quantités d'argent (p.e. 922 millions d'Euros pour la période 1995-2000) et dans les 10 premières années de son fonctionnement, environ 1,5 millions d'étudiants ont pu séjourner entre 3 et 12 mois dans une université d'un autre pays européen grâce à ce programme.

Quelque part l'hypothèse sous-jacente de l'UE pour le développement des échanges d'étudiants est celle de la psychologie sociale que le contact entre personnes de catégories sociales différentes (des nationalités différentes) favorise la solidarité et la compréhension entre catégories. Eventuellement - sous certaines conditions spécifiées par Allport (1954) : soutien social et institutionnel, potentiel de connaissance personnelle des membres de l'autre catégorie, égalité de statut, possibilités d'interactions coopératives - le contact peut même conduire jusqu'à la re-catégorisation des membres de catégories différentes dans une catégorie sociale englobante commune (l'appartenance commune à l'Europe).

Les données présentées ici proviennent d'une recherche évaluant l'incidence des relations créées lors des échanges Erasmus sur les attitudes vis-à-vis de l'Europe et des Européens. Cette recherche comporte deux volets. Le premier se consacre aux relations d'amitié et leur dynamique : qui sont les amis des étudiants Erasmus, comment se développent et évoluent les relations d'amitié, quelles relations durent au delà du séjour Erasmus etc. Ensuite il devient possible de s'intéresser aux effets de ces relations effectivement créées. Le deuxième volet étudie alors de forme systématique les effets des relations d'amitié sur les sentiments d'appartenance à l'Europe, sur les attitudes vis-à-vis des pays Européens et vis-à-vis de l'Union Européenne.

Les analyses présentées ici font partie du premier volet, c'est-à-dire l'étude du développement et de l'évolution des relations d'amitié des étudiants Erasmus. Avant de montrer l'application de SIENA pour expliquer l'évolution de ce réseau, nous caractérisons brièvement les relations d'amitié à partir de la littérature, puis nous décrivons succinctement le réseau étudié.

Les relations d'amitié

Paine (1969) et Allan (1979,1989) définissent l'amitié comme une relation personnelle, volontaire et principalement affective. Elles émergent sans doute car les personnes préfèrent avoir des amis que de pas en avoir. Les amis nous acceptent

« tels que nous sommes » (Allan 1989, Bidart 1991, 1993) et ils nous rendent la vie plus facile et agréable. Ils partagent nos intérêts et idées (Lazarsfeld et Merton 1954, Bidart 1991), ils nous aident avec les questions de la vie quotidienne (Wellman et al. 1988, Wellman et Wortley 1990) ainsi que dans les grands moments de crise (Ferrand 1993, Bidart 1993).

Les amitiés sont formées dans des contextes sociaux et sont soumises des normes et des modèles sociaux (Ferrand 1993). Au minimum, il est nécessaire d'être dans le même lieu au même moment, et de faire quelque chose qui permette une interaction minimale, pour pouvoir connaître quelqu'un, découvrir qu'il nous plaît et nous devenir amis. Par conséquent, pour la formation de relations d'amitié, il existe une structure d'opportunités liée à des aspects purement géographiques (où l'on passe son temps), mais aussi d'interdépendance fonctionnelle dans des *foci* d'activité (ce que nous faisons et quelles interactions permet l'activité) (Feld 1997, Flap et Volker 2002). À partir du moment où il existe la possibilité d'interaction, en général il paraît que les personnes préfèrent devenir des amis de personnes semblables. Les relations positives avec des caractéristiques semblables (valeurs, attitudes, position sociale) réduisent la dissonance cognitive (Heider dans Perlman et Fehr 1986).

Par ailleurs, les relations sont établies souvent à partir d'autres relations (Hallinan 1979, Ferrand et al. 1999). Nos réseaux personnels nous permettent d'avoir accès à de nouveaux contacts : les amis de mes amis deviennent mes amis (Granovetter 1973, Degenne et Forsé 1994). En même temps ils les limitent aussi : les ennemis de mes amis ne peuvent pas être mes amis (Lorrain et White 1971)

Jusqu'ici nous avons mentionné des facteurs classiques dans les théories de l'amitié (van de Bunt 1999) : en partant d'une préférence pour avoir des amis, nous avons mentionné l'importance de la *structure des opportunités* de contact, la supposée *préférence par la similarité*, et *l'accès conditionné* à de nouvelles relations à partir des précédentes.

Sur ce point, il convient aussi de faire une allusion au fait que, bien que dans nos sociétés occidentales l'amitié soit une relation personnelle, privée, volontaire et non ritualisée (Paine 1969), sa définition, contenu et attentes peuvent différer selon les pays. Il suffit de constater la différence de termes qui sont employés pour en faire allusion dans différents pays : en français il existe une distinction entre ami et copain. En Anglais britannique entre *friend* et *mate*, tandis que les Américains utilisent seulement *friend* ou éventuellement dans la littérature *friendly relation* (et de fait Fischer (1982b) a trouvé des différences systématiques entre les définitions des américaines et des britanniques, ces derniers étant bien moins enclins à qualifier leurs relations d'amis). En espagnol, on utilise seulement *amigo* pour se référer aux deux nuances, et en Grec (moderne), le terme pour désigner les amis est encore plus vaste : le mot φίλος comprend aussi le fiancé ou la fiancée. Le contenu spécifique des relations est à comprendre selon le contexte. Nous pouvons alors nous attendre à trouver des différences en ce qui concerne la façon de comprendre l'amitié selon les pays.

Le réseau d'amitié Erasmus

Le réseau étudié concerne 80 étudiants venant de 9 pays Européens(8) accueillis à partir du mois d'octobre à l'Université de Lille 1 en 1995 dans le cadre du programme d'échanges Erasmus. Nous avons effectué plusieurs mesures de ce réseau. La première mesure correspond au moment de leur arrivée en France au début du mois d'octobre. La deuxième a eu lieu en février de 1996 quatre mois après leur arrivée en France, juste avant le retour des premiers étudiants. Le générateur de noms utilisé avait une définition large de l'amitié : « Pouvez vous mentionner le nom de vos amis et copains lors de votre séjour Erasmus ». Cette question prétendait capturer toutes sortes d'amis d'intensités variées. Les relations d'amitié et les amis étaient ensuite caractérisés de façon assez détaillée.

Au moment de la première description du réseau, les relations d'amitié étaient obligatoirement peu nombreuses étant donné que les étudiants venaient d'arriver en France. Les quelques relations existantes correspondaient à des relations préalables à l'échange, c'est-à-dire, aux cas où les étudiants se déplaçaient accompagnés par d'autres amis ou leurs partenaires.

Il n'y avait pas à l'université de Lille d'arrangements institutionnels qui favorisaient particulièrement l'interaction exclusive parmi les étudiants Erasmus, et de fait nous avons employé un protocole d'analyse de réseaux personnels, cependant *80 % des amis des étudiants Erasmus étaient d'autres étudiants Erasmus*. Seulement 17 % des amis étaient des étudiants Français et 3 % avaient d'autres origines non Européennes.

L'auto sélection de cette population était flagrante si l'on considère qu'ils ne constituaient que 1,7 % de la population totale d'étudiants de l'Université de Lille 1 et qu'ils étaient dispersés dans des disciplines, diplômes et lieux de résidences différents où ils étaient minoritaires. Étant donné la constitution d'un cercle social d'étudiants Erasmus, il était pertinent d'un point de vue substantif de reconstruire le réseau total des étudiants Erasmus. Nous avons reconstitué alors l'état du réseau amical des étudiants à deux moments différents que nous allons appeler *T0* (moment de leur arrivée à Lille en octobre 1995) et *T1* (quatre mois plus tard en février 1996).

Le réseau a beaucoup changé entre ces deux moments comme il est possible de constater à partir des représentations réalisées au moyen de Pajek (voir figures 3 et 4).

Si nous examinons leurs caractéristiques de façon systématique, nous pouvons constater qu'à *T0* le réseau est quasiment vide. Le nombre moyen d'amis des individus dans le réseau n'est même pas de un. Le nombre de choix émis concernant les amis et le nombre de choix reçus varient de 0 à 7. La densité et la transitivité sont très basses. Par contre la réciprocité est très élevée. Il est clair que

dans une situation où si peu de personnes sont amis entre eux elles sont plus facilement d'accord sur quels liens d'amitié existent et sont mutuels (voir tableau 1).

Quatre mois plus tard, la situation est très différente. Le nombre moyen d'amis a augmenté jusqu'à quasiment 12 personnes, ce qui est un chiffre assez élevé quand on le compare avec d'autres réseaux d'amitié d'étudiants de caractéristiques similaires (voir van de Bunt 1999). Les relations émises varient de 0 à 35 et les relations reçues de 2 à 27. La densité est toujours assez basse (il est rare que des réseaux de relations personnelles si larges aient des densités élevées) mais elle a augmenté sept fois. La transitivité est aussi basse, mais elle a énormément augmenté, 34 fois. La réciprocité est toujours très élevée, mais un peu moins qu'avant. Avec une plus grande et plus diffuse sociabilité, il est plus facile qu'il y ait des désaccords ou des oublis quand on mentionne les amis de façon mutuelle.

Facteurs d'évolution du réseau d'amitié de t_0 à t_1

Nous présentons par la suite les facteurs intervenant dans l'évolution du réseau d'amitié parmi les étudiants Erasmus entre T_0 et T_1 grâce à SIENA. Pour analyser l'évolution du réseau depuis son état à T_0 jusqu'à son état à T_1 , nous allons mettre quelques hypothèses à l'épreuve au moyen de SIENA, ce qui permettra d'illustrer la façon de procéder.

Dans un *premier modèle*, nous pouvons supposer que l'évolution de ce réseau d'amitié est, en partie, due à des effets structuraux, dits endogènes au fonctionnement des réseaux. Notre premier modèle contiendra alors les effets structuraux de *densité* (il faut toujours l'inclure comme moyen de contrôle), *réciprocité* (il est plus probable de choisir ceux qui nous choisissent) et *transitivité* (les amis de mes amis deviennent mes amis) (9).

Étant donné que les trois effets sont significatifs, on les laisse dans le modèle et l'on ajoute encore 2-3 variables (voir tableau 2).

Le *deuxième modèle* teste alors les variables structurales étudiées, plus trois variables dyadiques. Ces variables sont dyadiques car leur valeur dépend de la combinaison des caractéristiques des deux acteurs impliqués. Leur distribution ne peut alors être exprimé que dans une matrice. Imaginons que l'acteur A est français, l'acteur B est français, l'acteur C est espagnol et l'acteur D est espagnol aussi. Les paires (dyades) A et B et C et D ont la même nationalité. Mais les paires A et C, A et D, B et C et B et D ont des nationalités différentes. Ceci peut être exprimé par une matrice où 1 symbolise la similarité sur la caractéristique donnée, ici la nationalité, et 0 la différence sur cette caractéristique :

	A	B	C	D
A	1	1	0	0
B	1	1	0	0
C	0	0	1	1
D	0	0	1	1

Concrètement les variables dyadiques que nous avons utilisées dans le modèle sont : la *similarité de sexe* (il est plus probable de devenir ami de quelqu'un de même sexe), la *similarité de nationalité* (il est plus probable de devenir ami de quelqu'un de même nationalité) et la *structure des opportunités* de rencontre liée à la proximité géographique (il est plus probable de devenir ami de quelqu'un qui habite près).

Ici on constate que la similarité de nationalité ainsi que la proximité des lieux de résidence jouent un rôle dans l'évolution du réseau (voir tableau 3).

Par contre, de façon différente à la plupart des populations où les relations d'amitié ont été étudiées y compris dans des tranches d'âge semblables, il n'y a pas d'effet d'homophilie de sexe. Puisqu'elle n'est pas significative, cette variable est éliminée du modèle.

Dans le *troisième modèle*, nous suggérons de tester des effets liées aux caractéristiques des acteurs ainsi qu'une variable étudiant des interactions. On constate dans la figure 4 que les étudiants britanniques semblent avoir bien moins d'amis que les autres nationalités, nous testons donc l'*activité* (choix émis) des britanniques. On constate aussi que les italiens se trouvent au centre d'un sous-réseau assez dense, nous testons alors, par exemple, la *popularité* (choix reçus) des italiens.

Comme variable d'interaction nous testons la réciprocité selon la similarité de nationalité. Ici il s'agit de voir si les relations d'amitié entre personnes de même nationalité ont tendance à être plus réciproques que les autres (voir tableau 4).

On constate que les Britanniques se montrent, en effet, significativement peu actifs dans l'établissement de liens d'amitié, par ailleurs les Italiens sont significativement plus populaires dans ce réseau.

Concernant la réciprocité par nationalité, nous retrouvons l'effet inverse de ce qu'on aurait pu penser : quand des personnes de nationalités différentes établissent des liens d'amitié (au départ plus rares) ceux-ci ont tendance à être plus réciproques, plus forts. Apparemment nous sommes face à une préférence pour les biens rares(10). Cet effet, surprenant et particulièrement intéressant concernant la population étudiée, aurait été difficile à trouver hors du cadre d'analyse de SIENA où les différents effets se contrôlent mutuellement.

Finalement, dans le *quatrième modèle*, nous suggérons d'ajouter un effet lié à la fonction de vitesse et un effet lié à la fonction de gratification, par exemple, *les femmes changent des relations d'amitié plus vite que les hommes et les relations réciproques ont plus de chances de survie que les autres* (voir tableau 5).

Nous constatons qu'aucune de ces deux variables n'a un effet sur le développement du réseau amical étudié.

À mode de résumé, nous pouvons dire que réseau amical des étudiants Erasmus à Lille en 1995 s'est développé à partir d'un petit nombre de mécanismes assez classiques : 1) Les amis habitent près. 2) Ils choisissent comme amis ceux qui les choisissent comme amis. 3) Les amis de leurs amis sont leurs amis.

Puis nous en avons découvert d'autres plus originaux : 1) Autant les études sur l'amitié montrent habituellement l'homophilie de sexe, nous ne la retrouvons pas du tout dans cette population. Ceci est un résultat surprenant et semble spécifique de cette population. Nous pouvons concevoir deux interprétations possibles et compatibles pour expliquer ce phénomène. Une possibilité est que le séjour à l'étranger ait un effet libérateur : les réseaux qui habituellement exercent le contrôle social sur les rôles liés au genre ne sont pas présents permettant des relations hétérophiles ou homophiles d'être créées indistinctement. Une autre possibilité est que les personnes soient limités dans le nombre de critères de similitude dont ils peuvent tenir compte quand ils créent des liens d'amitié. Etant donné qu'il s'agit d'une population internationale, le critère de similitude qui prend le dessus pour l'homophilie dans cette population est celui de la nationalité. Il serait alors cognitivement exigeant – ou difficile dans un contexte de choix restreint - pour les individus qu'ils aient des relations d'amitié homophiles concernant la nationalité et le sexe.

2) Un autre résultat particulièrement intéressant est que les relations d'amitié avec des personnes de nationalité différentes sont plus souvent des liens réciproques. Les liens créés avec des personnes ayant des origines différentes, *a priori* des liens plus rares, obtiennent le statut de réciprocité plus souvent. Ceci semble indiquer une valorisation particulière de ces liens rares qui possèdent une valeur ajoutée liée à la différence d'origine(11).

3) Finalement nous avons pu constater l'existence de différences dans les modèles nationaux de sociabilité ayant pour conséquence des comportements différenciés par nationalité. D'accord avec les recherches de Fischer (1982b), les Britanniques se sont montrés moins enclins à créer des liens d'amitié que les autres. Par ailleurs les Italiens semblent être préférés en tant qu'amis par rapport à d'autres nationalités.

CONCLUSION

Dans cet article, nous avons rendu compte des importants développements qui ont eu lieu depuis une vingtaine d'années dans le domaine des modèles statistiques permettant d'analyser les structures des réseaux sociaux en tant que variables dépendantes des conduites sociales des acteurs (voir le tableau 5). Ces modèles ont permis d'approfondir le point de vue proprement « structural » sur les réseaux sociaux, au sens « intentionnel » du mot selon Boudon (1968, p35) qui permet de « souligner qu'une méthode a pour effet de décrire un objet comme un système ».

Le passage aux modèles d'analyse dynamiques, dont SIENA et p^* font partie(12), a été d'une grande importance car ceci constitue une contribution méthodologique cruciale pour rendre compte empiriquement de théories génétiques sur la structure des réseaux sociaux totaux. Ils permettent de rendre compte de l'ensemble des *principes qui gouvernent le déploiement et la reproduction d'un réseau social* et donc de définir la structure d'un réseau dans ces termes-là. À la place de mettre l'accent sur la description de la totalité des relations, il devient possible d'examiner les agencements typiques de relations et, surtout, la reproduction de ces agencements comme le résultat de la décision des acteurs d'établir (maintenir ou rompre) une relation en tenant compte des relations préexistantes (voir tableau 6).

Ici une autre remarque devient importante car, paradoxalement, alors que les modèles statistiques pour analyser les réseaux sociaux totaux atteignent leur sommet de complexité théorique et technique, ils finissent par rejoindre des approches déjà proposées par les approximations d'analyse de réseaux personnels (voir Ferrand 1997) au moyens de sondages(13). Tel que Ferrand (2002) l'exprime, l'analyse de réseaux sociaux totaux est en train d'évoluer, en utilisant la distinction effectuée par Claude Levi-Strauss, d'une approximation « mécanique » à une approximation « statistique ». L'approximation mécanique prend les données, exhaustivement complètes, comme la véritable expression d'un ordre social prescrit, alors que la perspective statistique considère les données comme une expression approximative d'un ordre social fondé sur des préférences individuelles limitées par des règles proscriptives.

Ainsi, la description des propriétés d'un réseau total au moyen du dénombrement de triades et l'évaluation statistique de leur distributions (tel que fait p^*) ouvrent des nouvelles possibilités à la formalisation des structures des réseaux au moyen des *distributions statistiques différenciées de sous structures typiques*. Une fois familiarisés avec de telles descriptions, nous serons d'accord a) que nous pouvons penser aux structures en tant que modèles probabilistes ; b) que les proportions des différents types de réseaux personnels ou de sous-structures élémentaires typiques dans une population trouvées à partir d'échantillons aléatoires fournissent des indicateurs empiriques *suffisants* sur les *modèles des structures relationnelles* sous-jacentes inconnues des systèmes auxquels nous nous intéressons (Ferrand 1997, 2002, Spreen 1999). Au-delà de la simple description, l'approximation génétique pour rendre compte des principes à l'œuvre dans la dynamique des réseaux sociaux

a aussi été mise à l'épreuve avec pertinence dans les approches des réseaux personnels (Ferrand 1997, Ferrand et al. 1999).

L'approche à privilégier alors entre les réseaux personnels ou les réseaux totaux dépendra de la réponse aux questions suivantes : a) Dans quelle mesure les relations qu'on étudie sont circonscrites à des ensembles aux limites claires et pertinentes ou est-ce qu'il s'agit de réseaux potentiellement infinis ? Dans le premier cas, une approche en réseau total est tout à fait pertinente et plus riche, dans le deuxième, seule une approche en réseau personnel a un sens. b) Quelle est l'étendue pertinente du réseau autour de l'acteur pour expliquer les changements de liens qui se produisent dans son environnement ? Si on admet que, pour des relations de caractère plutôt volontaire, l'acteur maîtrise son réseau, il décide d'établir (maintenir ou rompre) ses relations, une étendue relativement limitée (par exemple à une distance 2) peut être suffisante, voir plus pertinente théoriquement, étant donné les limitations cognitives des acteurs concernant les relations indirectes dans les réseaux. À ce moment-là, une approche en réseau personnel peut être suffisante pour rendre compte de l'étendue pertinente en jeu pour la dynamique du réseau.

Par contre, pour des relations de caractère involontaire, ou si on admet que, même pour les relations de caractère volontaire, les acteurs ne maîtrisent pas complètement leurs réseaux et subissent les conséquences des actions des relations indirectes (ceci peut être particulièrement clair lors de conflits, tel que le montrent Lorrain et White 1971 ou Pizarro 2004) même au-delà de ce dont ils sont conscients, l'étendue pertinente peut être bien plus large. Aussi les approches en termes de réseaux personnels ne peuvent pas ou peuvent difficilement rendre compte d'effets structuraux tels que les avantages que peuvent tirer des acteurs dans des position intermédiaires entre des parties du réseau global non liées par ailleurs (ce qu'on appelle des « ponts » ou des « trous structuraux » (Burt 1992). Là où les effets proprement structuraux sont importants l'approche par des réseaux totaux serait préférable.

NOTES

(1) Je voudrais remercier Tom Snijders, Alexis Ferrand, et Zeline Lacombe pour leurs remarques sur des versions antérieures de ce texte.

(2) Les manuels pratiques, articles et explications complètes à propos de SIENA ainsi que le logiciel et quelques données à mode d'illustration sont disponibles sur les sites :

<http://statt.gamma.rug.nl/stocnet> et <http://statt.gamma.rug.nl/snijders/siena.html>.

SIENA est implémenté dans un programme qui a pour vocation d'accueillir d'autres programmes d'ARS. Le projet cadre s'appelle StOCNET. Ces programmes et leur documentation sont gratuits.

(3) Antérieurement Holland et Leinhardt (1976) ont proposé le modèle statistique U|MAN. Le dénombrement de triades montré dans la page 5 provient de leur proposition.

(4) Nous décrivons ici de façon principalement « verbal », les principes et fonctionnement du modèle SIENA. Pour des informations techniques plus détaillées voir Snijders (2004), Snijders (2001), Snijders et van Duijn (1997). La section qui suit d'ailleurs peut être considéré comme un résumé commenté et non technique de Snijders (2001).

(5) À présent SIENA permet de prendre en compte le renouvellement des acteurs dans le réseau dans ses analyses. Voir Huisman et Snijders (2003).

(6) De fait, si ce postulat était réaliste cela rendrait plus facile la vie des analystes de réseaux. Voir Ferrand et Snijders pour une discussion à propos des relations sexuelles.

(7) Plus précisément si $1,65 < t < 1,96$; $p < 0,10$; si $1,96 < t < 2,53$; $p < 0,05$; si $2,53 < t < 3,29$; $p < 0,01$ et si $t > 3,29$ $p < 0,001$.

(8) Allemagne, Autriche, Belgique, Espagne, Grande-Bretagne, Grèce, Italie, Portugal et Suède

(9) Le fait que le nombre de triplets transitifs soit un ingrédient plus courant de la notion de transitivité/fermeture du réseau ne devrait pas induire l'abandon d'autres effets de fermeture, tel que le nombre de relations indirectes, en tant que spécifications alternatives pour les effets de fermeture du réseau. Il se trouve que dans ce cas précis l'effet des triplets transitifs était plus fort que celui des relations indirectes, mais souvent c'est le contraire.

(10) Et qui d'ailleurs se maintient dans la durée. Un an après le retour de l'échange Erasmus, ce sont les relations transnationales qui sont maintenues, alors que les relations d'amitié avec des personnes de même nationalité disparaissent plus vite (de Federico 2003a).

(11) Tom Snijders affirme que les liens avec ceux qui sont faciles à approcher sont plus souvent asymétriques que ceux avec qui le contact est difficile. Pour les dyades formés entre deux personnes dont le contact ne serait pas évident ou facile (qui habitent près, avec des coûts faibles etc.) – dit il - soit ils n'ont pas du tout de lien, soit le lien est réciproque. Dans notre cas d'application aux étudiants Erasmus nous avons testé aussi si d'autres liens moins fréquents, plus difficiles, avaient aussi tendance à être plus réciproques (p.e. des liens entre personnes habitant des résidences différentes) mais ce n'était pas le cas. Il semblerait alors que a) soit la différence d'origines serait une barrière plus forte à surmonter que le fait d'habiter dans une autre résidence ce qui accentuerait l'effet proposé par Snijders, soit que la différence d'origines est particulièrement valorisée (par rapport à une variable plus

anodine comme le lieu d'habitation sur un campus). De fait ces deux interprétations ne sont pas incompatibles.

(12) Alors qu'au départ p^* n'avait pas été conçu pour l'analyse longitudinale des réseaux, il existe maintenant une version dynamique de ce modèle (Robins et Pattison 2001). La principale différence entre SIENA et la version dynamique de p^* est que la première applique l'idée qu'il y a des petits changements consécutifs des liens qui ne sont pas observés entre les moments de description du réseau à $t0$ et $t1$, tandis que le p^* dynamique suppose un changement « dans un pas » comme si tout changeait soudainement en même temps de $t0$ à $t1$. Robins et Pattison sont d'accord pour dire que SIENA est un approche supérieure à p^* .

(13) Rappelons la distinction méthodologique entre les réseaux totaux, où un ensemble d'acteurs limité est défini et toutes les relations possibles entre tous les acteurs sont examinées, et les réseaux personnels, où l'on choisit un échantillon d'acteurs dans une population, selon les méthodes standard des sondages, et on interroge ces acteurs sur l'étoile de relations qu'ils entretiennent et, éventuellement, les relations de leurs associés.

(14) Snijders, Steglich et d'autres travaillent pour incorporer d'autres sortes de triplets et d'autres effets triadiques. Cependant il faut tenir compte du fait que pour incorporer des triplets dans un modèle dynamique il faut avoir des hypothèses théoriques sur les transitions qui ont sens entre des types de triplet différents.

REFERENCES BIBLIOGRAPHIQUES

Allan, G. H. (1979) *A sociology of friendship and kinship*. London, George Allen and Unwin.

Allan, G. H. (1989) *Friendship: Developing a social perspective*. London, Harvester Neatsheaf.

Allport, G.W. (1954). *The nature of prejudice*. Addison-Wesley. Reading.

Bidart, C. (1991) « L'amitié, les amis, leur histoire. » *Sociétés contemporaines*, N° 5.

Bidart, C. (1993) *Les semblables, les amis et les autres: sociabilité et amitié*, Thèse de doctorat, Marseille EHESS, 1993.

Bloch F. et Buisson M. (1991) « Du don à la dette : la construction du lien social familial. » *Revue du MAUSS*, 11 : 54-71.

Boer, P., Huisman, M., Snijders, T.A.B., et Zeggelink, E.P.H. (2001). *StOCNET: an open software system for the advanced statistical analysis of social networks*. Groningen: ProGAMMA / ICS. <http://stat.gamma.rug.nl>

Boudon R. (1968) *A quoi sert la notion de "structure" ?*, Gallimard, Paris.

Boudon R. (1984) *La place du désordre : critique des théories du changement social*, Sociologies, PUF, Paris.

Bunt G. van de (1999) *Friends by choice* Amsterdam Thesis Publishers ICS series.

Burt R.S. (1992) *Structural holes*. Cambridge, Harvard University Press.

Cucó i Giner, J. (1995) *La amistad. Perspectiva antropológica*. Icaria. Institut Català d'Antropologia.

Degenne, A. et Forsé, M. (1994) *Les réseaux sociaux: une analyse structurale en sociologie*. Paris, A. Colin.

Duijn, M.A.J. van (1995). « Estimation of a random effects model for directed graphs ». En T.A.B. Snijders (Ed.) SSS'95 Symposium Statistische Software, nr. 7. Toeval zit overal: programmatuur voor random coefficient modellen (113-131). Groningen: ProGAMMA.

Duijn, M.A.J. van, Snijders, T.A.B., & Zijlstra, B.H. *p2: a random effects model with covariates for directed graphs*. *Statistica Neerlandica*, 58 (2004), 234-254.

Duijn, M.A.J. van; Zeggelink, E.; Huisman, M.; Stokman, F. N.; Wasseur, F.W. (2003) « Evolution of sociology freshmen into a friendship network » *Journal of Mathematical Sociology* 27 :153-191.

Everett, M. « Ego network betweenness » Sunbelt XIX International Conference on Social Networks, Charleston, USA, 18-21 février 1999.

Federico, A. de (1999). « Requena, F. Redes sociales y cuestionarios, Madrid, Centro de Investigaciones Sociológicas, Siglo XXI, Cuadernos metodológicos, 1996 » *PAPERS Revista de Sociología* N°58, p. 191-201.

Federico, A. de (2002) « Introduction à SIENA pour l'analyse longitudinale de réseaux complets » présenté dans le séminaire GARES, CLERSÉ, Université des Sciences et Technologies, Lille (France) 28 de février de 2002. Disponible à : <http://wwsympa.univ-lille1.fr/wws/arc/gares/2002-02/msg00003.html> sous le nom « Intro_SIENA.rtf »

Federico, A. de (2003a) *Réseaux d'identification à l'Europe. Amitiés et identités d'étudiants Européens*. Thèse en co-tutelle soutenue sous la direction d'Alexis

Ferrand (CLERSÉ, Université des Sciences et Technologies de Lille) et Mercedes Pardo (Universidad Pública de Navarra, Espagne).

Federico, A. de (2003b) « La dinámica de las redes de amistad. La elección de amigos en el programa Erasmus » *REDES. Revista hispana para el análisis de redes sociales*. Vol. 4, 2003.

Federico, A. de (2004) « Redes de amistad e identidades de estudiantes Europeos. Una articulación analítica macro-meso-micro de las identificaciones territoriales » à paraître dans *El analisis de redes sociales en la transdisciplina*.

Feld, S. L. (1997) « Structural embeddedness and stability of interpersonal relations ». *Social networks*, 19(1) : 91-95.

Ferrand, A. (1993) *L'Analyse des réseaux personnels*. Habilitation à diriger des recherches en sciences sociales et humaines auprès de l'Université de Lille.

Ferrand, A. (1997). « La structure des systèmes de relations » *L'Année Sociologique* 1997, 47(1) :37-54.

Ferrand, A. (2002) « Las comunidades locales como estructuras meso » *REDES. Revista Hispana para el análisis de redes sociales*. 3(4).

Ferrand, A. et Snijders, T. (1997) « Social networks and normative tensions ». In Van Campenhout L., Cohen M., Guizzardi G., Hauser D, (eds) *Sexual interactions and HIV risk : new conceptual perspectives in European research*, London, Taylor & Francis, 6-21.

Ferrand, A., Mounier, L., Degenne, A. (1999). « The diversity of personal networks in France ; social stratification and relational structures ». In Wellman B. (ed.) : *Networks in the global village* : 185-224. Westview Press. Boulder.

Fischer C.S. (1982a) *To dwell among friends : personal network in town and city*. Chicago, University Press.

Fischer C.S. (1982b) « What do we mean by friends ? An inductive study » *Social Networks*, n°3.

Flap, H. et Volker, B. (2002) « Occupational community and solidarity at work » presentado en la Second European Thematic Conference for Network Analysts. "Lilnet. Micro-macro relations : advances in the contribution of structural analysis", Lille, 30-31 mayo 2002.

Frank, O. (1991) « Statistical análisis of change in networks. » *Statistica Neerlandica*, 45, 283-293.

Frank, O. et Strauss, D. (1986) « Markov graphs. » *Journal of the American Statistical Association*, 81, 832-842.

Godbout J. et Charbonneau J. (1993) « La dette positive dans le lien familial ». *Revue du MAUSS*, « Ce que donner veut dire », 235-256.

Granovetter, M. S. (1973) « The strength of week ties » *American Journal of Sociology*, 81 : 1287-1303.

Hallinan, M. T. (1979) « The process of friendship formation », *Social Networks*, 1979, 1 : 193-210.

Hiramatsu, H. (1999) « Longitudinal data analysis on friendship network formation of Japanese students » Sunbelt XIX International Conference on Social Networks, Charleston, 18-21 febrero 1999.

Holland, P. W. et Leinhardt, S. (1976) « Local structures in social networks » En: D. Heise (ed.), *Sociological methodology*. San Francisco: Jossey Bass.

Holland, P. W. et Leinhardt, S. (1981) « An exponential family of probability distributions for directed graphs (with discussion) » *Journal of the American Statistical Association*, 1981, 76, 33-65.

Huisman, Mark, et Snijders, Tom A.B. (2003). « Statistical analysis of longitudinal network data with changing composition. » *Sociological Methods & Research*, 32 (2003), 253-287.

Kurth, S.B. (1970) « Friendships and friendly relations », in G.J. Mc Call, M.M. Mc Call, N. K. Denzin, G. D. Suttles, et S.B. Kurth (eds.) *Social Relationships*. Chicago, Aldine.

Lazarsfeld, P. et Merton, R. (1954) « Friendship as a social process », in Berger et al *Freedom and control in modern society*. Princeton, Van Nostrand.

Lazega, E. (2001) *The Collegial Phenomenon : The Social Mechanisms of Cooperation Among Peers in a Corporate Law Partnership*, Oxford, Oxford University Press.

Lazega, E. et van Duijn, M. (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 :375-397.

Leenders, R. (1995) *Structure and influence* Amsterdam, Thesis Publisher.

Lorrain, F. et White, H. (1971) « Structural equivalence of individuals in social networks », *Journal of Mathematical Sociology*, 1 : 49-80.

Mitchell, J. C., (ed.) (1969) *Social Networks in Urban Situations*, Manchester, Manchester University Press.

Paine R. (1969) « In search of friendship ». *Man*, vol. 4 p. 505-524.

Piaget, J. (1968) *Le structuralisme* PUF Que sais je?.

Pizarro, N. (2004) « Un nuevo enfoque sobre la equivalencia estructural : lugares y redes de lugares como herramientas para la teoría sociológica », *REDES. Revista hispana para el análisis de redes sociales*, 5 (2).

Robbins, H. et Monro, S. (1951) « A Stochastic Approximation Method. » *Ann. Math. Stat.* 22:400-407.

Robins, G. L., et Pattison, P. E. (2001). « Random graph models for temporal processes in social networks ». *Journal of Mathematical Sociology*, 25, 5[^]41.

Snijders, T.A.B. (1996) « Stochastic Actor-Oriented Models for Network Change » *Journal of Mathematical Sociology*, 21 (1-2) : 149-172.

Snijders, T.A.B. (1998), Methodological issues in studying effects of networks in organizations. *Computational and Mathematical Organization Theory*, 4, 205-215.

Snijders, T.A.B. (2001) « The statistical Evaluation of Social Network Dynamics » *Sociological Methodology*.

<http://statt.gamma.rug.nl/snijders/Siena.html>

Snijders, T.A.B. (2002) « Markov Chain Monte Carlo estimation of exponential random graph models » *Journal of Social Structure*.
<http://www.cmu.edu/joss/index.html>

Snijders, T.A.B. (2004). « Models for Longitudinal Network Data. » À paraître dans P. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis*. New York: Cambridge University Press.

Snijders, T.A.B. Duijn, M. van (1997) « Simulation for statistical inference in dynamic network models » In Conte, R. Hegselmann, R. Terna, P. (Eds.) *Simulating social phenomena* Berlin Springer. p. 493-512.

Sprien, M. (1999) *Sampling Personal Network Structures: Statistical Inference in Ego-Graphs*. Thela Thesis. ICS Dissertation Series. University of Groningen.

Wasserman, S. (1980) « Analyzing social networks as stochastic processes. » *Journal of the American Statistical Association* 75, p.280-294.

Wasserman, S. Faust K. (1994) *Social Network Analysis: Methods and Applications* Cambridge, Cambridge University Press.

Wasserman, S. et Iacobucci (1988) « Sequential social network data. » *Psychometrika* 53 : 261-282.

Wasserman, S. et Pattison, P. (1996) « Logit models and logistic regression for social networks : I. An introduction to Markov graphs and p^* . » *Psychometrika*, 61, 401-425.

Wellman, B. Carrington, P. J. et Hall, A. (1988) « Network as personal communities » en Wellman, B. Berkowitz *Social Structures: A Network Approach*, Cambridge, University Press, p 130-84.

Wellman, B. Wortley, S. (1990) « Different Strokes from different folks: Community ties and social support: Which ties provide what kinds of social support » *American Journal of sociology*, Novembre, 96 (3) : 558-88.

Zeggelink, E. (1993) *Strangers into friends. The evolution of friendship networks*. Amsterdam, Thesis Publisher.

Zeggelink, E. P. H. ; van Duijn, M. ; Hiramatsu, H. ; Stockman, F. N. ; van Oosten, R. ; Wasseur, F. (1999) « Determinants of friendship formation : a comparative analysis of friendship formation among freshmen in Japan and the Netherlands » Sunbelt XIX International Conference on Social Networks, Charleston, 18-21 febrero 1999.

Zeggelink, E. P. H. ; Hiramatsu, H. ; Stockman, F. N. ; van Duijn, M. ; Wasseur, F. (1997) « Friendship formation over time among sociology students : a comparison between Japan and the Netherlands » XVIII Conferencia Internacional y V Conferencia Europea de Análisis de Redes Sociales. Sitges 27-31 Mayo 1997.

Tableau 1. Caractéristiques du réseau d'amitié à t_0 et t_1 .

	t_0	t_1
Nombre moyen	0,79	11,68
Étendue du nombre d'entrée	0-7	2-27
Écart type du nombre d'entrée	1,73	5,56
Étendue du nombre de sortie	0-7	0-35
Écart type du nombre de sortie	1,67	7,61
Densité	0,020	0,148
Réciprocité	0,992	0,896
Transitivité	0,00071	0,02438

Tableau 2. Modèle 1. Effets structureaux.

Effet	Type	Paramètre	Écart type	t
Vitesse	Vitesse	10,884	0,331	-----
Densité	Structural	-1,40	0,543	-2,59
Réciprocité	Structural	2,769	0,560	4,94
Transitivité	Structural	1,068	0,215	4,96

Tableau 3. Modèle 2. Effets structureaux et dyadiques.

Effet	Type	Paramètre	Écart type	t
Vitesse	Vitesse	10,939	0,316	-----
Densité	Structural	-1,541	0,603	-2,55
Réciprocité	Structural	2,280	0,468	4,87
Transitivité	Structural	0,984	0,168	5,86
Sim. Sexe	Dyadique	-0,093	0,171	-0,54
Sim. Nationalité	Dyadique	1,192	0,183	6,51
Sim. Résidence	Dyadique	0,793	0,104	7,62

Tableau 4. Modèle 3. Effets structureaux, dyadiques et liées aux acteurs.

Effet	Type	Paramètre	Écart type	t
Vitesse	Vitesse	12,535	0,323	-----
Densité	Structural	-1,392	0,578	-2,41

Réciprocité	Structural	2,416	0,401	6,02
Transitivité	Structural	0,825	0,182	4,53
Sim. Nationalité	Dyadique	1,588	0,253	6,27
Sim. Résidence	Dyadique	0,699	0,099	7,06
Activité Britanniques	Acteurs	-2,553	0,446	-5,72
Popularité Italiens	Acteurs	0,286	0,142	2,01
Réciprocité X Sim. nationalité	Interaction	-2,377	0,759	-3,13

Tableau 5. Modèle 4. Effets structuraux, dyadiques, liées aux acteurs, vitesse et gratification.

Effet	Type	Paramètre	Écart type	<i>t</i>
Vitesse	Vitesse	12,535	0,323	-----
Densité	Structural	-1,452	0,553	-2,62
Réciprocité	Structural	2,416	0,401	6,02
Transitivité	Structural	0,825	0,182	4,53
Sim. Nationalité	Dyadique	1,588	0,253	6,27
Sim. Résidence	Dyadique	0,699	0,099	7,06
Activité Britanniques	Acteurs	-2,553	0,446	-5,72
Popularité Italiens	Acteurs	0,286	0,142	2,01
Réciprocité X Sim. nationalité	Interaction	-2,377	0,759	-3,13
Vitesse selon sexe	Vitesse	-0,148	0,086	-1,72
Gratif. selon réciprocité	Gratification	0,001	0,128	0,01

Tableau 6. Modèles statistiques pour les réseaux sociaux totaux.

Modèle	Effets	Dyades	Triades	Dynamique
<i>p1</i>	Endogènes	Oui	Non	Non
<i>p2</i>	Endogènes et exogènes	Oui	Non	Non
<i>p*</i>	Endogènes et exogènes	Oui	Tous les triplets.	Oui
SIENA	Endogènes et exogènes	Oui	Effets de fermeture(14)	Oui

Figure 1. Classes isomorphiques des triades selon Holland et Leinhard

Para ver esta película, debe
disponer de QuickTime™ y de
un descompresor Photo - JPEG.

Figure 2. Triplets transitifs et intransitifs

Triplet (ijk) en état transitif

Triplet (ikj) en état transitif

Para ver esta película, debe
disponer de QuickTime™ y de
un descompresor Photo - JPEG.

Triplet (ijk) en état intransitif

Triplet (ikj) en état intransitif

Figure 3. Réseau à t_0

Para ver esta película, debe
disponer de QuickTime™ y de
un descompresor Photo - JPEG.

Figure 4. Réseau à t_1

Para ver esta película, debe
disponer de QuickTime™ y de
un descompresor Photo - JPEG.

Légende pour tous les graphes :

Rouge :	Espagnols	Bleu foncé :	Grecs	Vert :	Allemands
Noir :	Britanniques	Bleu clair :	Suédois	Violet :	Belges
Orange :	Italiens	Vert clair :	Autrichiens	Jaune :	Portugais

Flèches noires : Relations orientées.

Lignes bleues : Relations mutuelles.
